Deep Learning for Automated Defect Detection in Sewer Systems Using Adaptive Image Pre-Processing

Doctor of Philosophy (PhD)

Wasiu Yusuf

This thesis is submitted in partial fulfilment of the requirements of the University of Hertfordshire for the degree of Doctor of Philosophy

Supervisors:

Professor Hafiz Alaka
Professor Neil Spencer

August 2025

Declaration

I declare that the work presented in this thesis is my own and has not been submitted, in whole or in part, for any degree at this or any other institution. All sources of information have been duly acknowledged in the text and listed in the references.

Signed: _____

Date: _____

Dedication

All praise is due to Allah, Lord of the worlds.

This work is dedicated to Allah, whose guidance and mercy made every step possible, and to my beloved family, whose unwavering support, patience, and prayers have been a great source of strength.

Acknowledgement

First and foremost, I thank Almighty Allah for His grace, guidance, and countless blessings throughout this journey. Without His will and mercy, the successful completion of this work would not have been possible.

I am deeply grateful to my principal supervisor, Professor Hafiz Alaka (Associate Dean, Research and Enterprise at the Hertfordshire Business School) for his invaluable guidance, encouragement, and expertise throughout this research. Likewise, I extend my sincere thanks to my second supervisor, Professor Neil Spencer, for his consistent feedback and insightful advice. Together, your constructive input and dedication to my progress have shaped the direction, depth, and quality of this work.

I am profoundly thankful to my family. To my parents, for instilling in me the values of hard work, integrity, and perseverance; to my wife, for her unending patience, support, and understanding during this journey; and to my children, for their joy and inspiration. Your sacrifices and support have been the foundation upon which this achievement rests. To my brothers and sister, thank you for your constant check-ins, encouragement, and prayers. You are all a blessing to me.

My sincere appreciation goes to UoH, especially Hertfordshire Business School for providing a supportive academic environment, and to the University's RDP and the Doctoral College team for equipping me with the skills, resources, and training essential for my doctoral journey. I also extend heartfelt thanks to my friends and colleagues at the Big Data Technologies and Innovation Lab and the wider Herts community, whose collaboration and encouragement have made the challenging moments lighter and the successes more rewarding.

Finally, I acknowledge the broader research community and our industry partners, whose cooperation and data made this study possible. Your openness to collaboration and commitment to innovation in sewer infrastructure have ensured that this work remains grounded in both academic rigour and practical relevance.

To all future readers of this thesis, I hope these pages not only share knowledge but also inspire you to build upon it, contribute to the advancement of this field, and continue the pursuit of solutions that create safer, more resilient, and sustainable infrastructure.

Abstract

Sewer systems are critical urban infrastructure for safeguarding public health, maintaining environmental integrity, and ensuring sustainable water management. However, their continuous operation is increasingly threatened by ageing assets, rapid urbanisation, and operational constraints that limit inspection frequency. Manual inspection methods including direct man-entry into sewer lines and prolonged review of CCTV survey footage are labour-intensive, costly, inconsistent, and unsuitable for timely large-scale monitoring. Recent advances in computer vision and deep learning offer opportunities to address these limitations, yet most research models remain confined to controlled environments, with little attention to reproducibility, defect-specific pre-processing, or deployment in operational settings. This research responds to these gaps by designing, validating, and deploying an automated deep learning defect detection tailored to real-world sewer inspection workflows.

A positivist, quantitative methodology guided the study, incorporating two systematic literature reviews to analyse prevailing defect detection approaches and identify key challenges in sewer condition assessment. Historical CCTV survey videos and inspection reports were obtained from multiple Sewer Management Companies (SMCs) in the UK, encompassing 16 defect categories across diverse urban environments. A defect-specific pre-processing strategy was developed, applying tailored enhancement techniques such as adaptive histogram equalisation, contrast stretching, and morphological operations, according to the visual and structural characteristics of each defect type. Cumulatively, 25 different defect-specific pre-processing operation was implemented and applied to the research data. This targeted approach aimed to mitigate issues of low contrast, illumination variation, and feature ambiguity, consequently improving data quality and supporting model robustness.

Multiple deep learning architectures, including YOLOv5 through YOLOv11, were comparatively evaluated under consistent hyperparameters, dataset splits, and hardware conditions. Metrics such as mean Average Precision (mAP), FLOPs, number of deep network layers, and latency were used to identify the most effective architecture for multi-class sewer defect detection. The YOLOv9 model was selected for its balance of detection accuracy, computational efficiency, and scalability, which achieves notable performance gains over both earlier YOLO versions and approaches documented in prior studies. Specifically, the optimised YOLOv9 achieved a mAP of 93%, underscoring its robustness and adaptability for diverse field conditions. Validation on unseen municipal datasets confirmed strong generalisation capability, supported by class-balanced augmentation techniques.

To ensure real-world applicability, the optimised YOLOv9 model was integrated into a Dockerised FastAPI service deployed on GPU-enabled AWS infrastructure. The system processes both uploaded survey videos and real-time image streams, delivering detection results and detailed defect reports through a user-friendly interface. While deployment results demonstrated operational readiness, the study also identified limitations including small dataset

size for defects and performance variation on texture-rich classes. Overall, the research advances the technical and practical readiness of AI-powered sewer defect detection, bridging the gap between academic prototypes and scalable operational tools.

Table of Contents

Declaration i

Dedication i

Acknowledgement ii

Abstract iii

Table of Contents v

List of Figures xviii

List of Tables xix

List of Abbreviations xix

List of Publications xxi

Published Journal Paper xxi

Published Conference Paper xxi

1.0 Introduction 1

1.1 Research Background 1

1.2 Definition of Popular Terms 5

1.3 Research Problem, Gaps in Knowledge, and Justification 6

1.4 Research Contributions 9

1.5 Research Aim and Objectives 10

1.6 Research Questions 11

1.7 Overview of Deep Learning 12

1.7.1 Image Classification 12

1.7.2 Object Localization 14

1.7.3 Object detection 15

1.8 Methodology Overview 16

1.8.1 Systematic Literature Review 16

1.8.2 Data Collection and Pre-processing 16

1.8.3 Model Development 17

1.8.4 Model Evaluation and Comparison 17

1.8.5 Implementation and Deployment 18

1.9 Research Scope 18

1.10 Thesis Structure 18

1.11 Chapter Summary 20

2.0 Automated Defect Detection in Sewer....21
2.1 Brief History and Overview....21
2.1.1 Stages of AuDD....22
2.2 Defects in Sewer Pipelines....26
2.2.1 Categories of Defect....26
2.2.1.1 Structural Defects....28
2.2.1.2 Operational Defects....30
2.2.1.3 Construction Defects....31
2.2.1.4 Miscellaneous Defects....31
2.2.2 Impact of Defects on Public Health and SMCs....32
2.3 Evolution of Sewer Inspection Techniques....33
2.4 Gaps in Current Defect Detection System....38
2.5 Existing Approaches in Developing AuDD....39
2.5.1 SLR Methodology....39
2.5.1.1 Identifying the literatures....42
2.5.1.2 Parameterized Screening....43
2.5.1.3 Eligibility and Inclusion....44
2.5.2 SLR Analysis and Result....44
2.5.2.1 Publication Trends Analysis....44
2.5.2.2 Citation Impact Analysis....46
2.5.2.3 Keyword Analysis of Selected Studies....47
2.5.2.4 SLR-RQ1: Recent approaches used in AuDD....49
2.5.2.5 SLR-RQ2: What type of defects were considered?....51
2.5.2.6 Detection Complexity, Methodological Alignment, and Comparative Performance Analysis...53
2.6 Selection of Model Development Method and Defect Types....55
2.7 Chapter Summary....56
3.0 Systematic Literature Review of Deep Learning Algorithms for Automatic Defect Detection in Sewers....58
3.1 Chapter Overview....58
3.2 Review Methodology....58
3.2.1 Identification....60
3.2.2 Screening....61
3.3 Inclusion....61
3.3 Overview of Deep Learning Applications in Sewer Defect Detection....62

3.4 Deep Learning Algorithms for Defect Detection in Sewers....63
3.4.1 Defect classification....63
3.4.2 Defect detection....67
3.4.3 Defect Segmentation....68
3.5 Analysis of Review....72
3.5.1 Keyword analysis....80
3.5.2 Publication Year....81
3.5.3 Data source....81
3.5.4 Problem type....82
3.5.5 Defect type....84
3.5.6 Pre-processing techniques....86
3.5.7 Annotation....89
3.5.8 Algorithm and the Present State of Knowledge in DL for AuDD....90
3.5.8.1 Efficiency....92
3.5.8.2 Speed....92
3.4.8.3 Robustness in Real-World Contexts....93
3.5.9 Performance Metrics....94
3.6 Enablers and Challenges of the AuDD Process....97
3.7 Conclusions from SLR....97
4.0 Research Methodology....99
4.1 Chapter Overview....99
4.2 Research Paradigm....99
4.2.1 Ontological Assumptions....101
4.2.2 Epistemological Assumptions....103
4.2.3 Research Design....107
4.3 Research Methods....107
4.4 Research Framework....108
4.5 Data Collection....110
4.5.1 Research Data....112
4.5.2 Data Collection Tools and Quality Assurance....115
4.5.3 Data Storage and Management....117
4.5.4 Defect Selection....117
4.6 Chapter Summary....119
5.0 Data Analysis and Preprocessing....121

5.1 Chapter Overview **121**
5.2 Need for Data Preprocessing **121**
5.3 Image extraction and analysis **122**
5.3.1 Duplicate removal 124
5.4 Data Augmentation **128**
5.4.1 Image-Level and Bounding-Box Augmentation 128
5.5 Image Annotation **131**
5.6 Grayscale and Contrast Enhancement **133**
5.7 Morphological Operations **135**
5.8 Image Segmentation **140**
5.8.1 Thresholding 140
5.8.2 Clustering 143
5.8.3 Edge Detection 145
5.9 Chapter Summary **147**
6.0 Development of Deep Learning Models for AuDDS **148**
6.1 Chapter Overview **148**
6.1 Experimental Setup for AuDDS **148**
6.1.1 Hardware Configuration 149
6.1.2 Software Configuration 150
6.2 Model Architecture Overview **151**
6.3 Model Training Procedure **154**
6.3.1 Data Preparation and Preprocessing 154
6.3.2 Model Training and Optimization. 157
6.3.3 Hyperparameter Tuning 159
6.4 Model Evaluation Metrics **161**
6.4.1 Selected Evaluation Metrics 161
6.4.1.1 Precision 161
6.4.1.2 Recall 162
6.4.1.3 F1-Score 162
6.4.1.4 mAP (Mean Average Precision) 163
6.5 Chapter Summary **164**
7.0 Result and Discussion **164**
7.1 Chapter Overview **164**
7.2 Performance of Different YOLO Models **164**

7.3 Improving AuDD through Contrast Enhancement 171
7.4 Effectiveness of Segmentation Techniques 174
7.5 Enhancing AuDD via Morphological Operations 181
7.6 Overall Comparison and Selection of the Best Models 188
7.7 Chapter Summary 193
8.0 Deployment of Deep Learning Model 194
8.1 Chapter Overview 194
8.2 Deployment Pipeline for Deep Learning Model 194
8.3 Containerization with Docker 197
8.4 API Development with FastAPI 199
8.5 Model Deployment on AWS 202
8.7 Features, Testing, and Evaluation of the Deployed Model 205
8.7.1 Feature Walkthrough: Survey Processing and Report Generation 205
8.7.2 Testing and Evaluation 212
8.9 Chapter Summary 215
9.0 Conclusion, Reflection, and Future Work 216
9.1 Chapter Overview 216
9.2 Summary of Key Contributions 216
9.2.1 Defect-Specific Pre-processing Strategies for Image Enhancement 216
9.2.2 Comparative Evaluation of YOLO-Based Models for Optimal Performance 217
9.2.3 Selection and Deployment of YOLOv9 for Real-Time Sewer Defect Detection 218
9.3 Addressing the Research Objectives 219
Objective 1: To review advancements in automatic defect detection, including defect characteristics and model development approaches. 219
Objective 2: To pre-process sewer survey videos using conventional and custom defect-specific techniques. 221
Objective 3: To develop and compare deep learning models using processed data for defect detection. 222
Objective 4: To deploy the optimised model as SaaS for research and practical use 223
9.4 Observed Limitations of the Study 225
9.5 Implications for Practice 225
9.6 Recommendations for Future Research 226
9.7 Chapter Summary 228
References 229

List of Figures

Figure 1.1: An example man-entry visual inspection being conducted (Environmental Waste Management, 2024) 3
Figure 1.2: Alignment of Research Objectives and Questions with the Research Workflow (Source: Author) 12
Figure 1.3: How computers understand digital images (Source: Author) 13
Figure 2.1: Generic AuDD framework with seven dependent phases (Source: Author)..... 23
Figure 2.2: PACP and MSCC defect categorization. Adapted from Guo et al. (2009) and Zayed & al Hawari (2017) 27
Figure 2.3: Sample structural defects like crack, deformation, and surface damage (Source: Research Data) 29
Figure 2.4: Sample operational defect such as deposit, infiltration, and obstacle. (Source: Research Data) 30
Figure 2.5: Taxonomies of inspection techniques. (Source: Author 36
Figure 2.6: Operational workflow of SewerBatt (Faram, 2014)..... 37
Figure 2.6: Proposed methodology for the SLR 41
Figure 2.7: Annual publication distribution of selected sewer defect detection studies..... 45
Figure 2.8: Annual citation trends for selected studies..... 47
Figure 2.9: Keywords and title analysis of selected Studies..... 48
Figure 2.10: Exploratory data analysis of the methodologies used in the primary studies 49
Figure 2.11: High-level overview of model development approaches. 50
Figure 2.12: Sample sewer defects 52
Figure 2.13: Analysis of defect considered in previous studies..... 53
Figure 3.1: Review methodology..... 59
Figure 3.2: Convolution layer in CNN. Adapted from (Kamil, 2023) 64
Figure 3.3: Max pooling of feature maps (source: Author) 65
Figure 3.4: Flattening pooled feature maps (source: Author)..... 65
Figure 3.5: Fully connected layer (source: Author)..... 66
Figure 3.6: Defect localization with bounding boxes (source: Research Data)..... 68
Figure 3.7: Object detection vs Segmentation (Xu et al., 2021d)..... 70
Figure 3.8: Mask R-CNN segmentation architecture using PAFPN (Xu et al., 2021a) 71
Figure 3.9: Word cloud for keyword ranking (source: Author)..... 80
Figure 3.10: Publication trend in selected primary studies (Source: Author)..... 81
Figure 3.12: Distribution of problem types (source: Author) 83
Figure 3.12: Defect type ranking from the literature (source: Author)..... 85
Figure 3.13: Preprocessing trends from literature (source: Author) 86
Figure 3.14: Distribution of annotation tools used in previous studies (source: Author)..... 90
Figure 3.15: Analysis of Algorithm 91
Figure 4.2: Robotic crawler with mounted CCTV camera 116

Figure 5.6: Multi-label image annotation using MakeSense (Piotr, 2023).....132
Figure 5.7: Different pre-processing techniques considered (source: Author).....134
Figure 5.8: Sample applications of eroded, dilated (source: Author).....136
Figure 7.1: Variations of YOLO results on two groups of dataset. 166
Figure 7.2: Operational performance of YOLO algorithms. 167
Figure 7.3: Pearson correlation heatmap for each evaluation metrics 169
Figure 7.4: Spearman Rank correlation heatmap for all metrics 170
Figure 7.3: Correlation analysis using heat map.....173
Figure 7.5: Image segmentation model performance. 179
Figure 7.6: Correlation analysis of image segmentation result.....180
Figure 7.7: Model performance for each defect class using morphological operations 183
Figure 7.8: Correlation analysis for morphological operations 184
Figure 7.9: Radar chart showing comparative performance of each operation 186
Figure 7.10: Line chart showing mAP^{50} trends across operations for each defect class 187
Figure 7.11: Contrast enhancement top model performance for each class 188
Figure 7.12: Image segmentation of top model performance for each class 189
Fig 7.13: Morphological operation of top model performance for each class.....190
Figure 7.14: How each pre-processing techniques performed for each defect.....192
Figure 7.15: Category-specific mAP50 comparison across defect categories.....193
Figure 8.1: High-Level Architecture of the Deployment Pipeline. 196
Figure 8.2: Example Docker file for YOLO Model Deployment.....198
Figure 8.3: FastAPI structure for model inference. 201
Figure 8.4: AWS Deployment Architecture. 204
Figure 8.5: The dashboard overview page displaying key metrics and quick access control205
Figure 8.6: Projects Overview page showing project metadata and summary statistics. 207
Figure 8.7: Project Detail view with surveys list and upload panel visible.....207
Figure 8.8: Survey Progress Modal showing different lifecycle stages.....208
Figure 8.9: Survey upload completed 209
Figure 8.10: Report Detail page defect table download buttons.....211
Figure 8.11: Sample Locust chart showing deployment performance via load testing 213

List of Tables

Table 2.1: Severity effect on pipe structure 25
Table 2.2: Research questions drafted to guide the SLR 42
Table 2.3: Screening parameters for inclusion/exclusion criteria definition. 44
Table 2.1: Average F1-score for selected defects using different techniques. 54
Table 2.1: Research questions 60
Table 3.2: screening parameters for inclusion/exclusion criteria definition 61
Table 3.2: Summary of primary studies 73
Table 3.3: Summary of different pre-processing techniques 87
Table 3.4: Different evaluation techniques in deep learning 95
Table 4.1: Research Paradigms, Methods, Ontological and Epistemological Assumptions 105
Table 4.2: Collected videos based on pipe materials 114
Table 4.3: List of all defects from both datasets 118
Table 5.6: Sample video conversion to images 123
Table 5.7: Categories of extracted images 123
Table 5.8: Final raw dataset after duplicate removal 127
Table 6.9: Comparison of YOLO Variants. 152
Table 6.10: Task coverage for different algorithms. 153
Table 6.11: Comparison of different YOLO algorithms using the COCO dataset. 154
Table 6.4: Model hyperparameters used for fine-tuning 160
Table 6.5: Summary of evaluation metrics 164
Table 7.1: Performance comparison across contrast enhancement techniques 172
Table 7.2: Summary of image segmentation results by defect class and techniques 174
Table 7.3: Summary of morphological operation results by defect class 181
Table 7.4: Comparative mAP^{50} Scores for Top 3 Best Techniques in each defect category 191
Table 8.1: Comparative Analysis of Cloud-Based versus Local Deployment 215
Table 9.1: Summary of key activities and outcomes tailored towards achieving the research objectives. 224

List of Abbreviations

AE – Adaptive Equalization
AI – Artificial Intelligence
API – Application Programming Interface
AWS – Amazon Web Services

BCE – Binary Cross Entropy
BI – Business Intelligence

CI/CD – Continuous Integration / Continuous Deployment
CNN – Convolutional Neural Network
CS – Contrast Stretching
CUDA – Compute Unified Device Architecture
CSV – Comma-Separated Values
CSPDarknet – Cross Stage Partial Darknet

DL – Deep Learning

E-ELAN – Extended Efficient Layer Aggregation Network
ELB – Elastic Load Balancing

FN – False Negative
FP – False Positive

GAN – Generative Adversarial Network
GDPR – General Data Protection Regulation
GELAN – Generalized Efficient Layer Aggregation Network
GPU – Graphics Processing Unit

HE – Histogram Equalization
HTTP – Hypertext Transfer Protocol

IoU – Intersection over Union

JWT – JSON Web Token

ML – Machine Learning
mAP – Mean Average Precision
mAP50 – Mean Average Precision at IoU threshold 0.50
mAP50–95 – Mean Average Precision averaged over IoU thresholds 0.50–0.95

NDA – Non-Disclosure Agreement
NMS – Non-Max Suppression
NMS-free – Non-Max Suppression-free training

OpenCV – Open-Source Computer Vision Library

PANet – Path Aggregation Network
PDF – Portable Document Format
PGI – Programmable Gradient Information
PNG – Portable Network Graphics

PyTorch – Python-based Deep Learning Framework

RDP – Research Development Programme
ReLU – Rectified Linear Unit
RGB – Red, Green, Blue colour model
RPN – Region Proposal Network

SaaS – Software-as-a-Service
SGD – Stochastic Gradient Descent
SMC – Sewer Maintenance Company
SMOTE – Synthetic Minority Oversampling Technique
SPP – Spatial Pyramid Pooling
SSD – Single Shot MultiBox Detector
SSE – Server-Sent Events
SSET – Sewer Scanner Evaluation Technology
SOTA – State of the Art
SWISH – Swish Activation Function

TP – True Positive

UI – User Interface
UoH – University of Hertfordshire
URL – Uniform Resource Locator
UX – User Experience

W&B / wandb – Weights & Biases (Machine Learning experiment tracking)

YOLO – You Only Look Once

List of Publications

Published Journal Paper

1. Deep learning for automated encrustation detection in sewer inspection. Intelligent Systems with Application.
   https://doi.org/10.1016/j.iswa.2024.200433

Published Conference Paper

1. Machine Learning Recognition Models in Construction: A Systematic Review. Environmental Design and Management.
   https://uhra.herts.ac.uk/id/eprint/14419/

CHAPTER ONE

1.0 Introduction

1.1 Research Background

Modern urban life depends fundamentally on the continuous and reliable operation of sewer systems, a critical infrastructure that functions silently beneath our cities to protect public health, safeguard the environment, and support sustainable water resource management (Krishnan & Devarajan, 2025). These underground networks play an essential role in transporting and treating wastewater, preventing disease outbreaks, and preserving the ecological integrity of surface and groundwater systems (Li et al., 2022). However, as cities expand and infrastructures age, sewer systems face increasing strain from rising population densities, land use intensification, and climate-induced stressors.

Compounded by historically infrequent inspections, limited automation, and reactive maintenance strategies, many urban sewer networks are now operating beyond their intended service life, leaving them vulnerable to structural failure, blockages, and environmental contamination (Aşchilean et al., 2018a). Over time, these systemic challenges contribute to the gradual deterioration of sewer pipelines, giving rise to a wide array of defects such as cracks, deposits, root intrusion, and joint displacement that may ultimately lead to severe failures such as blockages, sinkholes, or full structural collapse. These failures not only disrupt essential services but also pose long-term risks to public health and environmental quality (Ghinea et al., 2024a).

This growing mismatch between infrastructure capability, system demands, and inspection capabilities underscores the urgent need for intelligent, scalable solutions to monitor and maintain sewer infrastructure more effectively (Xu et al., 2021a). While the core of most monitoring and maintenance is tied to defects, the prompt identification and remediation of such defects in sewer systems offer multiple significant benefits. Foremost, it substantially lowers the associated repair costs and extends the functional lifespan of the sewer infrastructure (Tscheikner-Gratl et al., 2019). Likewise, early intervention plays a crucial role in enhancing overall public health by preventing potential environmental and health hazards associated with malfunctioning sewer systems, such as water pollutions and Sanitary Sewer Overflows (SSO).

With defect identification, the traditional assessment method of physical man-entry into sewer systems remains one of the most time-consuming and labour-intensive approaches to sewer condition evaluation (Noshahri et al., 2021). In many major metropolitan areas, inspection of large trunk sewers, often stretching several miles, requires significant logistical planning before any evaluation can commence. For example, a 4-mile-long sewer segment cannot be assessed

immediately because it must first be isolated from the rest of the network to ensure safety. This includes disconnection from inflow, mechanical airing of the enclosed space, and the setup of an intercommunication system to maintain contact with personnel during the inspection process (Environmental Waste Management, 2023).

During the actual inspection, trained human operators physically enter the sewer system, often under low-light and high-humidity conditions, to perform visual assessments of the internal surfaces. As shown in Figure 1.1, inspectors typically wear safety gear, carry headlamps or torches, and slowly traverse the corridors of the pipeline, visually checking for any signs of defect within the sewer linings (Syazana et al., 2019). In urban networks, where pipe diameters may vary and access is constrained, this method requires back-and-forth movement through confined spaces, significantly limiting the speed and efficiency of data collection.

The process also demands extensive preparation to meet safety standards, particularly due to the presence of harmful gases such as hydrogen sulphide (H_2S) and methane. These gases are both toxic and combustible, sometimes necessitating the use of gas detectors (nditv, 2023), forced ventilation systems, and escape protocols. Specialised crews must also be on standby, and all personnel are required to undergo rigorous training in confined-space entry. These safety requirements, while essential, add further delay and cost to each inspection, increasing the burden on municipal sewer maintenance teams.

As a result, the man-entry method is not only operationally slow and physically demanding but also resource-intensive and inherently hazardous since it involves considerable human risk and financial overhead, making it impractical for routine or large-scale monitoring. These limitations prompted a shift towards a review-based and semi-automated approach (Chen et al., 2019), which aim to mitigate the challenges of traditional method, and make long-term condition assessment more feasible for ageing sewer infrastructure.



Figure 1.1: An example man-entry visual inspection being conducted (Environmental Waste Management, 2024)

A widely used conventional method for sewer condition assessment is review based, which involves field experts manually reviewing hours of recorded inspection footage to identify structural or operational defects. These videos are typically captured using inspection technologies such as Closed-Circuit Television (CCTV), wheeled crawlers, robots, or drones, each fitted with cameras and lighting systems suitable for low-visibility environments. While this approach removes the need for physical man-entry, it still relies heavily on human oversight. Operators use software tools like WinCan (WinCan, 2024) to control camera navigation and interpret visual cues remotely, assessing pipe condition and defect severity from the video feed. It is important to clarify that while the introduction of CCTV hardware represented a major technological advancement over hazardous, traditional man-entry inspections, its reliance on manual human review makes it insufficient for modern, large-scale data acquisition and rapid defect analysis.

However, this method of reviewing hour-long video is both time-consuming and prone to human error. Extended footage reviews can lead to inspector fatigue, which compromises attention to detail and increases the likelihood of missed or misclassified defects (Kumar & Abraham, 2019a; Savino & Tondolo, 2021). Moreover, subjective interpretation varies across inspectors, introducing inconsistency into condition ratings and repair prioritisation. The

manual nature of the process also leads to significant delays in translating observations into actionable maintenance tasks. These limitations underscore the growing need for automated defect detection systems that can standardise interpretation, accelerate analysis, and improve the accuracy and reliability of sewer infrastructure assessments.

Broadly, automating the detection of sewer defects involves two primary tasks. The first task is data acquisition, which leverages a range of advanced technologies such as CCTV, drones, robots, lasers, and crawlers. These tools capture high-resolution images, videos, and even 3D spatial data, providing a comprehensive view of the sewer system's condition. The second task is processing and analysing this acquired data to automatically detect potential defects. Advanced algorithms, machine learning models, and computer vision techniques are typically employed in this stage to identify anomalies such as cracks, corrosion, blockages, and other structural issues, thereby reducing reliance on manual inspection and significantly enhancing both the accuracy and efficiency of defect detection (Ma et al., 2022).

Both phases have seen significant advancements over the past decade, driven by rapid improvements in sensing and imaging equipment, as well as enhanced computing capabilities. These developments have paved the way for a range of automated defect detection techniques that deliver more accurate and efficient monitoring of sewer conditions (Koch et al., 2015; Wang et al., 2023). These techniques now span statistical analysis, conventional image processing, sophisticated vision-based approaches, and hybrid techniques that integrate multiple strategies. For a detailed exploration of these inspection methods, refer to chapter two.

Despite the advancements in sensing technologies and computing power, which have led to the development of various models and algorithms for capturing and detecting different types of defects, significant challenges remain in creating robust and reliable detection models for sewer systems. One major challenge is the diversity and complexity of defects in sewer pipes, which requires specialized algorithms and data pre-processing techniques (Pasadas et al., 2020). Additionally, variations in operating conditions, such as differences in lighting and varying levels of moisture and humidity inside the sewer pipes adversely affect both the quality of acquired data, which affect the performance of detection models.

To mitigate the challenges posed by the complexity of defects and low-quality data, this research emphasizes a selective, defect-based image preprocessing strategy. By tailoring preprocessing techniques to enhance the clarity of each defect, considering their unique characteristics and diversity, the goal is to improve the visibility and definition of defects in the acquired images. This refined data then feeds into the development of a multi-label, robust deep learning model that can adapt to various conditions and defect types, ultimately enhancing detection accuracy and reliability in real-world sewer inspections.

1.2 Definition of Popular Terms

In sewer management, there are a couple of terms commonly used to denote or describe events and activities. The common terms peculiar to this research are summarised below.

Sewer survey or sewer inspection: A sewer survey is a general visual inquiry targeted at understanding the current operation of the sewer while an inspection is a systematic evaluation of the status of the sewer to understand whether corrective or maintenance measures are due by using commonly agreed-upon standards. These two terms are mostly used interchangeably in sewer management but the main aim of both is to identify if there are defects (or faults) within the sewer, thereby informing decision-making and quick intervention.

Defect: In sewers, a defect is any form of internal anomaly that could hinder the normal operations of an active sewer. These defects are discovered during the cause of an inspection of a sewer, either physically or using modern technologies. PACP has categorised this defect into four major groups.

I. **Structural:** Defects that affect the structure of the sewer pipes such as crack, fracture, collapse, and deformation among others.
II. **Operation and maintenance** are external objects found in pipes that may obstruct the sewage system's regular operation. This includes defects such as deposits, root intrusion, infiltration, and encrustation among others.
III. **Construction** is defects and conditions associated with the methodology employed in constructing and connecting pipes. This includes pipe joint, intruding seal materials, access point, and so on.
IV. **Miscellaneous:** Any defect or features that are not part of the other categories. This is a generic group that has no fixed type of defect.

Chapter 2 further provides a comprehensive analysis of sewer defects, including their classification, representative examples within each category, and their broader implications for both the environment and water management companies.

Defect Detection: This is a systematic way of examining the internal structure of a sewer for possible defects (or faults) based on acceptable standards. It involves thoroughly analysing a sewer either physically (Li et al., 2019a) or exploiting technologies like as Closed-circuit television (CCTV) (Cheng & Wang, 2018a) to evaluate the condition of a sewer and use standard coding mechanism (e.g., PACP coding standard) to denote the status of the different part of the sewer. The major aim of defect detection is to understand whether corrective or maintenance measures need to be taken to pre-empt any possible and avoidable blockage or collapse.

Artificial Intelligence (AI) vs. Machine Learning (ML): To properly contextualise the technological scope of this research, it is crucial to establish the hierarchical distinction between Artificial Intelligence and Machine Learning. In this thesis, AI is viewed as the overarching paradigm of creating systems capable of simulating human intelligence and

autonomous decision-making. ML, on the other hand, is positioned as a specific subset of AI where systems are not explicitly programmed with rules but rather learn to identify patterns autonomously from historical data (Myrans et al., 2018b).

To illustrate this distinction within the context of sewer management: an overarching AI system would encompass a fully autonomous robotic crawler capable of navigating pipe networks, avoiding physical obstacles, and deciding which pipes to inspect. Machine Learning, conversely, is the specific data-driven engine. Rather than being explicitly programmed with the exact mathematical geometry of a "crack" or "root intrusion," an ML model is fed thousands of CCTV images and autonomously learns the visual signatures of those defects. Therefore, this thesis positions itself firmly within the ML sub-domain.

ML approaches problems using different models depending on the nature of the task. The two most relevant to this field are:

I. **Supervised learning**: this is a method in ML that explicitly label training examples and map input variables (i.e., independent variables) to output variables (i.e., dependent or class variables) using the mapping function learned during training. An example of a supervised learning model is a model that classifies whether a defect exists in an image or not.

II. **Unsupervised learning**: For unsupervised learning, no labelling is done on the training data but tries to group data points based on similarities. These groups are referred to as *clusters*, and the goal of unsupervised learning is to maintain high intra-class similarities between data points in a cluster and low inter-class similarities between data points in different clusters. An example of an unsupervised learning model is a model that groups different sewers by pipe type and usage.

In this research, a more advanced and suitable subdomain of ML known as *deep learning* was explored while using supervised learning principles. Section 1.6 provides a high-level overview of deep learning, introducing its fundamental principles and relevance to automated defect detection. A more in-depth exploration of deep learning is presented in Chapter 3, where advanced deep learning architectures, key methodologies, and their applications in sewer defect detection are examined.

1.3 Research Problem, Gaps in Knowledge, and Justification

In the UK, many sewer systems are considered old when they have been in service for several decades, which were built starting from 1858 during “The Great Stink” of London (Porter, 2001). With an expected active lifespan ranging between 60 and 100 years (M. Wang, Luo, et al., 2021a), numerous sewers currently in use are either well past their prime or gradually ageing at the least (Guo et al., 2009; Moradi & Zayed, 2017; Wang et al., 2021). Over this extended lifespan, both internal factors and external influences contribute to the gradual

development of defects, compromising the overall integrity of the infrastructure. Specifically, factors such as deposits, general wear and tear, high vehicular traffic, harsh weather conditions, and external vibrations from nearby traffic, construction works, or natural events like earthquakes further accelerate this deterioration (Xu et al., 2021a).

Owing to these factors, structural and operational defects inevitably accumulate within the sewer network. As noted, the sheer age of the infrastructure is not the only concern; rather, it is the cascading consequences of these unresolved defects. The current problems and critical gaps resulting from this deterioration can be itemised into the following key areas (Gao et al., 2019a; Xu et al., 2021b):

- **Environmental and Ecological Damage:** Undetected defects lead to groundwater contamination via exfiltration and river pollution caused by sanitary sewer overflows (SSOs).
- **Catastrophic Infrastructure Failures:** Severe structural degradation results in sudden pipe collapses and urban sinkholes, causing massive disruption.
- **Public Health Risks:** The exposure of untreated wastewater to the environment poses enduring hazards to local communities.
- **Operational Inefficiencies:** Rapid urbanization combined with irregular maintenance schedules and insufficient inspection frequencies means defects are often only addressed reactively, rather than proactively.

Therefore, the timely identification and rectification of these defects can significantly decrease the exorbitant expenses associated with emergency repairs and extend the operational lifespan of sewer networks (Aşchilean et al., 2018a). However, traditional methods of defect detection, which involve physical man-entry for inspection or experts reviewing hours-long survey videos, have been identified as laborious, unsafe, costly, and resource-intensive.

To automatically detect defect in survey videos, traditional approaches rely on methods such as morphological segmentation, Support Vector Machines (SVM) (Xiang et al., 2012), and Hidden Markov Models (Pasadas et al., 2020), which involve manual inspection and feature extraction. Specifically, these techniques range from using Support Vector Machine (SVM) with Histogram of Oriented Gradients (HOG) for root intrusion detection (Halfawy & Hengmeechai, 2015) to employing SVM with wavelet and Fourier transforms for crack categorization and utilizing HMM for defect detection systems (Situ et al., 2023a). However, these methods are labour-intensive as they require manual tuning for each defect and are less efficient due to their complex pre-processing and feature extraction steps (Cheng & Wang, 2018b; Xie et al., 2019a).

Likewise, traditional approaches also struggle with imbalance and limited data, as well as handcrafted features (Shang et al., 2019). This imbalance and data limitation stem from the infrequent occurrence of many defect types relative to normal conditions, making it challenging to compile a sufficiently diverse and balanced dataset. Moreover, the laborious process of manually labelling defects, combined with the variability of defect appearances, further restricts the availability of high-quality annotated data (Dang et al., 2021a).

To overcome the limitations of handcrafted feature engineering, the industry has increasingly adopted Artificial Intelligence (AI) and, more specifically, Machine Learning (ML). While AI represents the broader concept of machines simulating human visual intelligence, ML is a specific subset of AI where algorithms autonomously learn patterns and rules from historical data rather than relying on human experts to explicitly program every parameter. By leveraging this data-driven approach, ML has significantly advanced automatic sewer defect detection beyond the capabilities of traditional methods (Hu et al., 2023; Xie et al., 2019d; Xu et al., 2021b).

Building upon ML, Artificial Neural Networks (ANNs) have been notably effective in analysing pipe images for defects with much better accuracy (Situ et al., 2023a; Zhou et al., 2021b). However, conventional neural networks are often limited by their inability to handle high-dimensional image data directly, requiring extensive pre-processing and feature selection. This limitation paved the way for Deep Learning (DL). As a highly advanced sub-field of ML, DL, particularly deep Convolutional Neural Networks (CNNs), has excelled not only in handling high-dimensional, grid-like datasets directly from raw images (Dang et al., 2021c; Hassan et al., 2022) but also in classifying and detecting various sewer defects simultaneously with superior accuracy. A foundational overview of these DL paradigms is provided subsequently in Section 1.6 of this chapter, while a comprehensive technical critique of specific DL algorithms, network architectures, and state-of-the-art models is explored in depth in Chapter 3

Recently, this modern approach has benefited from extensive, high-quality training datasets (Dang et al., 2021c; Haurum et al., 2021) and data augmentation, resulting in more accurate and dependable defect detection outcomes. For DL, it is crucial to utilize image pre-processing techniques to enhance data quality and model accuracy by mitigating issues such as contrast, noise, and lighting variations (Li et al., 2021a; Li et al., 2022). However, different defects often exhibit distinct attributes, which open a research gap necessitating specific pre-processing methods applied to specific defects to consistently improve defect detectability and feature extraction.

For instance, adjusting image intensity and converting colour images to grayscale can enhance image quality, but may result in the loss of information in a single channel (Ma et al., 2022). Additionally, different data augmentation techniques may be necessary for addressing various defects. For example, mirroring the image vertically may result in a misleading representation of defects such as pipeline obstruction (Gutierrez-Mondragon et al., 2020). Consequently, it is essential to consider the specific characteristics of each defect and customize the image pre-processing approach to develop a more robust detection model.

This targeted approach to pre-processing will not only improve the accuracy of defect detection but also enhance the robustness of the detection system, enabling it to consistently perform under the varied and challenging conditions typical of sewer systems. Therefore, this research focuses on employing conventional and customized pre-processing techniques tailored to addressing different types of defects under consideration. Likewise, despite the significant

advancements witnessed in the last decade, there remains a notable gap in the availability of open-source models deployed specifically for defect detection and reporting.

This lack of accessible and deployed models limits the ability of Sewer Management Companies (SMCs) or Water Companies (WCs) to leverage state-of-the-art models for real-time defect detection. This is because most existing models are either proprietary (Chataigner et al., 2020; Loupos et al., 2016) or confined to research settings, with limited practical deployment and validation in real-world environments. This gap underscores the need for a validated model that can be deployed in production environments and monitored over time to ensure scalability and long-term utility.

The rationale behind embarking on this research is rooted in the critical role that sewer systems play in maintaining public health, environmental safety, and urban infrastructure integrity (Li et al., 2022). As established above, the reliance on manual inspections and lengthy video reviews presents an unsustainable bottleneck. Given the aging nature of most sewer infrastructure (Guo et al., 2009; Moradi & Zayed, 2017; Wang et al., 2021b), there is an urgent need for innovative, efficient, and reliable defect detection solutions to prevent catastrophic infrastructure failures.

Advancements in sensing technologies and deep learning present an opportunity to revolutionize this process. By addressing the identified gaps, specifically the need for defect-specific pre-processing and practical deployment, this research aims to provide a robust multi-label deep learning model. The development of an open-source, deployable system will not only offer researchers a framework for translating academic output into real-world usage but also provide SMCs with a powerful tool to enhance operational efficiency. Ultimately, this research is timely and necessary to reduce maintenance costs, prevent environmental hazards, and ensure the resilience of sewer networks in the face of increasing urbanization (Gao et al., 2019a; Xu et al., 2021b).

1.4 Research Contributions

The need for research contributions in the domain of sewer defect detection is motivated by the critical role that sewer systems play in maintaining public health, environmental integrity, and urban infrastructure sustainability (Li et al., 2022). Traditional methods of defect detection, which rely on manual inspections or hours-long survey videos, are not only labour-intensive and costly but also prone to error and inefficiencies (Kumar & Abraham, 2019a; Savino & Tondolo, 2021). As urban areas continue to grow and infrastructures age, there is an increasing demand for innovative and automated solutions that can enhance the accuracy, efficiency, and reliability of defect detection processes.

This research addresses key limitations in the current practice of sewer defect detection by proposing a robust and scalable model grounded in deep learning and defect-aware pre-processing. Owing to this, this research makes three main contributions:

1. The design of defect-specific image pre-processing strategies tailored to different sewer anomalies.
2. A comparative evaluation of deep learning models for identifying sewer defects.
3. The development and deployment of a deep learning-based system for real-time identification by sewer operators and municipalities.

This research proposes to implement a selective image pre-processing techniques tailored to enhance data quality and model robustness. Previous studies (Kirstein et al., 2012; Myrans et al., 2018a; Wang et al., 2023; Yusuf et al., 2024) have employed different set of techniques to process the images but there were no considerations for specific characteristics of each defect. By mitigating issues such as contrast, noise, lighting variations, and adjusting image attributes to suit the specific characteristics of different defects, these pre-processing methods will improve the robustness of defect detection models, thereby ensuring a more accurate and reliable defect identification across the diverse and challenging conditions typical of sewer systems. Likewise, the pre-processing pipeline will strengthen the foundation for building models that are more sensitive to subtle variations and adaptable to a wide range of sewer conditions.

Secondly, the study provides a systematic comparative evaluation of modern deep learning models for multi-class sewer defect detection. Whereas prior research typically limits analysis to one or two object detection models, this research will benchmark several iterations of both the one stage and the two-stage family of deep learning algorithms under consistent experimental conditions. The evaluation framework will not only consider model complexity and architecture but also accounts for compatibility with sewer inspection constraints, such as real-time inference, low-light image handling, and varying defect sizes. This comparison will help identify the most suitable algorithms for practical sewer infrastructure monitoring and offers guidance to researchers and practitioners alike.

The third contribution lies in the design and deployment of an open-source, real-time defect detection model optimised for practical field use. The system will be packaged as a cloud-ready, containerised microservice that allows survey footage to be analysed automatically through a user-friendly interface. It will be built to integrate with existing sewer inspection workflows, supports secure deployment, and is designed for extensibility. By making the full pipeline openly accessible, deployment automation, this research reduces adoption barriers for utilities and contributes to a more reproducible and scalable ecosystem for intelligent infrastructure maintenance.

1.5 Research Aim and Objectives

The primary aim of this research is to develop an advanced deep learning model capable of automatic detection of multiple defect types within sewer infrastructures. To achieve this aim, a series of specific objectives has been established:

1. Review advancements in automatic defect detection, including defect characteristics and model development approaches.

2. Evaluate and apply conventional and custom defect-specific pre-processing techniques to enhance feature extraction in sewer survey videos.

3. Develop and compare various deep learning models using the adaptively processed data to determine the optimal architecture for defect detection.

4. Deploy the optimised model as an open-source Software-as-a-Service (SaaS) solution to facilitate practical industry use and future research.

1.6 Research Questions

Crafting research questions plays a vital role in enhancing the credibility of a research project (Kitchenham & Charters, 2007; Petticrew & Roberts, 2005) serves as a framework for aligning the research objectives with the study's overall aim. Therefore, the following research questions have been carefully developed to align with the overall aim and objectives of this research. Figure 2 further depicts the intersection between the research objectives and research questions.

1. What model development approaches have been applied in automatic sewer defect detection, and how effective are machine learning and deep learning algorithms in identifying common defect types?

2. Which conventional and custom image pre-processing techniques most effectively enhance data quality and detection performance in sewer defect imagery?

3. Among existing deep learning architectures, which offer the best trade-off between detection accuracy and computational efficiency for multi-class sewer defect classification?

4. What practical challenges and opportunities arise in deploying an optimised sewer defect detection model as an open-source Software-as-a-Service (SaaS) solution for research and industry use?

Figure 1.2 further illustrates the intersection between the research workflow and the specific research questions (RQs) that must be addressed to achieve the overarching objectives. As depicted, the methodology follows a structured, five-step process. Step 1 (Systematic Literature Reviews) and Step 2 (Data Collection) establish the foundational knowledge and dataset required to answer RQ1 and RQ2. Subsequently, Step 3 (Data Preprocessing) and Step 4 (Model Development & Evaluation) correspond to answering RQ3 and RQ4, focusing on the technical enhancement of images and the empirical benchmarking of the deep learning models.

Finally, Step 5 (Model Deployment) directly aligns with RQ5, emphasizing the practical transition of the developed model into a deployable architecture, thereby ensuring all research objectives are comprehensively met.

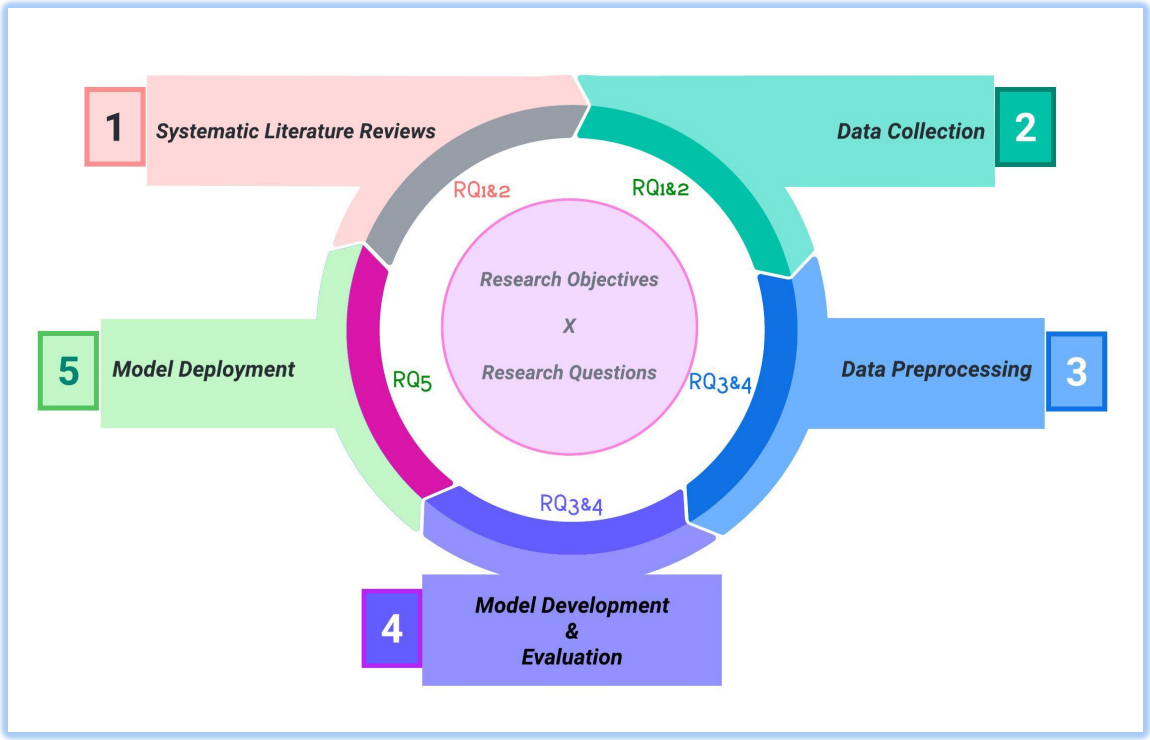


Figure 1.2: Alignment of Research Objectives and Questions with the Research Workflow (Source: Author)

1.7 Overview of Deep Learning

Deep learning, a subdomain in machine learning responsible for training a machine with the capability of mimicking how human's processes and understand information (Ajit et al., 2020; Fu & Menzies, 2017; Thompson et al., 2020). Like neurons in humans, deep learning uses a series of simulated neurons known as a neural network to learn distinctive features and complex patterns from a set of input data using algorithms like Multilayer Perceptron (MLP), Artificial Neural Network (ANN), Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM) among others. Popular applications of deep learning include image classification, object detection, chatbots, autonomous vehicles, and natural language understanding among others.

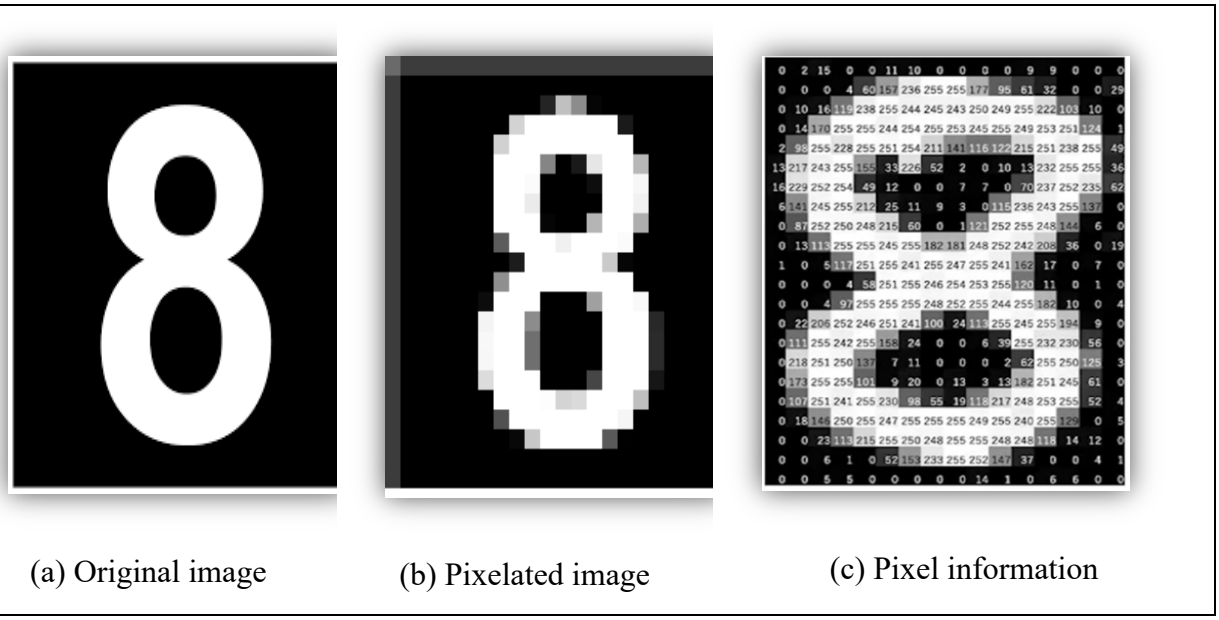
1.7.1 Image Classification

Image classification is a fundamental computer vision task that involves extracting and processing visual information from an image to categorize it based on its content. In the context

of sewer defect detection, image classification plays a crucial role in automatically identifying and labelling defects (Hu et al., 2023), such as cracks, root intrusions, encrustations, and structural deformations from inspection footage. Each digital sewer inspection image consists of a vast number of smaller, addressable elements known as pixels, where each pixel contains specific information about the intensity and colour composition of the image. These pixels collectively represent patterns, textures, and structural features that are essential for distinguishing between different defect types.

Computers, unlike humans, interpret images as multidimensional arrays of numerical values, where each pixel is assigned, a specific intensity ranging from 0 (black) to 255 (white) in grayscale images, or as combinations of Red, Green, and Blue (RGB) values in coloured images. As illustrated in Figure 1.3(a) to 1.3(c), this structured representation allows deep learning models to analyse variations in brightness, contrast, and edge patterns to detect and classify sewer defects accurately.

The ability of convolutional neural networks (CNNs) and other deep learning architectures to extract meaningful features from pixel-level data significantly enhances the accuracy and efficiency of sewer defect classification, eliminating the inconsistencies and subjectivity associated with manual inspections (Zhao et al., 2023). Through automated image classification, sewer infrastructure can be assessed more reliably, facilitating timely interventions and optimized maintenance strategies.



(a) Original image (b) Pixelated image (c) Pixel information

Figure 1.3: How computers understand digital images
(Source: Author)

Image classification leverages complex algorithms to extract meaningful features from pixel arrays, enabling the detection and differentiation of objects within an image. In sewer defect detection, deep learning models such as Convolutional Neural Networks (CNNs) analyse structural patterns, textures, and contrast variations to accurately classify defects like cracks,

root intrusions, and encrustations. Like facial recognition systems (Roundtree, 2021; Rowe, 2020) and other advanced computer vision applications (Pinheiro & Collobert, 2021; Yang et al., 2024), these models move beyond raw pixel values by identifying defects based on learned features.

1.7.2 Object Localization

While image classification is essential for identifying whether a defect is present in an image, it does not provide the precise location of the defect, which is crucial for automated sewer inspections. In real-world applications, merely knowing that an image contains a defect is insufficient for effective infrastructure assessment and maintenance. Object localization (Noothout et al., 2020) extends image classification by not only detecting the presence of a defect but also pinpointing its exact position within an image. This is achieved by drawing a bounding box around the detected defect, allowing inspectors to quickly identify and assess the severity and spread of the issue within a sewer pipe.

For instance, in a sewer inspection scenario, an image classifier may determine that a given frame contains a crack, but without object localization, it cannot indicate where in the pipeline the crack appears. Using localization techniques, such as those employed in YOLO (You Only Look Once) or Faster R-CNN, a bounding box can be drawn around the crack, enabling engineers to determine its exact dimensions, orientation, and position relative to other pipeline structures. This is particularly useful for identifying multiple defects within the same image (Zeshan et al., 2020), such as a crack near an encrustation or root intrusion, which ensures that maintenance teams can take targeted actions rather than relying on broader, less efficient interventions.

As illustrated in Figure 1.4, the position of a defect is as critical as its presence, making object localization a fundamental step in automated sewer defect detection.



Figure 1.4: Sample image showing the location of the defect via localization.
(Source: Research Data)

1.7.3 Object detection

Unlike image classification, which only determines whether a defect exists, or localization, which identifies the position of a single object, object detection can simultaneously classify and localize multiple defects within the same image (Chen et al., 2018). This technique is particularly useful in sewer defect detection, where a single image or video frame may contain multiple types of defects, such as cracks, encrustations, and root intrusions. This capability is crucial for automated sewer inspection systems, as it enables a more comprehensive assessment of pipeline conditions and allows for more precise maintenance planning (Moskalenko et al., 2020; Xie et al., 2019c).

Beyond sewer infrastructure, object detection has been widely adopted in various fields (Al-Azzo et al., 2018; Dionisi et al., 2012; Gardenier et al., 2018), with one of the most notable applications being in autonomous vehicles (Sligar, 2020). Self-driving cars rely on object detection to identify and track different objects in their environment, including road signs, traffic lights, pedestrians, and other vehicles, ensuring safe navigation. Similarly, in sewer inspections, object detection models such as Faster R-CNN, SSD (Single Shot MultiBox Detector), and YOLO (You Only Look Once) are employed to accurately detect and differentiate multiple sewer defects within an image or video sequence. This capability significantly enhances the automation of sewer defect detection by reducing manual inspection

efforts, improving detection accuracy, and enabling real-time analysis of pipeline conditions for timely maintenance interventions.

1.8 Methodology Overview

This section outlines the systematic approach to be employed in this research to achieve the stated aim and objectives through answering the research questions. Each component of the methodology is briefly defined, and its necessity justified, followed by the brief approach to be taken to address each objective and research questions (RQ1 – RQ5).

The methodology broadly comprises a systematic literature reviews to establish the current state-of-the-art and identify gaps in sewer defect detection research; data collection and pre-processing to gather and refine relevant data for algorithm adaptation; model development to develop and fine-tune a robust deep learning model; model evaluation and comparison using standardized metrics against benchmark datasets to ensure the model's effectiveness as a standalone, and also in comparison with existing models; deployment for real-world use cases to assess practical applicability. Figure 1 shows a high-level overview of the proposed methodology, while subsequent sub-sections will provide an overview of each component. A detailed explanation of each component is further provided in subsequent chapters

1.8.1 Systematic Literature Review

A systematic literature review (SLR) is a comprehensive survey of existing research on a particular topic (Kitchenham & Charters, 2007). Owing to this, a comprehensive survey of existing research on automatic defect detection will be conducted, which is critical for understanding the state-of-the-art, identifying gaps in the literature, and shaping the development of the research framework. This research involved two distinct SLRs to cover a wide area of defect detection in sewer systems. The first SLR focused on analysing existing methodologies for building defection models for sewer systems. Building upon these insights, the second SLR delved into the applications of ML and DL for sewer defect detection, encompassing effective image pre-processing techniques, defect types, annotation tools, object detection algorithms, and evaluation metrics among others. Consequently, this will address research questions *RQ1* and *RQ2* which will be detailed in chapters two and three.

1.8.2 Data Collection and Pre-processing

Data collection and pre-processing involved gathering raw data and preparing it for analysis, which was not only crucial for ensuring the quality and relevance of the data but also have direct impact on the model's performance (Harshini et al., 2022). The objective here was to collect and convert historical sewer survey videos into streams of images and extract

corresponding sewer metadata from inspection reports for both data-driven and selective image pre-processing (Meijer et al., 2019; X. Yin, Chen, Bouferguene, Zaman, et al., 2020a). Inspection reports and survey videos from sewers across the UK were collected to provide a rich data repository for this research. The targeted sewers spanned popular cities in the UK, and the survey videos were sourced from UK-based SMCs that handle a variety of industry-based services, which is not limited to CCTV surveying and drain maintenance.

Furthermore, inspection reports detailing the status of surveyed sewers at the time of the survey was also collected. Information extracted from these reports was used to form guidelines for data analysis and pre-processing, while streams of images was extracted from the survey videos to ensure a comprehensive dataset as model input. Furthermore, extracted images from each of the video underwent rigorous and selective pre-processing, which included various image-based conventional and custom pre-processing for quality enhancement. Likewise, the selective approach to pre-processing used in this research involved examining the specific characteristics of each defect and customizing the image pre-processing techniques on a defect-by-defect basis. This further addressed research questions *RQ1* to *RQ3* in chapters 4 and 5.

1.8.3 Model Development

Model development in this research involves designing and training deep learning models to perform a specific task of automatic detection of multiple defect types within sewer infrastructures. The pre-processed data was used for the model development, which began with the selection of suitable deep learning architectures identified from conducting a SLR on deep learning algorithms. Using PyTorch, a leading deep learning framework developed by Meta, these architectures were adapted, fine-tuned and customized to improve performance on the collected dataset while transfer learning (Situ et al., 2023a) techniques were leveraged to utilize pre-trained models in accelerating the training process and improving initial accuracy. This further addressed research questions *RQ3* and *RQ4 in chapters 5 and 6*.

1.8.4 Model Evaluation and Comparison

As a requirement for validity and reliability, model evaluation and comparison of the proposed models were essential to ensure their reliability and effectiveness in real-world applications. The models were assessed using standardized evaluation metrics such as accuracy, precision, recall, F1-score, mean Average Precision (mAP), and Intersection over Union (IoU) to provide a comprehensive view of their performance. Benchmark datasets such as Sewer-ML (Harum & Moeslund, 2021) were employed for the purpose of offering a consistent basis for comparing the developed models with state-of-the-art models in sewer defect detection. This part further addressed research questions RQ4 in chapter 7.

1.8.5 Implementation and Deployment

Implementation and deployment of the developed models involved encapsulating them within an Application Programming Interface (API) and deploying this API in real-world settings to assess their practical effectiveness. This step was crucial for ensuring that the models perform reliably outside the controlled research environment. An API was developed to serve as a wrapper for the developed models. This API enabled seamless interaction between the models and various applications to ensure efficient and scalable deployment. The deployment phase involves close collaboration with SMCs who are expected to provide practical insights by using the deployed models and reverting with detailed feedback. This iterative feedback loop will allow for the continuous refinement and enhancement of the models based on real-world performance and user experiences. This section addressed research questions *RQ5* in chapter 8.

1.9 Research Scope

This research focused on the development, evaluation, and deployment of advanced deep learning models for the detection of defects within sewer infrastructures. Thus, the scope included conducting systematic literature reviews to establish the current state-of-the-art in sewer defect detection by focusing on model development techniques, image pre-processing techniques, and the effectiveness of ML and DL algorithms among others. Historical sewer survey videos and inspection reports were gathered from SMCs across the UK, and selective pre-processing techniques were implemented to enhance image quality and data relevance.

Consequently, several deep learning models were developed and rigorously tested using standardized evaluation metrics and benchmark datasets to compare their performance against state-of-the-art models. Finally, an API was developed to serve as a wrapper and connection points to the models, which were deployed in real-world settings through collaboration with SMCs, while feedback was collected for model refinement.

1.10 Thesis Structure

This thesis is structured as illustrated in Figure 1.5, which includes the title of each chapter and a brief description of its focus. Each chapter is designed to progressively build on the previous one by outlining a logical flow from the introduction of the research problem to the detailed analysis and discussion of findings, which provides a clear roadmap of the research journey.



Figure 1.5: Overview of thesis structure and chapter descriptions
(Source: Author)

1.11 Chapter Summary

This chapter provides a comprehensive overview of the research, including the background studies on current practices in sewer inspection and defect detection. It identifies the research gap in existing defect detection methods and articulates the motivation and justification for this study. It further outlines the aim and objectives of the research while emphasizing the development of advanced deep learning models for sewer defect detection. A brief overview of the research methodology was provided, detailing the type and source of data used for developing the multi-label defect detection model and describing the pre-processing and model training processes.

Additionally, the intended contributions of the research were highlighted by revealing how this study's unique approach advances the defect detection in sewer infrastructure. Finally, the scope of the research was defined, and potential limitations were discussed, while providing a clear context for subsequent chapters and setting the stage for a thorough exploration of the research topic.

CHAPTER TWO

2.0 Automated Defect Detection in Sewer

2.1 Brief History and Overview

Defect detection involves the processes and techniques employed in uncovering different anomalies or defect in sewers by adopting different techniques and technologies. Though, these techniques can either be manual or automatic, the focus of this study will solely be on automatic defect detection. The evolution from manual to automatic defect detection in sewers has progressed in the last decade (Tscheikner-Gratl et al, 2019). This progression to can be traced back to the mid-20th century, when the initial efforts to improve sewer inspection efficiency through technological innovations began. Initially, sewer inspections were entirely manual, requiring inspectors to physically enter sewer systems to perform visual assessments. This method was not only hazardous but also prone to human error. The introduction of CCTV systems in the 1960s (Williams, 2003) marked a significant breakthrough, allowing for remote visual inspections and reducing the need for human entry into dangerous environments. However, the interpretation of CCTV footage remained manual, which continued to result in inconsistencies and errors in defect identification (Oh et al., 2022; Zhou et al., 2022a).

The 1990s piloted in the era of conventional computer vision techniques, which enabled the semi-automatic analysis of CCTV footage. These early systems could detect basic defects such as cracks and blockages but were limited by their reliance on handcrafted features and rule-based algorithms (Colyer et al., 2018). These approaches, while a step forward, lacked the flexibility and accuracy needed for more complex defect detection. The advent of machine learning, and particularly deep learning in the 2000s, marked a paradigm shift in automated defect detection. These advanced approaches enabled the automatic extraction of features from images, allowing for the development of more sophisticated models capable of detecting and classifying a wide range of defects with better accuracy and robustness. Likewise, the advancement and integration of robotics and sensor technologies (Chataigner et al., 2020; Ogai & Bhattacharya, 2018) has further enhanced these systems by laying the foundation for fully autonomous inspections of sewer networks, thereby revolutionizing the field of sewer defect detection.

The benefits of automated defect detection over manual methods are numerous. Automated systems can process vast amounts of data much faster than humans, reducing the time required for inspections and minimizing the likelihood of human error. Moreover, these systems can operate in hazardous environments where manual inspections would be risky or impossible,

thereby enhancing the safety of the inspection process. Automation also enables the consistent application of defect detection criteria, leading to more reliable and objective results. However, despite these advancements, several challenges and limitations remain. One of the primary hurdles is the variability in sewer environments, which can affect the performance of automated systems (Datta & Sarkar, 2016; Fu et al., 2022; Yu et al., 2021). Differences in lighting, pipe materials, and the presence of debris or obstructions can complicate defect detection. Additionally, while deep learning models have shown great promise, they require large amounts of labelled data for training, which can be difficult and costly to obtain and prepare (Harum & Moeslund, 2021; Kumar & Abraham, 2019a; Shen et al., 2023).

Furthermore, the complexity of sewer systems, with their numerous branches and varying conditions, presents ongoing challenges for the development of fully autonomous inspection systems (Rizzo, 2010). Nevertheless, the continuous evolution of technology holds the promise of overcoming these challenges, leading to even more effective and reliable automated defect detection systems in the future.

2.1.1 Stages of AuDD

The process of Automated Defect Detection (AuDD) in sewer systems is a multifaceted and complex task that involves a series of interdependent stages. Drawing from existing research and practical implementations, a high-level and generic AuDD framework was developed as depicted in Figure 2.1. This consists of seven sequential phases: *data acquisition, data pre-processing, feature extraction, defect identification, risk assessment, report generation,* and *decision making*. Each phase is essential to the overall effectiveness of the AuDD system by ensuring that defects are not only detected but also appropriately managed in terms of repair and maintenance. The first phase, *data acquisition*, serves as the starting point of the entire AuDD process. In this phase, data is collected using various techniques as detailed section 2.3. The quality and type of data acquired are critical, as they directly influence the subsequent stages of defect detection. This raw data, however, is rarely ready for direct use in defect detection models and requires substantial pre-processing

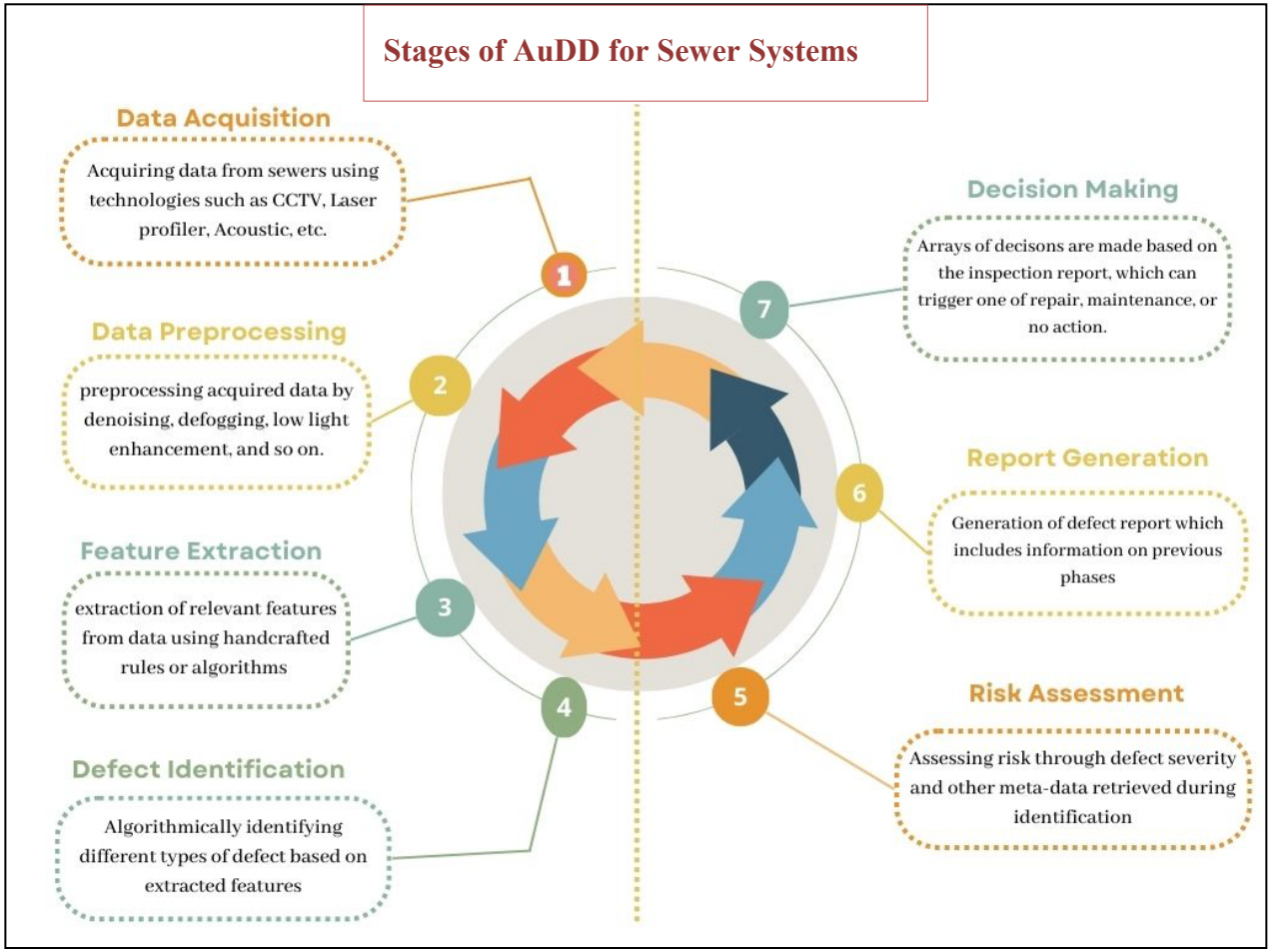


Figure 2.1: Generic AuDD framework with seven dependent phases
(Source: Author)

The *data pre-processing* phase enhances the quality of acquired data, making it suitable for analysis by applying specific techniques based on the data type, such as addressing missing information in tabular data (Haurum et al., 2021b; Zhou et al., 2022a) or performing grayscale conversion and noise removal in image data. Following pre-processing, the *feature extraction* phase identifies and extracts relevant features, either manually in traditional machine learning or automatically in deep learning, to ensure accurate defect identification in the subsequent phase. The core of the AuDD framework implementation in this research is in both data processing and defect identification. The *defect identification* phase, where the actual detection and classification of defects occurrence take place by leveraging the features extracted in the previous step to identify and categorize defects within the sewer system. However, merely identifying a defect is not sufficient for operational decision-making; the severity and potential impact of the defect must also be assessed.

The risk assessment phase quantifies the severity and potential impact of identified defects using standards like PACP or MSCC, guiding SMCs in prioritizing repairs based on the likelihood of a collapse or failure. High-risk defects require immediate intervention, while lower-risk issues may be monitored over time. A detail of severity grade assigned to defect is provided in colour-coded Table 2.1. Following this, a comprehensive *report* is generated, detailing the defects, their severity, and associated metadata. This serves as a crucial tool for informed decision-making in maintenance and repair strategies. The decision-making phase utilizes the generated report to determine the necessary actions, ranging from immediate repairs to monitoring or no action if the defect is low risk. This stage is crucial for ensuring the sewer system's efficiency and safety by prioritizing critical defects and optimizing resource allocation. The outcomes also feed back into future data acquisition and inspection strategies, closing the loop in the AuDD framework.

Table 2.1: Severity effect on pipe structure

Grade	Effect on Brick/Concrete	Effect on Sewer Pipe	Risk Assessment
1	No structural defect	No structural defect	Acceptable structural condition
2	Minor cracking, surface mortar loss, Spalling slight, wear slight	Circumferential crack, Moderate joint defects, Spalling slight, Wear slight	Minor collapse risk in long term but potential for further deterioration
3	Total mortor loss without other defects, single brick displaced, deformation up to 5%, spalling medium, Wear medium	Fractures with deformation up to 5%, Longitudinal cracking or multiple cracking, Minor loss of level	Collapse unlikely in near future but future deterioration likely
4	Total mortor loss with deformation greater than 10%, Deformation up to 10% and fractured, Displaced/hanging brickwork, small number of missing bricks	Broken, Deformation up to 10% and broken, fractured with deformation 5 - 10%, multiple fractures, Serious loss of level, spalling large, wear large	Collapse likely in foreseeable future
5	Already collapsed, missing invert, deformation over 10% and fractured, displaced/hanging brick work and deformation over 10%, Extensive missing bricks	Already collapsed, deformation over 10% and broken, Extensive areas of fabric missing, Fractured with deformation over 10%	Already collapsed or collapse imminent

2.2 Defects in Sewer Pipelines

Defects in sewer systems are any abnormalities or irregularities within the pipeline infrastructure that compromise its structural integrity or operational efficiency. These defects can range from minor issues, such as tiny cracks to major problems like collapses or severe blockages (Utepov et al., 2021). However, understanding what constitutes a defect is critical for the effective management and maintenance of sewer systems, as even minor defects can escalate into more significant issues if left unattended. The significance of detecting and addressing these defects lies in the essential role that sewer systems play in urban infrastructure. While they are responsible for the safe and efficient transportation of wastewater away from residential, commercial, and industrial areas, any defect that impairs these functionalities can lead to a cascade of problems, including reduced flow capacity, increased risk of overflows, and potential environmental contamination (Xu et al., 2021c).

To effectively identify and address these issues, sewer defects are classified into distinct categories, each representing different structural and functional problems that can occur within the pipeline network. Understanding these categories of defects is essential for implementing appropriate detection methods, prioritizing maintenance efforts, and preventing infrastructure failures before they become critical.

2.2.1 Categories of Defect

According to recognized standards established by bodies such as the Pipeline Assessment Certification Program (PACP) and Water Research Centre (WRc), defects can be broadly categorized into four main types:

- Structural
- Operational
- Construction
- Miscellaneous

Each category encompasses specific issues that affect the sewer system in unique ways. To provide a visual context, Figure 2.2 illustrates the four categories of defects along with examples for each category. This Figure will serve as a reference point for the detailed explanations that follow in the subsequent subsections, where each defect category is described, highlighting its causes, implications, and potential impact on sewer systems.

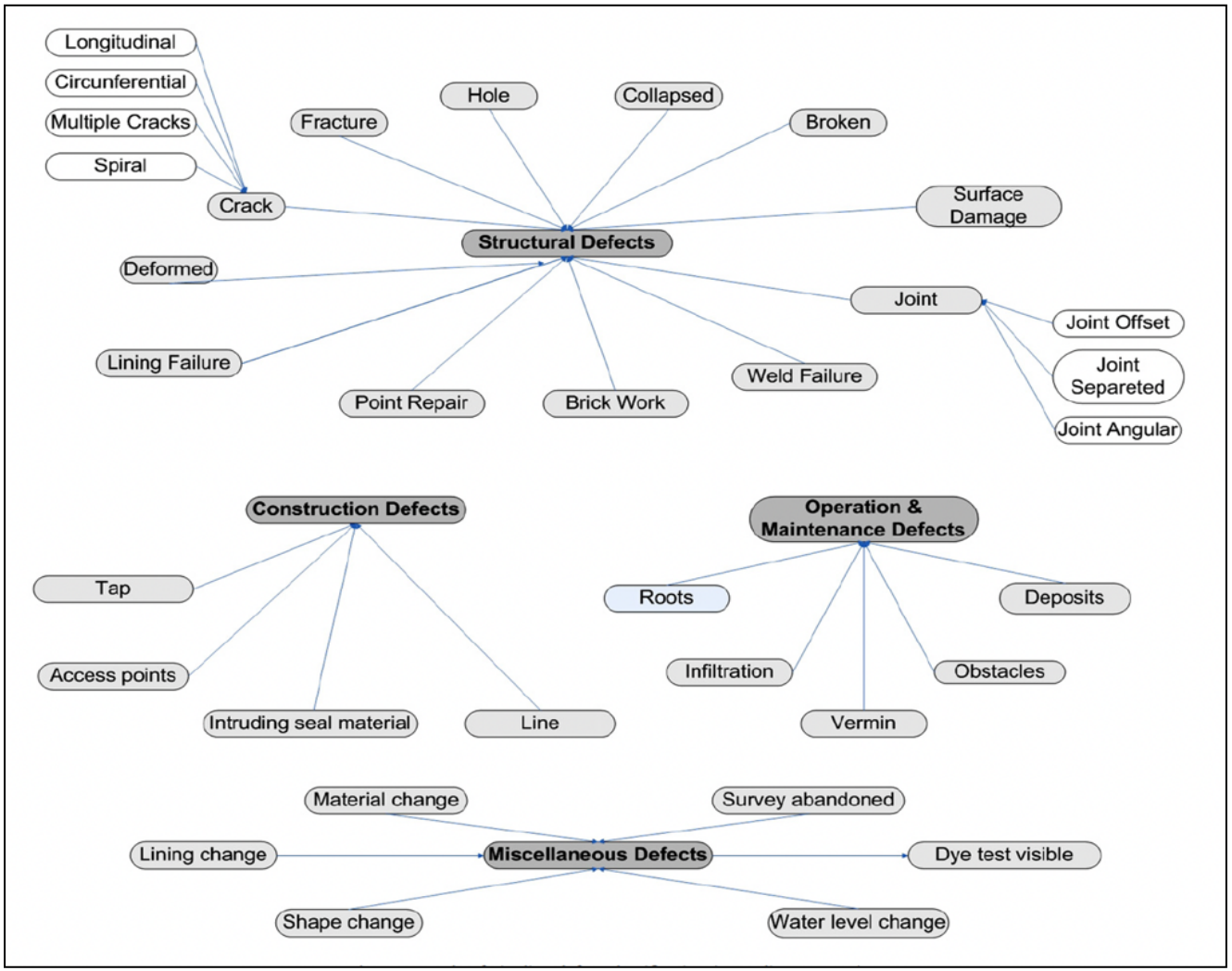


Figure 2.2: PACP and MSCC defect categorization.
Adapted from Guo et al. (2009) and Zayed & al Hawari (2017)

2.2.1.1 Structural Defects

Structural defects in sewer pipelines include physical damage that compromises the integrity and functionality of the pipes. *Cracks* are one of the most prevalent structural defects resulting from factors such as aging, ground movement, or heavy surface loads. They can be longitudinal, circumferential, or spiral, and weaken the pipeline structure and may lead to leaks or even collapse if not promptly addressed. As cracks worsen, they can evolve into more serious defects, further jeopardizing the sewer system's stability (Zuo et al., 2020). *Fractures* and *deformations* on the other hand represent more severe structural defects. Fractures occur when the pipe material experiences a complete break, often due to extreme pressure or untreated cracks, which can considerably disrupt the flow within the sewer system (Kaddoura, 2015a), requiring immediate attention to prevent catastrophic failures.

For *deformations*, it involves changes in the shape or alignment of the sewer pipes, typically caused by external pressures or soil movement, which can restrict flow capacity and create additional stress points or increasing the risk of further structural damage. The most critical of structural defect is collapse, where a section of the sewer pipe completely fails and collapse internally. This can be triggered by severe fractures, prolonged exposure to external pressures, or significant ground movement (Aşchilean et al., 2018b; Hughes et al., 2021a), leading to major service interruptions, environmental contamination, and costly emergency repairs. Figure 2.3(a) to 2.3(c) presents an example of structural defects such as crack, deformation, and surface damage



(a) Crack (b) Deformation (c) Surface damage

Figure 2.3: Sample structural defects like crack, deformation, and surface damage
(Source: Research Data)

2.2.1.2 Operational Defects

Operational defects refer to issues that impair the functionality of the system, often without causing immediate structural damage. These defects are particularly significant because they can lead to interruptions in the sewer system's operation, leading to blockages, overflows, and other complications that can escalate if not tackled. Common in this category is *blockages*, which are usually caused by the accumulation of debris, grease, or foreign objects within the pipes by restricting the flow of wastewater and resulting in increased pressure within the system and the risk of backups or overflows. *Root intrusions* present another critical operational issue, where tree roots penetrate sewer pipelines in search of moisture. As these roots grow and expand within the pipes, they can cause complete blockages and disrupt the flow of wastewater. Over time, root intrusions can lead to complete structural failures (Zheng, 2023) if the roots exert enough pressure on the pipe walls.

Likewise, an important defect to mention in this category are variations of *deposits* such as settled, attached, encrustation, and grease build-up. They also contribute to operational inefficiencies by reducing the effective diameter of the pipes, limiting flow capacity and increasing the risk of blockages. Accumulation of these deposits can lead to potential environmental hazards (Stanić et al., 2014). *Infiltration* and *exfiltration* are additional operational defects that pose serious issues to sewer systems. Infiltration refers to the unwanted entry of groundwater into the sewer system, which can overload the system during heavy rainfall and cause overflows (Pan et al., 2020). For exfiltration, it includes the leakage of wastewater from the sewer pipes into the surrounding soil, leading to environmental contamination and potential health risks (Nguyen et al., 2021). Figure 2.4(a) to 2.4(c) illustrates examples of these operational defects, such as settled deposit, infiltration, and Obstacle.

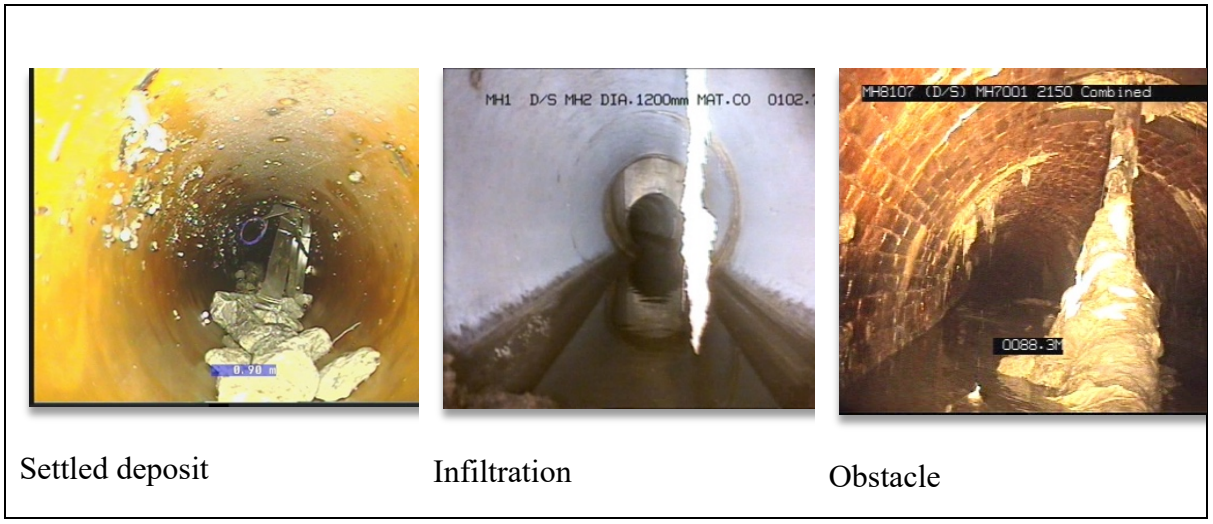


Figure 2.4: Sample operational defect such as deposit, infiltration, and obstacle. (Source: Research Data)

2.2.1.3 Construction Defects

These refers to the various issues and complications that can arise from the specific techniques, processes, and standards used during the installation and connection of sewer pipelines (Moselhi & Shehab-Eldeen, 1999). These defects are typically the result of flaws in the construction practices, or the methods chosen during the building phase of the sewer infrastructure. For example, if improper techniques are used when aligning and joining sections of sewer pipes, it could lead to *misalignment*, where the pipes do not line up correctly. This misalignment can cause uneven flow within the sewer system, leading to increased stress on certain sections of the pipes and eventually resulting in leaks, fractures, or even complete system failure. Additionally, if joints between pipes are not properly sealed or if inferior materials are used, it could result in *joint displacement* (Haurum et al., 2021a), where the connections between pipes shift or come apart over time, leading to infiltration of groundwater or exfiltration of wastewater.

Construction-related defects, particularly material defects, represent a significant threat to the long-term durability and functionality of sewer infrastructure. These defects typically arise when substandard, incompatible, or poorly specified materials are used during the construction phase. Examples may include the use of low-grade concrete, improperly sealed joints, or pipes that are not suited to the surrounding soil chemistry or hydraulic loading conditions. Over time, such material deficiencies degrade the physical integrity of the system, often leading to premature wear, joint displacement, or internal collapse. Unlike visible defects that emerge during routine operation, material-related issues may remain undetected for extended periods, silently undermining structural performance until major failures occur.

The implications of material defects are important because they are embedded during construction and masked by other operational layers, these issues are often only discovered after secondary problems such as infiltration, blockages, or surface collapses which have already developed. At that point, remedial work becomes not only technically challenging but also a financial burden. Moreover, undetected material degradation can accelerate the onset of other defect types, such as structural cracks or root intrusion, thereby compounding maintenance challenges. These cascading failures reduce the reliability of the sewer network, increase lifecycle costs, and elevate environmental risks due to uncontrolled discharges or contamination events. As such, detecting and mitigating construction-related defects, especially those involving materials remains important SMCs (Hu et al., 2023).

2.2.1.4 Miscellaneous Defects

Miscellaneous defects encompass a variety of issues that, while not common and directly related to structural or operational flaws, can still significantly impact the functionality and longevity of the infrastructure. Corrosion is a common defect in this category caused by chemical reactions between the materials of the sewer pipes and the surrounding environment,

which is usually accelerated by the presence of acidic or alkaline elements, gradually weakens the pipes, leading to leaks, fractures, pipe failure, and ultimately making it a critical concern for the long-term maintenance of sewer systems. Survey abandonment and water level changes during inspections are additional challenges that can complicate the accurate assessment of sewer conditions.

Abandoned surveys, often due to unexpected obstacles or safety concerns, leave sections of the sewer uninspected, potentially allowing defects to go unnoticed. Similarly, fluctuating water levels can obscure defects, making it difficult to identify and address issues timely and effectively. Other types of miscellaneous defects include animal intrusion, which pose unique challenges to sewer infrastructure. Animals, such as rodents or insects, can enter and inhabit sewer systems, causing blockages and damaging the internal structure of the pipes. Their presence can also lead to health and sanitation issues, as they may introduce pathogens into the wastewater.

While the implications of defect on the infrastructure itself have been explained, there are still bigger implications affecting not only the infrastructure's integrity but also posing significant risks to public health and the operational sustainability of SMCs. The next section will delve deeper into the specific impacts of sewer defects on both public health and SMCs, highlighting the urgent need for ongoing attention and innovation in this field.

2.2.2 Impact of Defects on Public Health and SMCs

Defects in sewer pipelines significantly impact the functionality of the sewer system (Gao et al., 2019b), leading to disruptions in wastewater management. Structural defects such as cracks, fractures, or deformations can cause leaks, reducing the pipeline’s capacity to transport wastewater efficiently. A major example occurred in Glasgow, UK, in 2026, where a large sewer collapse led to severe traffic disruption and required 2 months of emergency repairs (BBC, 2026). Similarly, in London, 2019, a ruptured sewer main in Islington caused raw sewage to flood streets and basements, leading to costly damage and public health concerns (Guardian, 2022). Over time, if left undetected, such structural issues can escalate into full-scale collapses, as seen in Manchester, 2017, where a 90-year-old sewer collapsed in Salford, creating a massive sinkhole that required weeks of repair (Manchester Evening News, 2022).

Outside the UK, a similar case was recorded in Baltimore, Maryland, in 2001, when the failure of a century-old sewer pipe caused widespread flooding and property damage, emphasizing the global risks posed by aging infrastructure (EPA, 2001). Operational defects, including blockages and root intrusions, also pose serious risks, often resulting in Sanitary Sewer Overflows (SSOs) and contamination of public spaces. A notorious example in the UK was the "Whitechapel fatberg" of 2017, a 130-ton mass of congealed fat, wet wipes, and waste that completely blocked a Victorian-era sewer in East London, requiring weeks to remove (Thames

Water, 2017). This event underscored the severe consequences of poor sewer maintenance and waste disposal habits.

Similarly, in Sidmouth, Devon, in 2019, another 64-meter-long fatberg was discovered, again highlighting the issue of sewer blockages caused by improper waste disposal (BBC, 2019). Beyond the UK, the 2013 Toronto floods resulted in widespread sewer overflows, contaminating water supplies and causing severe public health concerns, much like the 1993 Milwaukee cryptosporidium outbreak, which affected over 400,000 residents due to contaminated water linked to sewer system failures (CDC, 1993). All these points to the fact that the timely detection and repair of sewer defects are vital for maintaining public health and preventing large-scale environmental disasters. Incidents like these demonstrate the urgent need for automated defect detection systems that enable proactive sewer monitoring and maintenance, reducing the risks of infrastructure failures and their associated health hazards (Xu et al., 2021d). By leveraging advanced deep learning and image processing techniques, SMCs can prevent costly repairs, minimize public health risks, and ensure the sustainability of urban sewer systems.

For SMCs, the impact of defects extends beyond operational and public health concerns to significant financial and regulatory consequences. In recent years, regulatory bodies such as Ofwat and the Environment Agency (EA) have increased scrutiny and penalties for sewage mismanagement. For example, between 2023 and 2024, Ofwat recommended a £104 million penalty for Thames Water due to non-compliance with sewage management regulations, with additional, smaller fines suggested for Yorkshire Water and Northumbrian Water (Ofwat, 2024). Additionally, the EA is contemplating the removal of the existing £250,000 cap on civil penalties, potentially allowing for unlimited fines for sewage-related pollution (DEFRA, 2023).

Invariably, failing to detect and address sewer defects promptly could lead to revenue-crushing fines and severe reputational damage for SMCs. Consequently, as regulatory pressure intensifies, the financial and operational risks associated with consequences of sewer defects make the adoption of effective detection and maintenance strategies more critical than ever.

2.3 Evolution of Sewer Inspection Techniques

Sewer pipeline inspection is a critical aspect of maintaining the vital underground infrastructure. This has evolved from a traditional method to the use of modern technologies to ensure comprehensive coverage (Wang et al., 2021a). Owing to this, various inspection techniques have been previously developed to assess the condition of sewer systems. While each technique offers distinct advantages and limitations, the choice of method often depends on factors unique to each sewer, such as the type of sewer (sanitary or combined), sewer length and diameter, age, material (concrete or brick), location, and proximity to seismic zones (Hughes et al., 2021b; Yin et al., 2021). Traditional inspection methods, such as manual entry

and visual inspection (i.e., inspectors watching hour-long survey videos), have long been the standard for large sewer assessments. These methods involve physically entering the sewer system or using tools like smoke testing and dye testing to identify issues. However, these approaches are labour-intensive, time-consuming, and pose significant safety risks to inspectors.

The labour-intensive nature of traditional methods requires trained personnel to physically enter underground sewer systems or conduct extensive manual surveys. This process involves multiple workers, including engineers, technicians, and safety personnel, who must coordinate efforts to inspect, document, and report sewer conditions (Yusuf et al., 2024). For instance, in large urban sewer networks, workers may need to descend into tunnels using harnesses and ladders, manually examine pipe walls, and take note of defects. The sheer scale of the task, especially in cities with vast underground drainage systems, demands substantial human resources, making the process not only costly but also difficult to scale for frequent inspections. Additionally, preparatory work, such as ventilating confined spaces and setting up safety equipment, further prolongs the process, adding to the overall time and labour demands (Chen et al., 2019).

Moreover, time consumption in traditional sewer inspections is a major concern due to the slow and meticulous nature of manual assessments. Inspectors must move cautiously through dark, narrow pipelines, stopping frequently to examine walls and joints for damage. A prime example is London's aging Victorian sewer system, where manual inspections can take weeks or months to complete due to the complexity of navigating interconnected tunnels (MRSL, 2024). Safety risks are also a significant factor, as sewer environments expose inspectors to toxic gases such as hydrogen sulphide (H_2S), methane, and carbon dioxide, which can lead to asphyxiation or poisoning if proper ventilation is not maintained.

Likewise, workers face dangers such as flooding from unexpected surges, structural collapses, or exposure to harmful pathogens present in raw sewage. In 2017, an incident in India saw three workers suffocate and died due to toxic gas exposure while inspecting an underground sewer (NOV, 2023), highlighting the life-threatening risks associated with traditional inspection methods. These factors underscore the need for automated sewer inspection techniques, which can minimize human exposure to hazardous conditions while improving the efficiency and accuracy of defect detection. Moreover, the effectiveness of watching hour-long survey videos by inspectors is often limited by the subjective nature and fatigue that comes with the time-consuming activities.

The manual review of hour-long sewer inspection videos by inspectors is inherently limited by human factors such as fatigue, attention loss, and inconsistent interpretations from different inspectors. As inspections often require analysing long, monotonous footage captured by CCTV, drones, or robotic crawlers, inspectors may experience cognitive overload, leading to missed defects or inconsistent assessments (Kumar et al., 2020a). Additionally, the subjective nature of defect identification means that different inspectors may classify or grade the same defect differently, reducing the reliability and repeatability of assessments. Studies have shown

that prolonged video analysis leads to diminished accuracy over time (Trecroci et al., 2020), as fatigue impairs visual perception and decision-making abilities. These limitations highlight the inefficiencies of manual defect detection and underscore the need for automated approaches.

In contrast, modern sewer inspection methods have evolved significantly, incorporating advanced technologies to overcome the limitations of traditional techniques (Haurum & Moeslund, 2020). Advance visual inspection methods, such as Closed-Circuit Television (CCTV), have gained widespread adoption due to their ability to provide real-time video footage of the sewer interior (Romanova et al., 2013). CCTV inspection is particularly favoured by organizations like the National Association of Sewer Service Companies (NASSCO) for its ease of use and safety benefits. However, it is not without its drawbacks, including high operating costs, time-consuming inspection sessions, and the potential for subjective or erroneous interpretation by inspectors (Wang et al., 2021b). Additionally, CCTV is limited to inspecting only the sewer linings above the water level, leaving submerged sections unexamined (Halfawy & Hengmeechai, 2014b; Khan & Patil, 2019).

To address these limitations, a range of modern data acquisition methods has been developed and used, as depicted in Figure 2.5: A pentagonal mind map showing a high-level classification of sewer inspection techniques and highlighting the variety of techniques available.

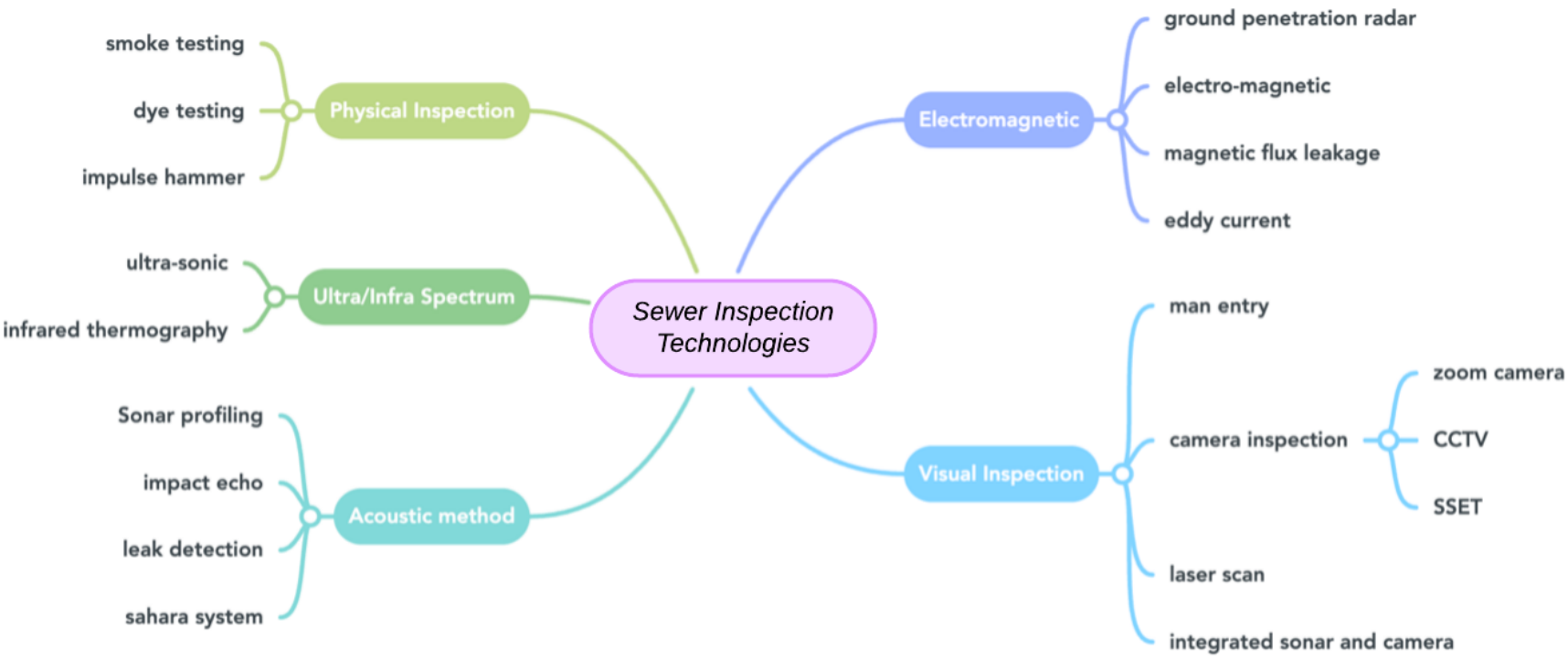
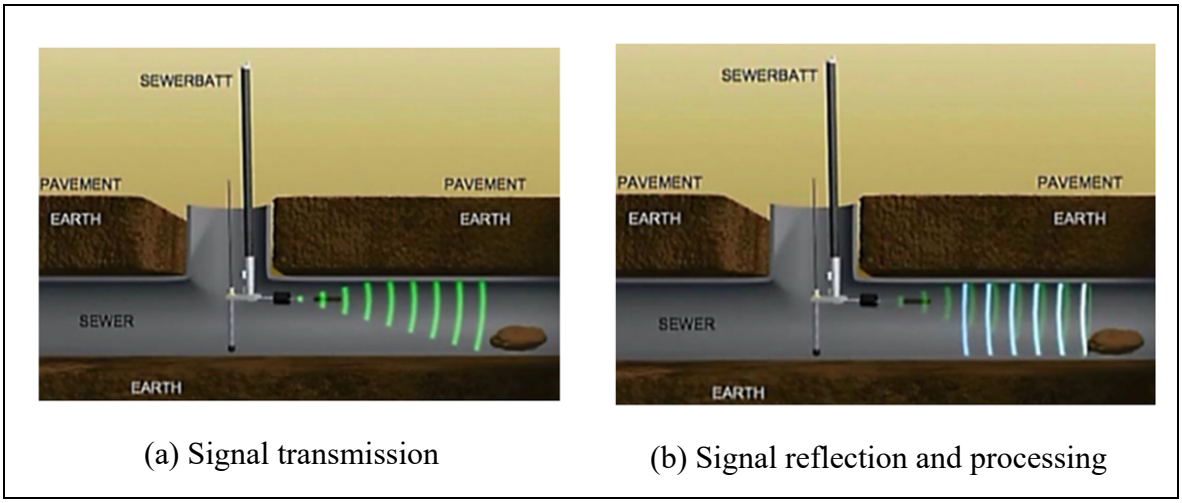


Figure 2.5: Taxonomies of inspection techniques. (Source: Author

For instance, *Electromagnetic* methods like Ground Penetrating Radar (GPR) and Magnetic Flux Leakage offer non-invasive ways to assess the structural integrity of pipes by detecting hidden defects that are not visible through traditional visual inspection alone (Karlovšek et al., 2011). Similarly, *Acoustic* methods such as Sonar Profiling and Leak Detection provide powerful tools for identifying cracks, voids, and leaks by analysing sound waves reflected from the pipe walls. Likewise, additional research in this regard has seen SewerBatt™ (Faram, 2014) built by researchers from University of Bradford, which uses electromagnetic induction technology to identify defects and abnormalities in sewer pipes by detecting changes in the electromagnetic field, as shown in Figure 2.6. These changes can indicate the presence of defects or abnormalities in the pipe.



(a) Signal transmission

(b) Signal reflection and processing

Figure 2.6: Operational workflow of SewerBatt (Faram, 2014)

As this research primarily focuses on defect detection rather than sewer inspection techniques, a comprehensive discussion of data acquisition techniques, including their advantages and limitations, is beyond its scope. However, studies such as (Moradi et al., 2019a) and (Yin, Chen, Bouferguene, & Al-Hussein, 2020) provide in-depth analyses of these methods, offering valuable insights into their applicability across different sewer environments. These studies highlight the importance of selecting an inspection technique that aligns with the specific characteristics of the sewer system, such as pipeline material, diameter, location, and structural condition, to ensure optimal data quality for defect detection.

2.4 Gaps in Current Defect Detection System

Despite the historical transition from hazardous man-entry to advanced non-destructive testing and high-resolution sensor technologies, the current system of sewer condition assessment remains fundamentally bottlenecked by its reliance on human analysis (Moradi et al., 2019d). Modern Sewer Management Companies (SMCs) possess the hardware to capture vast amounts of subterranean video footage, yet the industry standard still overwhelmingly dictates the manual, post-hoc review of this data (Gokhale et al., 2000). This reliance creates a severe operational paradox: while data collection is highly automated and scalable, the subsequent data analysis remains archaic, heavily labour-intensive, and inherently unscalable.

The inherent flaw in this manual review process is the physiological and psychological limitation of human operators. Reviewing hours of monotonous, poorly illuminated, and visually noisy sewer footage induces significant cognitive overload and fatigue-induced errors, leading to inconsistent and subjective assessments (Chaoyu & Zakaria, 2025). Consequently, the literature highlights that the current system is plagued by high degrees of subjectivity; while human operators can adapt to unexpected conditions, this reliance introduces severe variability in defect classification (Sinha et al., 1999). As noted in recent evaluations, this subjectivity severely weakens the reliability and reproducibility of condition grading across different inspections.

This operational subjectivity is further compounded by inconsistencies in industry standards. While established protocols like the Pipeline Assessment Certification Program (PACP) and SRM-4 provide structured frameworks, there is significant variability and a lack of consensus on the weighting of socio-economic and environmental factors (Simon, 2023). Comparative analyses of international condition grading systems reveal that differences in defect categorization and evaluation methods complicate benchmarking across regions (Cao et al., 2023). Current standards often rely heavily on manual data inputs, limiting their effectiveness in automated inspection contexts and introducing further uncertainty into intervention prioritization (Simon & Ehsan, 2022).

The financial, environmental, and structural importance of replacing this subjective assessment pipeline cannot be overstated. When manual methods delay data processing and decision-making, they directly hinder timely maintenance actions (J. Tang & Li, n.d.). Such delays lead to inefficient asset management, increased maintenance costs, and elevated risks of catastrophic infrastructure failure (Simon & Ehsan, 2022). Addressing these vulnerabilities is critical to ensuring system resilience and optimizing maintenance strategies before minor defects propagate into severe structural compromises (Kang et al., 2025).

Consequently, a definitive and critical gap remains in the literature and current practice: the glaring disconnects between high-throughput data acquisition capabilities and low-throughput, error-prone data analysis protocols. Although automated vision systems and deep learning models have been introduced to address labour-intensive inspections, gaps persist in handling diverse defect types and standardizing these methods across complex pipeline environments

(Khan et al., 2019; Laura, 2024). Until the manual review bottleneck is replaced with scalable, standardized, and uncertainty-aware automated pipelines, the massive volumes of visual data currently collected cannot be efficiently utilized for proactive asset management.

2.5 Existing Approaches in Developing AuDD

As earlier stated, the evolution of AuDD has significantly improved sewer inspection systems through advanced sensing technologies and computational models, but the increasing complexity of sewer environments highlights the need for more robust and adaptable models. Consequently, it is essential to delve deeper into the various methodologies that have been employed in the development of AuDD models. By reviewing over 90 studies from the past decade, trends, challenges, and opportunities in the development of AuDD models were identified, emphasizing how various techniques handle the complexities of sewer environments and the defects they aim to detect. Building on these insights, the following methodology outlines the systematic approach taken to gather, analyse, and synthesize the relevant literature.

2.5.1 SLR Methodology

The number of studies towards understanding defect detection in sewers has increased tremendously in the last decade, due to economic, environmental, social, and health & safety factor (Hawari et al., 2020; Patil et al., 2021) among others. To understand how existing studies have approached model development for defect detection problem, a systematic literature review (Kitchenham & Charters, 2007; Petticrew & Roberts, 2008) was conducted to recognize the recent methods explored in developing AuDD models. SLR is a scientifically focused and logical way of summarizing two or more primary studies on a precise topic using an unambiguous, pre-defined, and reproducible criteria, consequently eliminating arbitrary errors and reducing bias.

SLR has been previously applied to validate theories, summarize existing knowledge, and/or identify gaps in research, with its application cutting across other areas such as healthcare (Kuo et al., 2019; Rada et al., 2020), circular economy (Oluleye et al., 2022), deconstruction (Balogun et al., 2022), and bankruptcy prediction (Alaka et al., 2018) among others. To track the key processes in an SLR, Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol (Page et al., 2021a) was employed as a guideline. While PRISMA was primarily developed for the medical and health domain, its comprehensiveness, wide adoption, and applications to other fields (Moher et al., 2009; Page et al., 2021b) improves its general usability. Furthermore, PRISMA protocol encourages systematic reviews to be described in adequate detail that permit users to assess the reliability, repeatability, and relevance of the review discoveries. Consequently, it was adopted as the standard protocol to

scientifically review the recent model development approaches and defect consideration towards automated defect detection in sewers.

Using a top-down and stepwise elimination approach, the most relevant studies was retrieved and used in this systematic literature review. As diagrammatically shown in Figure 2.6, the PRISMA framework was adapted to categorize the systematic processes into 3 groups: *identifying the literatures, parameterized screening, and eligibility & inclusion.*

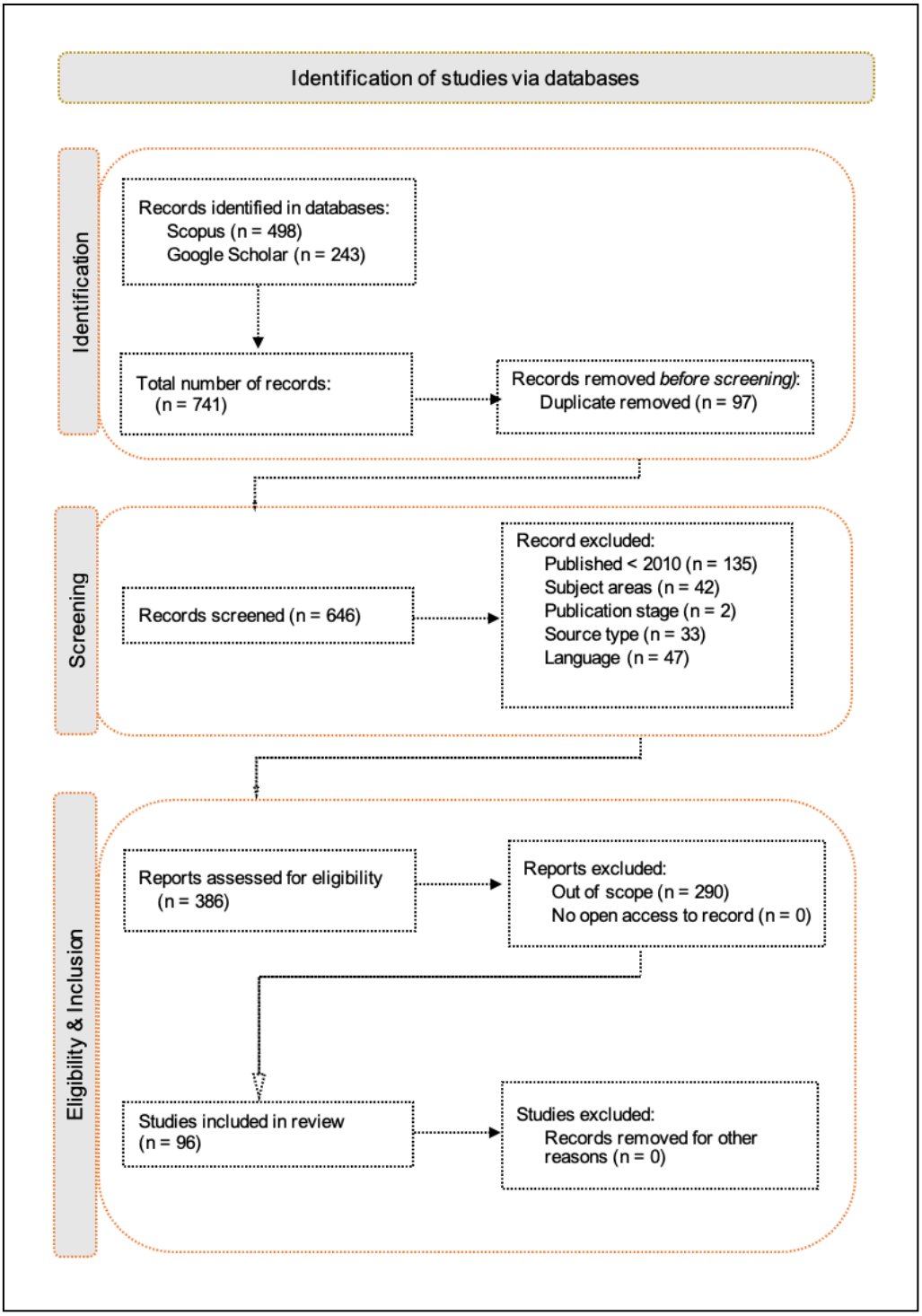


Figure 2.6: Proposed methodology for the SLR

2.5.1.1 Identifying the literatures

Defect detection in public sewers is a research area with numerous studies that uses different techniques and technologies; hence, a comprehensive search of the existing literatures was conducted. The research questions in Table 2.2 which formed the basis of our literature search was tokenized, and relevant keywords was extracted and used as the search criteria against 2 literature databases: *Scopus* and *Google Scholar*. Scopus is the leading intellectual and citation database with the highest number of peer-reviewed studies coupled with its easy retrieval, literature statistics, and other user-friendly functionalities (Oluleye et al., 2022; Walters, 2017) compared to other databases.

Table 2.2: Research questions drafted to guide the SLR

ID	Research Question	Motivation
SLR-RQ$_1$	What are the recent approaches used in developing defect detection models in sewers?	To identify the trends and challenges in using different model development techniques, including their pros and cons.
SLR-RQ$_2$	What type of defects were considered?	To understand the trends in defect consideration and overall challenges in their detection.

In addition to Scopus, Google Scholar is a widely accepted literature database (Haddaway et al., 2015; Meho & Yang, 2007) and was also used to retrieve relevant studies using the similar search strings that matches keywords, title, and abstract in the databases.

(Defect OR fault OR anomaly AND detection OR recognition OR segmentation AND sewers OR sewerage OR "sewer line" OR "sewer pipe")

Since the search was conducted on 2 different databases, there were high chances some literatures will be duplicated, hence, a merger of the search results was implemented, and removal of duplicate studies follows through. This guarantees no duplicate makes it to the next stage of screening. The preliminary search on the databases using logical search was able to retrieve 741 studies; 498 from Scopus and 243 from google scholar. These includes literatures from publishers such as Elsevier, IEEE, Willey, Springer, ASCE, and others. Unification of the search result revealed 97 duplicates which includes studies such as (J. Huang et al., 2017; C. Li et al., 2024; Myrans et al., 2019; Xie et al., 2019c.

2.5.1.2 Parameterized Screening

To retrieve the most relevant literatures, five parameters were defined: publication year, document type, publication stage, source type, and language. These parameters form the components of the inclusion criteria which was designed primarily to scale down irrelevant studies retrieved from the databases. The search was further limited to 2010 and 2022 since the goal was to understand recent trends and approaches to defect detection in public sewers. Furthermore, to validate originality of the studies, only literature in the final stage of publication was included. This means such studies have been peer-reviewed, validated for contribution to knowledge, and plagiarism. Lastly, all studies must be in English language as the cost of translating other languages was not included in the scope of this study. The final search string along with the screening parameters are given below.

Defect OR fault OR anomaly AND detection OR recognition OR segmentation AND sewers OR sewerage OR “sewer line” OR “sewer pipe”)
AND PUBYEAR > 2009 AND PUBYEAR < 2023 AND (LIMIT TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "cr")) AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (SRCTYPE , "j") OR LIMIT-TO (SRCTYPE , "p")) AND (LIMIT-TO (LANGUAGE , "English"))

A structured parameterized search was conducted using predefined inclusion and exclusion criteria to refine the selection of relevant literature. Initially, 506 publications were retrieved within the specified timeframe of 2010–2022. However, four documents were excluded as their types did not fall within the defined categories of conference or journal papers, reducing the count to 464. Further filtering based on publication stage, ensuring only final-stage publications were considered, resulted in 462 documents. Lastly, applying a language filter, restricting the selection to English-written papers, led to a final dataset of 386 studies for analysis.

As summarized in Table 2.3, the top-down filtering process systematically narrowed the dataset at each stage, ensuring the selection of only high-quality, relevant literature for the research. The search parameters included PUBYEAR, DOCTYPE, PUBSTAGE, and LANGUAGE, with each step refining the dataset by eliminating non-conforming papers. This structured screening process ensures that the final pool of studies is methodologically sound, relevant to the research scope, and aligned with the study's objectives.

Table 2.3: Screening parameters for inclusion/exclusion criteria definition.

Parameters	**Search Values**	**Search Code**	**Top-down Result Filter** ⇩
Publication year	2010 - 2022	PUBYEAR	506
Document type	Conferences, Journal	DOCTYPE	464
Publication stage	Final	PUBSTAGE	462
Language	English	LANGUAGE	386

2.5.1.3 Eligibility and Inclusion

To further validate the inclusion of eligible studies in the SLR, the title and abstract of each literature was carefully examined by looking for alignment with the research questions. This is primarily to eliminate studies that have the search keywords and falls within the inclusion criteria but are not relevant to the research questions. This approach narrowed down the screened studies to just the closely relevant ones. However, titles and abstract alone were only not sufficient to eliminate the not-so relevant studies. To finally scale down the literatures to only the relevant ones, each study that scaled through the title and abstract screening in the previous step was thoroughly assessed for relevance. This includes careful assessment of the background of study, methodology, case study, results, conclusion, and other relevant PRISMA checklist criteria.

Owing to this, only 96 studies satisfied all the inclusion criteria and passed all relevancy test. The final selection of primary studies then consists of conference papers and journal articles, with journal `articles making up 64% of the included documents and conference paper accounting for the remaining 36%. This distribution highlights the dominance of journal articles in the reviewed literature, suggesting a stronger emphasis on peer-reviewed, in-depth research contributions.

2.5.2 SLR Analysis and Result

2.5.2.1 Publication Trends Analysis

The distribution of publications between 2010 and 2022 reveals a clear path reflecting growing research interest in AuDD. As illustrated in Figure 4, research activity was relatively modest between 2010 and 2016, with only a few selected studies published each year. A particularly low point occurred in 2013, where no study made it through the rigorous inclusion process.

During this period (2010-2016), the research primarily focused on conventional image processing techniques such as edge detection and basic texture analysis methods, which were less computationally demanding and limited in terms of accuracy and scalability.

However, 2018 marked an essential shift with a sharp increase in research output, which continued its upward trajectory until reaching its peak in 2021. This increase may be attributed to advancements in deep learning techniques, particularly the rise of CNNs and other advanced machine learning architectures (Liu, 2020; Liu et al., 2019). These breakthroughs enabled more sophisticated defect detection with enhanced accuracy and efficiency, particularly in real-time applications. Figure 2.7 provides a summary of yearly publication frequencies between 2010 and 2022.

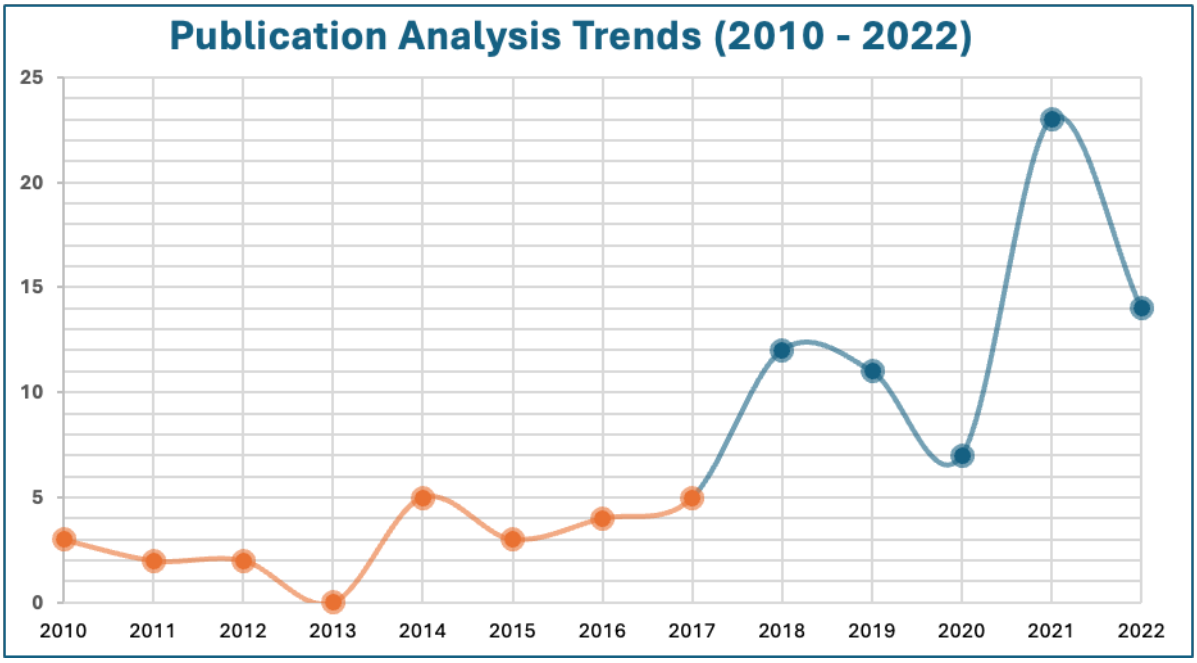


Figure 2.7: Annual publication distribution of selected sewer defect detection studies

The surge in publications post-2017 aligns with notable advancements in deep learning and computing capabilities (Dang et al., 2021b). Notably, 2021 saw the highest publication count, with 26 studies reflecting the broader adoption of more complex and computationally intensive methodologies, such as semantic segmentation, transfer learning, and hybrid approaches integrating CNNs with image processing techniques (Ghinea et al., 2024b; Goodfellow et al., 2016). The observed temporary decline in 2020 could be attributed to disruptions caused by the COVID-19 pandemic which impact research progress, funding availability, and logistical constraints.

The subsequent spike in 2021 likely represents a compensatory rebound, as researchers adjusted to new operational norms and resumed or intensified research efforts. Overall, the evolution in AuDD research can be divided into three distinct stages:

- **Exploratory Phase (2010–2014):** Publications during this period were limited, primarily focused on traditional image processing techniques and early machine learning applications, such as Support Vector Machines (SVM).

- **Transitional Phase (2015–2017):** Growing interest in deep learning approaches, particularly CNNs, marked the transitional period as researchers began to experiment with early powerful architectures and techniques.

- **Acceleration Phase (2018–2022):** This phase saw significant growth, with the introduction of hybrid methodologies, sophisticated deep learning architectures, and a focus on scalability and operational efficiency for real-world applications.

2.5.2.2 Citation Impact Analysis

The impact of research can often be measured not only by the number of publications but also by the citations each publication garners. The citation count reflects the significance and influence of research within the academic community and its potential to drive innovation and application. As the research field matures, the citation counts typically grow, indicating both the relevance and value of the work being conducted (Woods, 2024). Figure 2.8 presents show the year-by-year increase in citations based on the selected studies published between 2010 and 2022.

The chart clearly illustrates the accelerated influence of AuDD research, particularly in the years following the widespread adoption of deep learning methods. 2018 to 2022 stand out as the most significant years in terms of citations, with 2021 marking the peak in both publications and citations. This citation trend not only reinforces the increasing importance of AuDD but also highlights the growing recognition of research in this field across various sectors, including urban infrastructure, AI-based automation, and smart city technologies (Cumbajin et al., 2023; Moradi et al., 2019b).

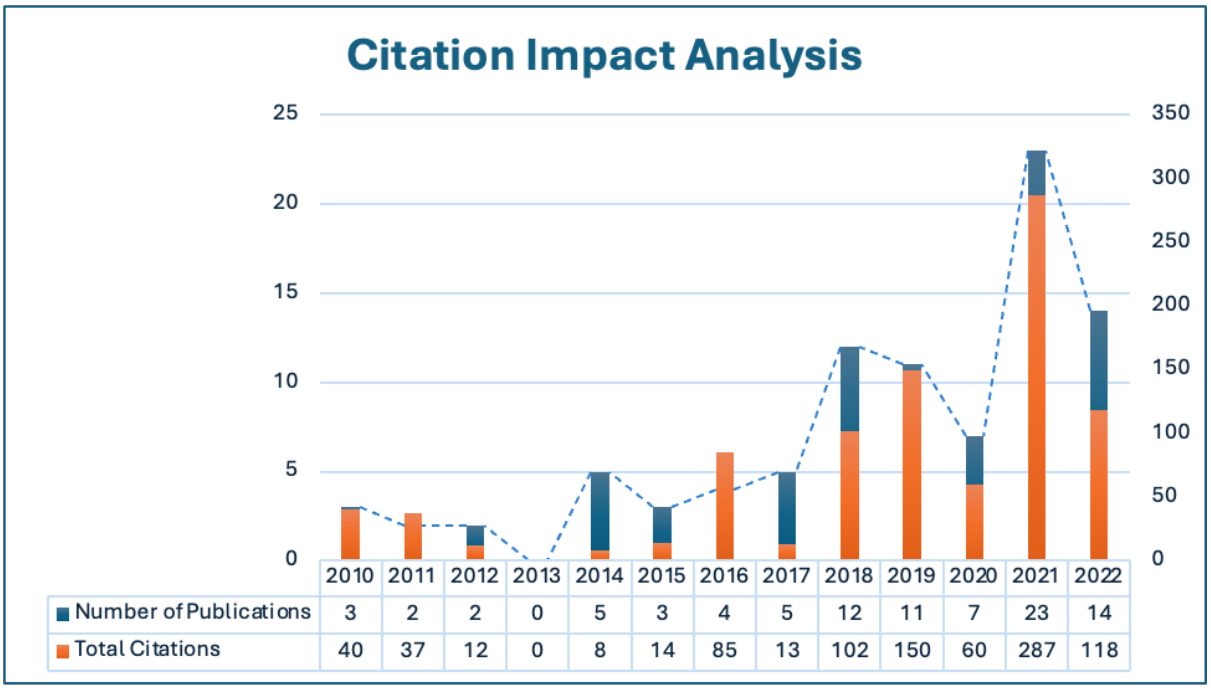


Figure 2.8: Annual citation trends for selected studies

2.5.2.3 Keyword Analysis of Selected Studies

Keyword analysis is essential for accurately identifying research trends by validating the completeness of the literature search and understanding predominant topics within a field (Woods, 2024). To ensure thorough coverage, a detailed analysis of keywords was performed by extracting the titles and author's keyword for the 96 selected primary studies. Figure 2.9 presents a word cloud visualization, highlighting the most frequently occurring terms, with larger words indicating higher usage frequency.

Prominent terms identified include "defect detection", "sewer pipelines", "condition assessment", "sewer pipes", and notably algorithmic terms such as "deep learning" and "convolutional neural network". The dominance of these keywords aligns with the growing trend toward advanced AI-based approaches observed in recent sewer inspection studies. For example, "deep learning" and "CNN" frequently appeared in studies published post-2017, coinciding with advancements in machine learning technology and increased computational power (Zhang et al., 2023). Recent influential studies, such as those utilizing YOLO (You Only Look Once) and Faster R-CNN architectures, have prominently featured these keywords, which further solidifies their importance within the field.

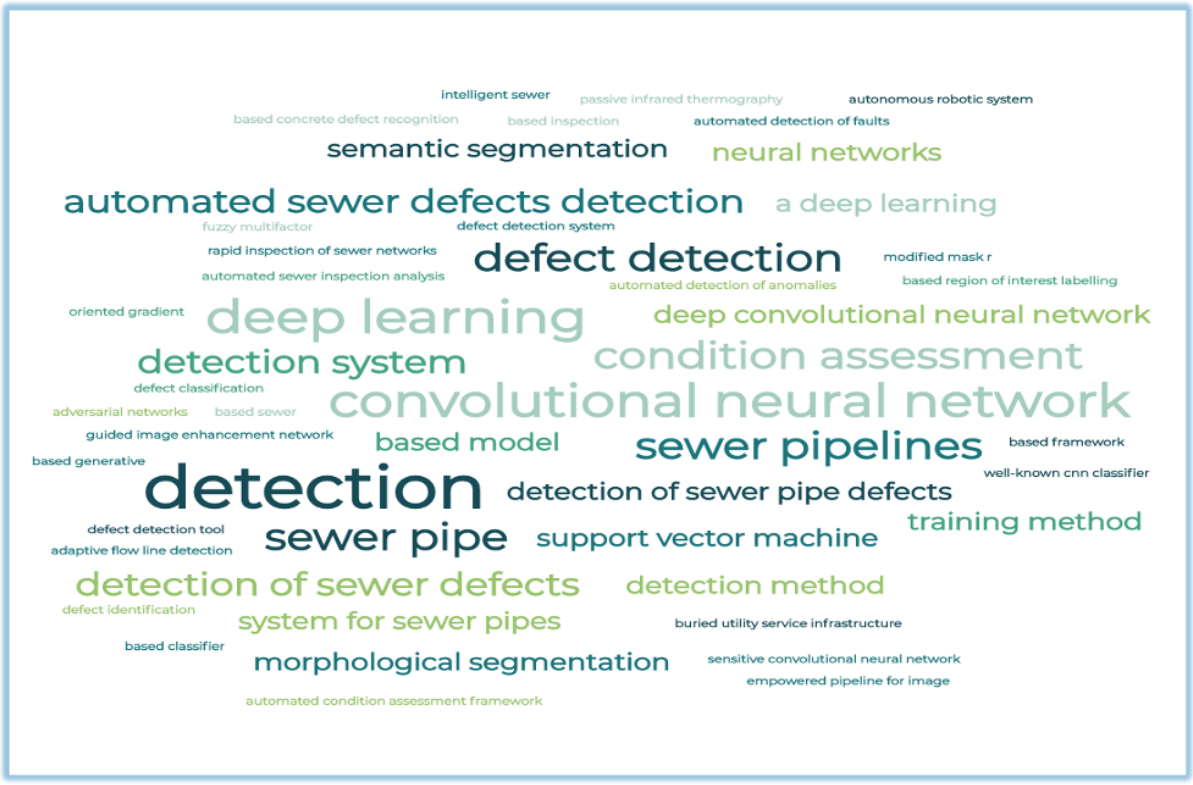


Figure 2.9: Keywords and title analysis of selected Studies

Likewise, the coexistence of domain-specific terms (e.g., “sewer”, “pipeline”, and “condition assessment”) alongside algorithmic terminology (“CNN” and “deep learning”) underscores the interdisciplinary nature of automated sewer defect detection research. The integration of civil infrastructure management expertise with advanced computer vision and AI technologies highlights the cross-domain collaboration essential in this research area. However, this interdisciplinary character introduced challenges during the literature search process. Terms widely used in computer vision (such as segmentation, classification, detection) often overlap with general technical terms in civil engineering, complicating keyword-based searches. This required carefully balancing specificity and breadth in the search queries to capture relevant studies without retrieving excessive unrelated literature.

Notably, the keyword analysis validated the effectiveness of the initial search strategy used. The presence of both domain-specific and algorithmic keywords, such as “sewer pipelines”, "defect detection", "condition assessment", “deep learning”, and “convolutional neural network” confirmed alignment with the foundational objectives of sewer infrastructure research. Additionally, the frequent occurrence of algorithm-specific terms assured comprehensive coverage of the recent literature that utilizes cutting-edge AI and deep learning techniques.

2.5.2.4 SLR-RQ1: Recent approaches used in AuDD

Recognizing the recent trends in AuDD model development and condition assessment of sewers requires extracting relevant meta-data from all primary studies. A key piece of meta-data is the method or approach employed in building the detection model. An exploratory data analysis (EDA) of the methodologies used in the primary studies, as depicted in Figure 2.10, revealed that more than half (54.1%) of the studies applied various deep learning techniques, while 10.6% used conventional machine learning approaches. Other notable approaches include a combination of conventional image processing techniques, including image processing, morphological segmentation, and histogram of oriented gradients (14.12%), as well as hybrid approaches (3.53%) and acoustic techniques (5.88%).

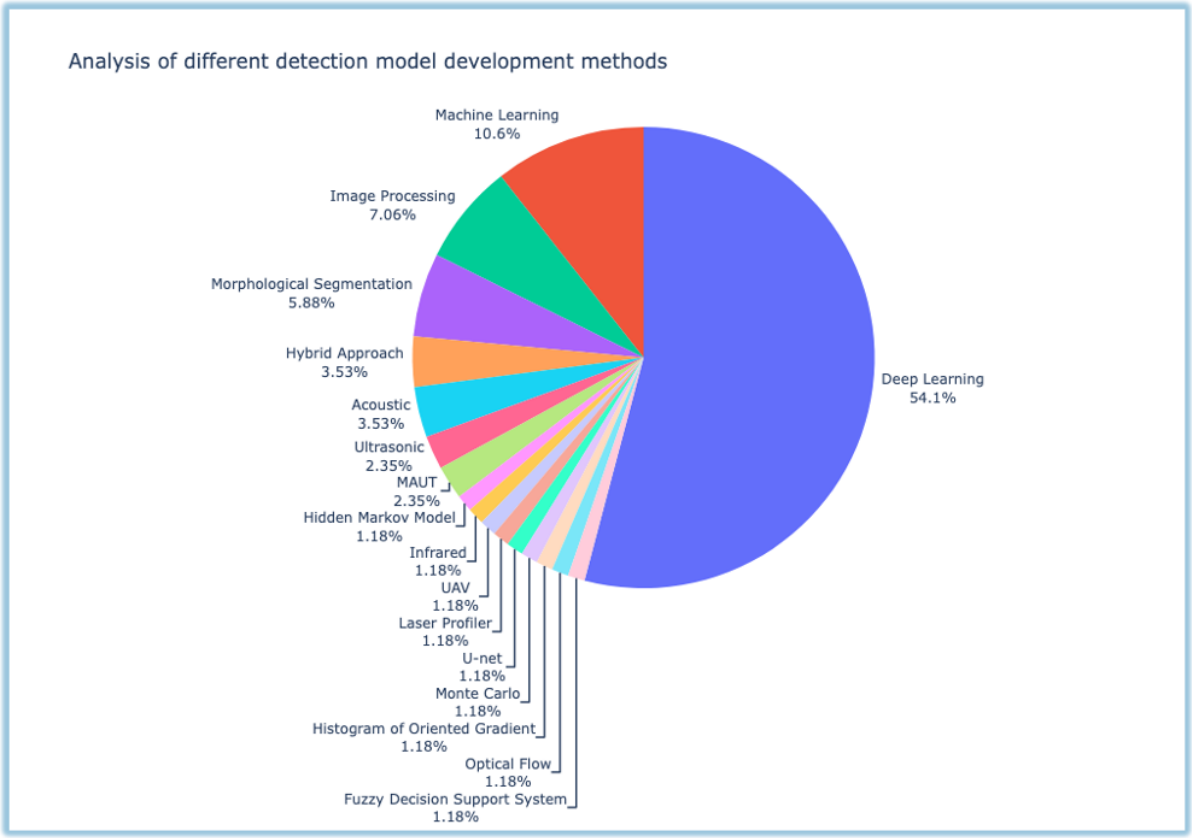


Figure 2.10: Exploratory data analysis of the methodologies used in the primary studies

To ensure a comprehensive understanding of the methodologies, the techniques were examined holistically and grouped based on related approaches to avoid bias in interpreting the data. The high-level analysis of the different AuDD model development approaches revealed clear trends in the methodologies used for sewer defect detection. When related techniques are grouped together, as shown in Figure 2.11, deep learning stands out as the dominant method, accounting

for 54.7% of all studies. This dominance may reflect its adaptability and effectiveness in managing large datasets and complex problems(Gupta & Kaler, 2018; Long et al., 2017), such as multi-label defect detection, where both the defect type and location are crucial. Deep learning techniques also benefit from advancements in computational power and the availability of extensive pre-trained models, enabling the rapid development of models through techniques like *transfer learning* and *fine-tuning*. Machine learning methods, which account for 10.6%, remain relevant in specific use cases where data complexity is lower or computational resources are limited.

Combining image processing techniques, such as morphological segmentation and histogram of oriented gradients, shows a considerable representation at 14.12%. This may suggest that despite the growing reliance on more complex deep learning models, traditional methods still offer valuable insights, particularly when dealing with simpler or well-structured data. Other approaches, such as hybrid techniques (3.53%) and acoustic methods (5.88%), highlight niche areas where multiple modalities or non-visual data are required to detect defects not visible in regular imagery, such as internal pipe degradation. Additionally, methods like MAUT and decision support systems, as well as other specialized techniques, collectively contribute 7.06% to the overall studies. This indicates their value in handling unique sewer inspection challenges that require a combination of different data types or decision-making frameworks.

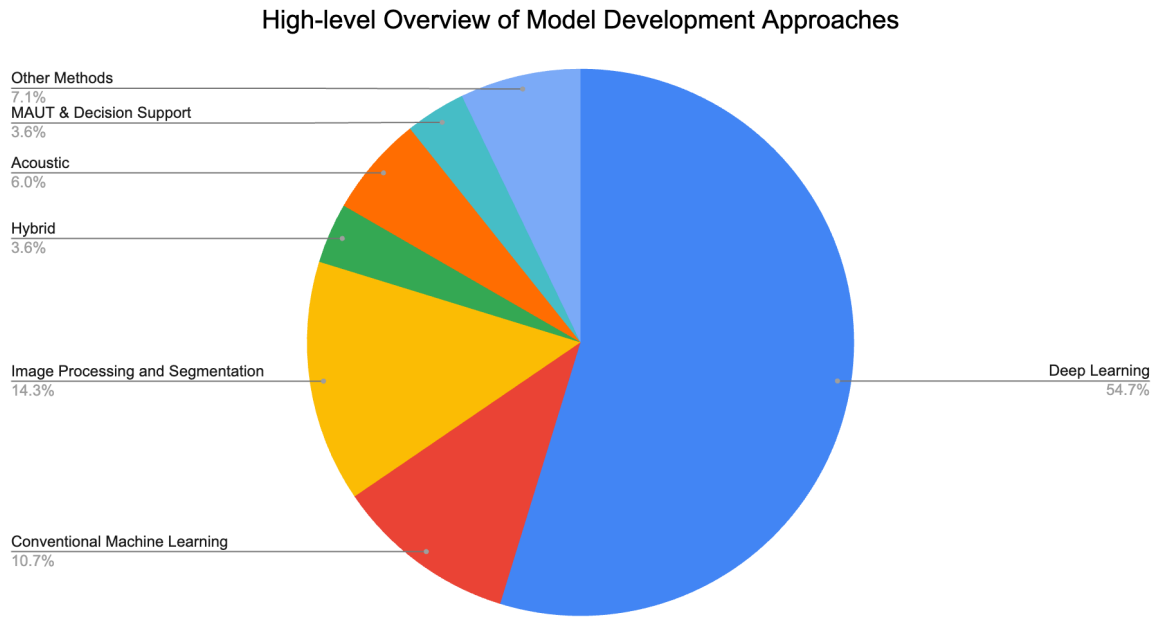


Figure 2.11: High-level overview of model development approaches.

Specifically with deep learning, as the leading methodological approach, Convolutional Neural Networks (CNNs) are the most prevalent algorithms, accounting for approximately 65% of

deep learning implementations. Popular CNN architectures include ResNet, VGGNet, and custom-designed CNNs tailored specifically for sewer inspection tasks, including defect classification. Region-based CNNs such as R-CNN and Faster R-CNN represent about 15% of studies and are particularly effective for precise defect localization tasks, accurately identifying and defining defects within sewer images. Segmentation-focused architectures, notably U-Net and other segmentation networks, account for roughly 12% of deep learning studies, delivering detailed pixel-level defect segmentation, crucial for identifying specific sewer defects like cracks, corrosion, and root intrusions. Additionally, advanced architectures such as Generative Adversarial Networks (GANs), autoencoders, and transformer-based networks collectively comprise approximately 8% of deep learning applications, offering innovative approaches for anomaly detection, feature enhancement, and improved generalization capabilities.

Overall, this analysis shows a strategic application of these methods across different contexts, showing a clear trend towards data-intensive, scalable solutions like deep learning, while still recognizing the importance of traditional image processing and niche methods in addressing specific sewer defect detection challenges. Consequently, a hybrid approach integrating conventional image processing techniques at the data pre-processing stage before utilizing deep learning models for defect detection will be further explored with a detailed analysis of deep learning presented in the next chapter.

2.5.2.5 SLR-RQ2: What type of defects were considered?

Over the years, different types of defects have been considered in several efforts to develop an efficient defect detection model. Defects are basically fault or anomalies that affect the normal day-to-day operation of the sewer systems. These anomalies are caused by different factors such as sewer lifespan, pipe material, length, sewer type, soil type, and surrounding trees among others (Dang et al., 2018; Khan et al., 2019a; Yin et al., 2020). Since these factors are both generic in features to most sewer systems, they result in different types of defects, hence, classifying these defects based on either the root cause or the effect they have on sewers, provides a more reasonable taxonomy to manage these defects. Coding standards such as Manual of Sewer Condition Classification (MSCC), and PACP (Yin, Chen, Bouferguene, & Al-Hussein, 2020) has grouped most known defects into notably four groups: structural, construction, operation & maintenance (O&M), and miscellaneous. Some of these defects are further subdivided into different subtypes based on their appearance. For example, crack can be longitudinal, circumferential, multiple, or spiral, while joints can be offset, separated, or angular. However, most studies deal with defect based on the high-level classification (i.e., crack), and not the subtypes (i.e., longitudinal crack). Some of these defects are depicted in Figure 2.12.



(a) Deposit

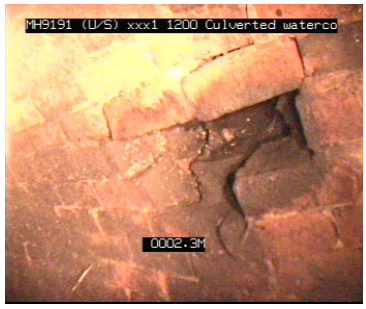
(b) Fracture



(c) Crack



(d) Encrustation



(e) Missing bricks



(f) Infiltration

Figure 2.12: Sample sewer defects

Frequent occurrences of certain defect in previous studies can either indicates the prevalence of such defects or growing interest of SMCs and researchers in mitigating the effect of such defect. An analysis of different defect considered by previous studies was conducted, as seen in Figure 2.13, while carrying out defect detection on different sewer system across different cities. Crack constitutes almost one-quarter (25%) of the defect recently investigated in our primary studies, and it is a popular defect caused by environmental influences and tree roots ingression among others (S. Khan et al., 2020).

Several studies including (Ganegedara et al., 2012; Künzel et al., 2018a; Tung-Ching, 2015; Wang & Cheng, 2020a; Zaheer et al., 2021b) have explored crack in sewers as they are prone to inflow of groundwater that can fill up the sewer and causes incessant SSO. When crack is left untreated, wear and tear from sewage coupled with external influence could degrade it to a more severe defect such as fracture. While approximately 4% of the primary studies (Su et al., 2011a; Viacheslav et al., 2020; Yin, Chen, Bouferguene, Zaman, et al., 2020b) investigated fracture, more than 10% examined deposit (Huang et al., 2020; Myrans et al., 2018a; Tennakoon et al., 2018) and tree roots intrusion (Kumar et al., 2020b; Situ et al., 2021b; Wang, Kumar, et al., 2021c) respectively. Figure 2.13 represents the distribution of different defect investigated according to the primary studies.

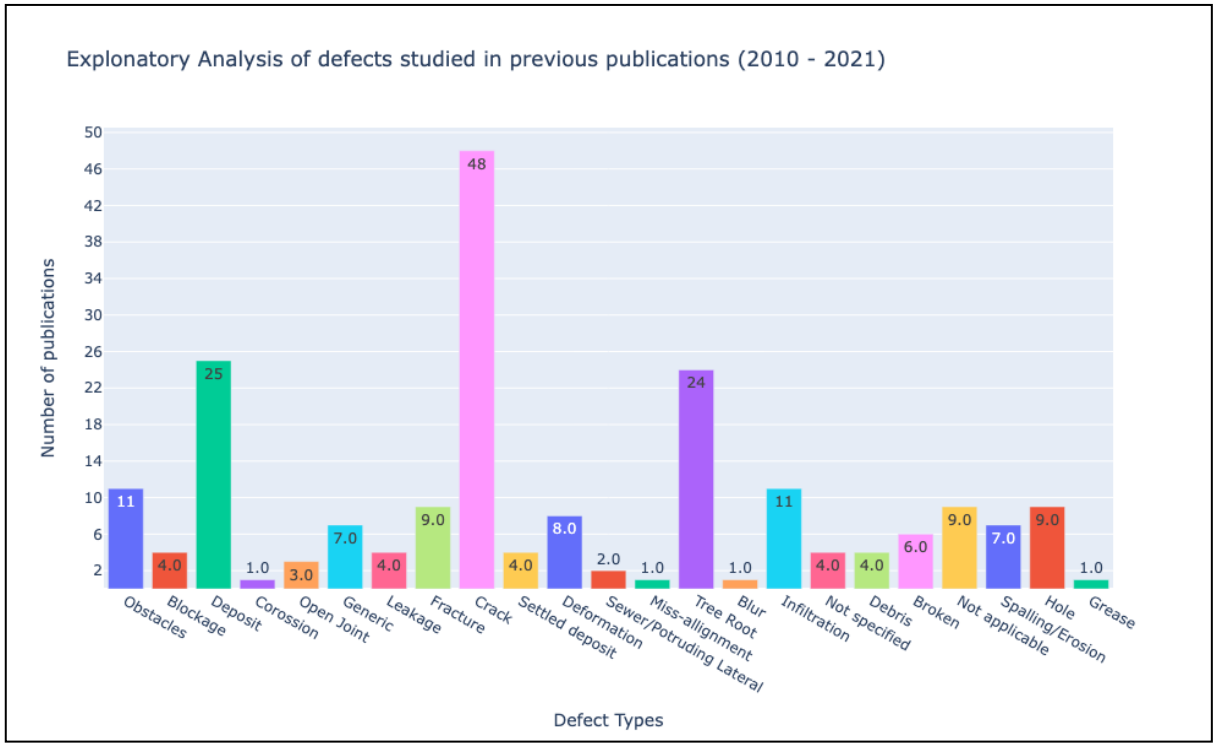


Figure 2.13: Analysis of defect considered in previous studies

Obstacles (Künzel et al., 2018b; Situ et al., 2021c; Zhou et al., 2021a) and infiltration (Kaddoura, 2015b; Myrans et al., 2018a; Phyu et al., 2021) are another set of defects with notable recent research effort. While obstacles include internal or external objects that prevent normal flow of sewage in the sewer, infiltration is generally external groundwater that enters the sewer through leaks and cracks from nearby soil. Like crack, obstacles and infiltration also result in several SSOs. Spalling or erosion, hole, broken, debris, and open joint are other types of defects that has been occasionally explored in recent years. This might be due to either non-frequent occurrence of such defect or unavailability of required data to investigate them. However, defect such as grease, blur, miss-alignment, and corrosion are the least to have been investigated in the last decade as evident from the primary studies. Although, corrosions are mostly common in metallic pipes.

2.5.2.6 Detection Complexity, Methodological Alignment, and Comparative Performance Analysis

Owing to further correlation and comparative analysis conducted at defect level, significant relationships exist between defect types and the corresponding preferred methodological approaches. Deep learning techniques again emerged as the dominant methodology across a

wide range of defect types, particularly excelling in the detection of cracks , which benefit from deep learning's superior ability to capture complex visual features. Defects such as settled deposits, root intrusions, and other visually prominent issues also align strongly with deep learning methodologies (Stanić et al., 2014), due to their clear and distinct visual characteristics that are well-suited for visual detection systems. These defects present patterns that deep learning models can easily interpret, making them highly effective for these specific cases.

On the other hand, defects characterized by more subtle or dynamic characteristics, such as deformation and infiltration, have shown a notable preference for other techniques such as conventional image pre-processing and acoustic detection techniques (Kaddoura et al., 2019). Acoustic methods have shown to work well for capturing temporal and dynamic defect features that are difficult to detect using traditional visual methods(Romanova et al., 2013; Yu et al., 2021). These defects may involve changes that occur over time or under certain conditions, which visual techniques might miss. As a result, integrating acoustic and visual methodologies could offer a more comprehensive approach to sewer defect detection, ensuring that both visually prominent and more nuanced defects are accurately identified.

When comparing the performance of various methodologies across defect types, deep learning consistently outperforms conventional methods. The colour-coded Table 2.4 summarizes the average F1-scores for the most common defect types, revealing that deep learning techniques achieve superior performance in detecting complex defects. For example, the detection of cracks using deep learning achieved an F1-score of 0.83, significantly higher than other methods, such as machine learning (0.68) and image processing (0.68).

Table 2.1: Average F1-score for selected defects using different techniques.

Defect Types	**Deep Learning**	**Machine Learning**	**Image Processing**	**Others**
Cracks	0.83	0.68	0.63	0.79
Deposit	0.80	0.71	0.68	0.78
Root Intrusion	0.78	0.65	0.60	0.76
Obstacles	0.75	0.66	0.62	0.74
Infiltration	0.63	0.64	0.58	0.75

This pattern of superior performance is consistent across other defect types such as obstacles, root intrusion, and deposits. Other techniques such as acoustic, while often showing less accuracy for visually obvious defects, still demonstrate good results for defect such as infiltration and deformation, indicating their potential as a complementary approach. However, it is important to consider that publication biases may affect these results, as studies with

weaker outcomes may be underrepresented in the literature, potentially skewing the reported performance metrics.

2.6 Selection of Model Development Method and Defect Types

While various techniques, such as acoustic methods, conventional image processing, and machine learning (ML), have been explored for AuDD in sewers, each comes with limitations. Acoustic methods are sensitive to environmental noise and can only detect one defect type at a time, while image processing relies heavily on manual feature extraction, making it less suitable for multi-class problems. Conventional machine learning algorithms has also been applied to sewer defect detection, using algorithms like SVM and RF (Decor et al., 2019; Pasadas et al., 2020) to classify defects. However, these approaches often require manually engineered features, which limits their ability to generalize across different defect types.

Furthermore, conventional ML techniques can struggle with multi-class detection problems, requiring additional pre-processing steps like feature selection, which increases computational complexity, and cannot perform object detection end to end. Despite their usefulness, conventional ML methods are not efficient for handling large datasets and multi-label classification.

In contrast, deep learning presents a robust solution for AuDD due to its ability to automatically learn and extract features from raw image data without manual intervention. Models such as CNNs, YOLO, and Faster R-CNN have proven highly effective in multi-class object detection, allowing for both the classification and localization of sewer defects in a single pass. Another key pro of deep learning is the ability to leverage techniques such as transfer learning (Huang & Zhang, 2019; Li et al., 2022; Situ, Teng, Liao, et al., 2023b), where models pre-trained on large datasets are fine-tuned to specific tasks like sewer defect detection.

This consequently reduces the amount of data needed for training and speeds up the model development process. Additionally, model fine-tuning enables the adaptation of these pre-trained networks to handle the variability of defects, improving accuracy and performance. Given these benefits, deep learning is a highly scalable and flexible approach, making it the most viable option for addressing the complexities and variability of sewer defects. Nonetheless, there are still some drawbacks in its adoption (Romanova et al., 2013).

First, it requires huge amount of data for effective model development, which directly leads to increment in the computational cost of running such models and thereby elevate the minimum standard for hardware compatibility. Similarly, data with low quality have reportedly (Dang et al., 2021c; Du et al., 2021) perform poorly with this approach. Furthermore, preparing training data for such models requires annotating huge number of images using annotating tools like *LabelImg* (Darren, 2022) *and makesense.ai* (Piotr, 2023), which will form the basis of the ground truth, hence, it requires a lot of time to complete data annotation.

Lastly, specialized expertise may be required to understand and implement such method, hence, experimentation and repeatability of this approach in research environment is not straight forward. Nevertheless, this approach will be adapted in this research due to its many benefits over other approaches and different workaround will be provided for its limitation, as there is no one approach fit all.

The selection of specific sewer defects for this research is informed by their relevance, prevalence, and impact on sewer systems, as noted in the literature (Moradi et al., 2019c). However, the adoption of defects such as cracks, variants of deposits, root intrusions, obstacles, and infiltration for this study remains subject to the availability of relevant data. While these defects are frequently highlighted in recent studies as critical issues affecting sewer functionality and safety, their inclusion in this research is contingent upon accessible datasets. This issue is explored in greater detail in Chapter 4, where the final selection of defects is made based on data availability.

Furthermore, the inclusion of these defects is also supported by their representation in industry coding standards such as PACP and MSCC (BSI, 2006), which classify them as key defects requiring attention. Their alignment with the operational priorities of SMCs (Tscheikner-Gratl et al., 2019) underscores their significance in the broader context of maintenance and operational challenges.

2.7 Chapter Summary

This chapter established the foundational context for Automated Defect Detection (AuDD) by comprehensively reviewing defect taxonomies, the evolution of inspection technologies, and the current methodological landscape. To ensure a rigorous evaluation, a Systematic Literature Review (SLR) of 96 primary studies was conducted. By mapping the transition from hazardous physical man-entry to advanced, non-destructive data acquisition methods, such as high-resolution CCTV and robotic crawlers, the chapter defined the severe implications that structural and operational anomalies pose to public health, environmental safety, and infrastructure sustainability.

The present state of knowledge regarding sewer inspection reveals a distinct operational paradox. While modern Sewer Management Companies (SMCs) possess highly advanced hardware capable of capturing vast amounts of subterranean video footage, the industry standard still overwhelmingly dictates the manual, post-hoc review of this data. The literature confirms that condition grading systems are heavily compromised by the physiological limitations of human operators, leading to high degrees of subjectivity and inconsistent defect classification across regions. Simultaneously, the state of knowledge establishes a definitive paradigm shift: the field is rapidly transitioning away from manual coding toward deep learning and artificial intelligence, which are now recognized as the primary enablers for managing this massive influx of visual data.

Consequently, this chapter exposes critical research gaps in the current literature and operational practice. Foremost is the analytical bottleneck: there is a glaring disconnect between high-throughput data collection and low-throughput, error-prone manual data analysis. Furthermore, while initial automated models exist, a gap remains in developing systems that can reliably classify diverse, multi-label defects within the complex, highly degraded visual environments typical of active underground sewers, free from subjective human input. Addressing these gaps requires a paradigm shift toward intelligent, uncertainty-aware diagnostic algorithms. Thus, this chapter sets the critical stage for Chapter 3, which explores specific deep learning architectures and their application in solving the automation bottleneck of sewer defect detection

CHAPTER THREE

3.0 Systematic Literature Review of Deep Learning Algorithms for Automatic Defect Detection in Sewers

3.1 Chapter Overview

The purpose of this chapter is to critically examine the evolution, performance, and applicability of deep learning techniques for automated sewer defect detection. Building on the broader systematic literature review conducted in Chapter 2, this chapter narrows its focus to the subset of object detection models, particularly convolutional neural networks (CNNs) and their derivatives that have shown relevance in sewer inspection tasks. The chapter aims to provide a detailed understanding of the model architectures, performance trade-offs, and design trends that influence the effectiveness and deploy ability of deep learning-based detection systems in real-world civil infrastructure contexts.

This chapter directly contributes to further addressing Objective 1 of the research by synthesizing more than a decade of deep learning research. These insights establish the empirical and architectural foundation for model selection and experimental design in Chapters 6 and 7, while also exposing significant research gaps that are further investigated and addressed throughout the thesis.

3.2 Review Methodology

Over the past decade, there has been a notable increase in research focused on defect detection in sewer systems, driven by growing concerns over economic losses, environmental impact, public health, and safety risks associated with deteriorating underground infrastructure (Hawari et al., 2020; Patil et al., 2021). To critically examine how recent studies have approached the development of models for automated defect detection (AuDD), this chapter employed a Systematic Literature Review (SLR) methodology. Guided by the principles outlined by (Kitchenham & Charters, 2007; Petticrew & Roberts, 2008), the SLR offers a structured, transparent, and replicable means of synthesising primary research on a clearly defined topic. By adhering to predefined inclusion criteria and minimising subjective bias, this review process ensures a rigorous and comprehensive understanding of current deep learning approaches applied to sewer inspection.

As diagrammatically shown in Figure 3.1, the systematic processes in this review were categorized into 3 phases: *identification, screening, and inclusion*

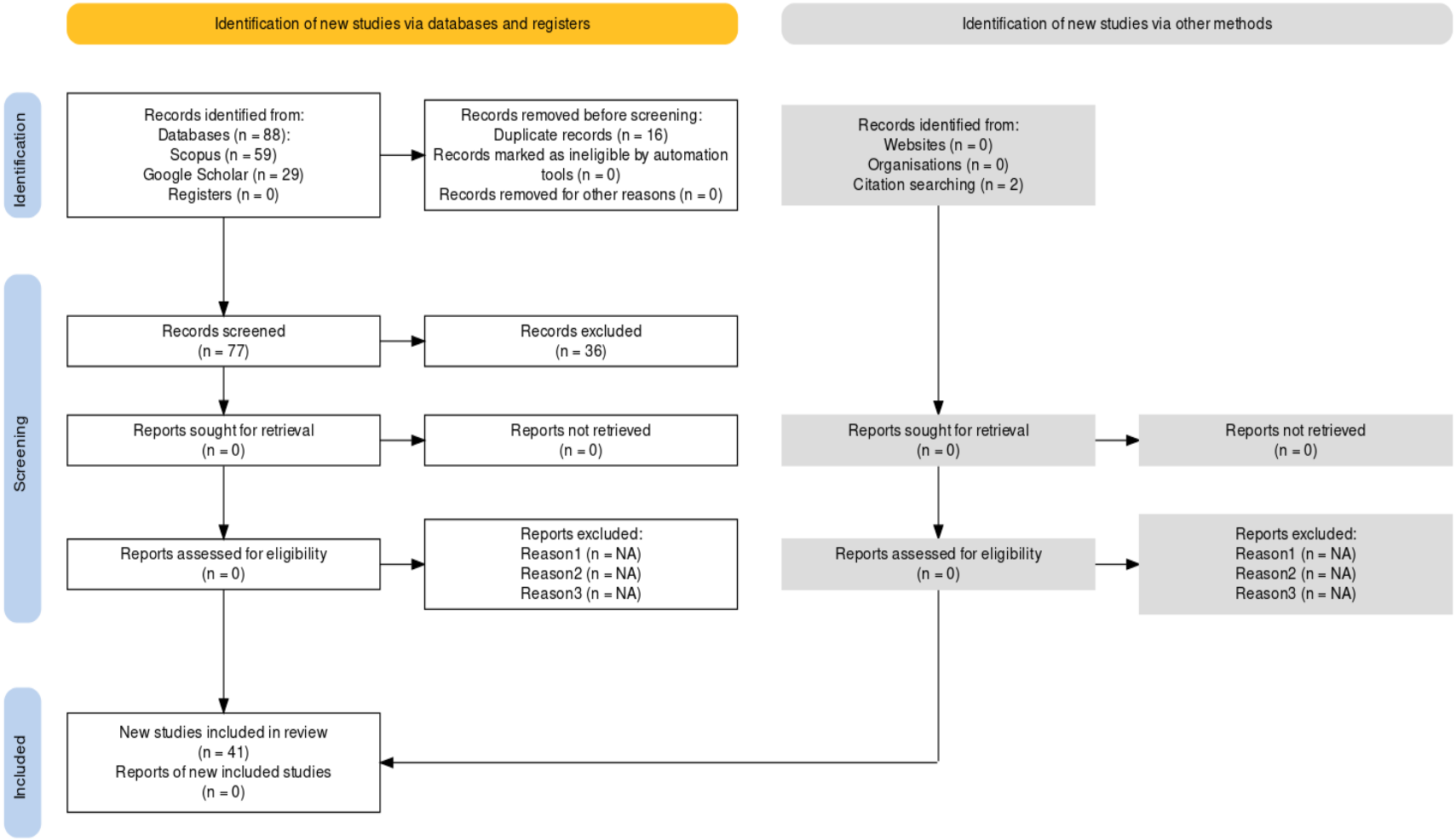


Figure 3.1: Review methodology

3.2.1 Identification

To guide the review and ensure its alignment with the broader research objectives, the systematic literature review was structured around a core research question. The research questions in Table 1 formed the query strings of this literature search. The research question was tokenized, and relevant keywords were extracted and used as the search criteria against 2 literature databases: *Scopus* and *Google Scholar*. This question was formulated to explore the breadth of state-of-the-art deep learning models applied in the context of sewer inspection. The motivation behind this inquiry was two-fold: first, to identify dominant architectural paradigms and second, to uncover prevailing trends, methodological challenges, and performance bottlenecks reported in recent studies. This investigation was essential to establish the technical landscape from which this thesis would later select, evaluate, and deploy appropriate models for sewer defect detection.

Table 2.1: Research questions

Research Question	**Motivation**
What are the deep learning algorithms and architectures used in automatic defect detection in sewers?	To explore and understand the current state-of-the-art deep learning algorithms employed in sewer defect detection.
	To identify the trends and challenges in using different deep learning algorithms.

To retrieve studies relevant to this research question, a set of carefully constructed search strings were applied across the selected academic databases. These strings were matched against keywords, titles, and abstracts to maximise recall without sacrificing precision. Boolean operators were used to combine terms such as “deep learning,” “object detection,”, “sewer inspection,” “pipeline defect,” and “automated visual inspection.” This ensured that the search captured a comprehensive and relevant set of publications spanning multiple model types and application domains within infrastructure monitoring. The goal was to assemble a sufficiently diverse and representative sample of literature to support a robust and generalisable analysis of model development techniques used in the sewer defect detection domain. Below is the full search strings used.

(("automatic sewer inspection" OR "defect classification" OR "fault classification" OR "defect detection" OR "fault detection" OR "defect segmentation" OR "condition assessment")
AND ("sewer" OR "Wastewater") AND ("deep learning" OR "CNN"))

3.2.2 Screening

Following the initial search across selected databases, duplicate records were identified and removed, resulting in a refined pool of 77 potentially relevant studies. These papers were then subjected to a systematic screening process based on predefined inclusion and exclusion criteria, as summarised in Table 4. To ensure alignment with the research objectives, inclusion criteria required that papers explicitly address the application of deep learning algorithms for automatic defect detection in sewer systems. Studies were only considered if they were published between 2010 and 2024, written in English, peer-reviewed, and categorised as journal articles or conference proceedings in their final publication stage. This filtering ensured that the dataset reflected both methodological rigour and relevance to the research domain.

Table 3.2: screening parameters for inclusion/exclusion criteria definition

Parameters	Search Values	Search Code
Publication year	2010 - 2024	PUBYEAR
Document Type	Conferences, Journal	DOCTYPE
Publication stage	Final	PUBSTAGE
Language	English	LANGUAGE

During the screening process, several papers were excluded for not meeting these standards. Specifically, two studies were removed due to incompatible document types, as they were either preprints or incomplete reports, while one study was excluded due to being published in a non-English language. This vetting process helped maintain a high standard of quality and ensured that the final selection of studies provided direct insights into the technical design, training pipelines, and evaluation of deep learning models specifically applied to sewer defect detection.

3.3 Inclusion

After the initial screening phase, 57 studies were identified as potentially relevant based on a review of their titles and abstracts. These studies were subjected to further scrutiny through full-text analysis, during which each paper was assessed against a refined set of criteria aligned with the PRISMA checklist. Evaluation parameters included the clarity of the research background, the robustness of the methodology, the relevance and depth of case studies, and the transparency of results and conclusions. Through this secondary filtering, only studies that

presented substantive empirical contributions to deep learning-based sewer defect detection were retained. This led to a provisional inclusion of 48 papers for detailed review.

However, despite meeting the eligibility requirements, seven studies (Kumar et al., 2020c; Moradi et al., 2018; Kumar et al., 2018; Wang et al., 2020; Wang et al., 2023; Zhou et al., 2021c)) could not be accessed in full text at the time of this review and were therefore excluded. As a result, the final pool of primary studies consisted of 41 peer-reviewed papers. These were systematically reviewed, and relevant data points were extracted, including the type of deep learning architecture employed, the type of defect explored, annotation strategy, pre-processing techniques employed, performance metrics, and major outcomes or contributions. The findings were then categorised to identify prevailing patterns, methodological trends, and knowledge gaps, forming the analytical backbone for subsequent sections of this chapter.

3.3 Overview of Deep Learning Applications in Sewer Defect Detection

Deep learning, a subfield of machine learning, includes a set of algorithms and techniques modelled around how human cognitive works. This involves learning and extraction of high-level features from complex datasets. This has been applied in many areas such as autonomous vehicle (Ni et al., 2020), construction (Mostafa & Hegazy, 2021; Yu et al., 2022), healthcare (F. Jiang et al., 2017; Liang et al., 2014), and manufacturing (Bartsch et al., 2021; Y. Chen et al., 2021) among others. Its success has increased its adoption for defect detection problems. At its core is a popular algorithm, Convolutional Neural Network (CNN), which has gathered much success in image-based problems (Gupta et al., 2018; Long et al., 2017). In the context of sewer defect detection, deep learning algorithms can be trained on large volumes of labelled data, primarily to recognize and classify different types of defects such as cracks, deposit, root intrusions, and hole. Unlike conventional ML method, automatic feature extraction and multi-label problems are well handled end-to-end using this algorithm, hence, it further aids the adoption of deep learning for defect detection in recent years.

By leveraging deep learning capabilities, automatic extraction of relevant features can help to effectively identify subtle defects that may go unnoticed in manual inspections, thereby improving the accuracy and reliability of the defect detection process. At the early stage, DL applications to AuDD problems were formulated as a classification problem (Hawari et al., 2020), which focused majorly on the existence of defect in a frame, irrespective of the defect severity or location in the pipe. However, to gain more understanding of the defect condition and specific location, deep learning has evolved from classical defect classification to detection (classification plus localization) and segmentation (M. Wang, 2019), using state-of-the-art (SOTA) detection algorithms. Overall, the ability of deep learning algorithms, such as CNNs, object detection methods, and segmentation networks, to learn complex patterns and features

from labelled data enables accurate and efficient identification of sewer defects(Tscheikner-Gratl, Caradot, …, et al., 2019). By automating the defect detection process, deep learning algorithms can improve the overall efficiency of sewer maintenance operations, thereby reducing manual inspection efforts, and aid in more proactive and targeted interventions.

This review explores the current state of deep learning algorithms in AuDD, by identifying various deep learning techniques and architectures employed in this domain while simultaneously looking at their pros and cons. The following sections will delve into specific review methodology, deep learning techniques, performance evaluation measures, and challenges associated with automatic defect detection in sewers, thereby providing a comprehensive overview of the current state-of-the-art in AuDD.

3.4 Deep Learning Algorithms for Defect Detection in Sewers

Due to the diverse requirements of AuDD in sewer infrastructure, a range of deep learning algorithms has been employed or proposed in recent literature, each addressing the task through different problem formulations such as classification, object detection, and image segmentation. These categories reflect varying levels of complexity and spatial precision in defect localisation and characterisation. Classification approaches treat the problem as identifying the presence or absence of specific defect types in an image (Haurum et al., 2021b); detection-based models extend this by not only recognising defects but also locating them via bounding boxes (Arjun et al., 2024); while segmentation techniques provide pixel-level accuracy, offering fine-grained delineation of defect boundaries (Fang et al., 2022a). The choice of algorithmic strategy often depends on the application context, dataset characteristics, and computational constraints. The following sub-sections provide a structured review of the most prominent deep learning architectures within each category, highlighting their design, typical use cases, and relevance to sewer defect detection tasks.

3.4.1 Defect classification

At the early stage, deep learning applications to AuDD problems was formulated as a classification problem which focused majorly on the existence of defect in a frame, irrespective of the defect severity or location in the pipe. At the core of defect classification is a popular algorithm, Convolutional Neural Network (CNN), which has gathered much success in image-based problems (Gupta et al., 2018; Long et al., 2017). Unlike conventional ML method, automatic feature extraction and multi-label problems are automatically handled using CNNs. Consequently, it further aids its adoption for defect classification in recent years.

Convolutional Neural Networks (CNNs)

CNN is a type of deep neural network algorithm that has gained much traction in AuDD due to its effectiveness on problems involving grid-like dataset, such as image classification (Li et

al., 2021b; Wang et al., 2021d). Similarly, it is a popular feed-forward neural network that builds a hierarchical layer of interconnected weighted neurons through its hidden layers to produce an output. The architecture of CNN is divided into four dependent stages: *convolution, activation, pooling,* and *fully connected layer*.

At the convolution layer, a *filter* (also called the kernel, weight matrix, or the receptive field) slides through the pixels of the input image from left to right in a bid to extract relevant features from the image by conducting an element-wise dot operation of the filter weights against each subarray of the image pixel. As shown in Figure 3.2, this dot operation results in a stack of *feature maps* $(x * w)$, which is a matrix-like representation of the relevant features extracted from the image. The weight of the filter is randomly generated at the initial stage, and we a *stride* (i.e., the vertical movement of the filter across the pixels) of one is set to aid in more feature extraction. Figure 3.2 show the processes within a typical CNN model during defect classification.

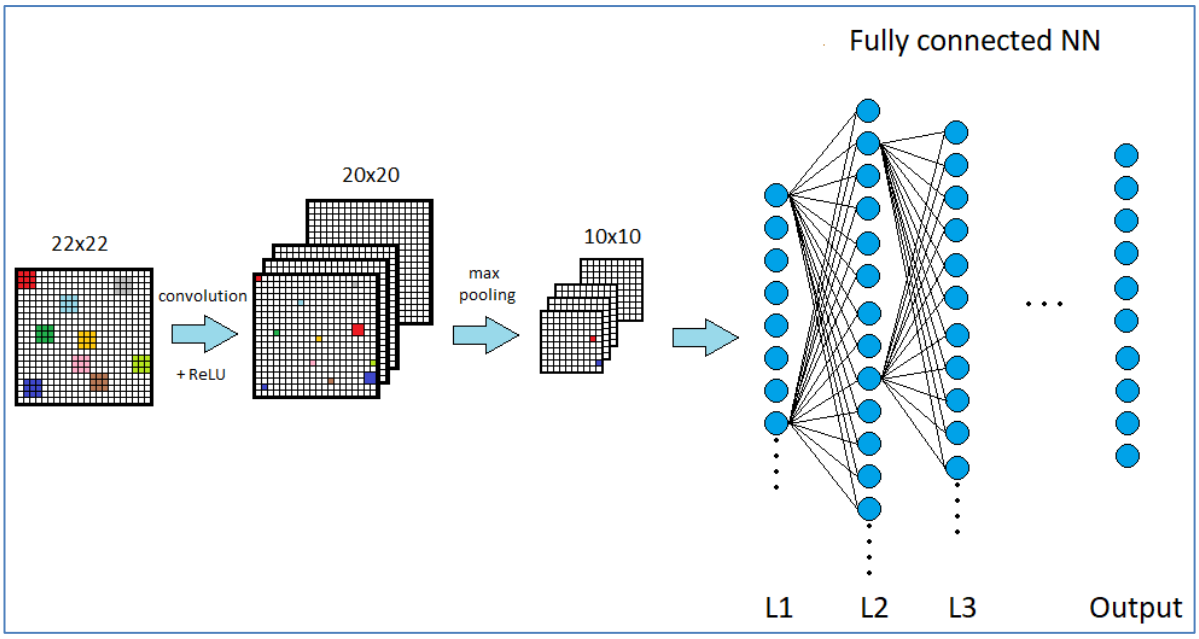


Figure 3.2: Convolution layer in CNN. Adapted from (Kamil, 2023)

Since images are non-linear data, a common way to add non-linearity is the activation function on feature maps. Rectified Linear Unit (ReLU) (Javid et al., 2021) is a popular activation function that takes each detail the network is looking at and decides if it's important or not. If a detail is confusing or negative (like noise), it simply ignores it, setting it to zero. This helps the network focus only on useful information, making it better at recognizing patterns and understanding what's in the image. Other common activation function includes *sigmoid, softmax,* and *tanh.* The popularity of ReLU is due to its low time complexity in training the network without serious consequence to generalisation accuracy (Javid et al., 2021).

Simultaneously, the pooling layer is introduced between subsequent convolution layers to progressively reduce the spatial size of the image. Pooling is like summarizing the important parts of an image to make it smaller without losing what matters. This is like looking at a group of pixels and picking just the brightest one or the average value. Thus, it helps the computer to understand a simpler version of the image and speeds up the learning process, while keeping the relationships between parts of the image. Max pooling is common in image classification for pooling the maximum value from the pixel subarray in the feature maps (Ajit et al., 2020; Mushtaq & Su, 2020). Other pooling techniques include *average pooling* and *L2 norm.* Figure 3.3 illustrate how max pooling reduces the spatial size of feature maps.

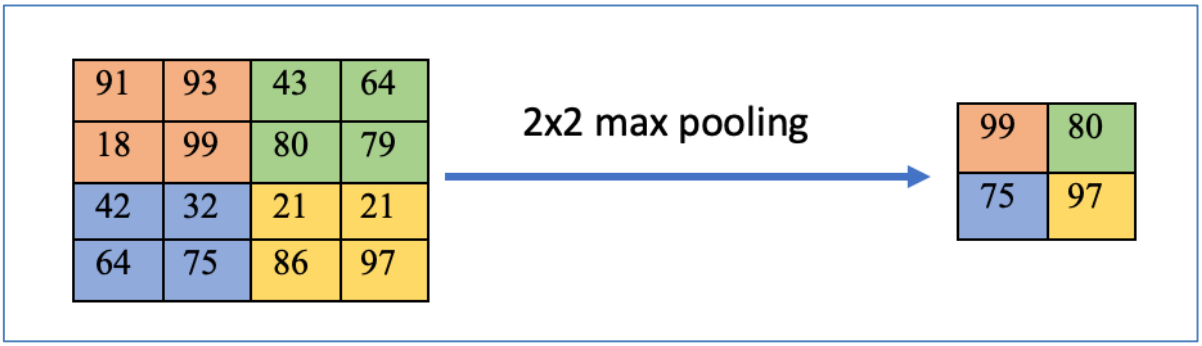


Figure 3.3: Max pooling of feature maps (source: Author)

Pooled feature maps are usually in a multi-dimensional array of features. This must be transformed into a form that is executable by the fully connected layer. To achieve this, *flattening* is employed in converting multi-dimensional array into a sequential 1D array (i.e., a long vector) to prepare these features for further processing, as seen in Figure 3.4.

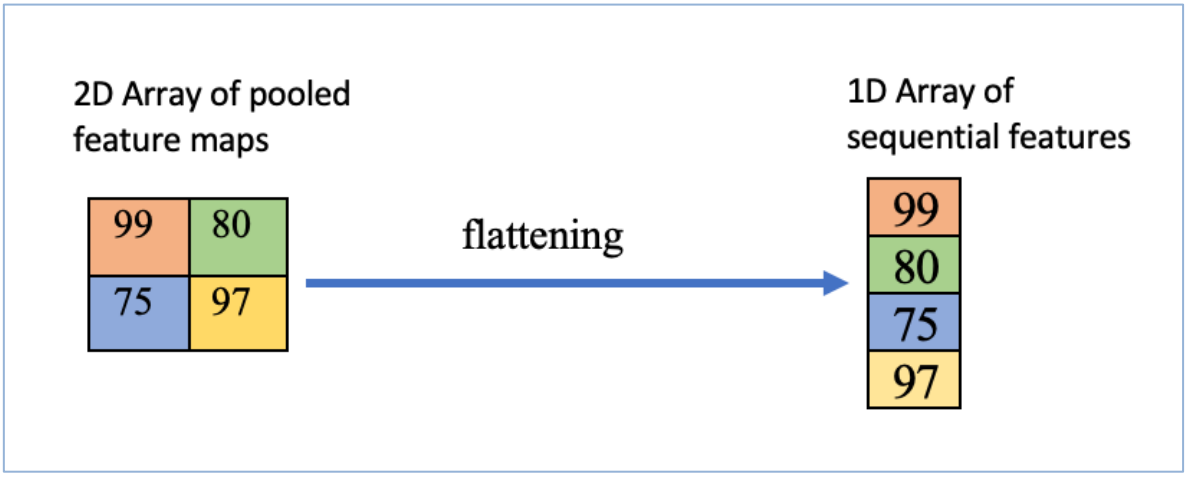


Figure 3.4: Flattening pooled feature maps (source: Author)

The fully connected layer is a typical Multilayer Perceptron (MLP) sitting at the edge of the architecture. Individual values in the flattened array are treated as separate features that denote the images, by combining the features into more attributes. This attribute is used to predict the class of an image (i.e., defective or non-defective) by using a classifier at the output layer. Likewise, the error rate is calculated at this layer and sent backward (backpropagation) to the network to adjust the weight and the feature detectors to improve and optimise the model's performance, as depicted in Figure 3.5. This backpropagation is conducted several times to predict the class of an image using the features, a linear and non-linear transformation.

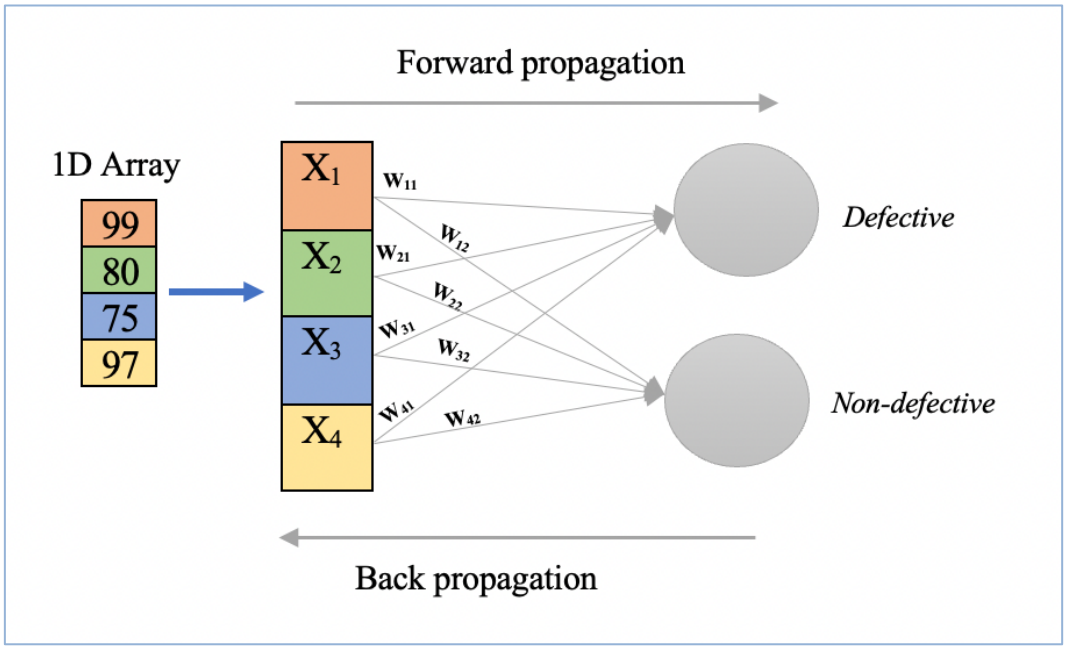


Figure 3.5: Fully connected layer (source: Author)

Owing to this algorithm, several studies have attempted the defect classification in sewers, Kumar et al. (2018) proposed a framework for developing a multi-layered CNNs specifically for detecting different defect type, however, this approach is not practical in real word scenario based on computational cost of each CNN. Additionally, stack of consecutive convolution and pooling in a typical CNN has been reported to be largely down sampling the feature maps with a huge factor, thereby leading to large spatial information loss and subsequently generating feature maps with low resolutions (Wang & Cheng, 2019a). Consequently, different modifications and adaptations to CNN have been proposed to mitigate this problem. However, recent studies using deep learning method has shifted their focus from classification to detection and generally defined the problem as either *detection* (i.e., classification and localization) or *segmentation*.

3.4.2 Defect detection

Defect detection or localization using deep learning algorithms is a more granular technique that classify a defect not only into a predefined class but correspondingly draws a bounding box around the defect location on the frame, while also predicting the confidence score as seen in Figure 3.6. While its adoption in image-based problems has become more popular in many areas (Long et al., 2017; Ni et al., 2020), it has also been employed in sewer defect detection to extract more specific information about defect location using variants of CNN and object detection networks.

To experimentally apply a deep learning framework in detecting structural defect such as fractures. Kumar & Abraham, (2019) proposed a deep learning detection model using a multi-layered CNN and YOLO detection network. They used the multi-layered CNN as a classifier to separate defective from non-defective images, while attempting to localize the defect area only in the defective images using YOLO because only smaller percentage of the sewer is usually defective. This way, computational cost is reduced by eliminating unnecessary processing of non-defective images by the YOLO network.

In a similar study, to further reduce the number of redundant images to be processed, (Khan et al., 2020) proposed extracting only the *nth* frame from a sewer video, where the frequency rate (i.e., FPS) of the video determines the value of *n*. For example, if a survey video is captured at 25 images per second, only the 25^{th} image is extracted from the video since the difference with the other 24 frames are mostly insignificant. However, a more robust approach was taken to solve this problem by Wang et al., (2021), where they proposed defect detection and tracking in sewer pipes using Faster R-CNN and metric learning. This is primarily to track defective objects and eliminate analogous and/or recurring defect across consecutive frames in a video.

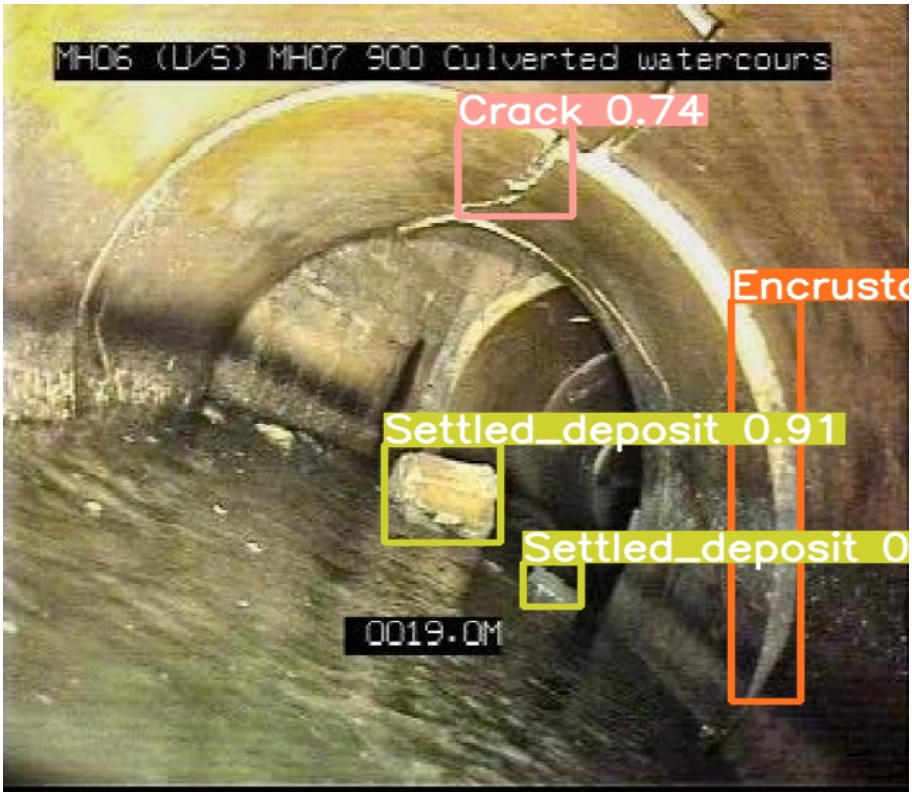


Figure 3.6: Defect localization with bounding boxes (source: Research Data)

While having an excellent detection accuracy is vital in building required models for sewer defect detection, providing more relevant features for the algorithm to learn from is more important. In this regard, Li et al. (2021) used Faster R-CNN with feature fusion to detect defect in sewer pipes by concatenating the proposed region features and the global contextual features from the corresponding image which provides the algorithm with additional features to learn from. There is obviously a trend of a choice between YOLO and Faster R-CNN as the network of choice, however, other detection network that is worth mentioning includes Single Shot Detector (SSD) (Wei et al., 2016), R-CNN (Chenyi et al., 2017), MobileNet (Younis et al., 2020), and SqueezeNet (Iandola et al., 2016) among others.

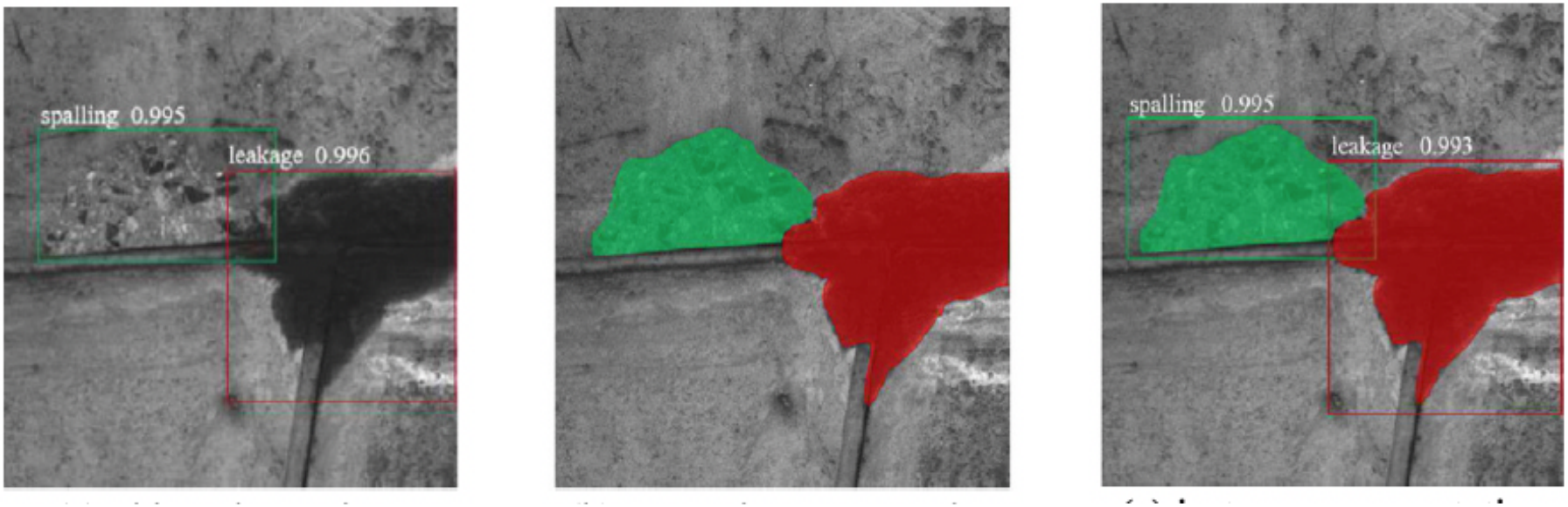
3.4.3 Defect Segmentation

Defect segmentation provides a more granular details about pixel level information of a defect by annotating or masking the affected pixel. Previous segmentation approaches use mathematical morphology by employing techniques such as open top-hat operation (OTHO), closing bottom-hat operation (CBHO), and the more popular morphological segmentation based on edge detection (MSED). While MSED has achieved more success in detecting cracks and leaking (Su et al., 2011b; Su & Yang, 2014), OTHO and CBHO have demonstrated a good prospect in open joint (Tung-Ching, 2015). Generally, these segmentation methods require an

additional feature extractor such as SIFT or HOG to extract relevant features from small areas of the input image (Halfawy & Hengmeechai, 2014a). While this segmentation approach is effective in detecting defects in images with high contrast edges, the feature extractors introduce additional computational complexity that is avoidable (Thoma, 2016). Additionally, it requires manual parameter tuning to achieve optimum performance, hence, not suitable for task with many parameters. Moreover, these segmentation methods are most suitable for identifying a single defect in a frame (i.e., single label per frame). Consequently, a more robust and low-complexity segmentation method suitable for multi-label problems became increasingly necessary.

Variants of deep learning algorithms such as CNN, coupled with Region Proposal Networks (RPNs) have shown more capability not only to accurately perform pixel-level detection but implementing it without additional feature extractors like previous methods. Such algorithm formulates pixel-level problem as either semantic or instance segmentation. As depicted in Figure 8, while semantic segmentation basically annotates or mask defective pixels for more visibility of the defect area, instance segmentation classifies the defect, annotate the defective pixels, and simultaneously draw a bounding box around the defect.

Both methods are applicable to multi-label problem by training the model with different classes of pre-annotated frames, where features are automatically extracted and learned. (M. Z. Wang & Cheng, 2019b) proposed a deep CNN called DilaSeg, for segmenting cracks, tree root, and deposit. It comprises a dilated convolution to up-sample the convoluted feature maps, thereby improving the generic segmentation of sewer pipes. A further work by the same authors a year later birthed DilaSeg-CRF, which fully integrated a deep CNN with dense Conditional Random Field (CRF) by step-wisely formulating the CRF as a Recurrent Neural Network (RNN)



(a) Object detection (b) Semantic segmentation (c) Instance segmentation

Figure 3.7: Object detection vs Segmentation (Xu et al., 2021d)

Obviously, both instance and semantic segmentation using deep learning is still at its early stage in sewer defect detection, with fewer research adopting this method so far, however, (Xu et al., 2021d) offered a unique detection method by using Mask R-CNN to detect and segment leakages and spalling in concretes. In their method as shown in Figure 3.8, they employed Path Augmentation Feature Pyramid Network (PAFPN) at the backbone network to generate multi-scale features rich in localization information and passing it to the RPN to generate region proposals. They reported PAFPN has been useful in aiding smooth information distribution within the backbone network of Mask R-CNN, thereby improving the accuracy of the network. However, they noted the extra memory and computational overhead to be incurred in adopting this method.

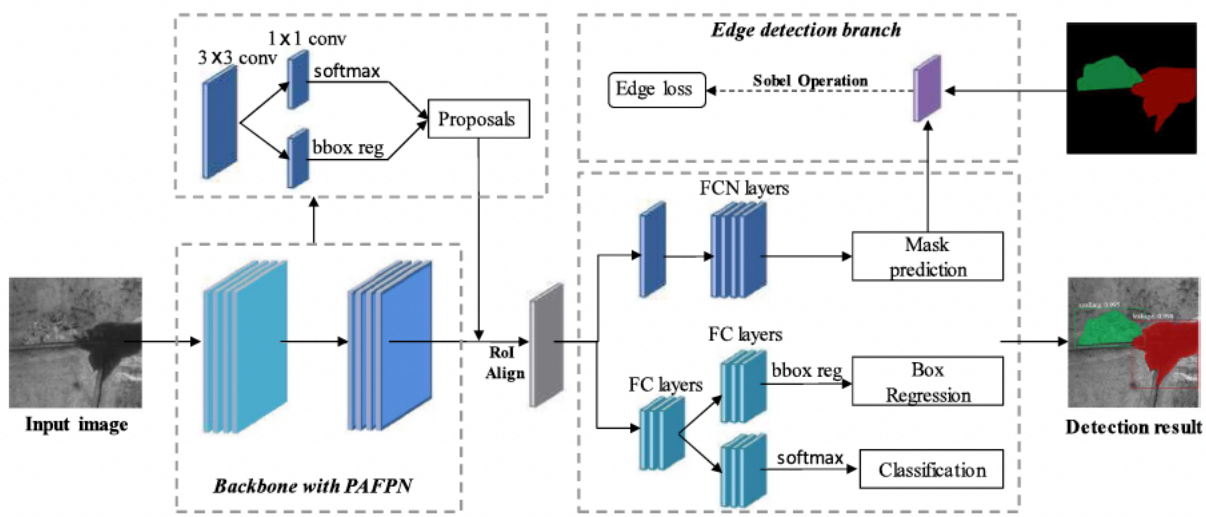


Figure 3.8: Mask R-CNN segmentation architecture using PAFPN (Xu et al., 2021a)

Since deep learning algorithms have demonstrated continuous success in many areas, they have achieved state-of-the-art result in sewer defect detection as they are able to learn complex defect patterns. Nonetheless, there are still some drawbacks in its adoption. First, it requires huge amount of data for effective model development, which directly leads to increment in the computational cost of running such models and thereby elevate the minimum standard for hardware compatibility. Similarly, data with low quality have reportedly perform poorly with this approach. Furthermore, preparing training data for such models requires annotating huge number of images using annotating tools like *LabelImg* (TzuTa Darren, 2022) *and makesense.ai* (Piotr, 2023), which will form the basis of the ground truth, hence, it requires a lot of time to complete data annotation. Lastly, specialized expertise may be required to understand and implement such method, hence, experimentation and repeatability of this approach in research environment might be difficult.

3.5 Analysis of Review

To gain a deeper understanding of the methodological landscape and technical variations across the reviewed literature, a structured synthesis was conducted on all 41 primary studies selected through the systematic review process. Each study was analysed across a predefined set of attributes to capture not only the modelling approaches but also the contextual, architectural, and experimental details that influence replicability and performance. This multi-attribute framework supports the identification of trends, patterns, and methodological gaps across the body of work. The extracted attributes included country of study, publication year, and the scope of the task, whether focused on defect detection or broader condition assessment.

The nature of the problem was categorised into classification, object detection, or segmentation, and each model was mapped to its corresponding detection architecture. Additional details such as the specific algorithm used, defect types addressed, dataset source and sizes, and pre-processing techniques were also recorded to assess experimental rigour and domain fit. Other important factors included image resolution, annotation tools, and frameworks used (e.g., TensorFlow, PyTorch), alongside whether edge detection or segmentation was applied. Likewise, the presence or absence of data augmentation was clearly specified.

Furthermore, the studies were evaluated on whether they addressed defect severity, the metrics used for evaluation (e.g., accuracy, recall), and whether the reported results achieved state-of-the-art (SOTA) performance within their respective tasks. These extracted variables are summarised in Table 3.2, which provides a comprehensive side-by-side comparison across all reviewed studies. Due to the number of extracted insights which cannot fit into the whole table below, only some were selected. To view the full table, see Appendix B.

Table 3.2: Summary of primary studies

Authors	Title	Problem Type	Architecture	Algorithm	Defect type	Pre-processing technique	Evaluation metrics
Hassan et al (2019)	Underground sewer pipe condition assessment based on convolutional neural networks	Classification	CNN	CNN	longitudinal defect, debris silty, joint faulty, open joint, surface damage, protruding lateral	image extraction, grayscale conversion, histogram equalizer, binarization, morphological operations, edge detection	accuracy
Fang et al. (2022)	Sewer defect instance segmentation, localization, and 3D reconstruction for sewer floating capsule robots	Semantic segmentation	Mask R-CNN	Mask R-CNN, YOLOv3, SSD	crack, break, deformation	image extraction, annotation	AP, mAP, IOU
Meijer et al. (2019)	A defect classification methodology for sewer image sets with convolutional neural networks	Classification	CNN	CNN	Fissure, Surface damage, Intruding connection , Defective connection, Intruding sealing material , Displaced joint.	oversampling	AUROC, AUPR, specificity at sensitivity, precision at recall
Khan et al (2019)	Robotics and Deep Learning Framework for Structural Health Monitoring of Utility Pipes	Classification	CNN	CNN	Crack	grayscale conversion, image resizing,	accuracy
Kumar et al (2019)	Automated defect classification in sewer closed circuit television inspections using deep convolutional neural networks	Classification	CNN	CNN	tree root, deposits, cracks, infiltration, debris, connections, material change	image extraction	accuracy, precision, and recall

Authors	Title	Problem Type	Architecture	Algorithm	Defect type	Pre-processing technique	Evaluation metrics
Pan et al (2020)	Automatic sewer pipe defect semantic segmentation based on improved U-Net	Semantic segmentation	PipeUnet	U-Net	crack, joint offset, infiltration and intruding lateral	image resizing, annotation	MIoU
Cheng &Wang (2018)	Automated detection of sewer pipe defects in closed-circuit television images using deep learning techniques	Detection	Faster R-CNN	Faster R-CNN	tree root, cracks, infiltrations and deposits as	image resizing, annotation	mAP, missing rate, detection speed and training time
Wang et al (2019)	Semantic segmentation of sewer pipe defects using deep dilated convolutional neural network	Semantic segmentation	DilaSeg	CNN	cracks, deposits, tree root	image extraction, annotation	AP, mAP, MIoU, fwIoU
Hu et al (2023)	Toward Purifying Defect Feature for Multilabel Sewer Defect Classification	Classification	SPM+CNN	CNN	sewer-ml defect	-	F1, F2
Yin et al (2020)	A deep learning-based framework for an automated defect detection system for sewer pipes	Detection	YOLO	YOLOv3	broken, hole, deposits, crack, fracture, tap, and root	image extraction, annotation	mAP
Chen et al (2019)	Automated detection of sewer pipe defects based on cost-sensitive convolutional neural network	Classification	CNN	CNN	-	image extraction, image resizing	FNR, ATC

Authors	Title	Problem Type	Architecture	Algorithm	Defect type	Pre-processing technique	Evaluation metrics
Dang et al (2022)	DefectTR: End-to-end defect detection for sewage networks using a transformer	Detection	DETr	DETr	broken pipe, longitudinal crack, circumferential crack, debris silty, displaced joint, faulty joint, separated joint, protruding lateral, surface damage, root intrusion.	histogram equalization, GCANET	mAP, precision, recall
Dang et al (2021)	Deep learning-based sewer defect classification for highly imbalanced dataset	Classification	CNN	CNN	crack, debris silty, joint faulty, open joint, surface damage, protruding lateral, pip broken	-	accuracy
Wang et al (2021)	Automated sewer pipe defect tracking in CCTV videos based on defect detection and metric learning	Detection	Faster R-CNN	Faster R-CNN	fracture, root intrusion, protruding lateral	-	mAP, IoU
Li et al (2021)	Sewer pipe defect detection via deep learning with local and global feature fusion	Detection	CNN	CNN + SRPN	Barrier, Deposition, Distortion, Fraction and Foreign-body	histogram equalization, bilateral filter	Recall
Zhou et al (2022)	Sewer defect detection from 3D point clouds using a transformer-based deep learning model	Classification	CNN	CNN	brick, rubber ring, and displacement	-	precision, recall, F1-score
Kumar & Abraham (2019)	A Deep Learning Based Automated Structural Defect Detection System for Sewer Pipelines	Detection	YOLO	CNN	fractures, root intrusion, deposits	-	mAP

Authors	Title	Problem Type	Architecture	Algorithm	Defect type	Pre-processing technique	Evaluation metrics
Haurum et al (2021)	Sewer defect classification using synthetic point clouds	Classification	DGCNN	DGCNN + PointNet	displaced joints, defective rubber rings, and obstacless	Statistical Outlier Removal (SOR)	F1
Siu et al 2022	A framework for synthetic image generation and augmentation for improving automatic sewer pipe defect detection	Detection	Faster R-CNN	Faster R-CNN	cracks, deposits, or tree roots	style transfer	AP
Li et al (2019)	Sewer damage detection from imbalanced CCTV inspection data using deep convolutional neural networks with hierarchical classification	Classification	Resnet	CNN	deposits settlement, joint offset, broken, obstacles, water level stag, and deformation	oversampling, image image resizing	AUC
Ma et al (2021)	A multi-defect detection system for sewer pipelines based on StyleGAN-SDM and fusion CNN	Detection	StyleGAN	CNN	misalignment, leakage, corrosion, obstacle	image resizing, annotation	accuracy
Chen et al (2018)	An intelligent sewer defect detection method based on convolutional neural network	Classification	CNN	CNN	Deposit, obstacle, blur, intrusion	-	accuracy
Tan et al (2021)	Automatic detection of sewer defects based on improved you only look once algorithm	Detection	YOLO	YOLOv3	crack, deposit, root, infiltration	-	recall rate, AP, mAP

Authors	Title	Problem Type	Architecture	Algorithm	Defect type	Pre-processing technique	Evaluation metrics
Wang et al (2023)	PIPE-CovNet: Automatic In-Pipe Wastewater Infrastructure Surface Abnormality Detection Using Convolutional Neural Network	Classification	PIP-CovNET	CNN	cracks, breaks, and collapses; intruding sealing material; roots; settled deposits; attached deposits; connection with construction changes; surface damage; production error; deformation;	-	accuracy, F1
Wang & Cheng (2020)	A unified convolutional neural network integrated with conditional random field for pipe defect segmentation	Semantic segmentation	DilaSeg-CRF	RNN + CRF	cracks, deposits, and tree root,	-	MIoU
Oh et al (2022)	Robust Sewer Defect Detection with Text Analysis Based on Deep Learning	Detection	YOLO	YOLOv3	Crack	image resizing, annotation	precision, recall, mAP, FPS
Situ et al (2021)	Automated Sewer Defects Detection Using Style-Based Generative Adversarial Networks and Fine-Tuned Well-Known CNN Classifier	Classification	StyleGAN	CNN	disjoint (DJ), obstacle (OB), residential wall (RW) and tree root	image resizing	mAP
Haurum et al (2022)	Multi-scale hybrid vision transformer and Sinkhorn tokenizer for sewer defect classification	Classification	Resnet	CNN	sewer-ml defect	-	F1, F2

Authors	Title	Problem Type	Architecture	Algorithm	Defect type	Pre-processing technique	Evaluation metrics
Li et al (2022)	A robust instance segmentation framework for underground sewer defect detection	Instance segmentation	Pipe-SOLO	SOLOv2	open joint, faulty joint, protruding lateral, crack, broken pipe, and surface damage	image extraction, defogging, dehazing, annotation.	mAP
Xie et al (2019)	Automatic Detection and Classification of Sewer Defects via Hierarchical Deep Learning	Classification	CNN	CNN	deposition, stagger, fracture, high water level, disjunction, barrier	image resizing, normalization	F1
Wang et al (2021)	Towards an automated condition assessment framework of underground sewer pipes based on closed-circuit television (CCTV) images	Semantic segmentation	Faster R-CNN	Faster R-CNN + SSD + YOLO	crack, root, deposit	grayscale conversion, image resizing, edge detection, non-maximum supression	AP
Zhou et al (2022)	Comparison of classic object-detection techniques for automated sewer defect detection	Detection	Faster R-CNN	Faster R-CNN + YOLOv2	crack, displaced joint, tree root, obstacle	image resizing, annotation	AP
Ma et al (2022)	Training a Lightweight CNN Model for Fine-Grained Sewer Pipe Cracks Classification Based on Knowledge Distillation	Classification	Resnet	Resnte50 + CMI	crack	-	F1
Zhang et al (2023)	Automatic Detection Method of Sewer Pipe Defects Using Deep Learning Techniques	Detection	YOLO	YOLOv4	crack, deposition, root and stagger	image resizing, annotation	mAP

Authors	Title	Problem Type	Architecture	Algorithm	Defect type	Pre-processing technique	Evaluation metrics
Zhao et al (2023)	Towards Trustworthy Multi-Label Sewer Defect Classification via Evidential Deep Learning	Classification	TMSDC	CNN + EDL	sewer-ml defect	-	F1
Ma et al (2022)	Automatic defogging, deblurring, and real-time segmentation system for sewer pipeline defects	Instance segmentation	Pipe-Defog-Net	Mask R-CNN + SOLO + SOLOv2 + FCN	misalignment, obstacle, and fracture	-	mAP
Dang et al (2023)	Lightweight pixel-level semantic segmentation and analysis for sewer defects using deep learning	Semantic segmentation	DeepLabV3+	DeepLabV3+ + Resnet-152	broken pipe, crack, displaced joint, lateral, surface damage, root intrusion, hole, obstacles	image resizing, annotation	mAP, IoU
Li et al (2022)	Intelligent identification and classification of sewer pipeline network defects based on improved RegNetY network	Classification	RegNetY + Grad-CAM	RegNetY + Grad-CAM	root, disjointedness, deformation, deposit, rupture, misalignment, scaling, and interface material failure	Image enhancement Canny edge detection Closed operation to obtain mask graph Image restoration	F1, Recall
Situ et al (2023)	Real-time sewer defect detection based on YOLO network, transfer learning, and channel pruning algorithm	Detection	YOLO	YOLOv5	disjoint, obstacle, residential wall, tree root	-	precision, recall, mAP, and inference time, IoU

3.5.1 Keyword analysis

To complement the systematic review, a keyword frequency analysis was conducted across the titles and abstracts of the 41 primary studies. The goal was to uncover dominant themes and terminologies shaping the research landscape in automated sewer defect detection. As expected, the most frequently occurring terms included "sewer", "defect", "detection", and "deep learning", reflecting the central focus of the field on applying deep learning to sewer defect detection. Other frequently cited keywords such as "pipe", "classification", and "automatic" highlight the common problem formulations and data modalities in use.

To visually represent these patterns, a word cloud (Figure 3.9) was generated programmatically using a custom Python script utilizing the `wordcloud` library. The methodological process involved compiling the text from the titles of the reviewed studies, filtering out standard English stop words (such as 'the', 'and', 'in'), and calculating the frequency distribution of the remaining core terms. In the resulting visualization, larger font sizes represent a higher frequency of occurrence, providing an at-a-glance view of key concepts emphasised across the literature. This type of analysis reinforces the alignment between the reviewed studies and the research objectives of this thesis, especially in identifying which modelling paradigms and terminologies dominate the discourse on deep learning-based defect detection in sewer systems.

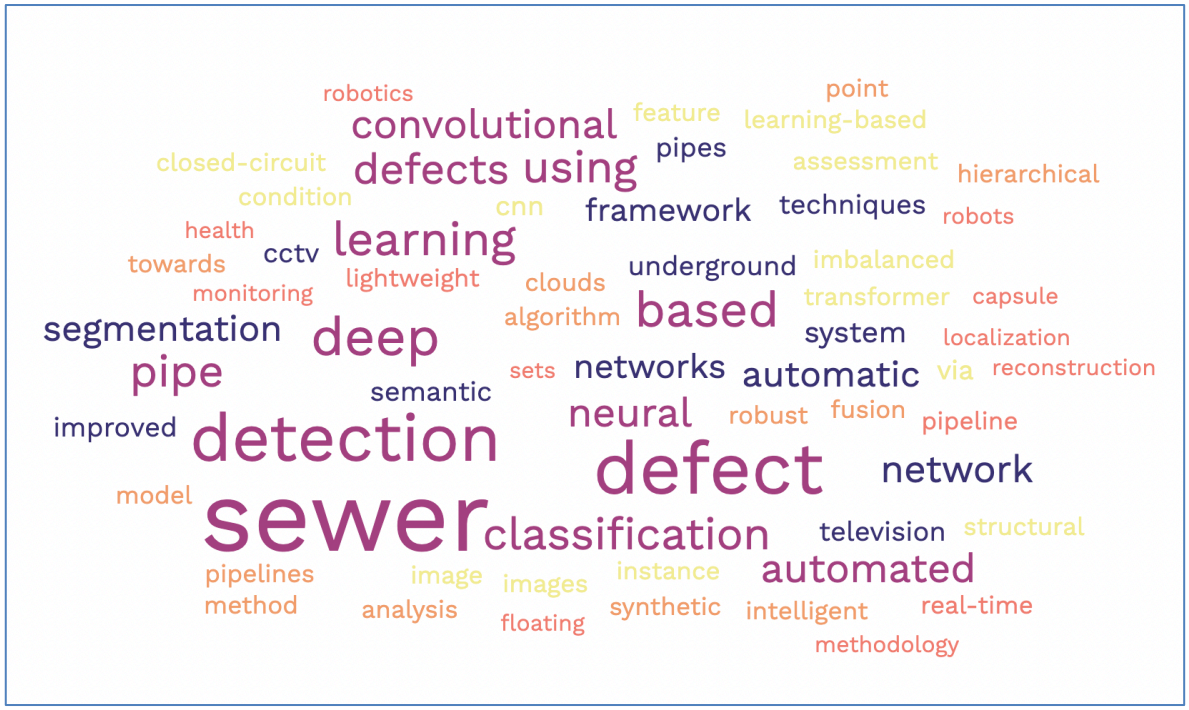


Figure 3.9: Word cloud for keyword ranking (source: Author)

3.5.2 Publication Year

Evolving from conventional ML approach, application of deep learning techniques to AuDD was realised starting from 2018, including studies such as (Cheng & Wang, 2018b; S. S. Kumar et al., 2018). Even at this stage, majority of the studies focused on defect classification. As represented in Figure 3.10, there was a progression in studies using deep learning from 2018 to 2019, however, 2020 saw witnessed a fall of this progression to the level of the maiden year, which could be due to the impact COVID19. 2022 saw the highest number of studies applying deep learning to AuDD, although studies from the current year (i.e., 2024) is still inconclusive as more studies are expected to be published. Overall, there has been a positive progression in the application of deep learning to AuDD, which span different approach such as classification, detection, and segmentation.

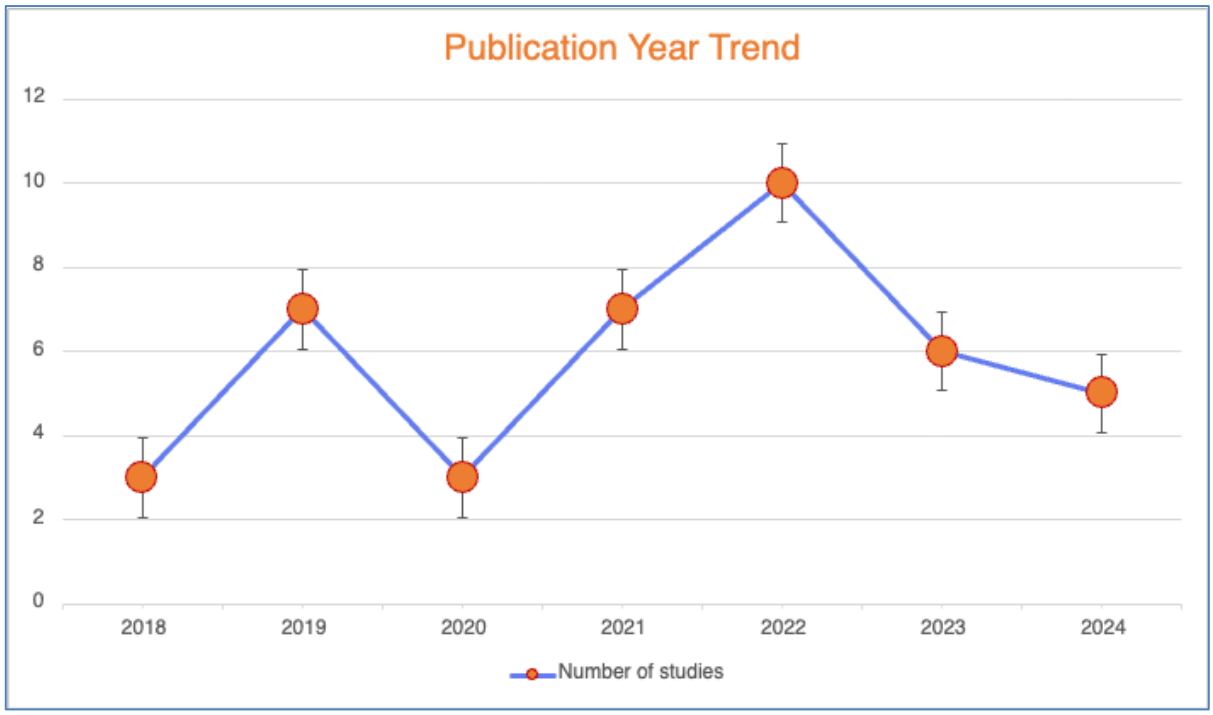


Figure 3.10: Publication trend in selected primary studies (Source: Author)

3.5.3 Data source

The reviewed studies employed a range of data types to address the problem of Automated Defect Detection (AuDD) in sewer systems, primarily relying on visual data such as images and videos. Out of the 39 accessible primary studies, 19 (49%) used image datasets, while 18

(46%) utilised video footage as their initial data source, as shown in 3.11. This nearly even distribution suggests that both data types remain central to model training and evaluation within this domain. However, it is important to note that many of the studies that began with video data ultimately converted these into frame-level image sequences for downstream processing, such as labelling, augmentation, and inference. This practice reflects a common trend in the field where video is often used for acquisition, but image-level analysis dominates the modelling workflow due to ease of annotation and computational efficiency.

Only 2 studies, notably (Haurum et al., 2021b; Zhou et al., 2022b), incorporated point cloud data, reflecting an emerging interest in leveraging 3D data for more detailed condition assessment. The low uptake of point cloud methods may be attributed to higher acquisition costs, complex processing requirements, and limited availability of labelled datasets in this format. Nevertheless, point clouds offer the potential for spatially accurate defect localisation, especially in scenarios involving structural deformation or volumetric anomalies. The dominance of image-based approaches in the literature underlines a trade-off between data richness and practicality, where 2D data remains the most accessible and scalable format for current deep learning applications in sewer inspection.

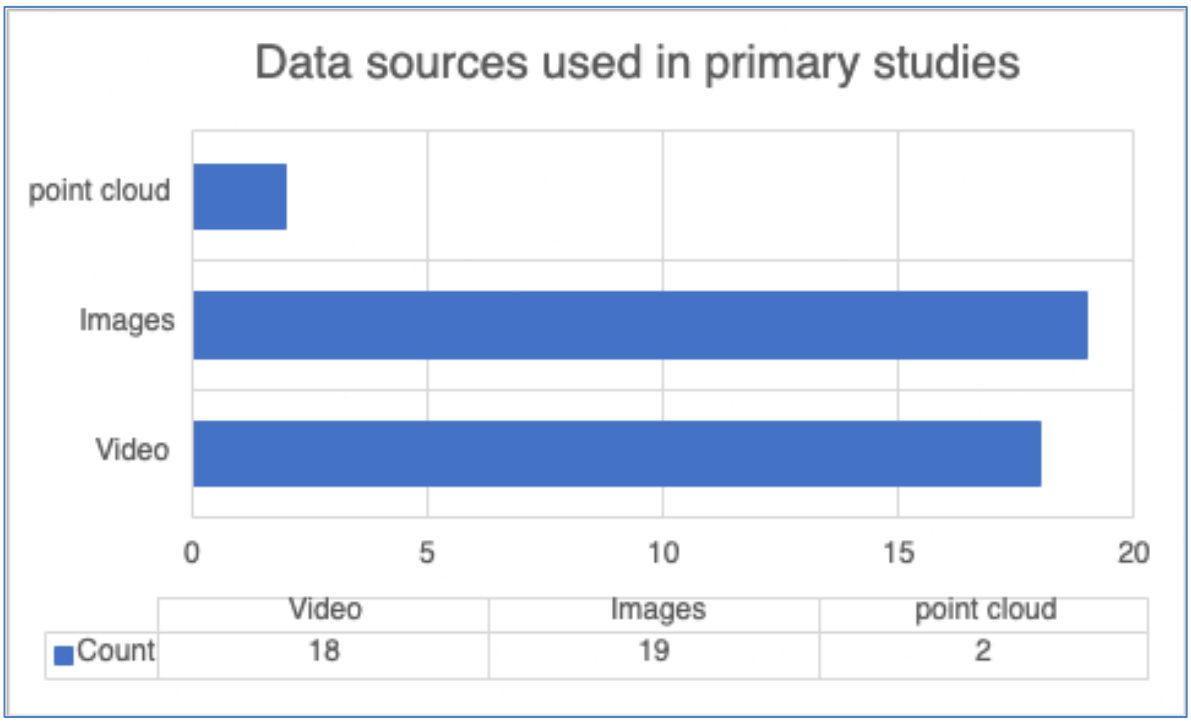


Figure 3.11: Data Source Distribution from primary studies (source: Author)

3.5.4 Problem type

The primary studies reviewed exhibit a notable diversity in how the defect detection task is formally structured. Classification remains the most prevalent formulation, adopted in 46% of studies, reflecting the field's early focus on broad defect recognition without spatial localisation. Object detection approaches were employed in 33%, indicating a growing emphasis on identifying defect positions within images. More spatially detailed formulations like semantic segmentation appeared in 6 studies (15%)), while only 2 studies (5%) utilised instance segmentation.

As seen in Figure 3.12, this distribution suggests that while researchers are beginning to adopt more granular and spatially aware problem types, a significant proportion of the literature still focuses on simpler classification models. The relatively low uptake of segmentation methods may be attributed to the high cost of pixel-level annotation and the lack of publicly available datasets with sufficient granularity (Dang et al., 2023). However, the trend towards detection and segmentation indicates a methodological shift that aligns with practical demands for more interpretable and actionable defect localisation in real-world sewer inspection tasks.

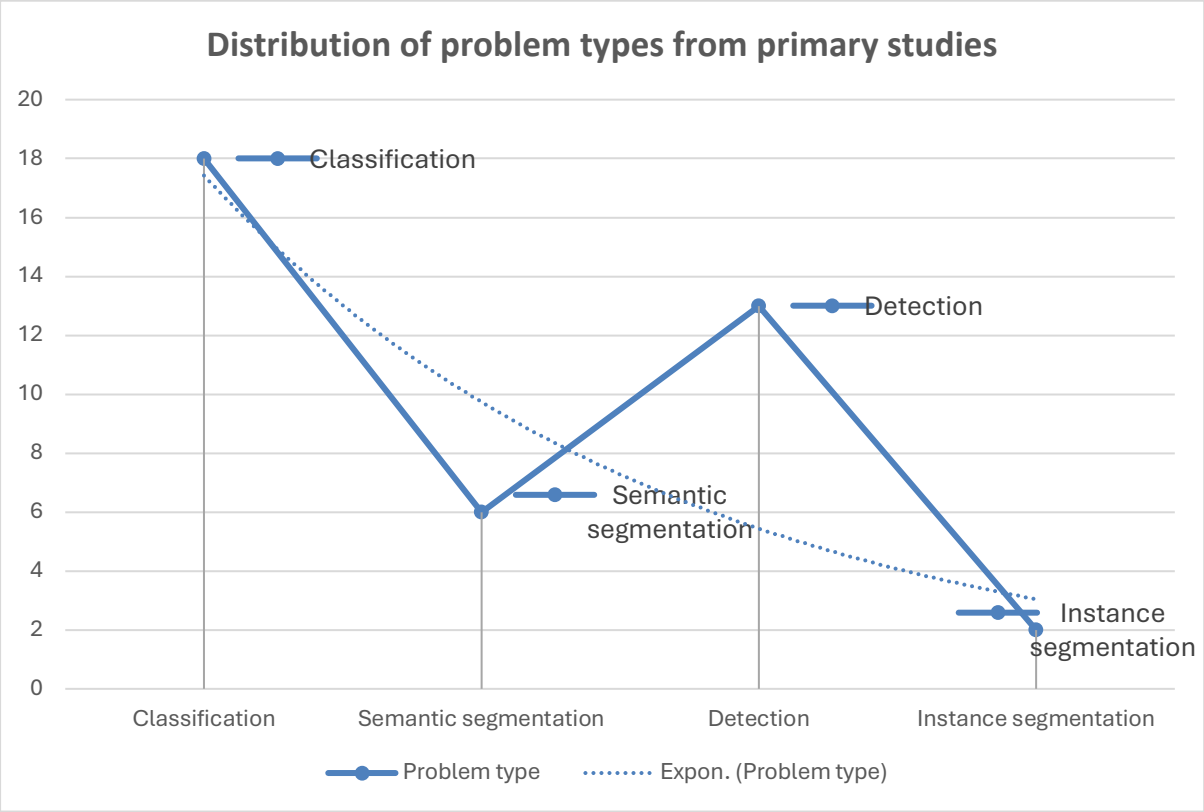


Figure 3.12: Distribution of problem types (source: Author)

3.5.5 Defect type

More than 20 different sewer defects has been explored in previous studies that employed deep learning techniques. Detecting these defects accurately is crucial for timely repairs and maintenance. Out of this, defect such as crack, root intrusion, and variants of deposits has been mostly researched in the last decade as represented in Figure 3.12. Although, surface damage and infiltration were also moderately investigated, however, scaling and blur are among those rarely researched. Analytically, more than 90% of these defects are either structural or O&M, with very few been constructional defects. For structural defects, cracks and fractures are common defects in sewer systems which can lead to leakage and physical issues. Their detection is challenging due to changing sizes, shapes, and orientations. To pre-process crack or fracture images, edge detection techniques like as Canny edge detection is used to enhance the visibility of cracks. Likewise, thresholding techniques such as Otsu's thresholding can also be applied to segment cracks and fractures from background noise. Furthermore, enhancement techniques like adaptive or histogram equalization, are used to improve the visibility of these defects.

Surface damage is another structural defect that can originate from corrosion or wear on the inner surface of the pipes. Its detection can be become more difficult due to its elusive appearance and resemblance to the nearby pipe material. Pre-processing techniques might include texture analysis methods using Gabor filters or Local Binary Patterns (LBP), to identify the textural disparities between defective and non-defective areas.

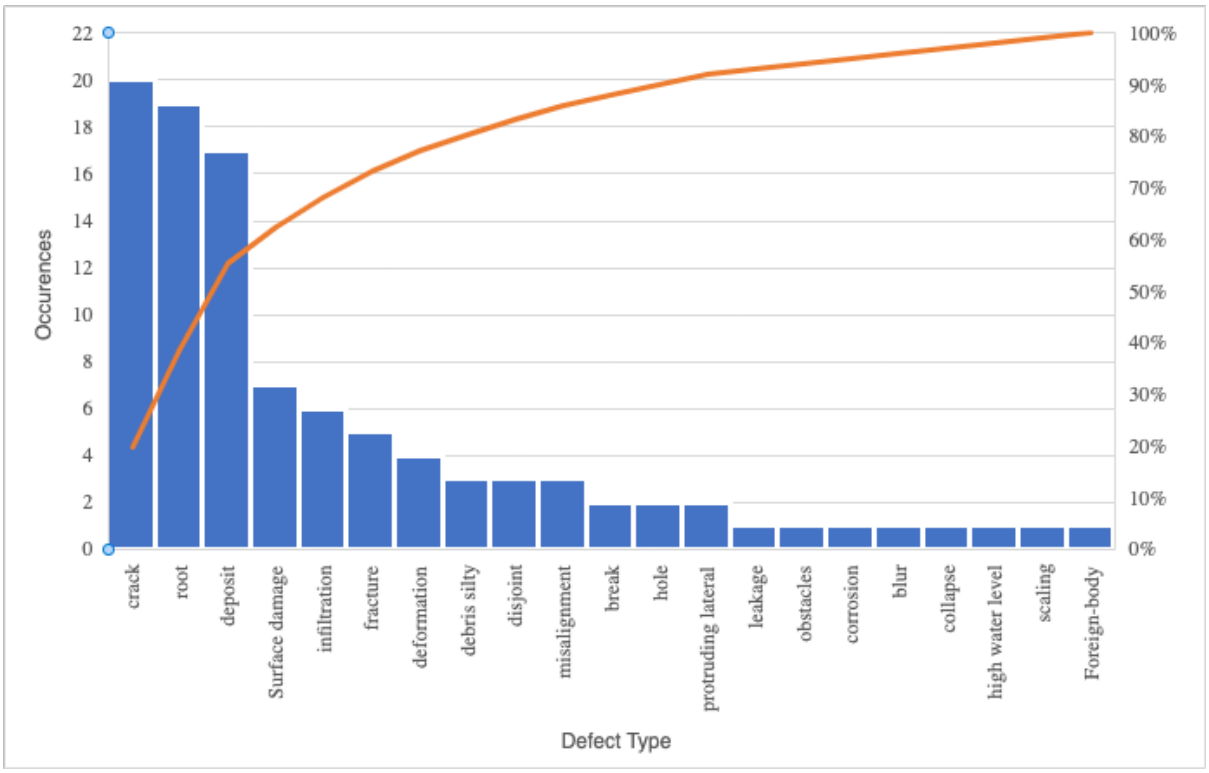


Figure 3.12: Defect type ranking from the literature (source: Author)

O&M defect such root, deposit, and infiltration have received increased attention in the last decade. These defects are common averagely across most sewers in Asia, Europe, and North America. Root intrusion occurs when tree roots infiltrate sewer pipes, causing blockages and structural damage. Its detection is challenging due to the irregular and complex nature of root structures, however, pre-processing techniques such as image enhancement (histogram or adaptive equalization) and edge detection algorithms, like Sobel or Laplacian can be used to identify the boundaries of root structures, thereby improving the visibility of roots against the background. For deposits in sewer pipes, it refers to the accumulation of residue, remains, or other materials that can cause blockages and hinder flow. A common challenge with this defect is variability in size, texture, and appearance, and previous studies have employed morphological operations such as erosion and dilation to enhance the contrast between the deposits and the pipe surface. Similarly, thresholding techniques like adaptive thresholding can help segment deposits from the background.

Consequently, defect such as crack, root, and deposit were considered in this research due to their prevalence in most sewers. With this, the model has a better and wider coverage of applications.

3.5.6 Pre-processing techniques

Sewer image preprocessing techniques plays an important part in improving the quality of sewer images and enhancing the accuracy of subsequent analysis and defect detection. Preprocessing involves different techniques and operations performed on the image to prepare it for further analysis or modelling. Over the years, various preprocessing techniques were used to enhance the quality and extract important information from the images. As depicted in Figure 3.13, preprocessing techniques such as resizing, annotation, extraction, edge detection, grayscale conversion, and histogram equalization are among the most widely used in AuDD for sewers according to the primary studies.

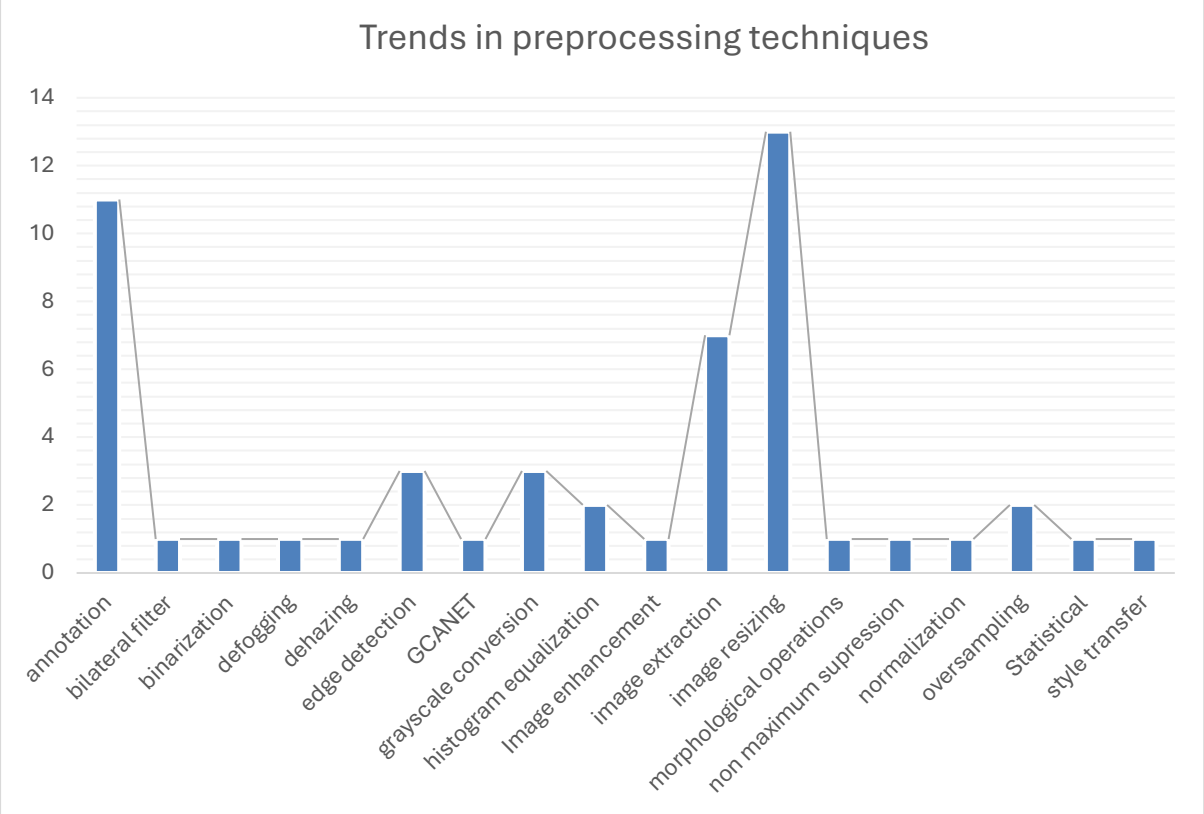


Figure 3.13: Preprocessing trends from literature (source: Author)

While suitability of the techniques depends on the data and the type of task – *classification, detection,* or *segmentation*, techniques such as image extraction from video and resizing are generic preprocessing steps for AuDD, irrespective of the task. A description of each technique, its characteristics, and suitable task type is summarized in Table 3.3.

Table 3.3: Summary of different pre-processing techniques

Technique	Description	Suitable task
Annotation	- labelling or marking specific regions or objects of interest. - identify and label different types of defects - provides ground truth for deep learning models.	- Detection - Segmentation
Bilateral filtering	- Edge smoothing technique - reduces noise and preserves important structural details.	- Classification - Detection - Segmentation
Binarization	- convert grayscale or colour image into a binary - separate foreground object from background - isolate or highlight ROI.	- Classification - Detection - Segmentation
Defogging	- remove or reduce fog effects in images - enhance clarity of defect for detection	- Classification - Detection - Segmentation
Dehazing	- remove atmospheric haze from images - restore the true colours and contrast of hazy images. - improves detection of subtle defects	- Classification - Detection - Segmentation
edge detection	- identify boundaries and edges in images. - locate areas of rapid intensity changes - identify and segment important features	- Classification - Detection - Segmentation
GCANET	- Gated Context Aggregation Network - designed for image segmentation tasks - Can segment sewer defects from the background	- Segmentation
Grayscale conversion	- convert colour image into a grayscale. - represents image in single intensity – white or black.	- Classification - Detection

Technique	Description	Suitable task
	- removes insignificant colour variations.	- Segmentation
Histogram equalization	- enhance image contrast by redistributing pixel intensities - improve the visibility of defects - enhance overall image quality	- Classification - Detection - Segmentation
Image enhancement	- aim to improve the overall image quality - adjust attributes such as brightness, contrast, and sharpness. - aim to improve detection and analysis results.	- Classification - Detection - Segmentation
Image extraction	- extract images from sewer videos or reports - makes it easier to work on sewer data	- Classification - Detection - Segmentation
Image resizing	- changes the dimensions of an image. - standardize the image dimensions - reduce computational complexity	- Classification - Detection - Segmentation
morphological operations	- deals with shape and structure of image. - Includes dilation, erosion, opening, and closing. - enhance the shape and structure of defects.	- Detection - Segmentation
Non maximum suppression	- suppress redundant or overlapping detections. - retains salient or representative image feature. - removes duplicate or overlapping defect detections.	- Detection - Segmentation
Normalization	- standardize the pixel values to a common scale or range. - provides consistent intensity distributions	- Classification - Detection - Segmentation

Technique	Description	Suitable task
Oversampling	- addresses class imbalance problem. - generates additional samples of minority classes. - handles defect under-representation.	- Classification - Detection - Segmentation
Style transfer	- transfer of artistic style between images - enhance visual appearance of sewer images	- Detection - Segmentation

3.5.7 Annotation

Image annotation plays a crucial role in identifying and labelling specific regions or objects of interest within sewer images. It is essential for creating the ground-truth datasets used to train and evaluate deep learning models, particularly for object detection and semantic segmentation tasks (Dang et al., 2023). In the systematic review of primary studies, several annotation tools were identified, including LabelME, LabelImg, and MATLAB.

As represented in Figure 3.14, the annotation tool usage was "Not specified" in most cases (56%). There are two primary reasons for this large percentage. First, a significant portion of these "Not specified" studies focused solely on whole-image *classification* (e.g., simply categorising an entire image as 'defective' or 'normal'). For purely classification-based tasks, spatial annotation tools (such as drawing bounding boxes) are not required; labels are typically assigned simply by organising images into named folders or logging them in a spreadsheet. Second, among the studies that did perform object detection or segmentation, several simply failed to report the specific software tool used to create their bounding boxes or polygon masks. This omission highlights a continuing gap in methodological transparency and reporting standards within the current literature.

For the studies that did specify their spatial annotation tools, LabelME and LabelImg were the most popular, each being employed in 21% of the reviewed studies. This indicates their status as standard open-source tools for automated defect detection (AuDD) and general image annotation. Conversely, only a marginal 3% of the studies (e.g., Zhou et al., 2022b) reported using MATLAB to annotate their dataset.

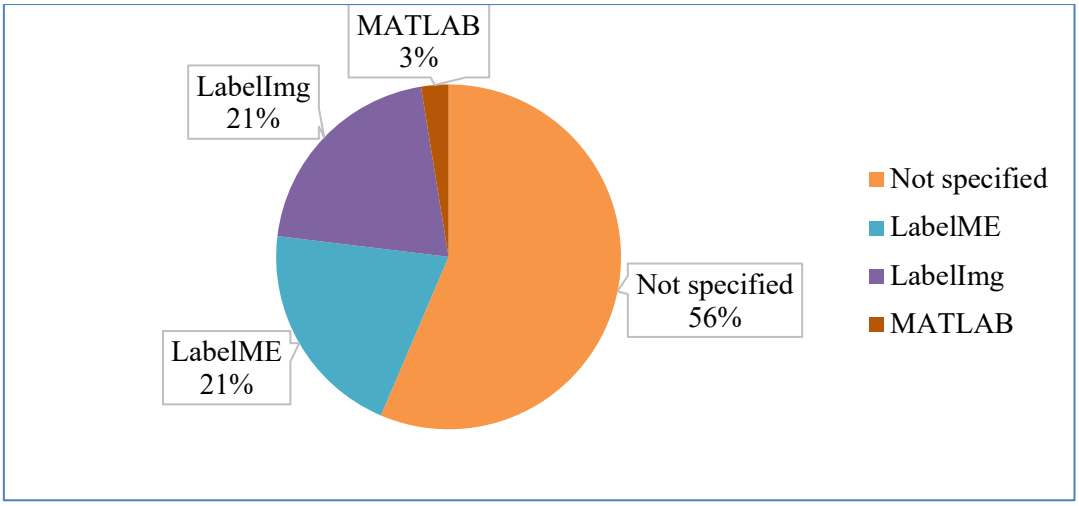


Figure 3.14: Distribution of annotation tools used in previous studies (source: Author)

3.5.8 Algorithm and the Present State of Knowledge in DL for AuDD

The present state of knowledge regarding AuDD establishes that deep learning, specifically CNNs, has definitively superseded traditional machine learning paradigms (Li et al., 2022b). Early automated systems relied heavily on conventional image processing and traditional classifiers that required brittle, mathematically rigid handcrafted feature extraction (Moradi et al., 2019d). Current literature demonstrates that deep learning architectures overcome these limitations by autonomously learning optimal, high-level visual feature representations directly from raw CCTV data. However, the state-of-the-art also confirms that no single, out-of-the-box architecture provides a universal solution, necessitating a critical evaluation of model trade-offs.

Within this modern deep learning landscape, a critical architectural contrast defines the current state of knowledge: the trade-off between accuracy-focused two-stage detectors and speed-focused single-stage detectors. Single-stage models, most notably the YOLO (You Only Look Once) series, currently dominate the literature due to their rapid inference speeds, which are essential for processing high-framerate CCTV datasets (Tang et al., 2024). Recent benchmarking studies continually push this frontier, comparing YOLO variants against newer architectures like RT-DETR to optimize real-time detection (Chaoyu & Zakaria, 2025; Oğurlu et al., 2025). However, this speed historically comes at the cost of reduced feature granularity; the literature indicates that single-stage architectures frequently struggle with small-object detection, leading to higher rates of false negatives when identifying subtle, fine-grained textures like minor hairline cracks against complex brickwork. Figure 3.15 shows the distribution of the algorithm's usage in the literature.

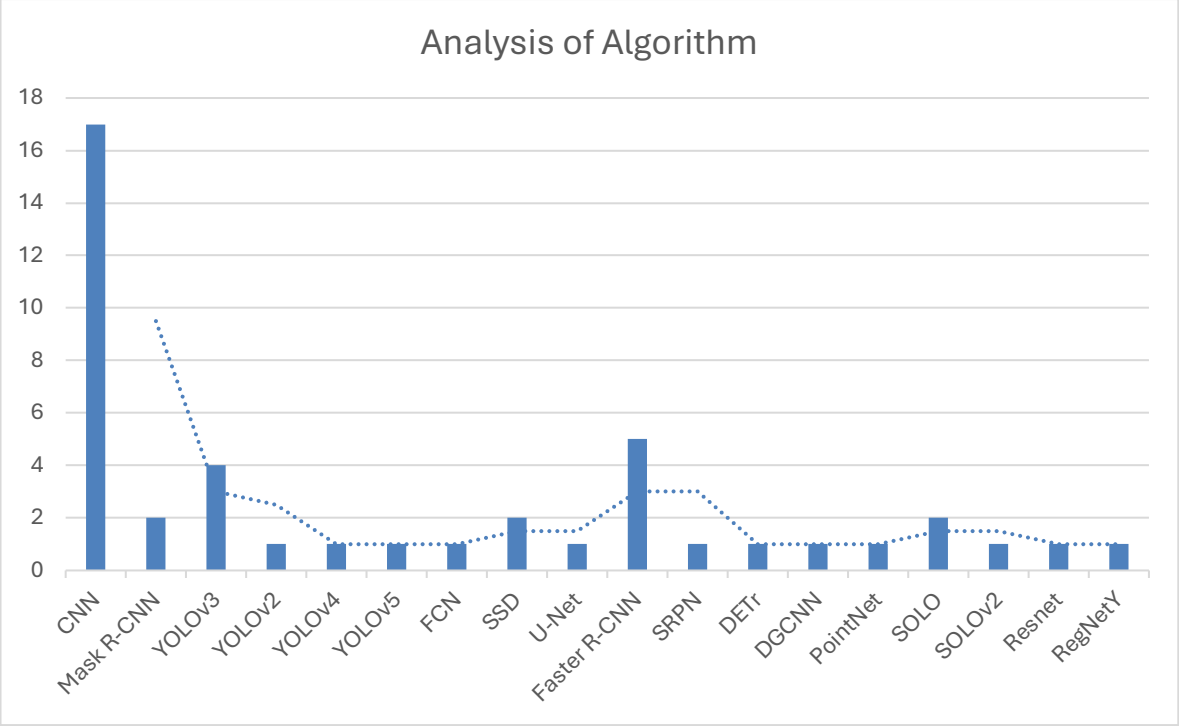


Figure 3.15: Analysis of Algorithm

To overcome these limitations, the absolute frontier of current research is shifting toward complex, multi-modal, and evidential deep learning frameworks. Recent studies are exploring Trustworthy Multi-Focus Fusion (TMFF) frameworks for multi-label classification (Hu et al., 2024) and Pareto-optimal evidential networks to manage the inherent uncertainty in video-based inspections (Zhao et al., 2024). Furthermore, the introduction of Vision Transformers (ViTs) applied to 3D point clouds attempts to capture long-range global dependencies within a pipe, rather than just localized pixel textures(Jing et al., 2025). While theoretically promising for highly accurate defect classification, these state-of-the-art models are notoriously computationally expensive, meaning their practical application in standard, resource-constrained SMC systems remain largely experimental.

The practical impact of this rapid algorithmic advancement is a growing disconnect between theoretical model complexity and actual data quality. As Jang & Kim (2025) critically highlight, the performance of even the most advanced CNN-based models is fundamentally bottlenecked by the quality of the input data. A pervasive flaw in the current literature is the hyper-focus on architectural tweaking (e.g., adding more network layers or attention mechanisms) while ignoring the harsh physical realities of sewer environments. An algorithm's mathematical robustness is of little practical value if it cannot separate the visual signature of a structural defect from pervasive environmental noise like flowing water or lens fog.

Therefore, the present state of knowledge reveals a distinct research gap: the field has not yet solved the challenge of deploying highly accurate models that can operate reliably within the computational and environmental constraints of real-world infrastructure management. To

advance the state-of-the-art, research must shift away from merely applying heavier, generic algorithms to benchmark datasets. Instead, as proposed in this thesis, the focus must pivot toward extensively optimizing, hybridizing, and rigorously pre-processing data for these architectures to achieve a viable, operational balance between high-speed real-time inference and high-fidelity defect detection.

3.5.8.1 Efficiency

Efficiency refers to the capability of an algorithm to achieve accurate results while minimizing computational resources. Among these algorithms, CNNs have been widely employed in sewer defect classification (H. Gupta et al., 2018; Long et al., 2017), with a good balance between accuracy and computational efficiency due to their simple architecture compared to others (Kefan Chen et al., 2018), making them efficient for real-time or near real-time applications. However, modified CNN architectures such as ResNet are not as simple as conventional CNNs (B. Ma et al., 2022). Likewise, one-stage algorithms such as YOLO and SSD are also efficient and designed for single-shot object detection, eliminating the need for complex Region Proposal Networks (RPN) (Redmon et al., 2015).

In contrast, two-stage algorithms such as Faster R-CNN and SOLO have higher computational requirements due to their two-stage detection process, which includes both RPN and classification (Ren et al., 2015), (M. Wang & Cheng, 2018). Although, both one-stage and two-stage algorithms have had great success in AuDD by detecting different defect with good degree of accuracy, the RPN embedded in two-stage algorithms' architecture like Faster R-CNN has made it computationally less efficient compared to single shot detectors.

3.5.8.2 Speed

Speed is a critical consideration in real-time or inferencing applications, especially when dealing with large-scale sewer inspection. One-stage algorithm such as YOLOv3 and SSD are popular for their high inference speed on images or video (Redmon et al., 2015), (Fang et al., 2022b). It is particularly called "one-stage" because it directly predicts the bounding boxes and class labels of objects in a single pass. Hence, they can process images rapidly by using a single neural network for bounding box prediction and classification. This makes them suitable for applications, such as AuDD , where speedy detection is vital. In contrast, two-stage algorithms such as Faster R-CNN and SOLO uses a different stage (i.e., RPN) for proposals or candidate regions and another stage for classification. Owing to this two-stage approach, Faster R-CNN and SOLO are usually slower than YOLO or SSD but provide higher accuracy (C. Siu et al., 2022; M. Wang, Luo, et al., 2021c). For CNNs, they offer a range of speeds depending on the model architecture and implementation (Dang et al., 2021c).

3.4.8.3 Robustness in Real-World Contexts

In the domain of deep learning, algorithmic robustness refers to a model's capacity to maintain high detection accuracy and low variance when subjected to degraded or out-of-distribution input data. When critically evaluating the current AuDD literature, it is imperative to contextualize these deep learning models strictly within the physical location they are designed to operate: the harsh, highly unpredictable environment of an active underground sewer pipe. An algorithm's mathematical robustness is of little practical value if it has only been validated in sterile, academic conditions that fail to replicate the extreme environmental hostility of subterranean infrastructure (Moradi et al., 2019d).

A significant and pervasive flaw in current literature is the widespread reliance on highly curated, artificially clean datasets to train and validate AuDD models. Many studies claim state-of-the-art robustness based on test sets comprising newly constructed, dry, or well-lit concrete pipes. However, the visual data acquired from active sewers is fundamentally different. Real-world CCTV footage is consistently degraded by severe environmental artifacts: thick camera lens fogging from humidity, sudden illumination drop-offs, severe motion blur, and overlapping organic debris (Li et al., 2022b). Furthermore, active wastewater flow creates highly reflective, dynamic surfaces that easily confuse optical sensors.

When models are removed from their curated academic context and deployed into this real-world location, their theoretical robustness frequently collapses. As Jang & Kim (2025)critically demonstrate, the performance of CNN-based sewer defect detection is heavily bottlenecked by input data quality. Computationally lightweight single-stage models (like YOLO) tend to lose feature fidelity entirely in low-light contexts, outputting unacceptable rates of false negatives as subtle hairline cracks become occluded by sludge or poor pixel contrast. Conversely, while two-stage detectors (like Faster R-CNN) are slightly more resilient to occluded scenes, they are highly susceptible to hallucinating false positives, often misclassifying thick grease deposits or hanging cobwebs as structural deformations.

The practical impact of ignoring location and environmental context is the proliferation of academically successful models that are entirely operationally unusable by Sewer Management Companies (SMCs). The current literature demonstrates a critical oversight wherein researchers treat poor image quality as a failure of the physical hardware rather than a fundamental environmental variable that the software must actively manage. An algorithm cannot be considered truly robust if it cannot mathematically separate the visual signature of a critical fracture from the pervasive visual noise inherent to aging pipeline networks. Consequently, there is a pressing need for trusted, uncertainty-aware systems capable of managing this noise (Zhao et al., 2024).

This methodological oversight underscores the critical necessity of the approach proposed in this thesis. Robustness in sewer defect detection cannot be achieved solely through the architectural tweaking of deep neural networks; it mandates a holistic pipeline that actively addresses the environmental context. By integrating dynamic, defect-specific image enhancement and pre-processing strategies, this research aims to strip away location-specific

environmental noise before the data ever reaches the classifier, bridging the gap between theoretical algorithmic robustness and real-world operational viability.

3.5.9 Performance Metrics

Evaluation metrics play a crucial role in assessing the performance of deep learning models for automatic sewer defect detection. By quantifying the accuracy and reliability of these models, evaluation metrics provide valuable insights into their effectiveness and suitability for real-world applications (Situ et al., 2023b). In this analysis, commonly used evaluation metrics, including accuracy, precision, recall, average precision (AP), mean average precision (mAP), intersection over union (IoU), mean intersection over union (MIoU), and F1 score. Table 3.4 provides a summary of these common metrics, including their suitable task category, description, and possible limitations.

Table 3.4: Different evaluation techniques in deep learning

Metrics	Task	Description	Limitation
Accuracy	- classification - detection - segmentation	- measures the overall correctness of the model's predictions - calculates ratio of correctly predicted instances to the total number of instances	- Not be suitable for imbalanced datasets
Precision	- classification - detection - segmentation	- measures the proportion of correctly identified positive predictions among all predicted positives	- Does not consider false positives or false negatives.
Recall	- classification - detection - segmentation	- measures the proportion of correctly identified positive predictions among all actual positive instances	- Do not consider false positives or false negatives.
F1	- classification - detection - segmentation	- combines precision and recall into a single metric - provides a balanced measure of model performance	- Not be suitable for scenarios with trade-off between precision and recall
AP	- detection - segmentation	- measures the precision-recall trade-off - calculate the area under the precision-recall curve	
mAP	- detection - segmentation	- average AP across different defect types	
IoU	- detection - segmentation	- measures overlap between predicted bounding boxes and ground truth annotations - calculates the intersection area divided by the union area of the two regions	- Sensitive to slight variations in the size of the predicted regions
MIoU	- detection - segmentation	- the mean IoU across all defect classes - provides overall measure of segmentation accuracy	- Sensitive to slight variations in the size of the predicted regions

Analysing these metrics using the primary studies, varying number of each evaluation metric was revealed as represented in Figure 3.16. AP and mAP have the highest proportion with 22% and 18% respectively, emphasizing their importance in sewer defect detection (Tan et al., 2021a). This reveals the importance of assessing precision-recall trade-offs (Fang et al., 2022b) and localization accuracy. Recall, with 11% of the studies applying this metric reflect the emphasis on minimizing false negatives in defect detection. Precision and F1 account for 7% and 11% respectively, indicating the need for accurate positive predictions while considering the balance between precision and recall (Hu et al., 2023). Accuracy and IoU have counts of 8% each, highlighting their relevance in evaluating overall correctness (Wang & Cheng, 2020b) and spatial agreement.

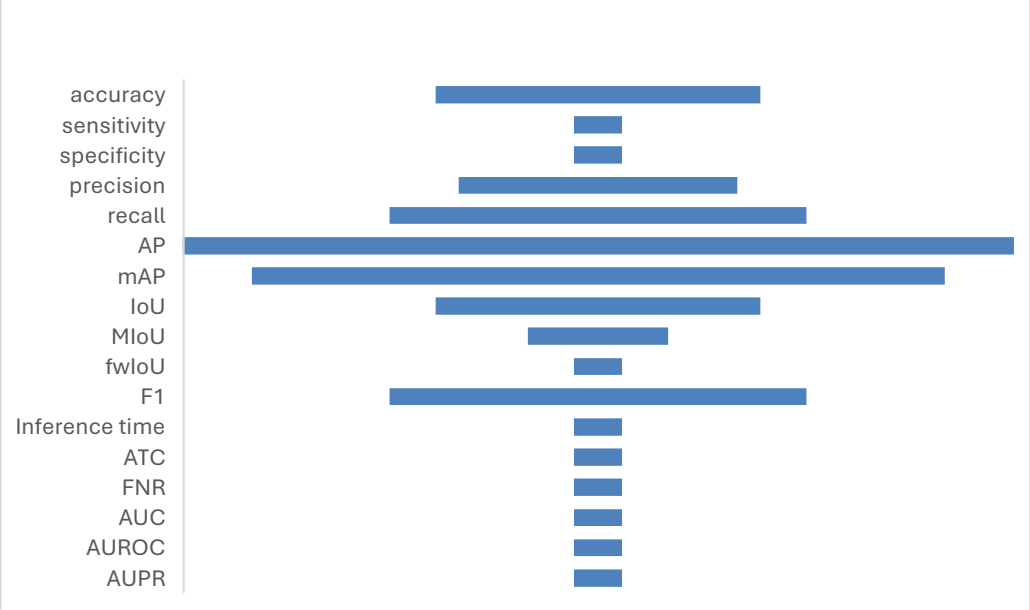


Figure 3.16: Distribution of performance metrics (source: Author)

By considering these metrics, broad evaluation and comparison of different models becomes feasible, including capturing various aspects of model performance, such as correctness, completeness, localization accuracy, and trade-offs between precision and recall. Understanding the strengths and limitations of each metric is crucial for selecting appropriate evaluation approaches and interpreting results accurately in the context of automatic defect detection in sewers. Since this research is proposing a defect detection using Faster R-CNN, metrics such as AP, mAP, IoU, and MIoU will be adopted to cater for precision-recall trade-off and overall measure of correctness across different defect types, thereby improving the robustness of the proposed model.

3.6 Enablers and Challenges of the AuDD Process

The transition from manual CCTV review to Automated Defect Detection (AuDD) is driven by several critical enablers. Foremost among these is the rapid advancement of deep learning architectures, specifically Convolutional Neural Networks (CNNs). Unlike traditional machine learning that relies on rigid, manual feature extraction, deep learning algorithms act as a primary enabler by autonomously learning complex, high-level visual representations directly from raw data (Li et al., 2022b). Furthermore, the democratization of high-performance computing, such as multi-GPU environments, and the availability of advanced data balancing techniques (e.g., SMOTE and geometric data augmentation) enable researchers to train highly complex models like YOLO and Faster R-CNN on massive, previously unmanageable municipal datasets.

Despite these technological enablers, the AuDD process is fundamentally constrained by severe operational and environmental challenges. The primary challenge, as identified throughout the literature, is the extreme degradation of input data quality (Jang & Kim, 2025). The subterranean environment introduces pervasive visual noise, such as severe lens fogging, motion blur, poor illumination, and active wastewater flow, which drastically reduces algorithmic feature fidelity. Additionally, the field faces the challenge of severe class imbalance; critical structural failures (e.g., fractures or collapses) occur far less frequently in training data than minor operational issues (e.g., attached deposits), leading to biased models that struggle with minority class detection.

Finally, a significant deployment challenge remains in balancing the accuracy-efficiency trade-off. While heavy, two-stage architectures provide the theoretical accuracy needed for complex defect segmentation, their computational burden makes real-time edge deployment on active sewer crawlers highly difficult. Therefore, overcoming these challenges requires moving beyond simply applying standard algorithms to raw data. It necessitates the integration of targeted, defect-specific pre-processing pipelines such as custom thresholding and contrast enhancement to actively strip away environmental noise and leverage the deep learning enablers to their full potential.

3.7 Conclusions from SLR

One of the most critical limitations identified is the lack of reproducibility in model development pipelines. Many studies do not fully report key methodological components such as data preprocessing, augmentation strategies, or training configurations. For example, details on contrast enhancement thresholds, morphological kernel sizes, or even model hyperparameters like learning rates and anchor boxes are frequently missing. This opacity significantly hinders replication and weakens scientific rigour. Without transparent methodologies, other researchers' and industry partners cannot validate or build upon prior

work, which perpetuates siloed experimentation and undermines collaborative advancement in the field.

Furthermore, a major gap exists in the application of defect-specific preprocessing. Most reviewed studies adopted generic pipelines such as grayscale conversion or histogram equalisation without regard to the morphological diversity of sewer defects. This approach introduces performance bias, where certain pre-processing techniques disproportionately favour visually prominent defects like cracks or deposits, while diminishing subtler anomalies like encrustation or root intrusion. Failing to tailor preprocessing strategies risks weakening model generalisability and undervaluing defect classes critical for structural health assessments. In real-world use, this could lead to missed or misclassified defects with serious infrastructure and safety implications.

Another key finding is the limited attention given to computational efficiency and model benchmarking. Fewer than 15% of studies reported inference speed, FLOPs, or training time. Yet these factors are critical for deployment in time-sensitive, resource-constrained environments like robotic crawlers or real-time CCTV platforms. Several studies pursued marginal accuracy gains like 1 to 4 percentage gain at the expense of drastically increased model complexity or runtime costs. This overemphasis on marginal metric improvements, without regard for scalability or latency, highlights a fundamental disconnect between academic benchmarking and operational deployment.

Another noticeable absence in the literature is any practical roadmap toward deployment. None of the studies reviewed provided a full-stack deployment strategy. Moreover, discussions of API design, containerisation, or system integration were entirely absent. It can be inferred from the data gathered from the reviews that this field remains overly focused on algorithmic experimentation and has yet to embrace deployment-aware design. As a result, many promising models remain trapped within the confines of research labs, unable to scale or integrate into the workflows of Sewer SMCs. This limitation severely restricts the real-world impact of otherwise capable models.

In summary, the review identifies a fragmented but promising landscape. Algorithmic innovation is evident, yet the field struggles with foundational issues of reproducibility, domain adaptation, and practical deployment. These gaps shaped the core research objectives of this thesis and informed the design of its experimental pipeline. By introducing defect-sensitive preprocessing, benchmarking multiple YOLO variants under consistent conditions, and proposing a scalable Software-as-a-Service (SaaS) deployment model, this study directly responds to the limitations uncovered in the literature. If the research community fails to address these systemic issues, sewer AI models may continue to excel in academic metrics but fail to deliver real-world value.

CHAPTER FOUR

4.0 Research Methodology

4.1 Chapter Overview

This chapter provides an in-depth exploration of the research methodology utilized to achieve the objectives of developing and evaluating an Automated Defect Detection (AuDD) system for sewer infrastructures. It begins by discussing the foundational research paradigm, outlining critical ontological and epistemological assumptions that justify the quantitative research strategy adopted. The chapter further emphasizes how philosophical beliefs about the nature of reality (ontology) and knowledge acquisition (epistemology) directly informed methodological choices, such as the use of deep learning models.

Subsequently, the chapter details the research methods, highlighting the systematic data collection process from various Sewer Management Companies (SMCs) across multiple UK cities. It also explains the robust measures implemented to maintain data quality, privacy, and standardization in line with European norms (EN 13508-2). Finally, the chapter addresses defect selection criteria by outlining the rationale behind focusing on specific defects based on their prevalence, operational impact, and detectability, thereby ensuring the effectiveness and applicability of the proposed automated detection models.

4.2 Research Paradigm

The research paradigm, sometimes referred to as the research philosophy, is a crucial element of the research process that significantly influences the study's design, methodology, and overall approach (Saunders, 2009). As outlined by (Guba & Lincoln, 1994), a research paradigm establishes a framework that defines the nature of knowledge, reality, and how they should be understood within a given field. Understanding and articulating the research philosophy is crucial, as it underpins every aspect of research design, from the formulation of research questions to the selection of data collection and analysis methods (Crotty, 2020). For example, in the context of Automated Defect Detection (AuDD) in sewer systems, the research paradigm not only dictates how defect types are prioritized but also determines the analytical techniques chosen to assess defects and the interpretation of findings regarding the performance of various detection methodologies (e.g., deep learning vs. traditional methods).

According to Kuhn (1962), paradigm is defined as “a constellation of beliefs that influence what should be studied by a scientist, how the research is conducted, and how the result is interpreted”. This definition emphasizes the significance of research paradigms in guiding not only the topics and questions explored but also the way research is approached. In the context of AuDD, the choice of paradigm (whether positivist, interpretivist, or constructivist) influences the entire research design. For instance, a positivist paradigm, which relies on objective measurements and data-driven conclusions, is often associated with studies that emphasize empirical testing of automated algorithms and their performance in defect detection (e.g., using deep learning models for crack detection in sewer systems). Conversely, an interpretivist paradigm might focus more on understanding the human and socio-technical factors that influence the adoption of these automated systems in real-world infrastructure management.

The research process, as noted by (Saunders, 2009), is shaped by assumptions about both the nature of reality (ontological assumptions) and the nature of knowledge (epistemological assumptions). Ontological assumptions concern the nature of reality and the existence of phenomena, while epistemological assumptions focus on how knowledge about those phenomena can be acquired and validated. In AuDD, these assumptions are fundamental in shaping the methods used for data collection and analysis. For example, an ontological assumption may involve the belief that sewer defects (e.g., cracks or root intrusions) exist as tangible and measurable entities that may be identified through automated systems. Epistemologically, researchers may assume that knowledge of defect detection can be best acquired through objective data analysis (e.g., using deep learning models trained on labelled images of sewer systems).

The differentiation between ontological and epistemological assumptions is particularly important in AuDD research. On the ontological side, the assumption that sewer defects are physical, observable objects impacts the choice of research techniques, such as the use of computer vision technologies to visually identify defects. From an epistemological standpoint, researchers may choose to rely on quantitative data analysis (e.g., F1 scores) to assess the accuracy of automated defect detection methods by leveraging an assumption that the validity of the detection systems is grounded in statistical measures of performance. This assumption is closely tied to the positivist research philosophy, which aims for objective validation of knowledge through empirical testing (e.g., evaluating the accuracy of deep learning models in detecting cracks in sewer pipes).

Importantly, these assumptions and paradigms also play a vital role in shaping the course of research in AuDD. The adoption of deep learning models, for instance, is not simply driven by technological capability but also by the epistemological belief that automated systems can provide more accurate, scalable, and reproducible results than manual inspection methods (Huang et al., 2023). Moreover, the methodological alignment with certain paradigms, whether qualitative, quantitative, or mixed methods, informs how researchers approach the analysis of

results. For instance, an epistemological commitment to a quantitative approach leads to a focus on performance metrics (Godwin et al., 2021), such as precision, recall, and F1-score when evaluating model effectiveness in defect detection.

The critical role of understanding and acknowledging these paradigms, assumptions, and philosophical underpinnings cannot be overstated, as they shape the course of research investigations and influence the ultimate outcomes of studies. This is particularly true in AuDD research, where methodological choices which is driven by both ontological and epistemological beliefs determines the effectiveness of defect detection systems and the future directions of technological advancements in the field. As highlighted by (Denzin, 2009), the theoretical framework supporting research must be critically examined to ensure that the chosen methods align with the research questions and the broader philosophical approach adopted. This comprehensive understanding of the research paradigm is essential for designing robust studies that effectively address the complex issues surrounding automated sewer defect detection.

4.2.1 Ontological Assumptions

In academic research, ontological assumptions play a crucial role in determining what is considered a valid subject of analysis and the type of relationships that exist among the elements being studied (Ylönen & Aven, 2023). An example of valid subject of analysis can include sewer systems, defects, and detection models. There are generally two main ontological viewpoints: *objectivism* and *constructivism*. Objectivism proposes that entities have an existence that is separate from human perception (Al-Ababneh, 2020). This insinuates that reality is separate and objective, and can be observed, evaluated, and quantified without being influenced by the observer. This standpoint is closely associated with paradigms such as positivism, which forms a part of quantitative research methods.

Positivism emphasizes empirical observation and the use of statistical methods to derive objective truths, relying heavily on measurable data to validate hypotheses (Bryman, 2016). This objectivist view also aligns with other paradigms such as post-positivism, which retains the idea of an objective reality but acknowledges the limitations of human observation and the potential for error (Saunders, 2009). Similarly, critical realism, while more nuanced, maintains that reality exists independently of human thoughts, and that research should aim to uncover the underlying structures that cause observable phenomena (Bhaskar, 2013). All these paradigms share the foundational belief that reality can be studied independently of personal biases, often with scientific tools and quantitative methods to gain objective, verifiable knowledge (Guba & Lincoln, 1994).

Specifically, when examining sewer defect, an objectivist standpoint would suggest that defects in sewer systems are independent of the SMCs and other stakeholders and can be objectively identified and quantified using automated methods. For example, an automated system using

deep learning models can objectively identify and quantify the severity of cracks in sewer pipes based on predefined criteria, such as crack width or length, regardless of the inspector's experience or interpretation. The system's assessment of these defects is based purely on measurable data, such as pixel intensity and geometric patterns, ensuring that the identification process remains consistent and independent of any subjective influence from the SMCs or stakeholders involved.

Constructivism, on the other hand, suggests that entities are formed from the perceptions and actions of social actors (Sardana et al., 2023). This means that reality is subjective and is shaped by individual experiences and interpretations. According to this view, reality is not an objective entity to be discovered, but rather a construct that is co-created through human interactions and perceptions (Berger & Luckmann, 1991). Constructivism is aligned with different paradigms, including *interpretivism*, which underpins qualitative research methods. Interpretivism emphasizes understanding the meanings and experiences that individuals assign to social phenomena, often through in-depth, qualitative methods such as interviews or ethnography (Schwandt, 1998). Additionally, *symbolic interactionism*, a paradigm often associated with subjectivism, highlights the role of social interactions and symbols in constructing reality (Blumer, 1980). These paradigms share the fundamental belief that reality is not something fixed but is interpreted and constructed through human experience and interaction.

In sewer defect detection, a constructivist perspective would imply that the interpretation of defects is influenced by the perceptions and experiences of the inspectors. For instance, when inspecting sewer pipes for a defect, two inspectors might interpret the same visual data differently based on their previous experiences, training, or even the specific tools they use. One inspector might classify a small, shallow crack as a major defect, while another might consider it a minor issue. Furthermore, an inspector's interpretation might be influenced by contextual factors, such as prior experiences with similar defects in other infrastructure projects. Such variations in interpretation illustrate the constructivist view that the reality of a defect is not fixed but is shaped by the individual's knowledge, beliefs, and perceptions. This subjective interpretation can significantly affect the decision-making process regarding the urgency of repairs or replacements in the sewer system.

However, given the objective nature of the data (i.e., sewer images and videos), this research adopts an objectivist stance, assuming that defects can be objectively identified and measured using automated detection models. Overall, the ontological assumptions of this research are rooted in the belief that defects in sewer systems are real, observable facts that can be detected and quantified independently of human interpretation. This aligns with the research aim of developing a data-driven methods that can objectively identify and quantify defects in sewer systems based on consistent and measurable data, independent of subjective interpretations or external influences. By adopting an objectivist perspective, this research ensures that the detection models are based on empirical evidence and can be generalized to various sewer systems.

4.2.2 Epistemological Assumptions

For Epistemology, it deals with the nature and scope of knowledge by focusing on how knowledge is acquired and what constitutes valid knowledge in a particular field (Bradshaw et al., 2017). It is concerned with understanding the processes through which we gain knowledge and how we can justify the truth of what we know. The epistemological viewpoint adopted by a researcher profoundly influences the methods used for data collection, analysis, and interpretation. Different epistemological perspectives frame how we approach questions of what can be known, how we can know it, and how knowledge can be validated (Guba & Lincoln, 1994). Within the field of epistemology, two dominant paradigms are positivism and interpretivism.

Positivism emphasizes the need for objective measurements and observable facts. Grounded in the belief that reality exists independently of human thoughts and perceptions, positivism asserts that reality can be measured and quantified using scientific methods (Kotronoulas et al., 2023; Saunders, 2009). This approach maintains that knowledge is derived from sensory experiences that can be systematically observed and tested (Crotty, 2020). Other paradigms that align with this view include post-positivism, which retains the idea of an objective reality but acknowledges the limitations of human observation and the potential for error (Al-Ababneh, 2020), and critical realism, which also holds that an objective reality exists, but that our understanding of it is mediated through human perception and social influences.

In AuDD, a positivist approach would involve collecting numerical data on the accuracy of deep learning models in detecting specific defects such as cracks or root intrusions in sewer systems. This data is then analysed statistically to identify patterns and validate the effectiveness of the model. The paradigm strongly supports the use of quantitative research methods, which involve the collection of numerical data, statistical analysis, and the application of objective, measurable metrics (Creswell, 2009). In AuDD research, this could include metrics such as precision, recall, F1-score, IoU, and mAP to assess the performance of defect detection models. Figure 4.1 represents a decision tree illustrating how the assumptions of objectivism and positivism guide the selection of these paradigms for objective data collection and statistical analysis in this research.

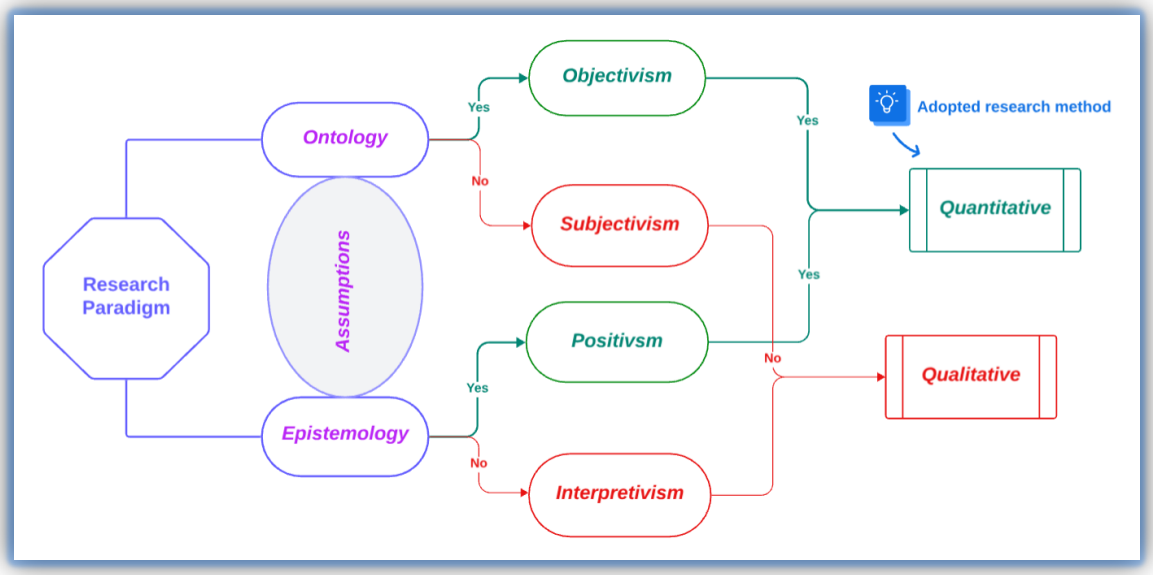


Figure 4.1: Decision tree representing the choice of research method (source: Author)

On the other hand, *interpretivism* focuses on the subjective interpretation of human experiences and actions. It posits that reality is socially constructed and cannot be fully understood without considering the perspectives of individuals involved in the phenomenon being studied. Interpretivism emphasizes the importance of context and seeks to understand how people make sense of their world . This approach is supported by other paradigms, such as constructivism, which asserts that knowledge is actively constructed by individuals through their interactions with their environment, and phenomenology, which focuses on understanding lived experiences and the essence of human consciousness (Guba & Lincoln, 1994).

According to interpretivism, knowledge is gained through subjective experiences, and meaning is constructed from those experiences rather than being discovered as an objective truth. Interpretivism supports the use of qualitative research methods, such as interviews, ethnographies, and content analysis, which focus on non-numerical data, such as text, audio, and video, to explore complex social phenomena and gain deeper insights into human perceptions, motivations, and actions (Denzin, 2009). While interpretivism is invaluable for understanding human experiences, motivations, and social contexts, it is less suitable for research that requires objective measurements and statistical analysis. In the case of AuDD, the primary aim is to measure, evaluate, and validate the performance of automated systems, which requires the use of quantitative methods to ensure reliability and validity.

Consequently, positivism is the most appropriate epistemological stance for this research. Positivism enables the objective measurement of sewer defects (e.g., through image analysis algorithms) and supports the statistical analysis required for developing and validating detection models (Bai et al., 2014). For example, when using deep learning models to detect cracks in sewer pipes, positivism would guide the use of large datasets, computational performance metrics, and algorithmic evaluations based on objective, measurable data. This

approach provides clear, repeatable results that are essential for validating the effectiveness of automated systems in real-world applications.

Beyond the adopted paradigm and associated viewpoint in this research, Table 4.1 presents an overview of various research paradigms, their associated methods, ontological and epistemological assumptions, as well as their viewpoints. These highlight the diverse philosophical foundations guiding research approaches and how each paradigm's methodology reflects the underlying beliefs about the nature of reality and the acquisition of knowledge.

Table 4.1: Research Paradigms, Methods, Ontological and Epistemological Assumptions. Source: Author's synthesis based on Al-Ababneh (2020), Crotty (2020), and Saunders (2009)

Paradigm	**Methods**	**Associated Viewpoints**	**Ontological Assumptions**	**Epistemological Assumptions**
Positivism	Quantitative	Objective reality exists, can be measured and quantified using scientific methods	Reality is objective and independent of human perception	Knowledge is derived from empirical observation and measurable data
Post-positivism	Quantitative	Objective reality exists, but our understanding of it is always fallible and limited	Reality exists but can only be imperfectly understood	Knowledge is derived from empirical observation but acknowledged as potentially fallible
Interpretivism	Qualitative	Reality is socially constructed, understood through subjective human experiences	Reality is socially constructed through human experience	Knowledge is derived from subjective interpretation and meaning making
Critical Realism	Mixed methods	There is an objective reality, but it can only be understood through a combination of subjective and objective approaches	Reality exists, but our understanding is mediated through both subjective and objective means	Knowledge is a combination of objective data and subjective interpretation

Paradigm	Methods	Associated Viewpoints	Ontological Assumptions	Epistemological Assumptions
Pragmatism	Mixed methods	Knowledge is derived from both objective and subjective sources, focused on practical solutions and real-world applications	Reality is both objective and subjective, context-dependent, and truth is relative	Knowledge is derived from both objective and subjective sources, using both qualitative and quantitative methods
Feminist Research	Qualitative	Reality is shaped by gender and power dynamics, focusing on marginalized voices and experiences	Reality is shaped by social constructs, especially gender and power dynamics	Knowledge is subjective and rooted in experience, particularly of marginalized groups
Action Research	Mixed methods	Knowledge is co-constructed through collaboration between researchers and participants to create change	Reality is dynamic, shaped by interactions within a specific context	Knowledge is co-constructed through collaboration with participants, aiming to bring about change
Post-structuralism	Qualitative	Reality is constructed through language, culture, and social systems, fragmented and subjective	Reality is constructed through language, culture, and social systems	Knowledge is subjective, shaped by power relations, language, and societal norms
Phenomenology	Qualitative	Knowledge is gained through lived experiences and subjective perceptions	Reality is experienced subjectively, and meaning is given to experiences by individuals	Knowledge is subjective and gained through individual lived experiences and perceptions

4.2.3 Research Design

While the research paradigm establishes the philosophical foundation, the research design provides the actionable framework for data collection and analysis. This study adopts a cross-sectional research design. A cross-sectional design involves the collection of data on multiple variables from a specific population at a single, defined point in time to provide a snapshot of the current state of a phenomenon (Saunders, 2009)

In the context of this research, a cross-sectional approach was deemed the most appropriate strategy. The objective was not to longitudinally track how individual sewer defects degrade over several years (which would require a longitudinal design), but rather to assess the current performance and accuracy of deep learning models in detecting a wide variety of existing defects across different pipes. Data in the form of CCTV survey videos and images were collected from the two participating UK Sewer Management Companies (SMCs) as a singular dataset representing the current state of the infrastructure. This cross-sectional data provided the necessary "snapshot" to train, test, and benchmark the proposed Automated Defect Detection (AuDD) models, thereby fulfilling the quantitative, objectivist methodology established in the preceding sections.

4.3 Research Methods

The methodical research approach refers to the systematic plan for conducting research (Saunders, 2009). It encompasses the strategies, tools, and techniques used to collect and analyse data. Methodical choices are informed by both ontological and epistemological assumptions by ensuring that the chosen methods are consistent with the researcher's philosophical viewpoint and the nature of the research problem (Pereira et al., 2015). In this research, a quantitative approach is adopted based on the need for precise, measurable, and objective data to develop and validate statistical and deep-learning algorithms for defect detection in sewer systems. The quantitative approach involves the collection and analysis of numerical data (Kotronoulas et al., 2023), which is essential for identifying patterns and testing hypotheses related to defect detection in sewer systems.

In this research, a quantitative approach is favoured due to the need for precise, measurable, and objective data to develop and validate statistical and deep-learning algorithms for defect detection in sewer systems. The adoption of a quantitative approach aligns with positivist epistemology, which assumes that reality is independent of human perception and can be quantified using empirical methods (Pereira et al., 2015). In AuDD, the data collected from survey videos of sewer pipes can be objectively analysed using numerical methods. This video data contains measurable features, such as pixel intensity, edge sharpness, or the width of cracks, which can be extracted and quantified into numerical data.

Not only does the quantitative approach facilitate the identification of patterns (Kotronoulas et al., 2023), but it also allows researchers to employ advanced statistical models and deep learning algorithms. Since deep learning, uses mathematical models to automatically extract

features from raw data and analyse patterns within the data, the algorithm is trained on labelled datasets to learn relationships between input data (e.g., video frames of sewer pipes) and the desired output (e.g., detecting cracks or root intrusions). These models rely heavily on mathematical optimization techniques and statistical analysis, such as gradient descent (Fjellström & Nyström, 2022), to minimize errors in defect classification, ensuring high accuracy in predictions (Goodfellow et al., 2018).

The ability to use deep learning models in conjunction with statistical techniques provides a robust framework for analysing large volumes of video data. For instance, deep learning networks can quantify various types of defects in sewer pipes by recognizing specific patterns, textures, and shapes that correspond to known defects. This allows the automated system to analyse hundreds of kilometres of sewer footage, something that would be impractical for human inspectors to do manually. Moreover, statistical validation, such as F1-scores or precision-recall analysis, allows the evaluation of the performance of defect detection algorithms based on objective numerical measures, providing clear, reproducible results (Bai et al., 2014)

Consequently, the quantitative approach, which allows the use of mathematical and statistical models, is essential for achieving the research goal of developing automated systems that can efficiently and accurately detect defects in sewer systems. This approach not only ensures that the data can be systematically analysed but also aligns with the research's philosophical stance that emphasizes objectivity, accuracy, and reproducibility in the detection process.

4.4 Research Framework

Research framework encompasses the complete approach to combine the various elements of the study cohesively and logically by ensuring that the research problem is addressed effectively. In this research, the research framework is specifically tailored towards addressing the complexities of automatic defect detection in sewer systems. As depicted in Figure 4.2, this framework encompasses a systematic approach that includes two comprehensive SLRs, data collection, rigorous pre-processing techniques, advanced model development, thorough evaluation, and deployment. This research leverages historical sewer survey videos and inspection reports to develop and validate deep-learning models aimed at automating the detection of defects in sewer pipelines. The iterative nature of the framework allows for continuous refinement of the models based on performance feedback and real-world application insights.

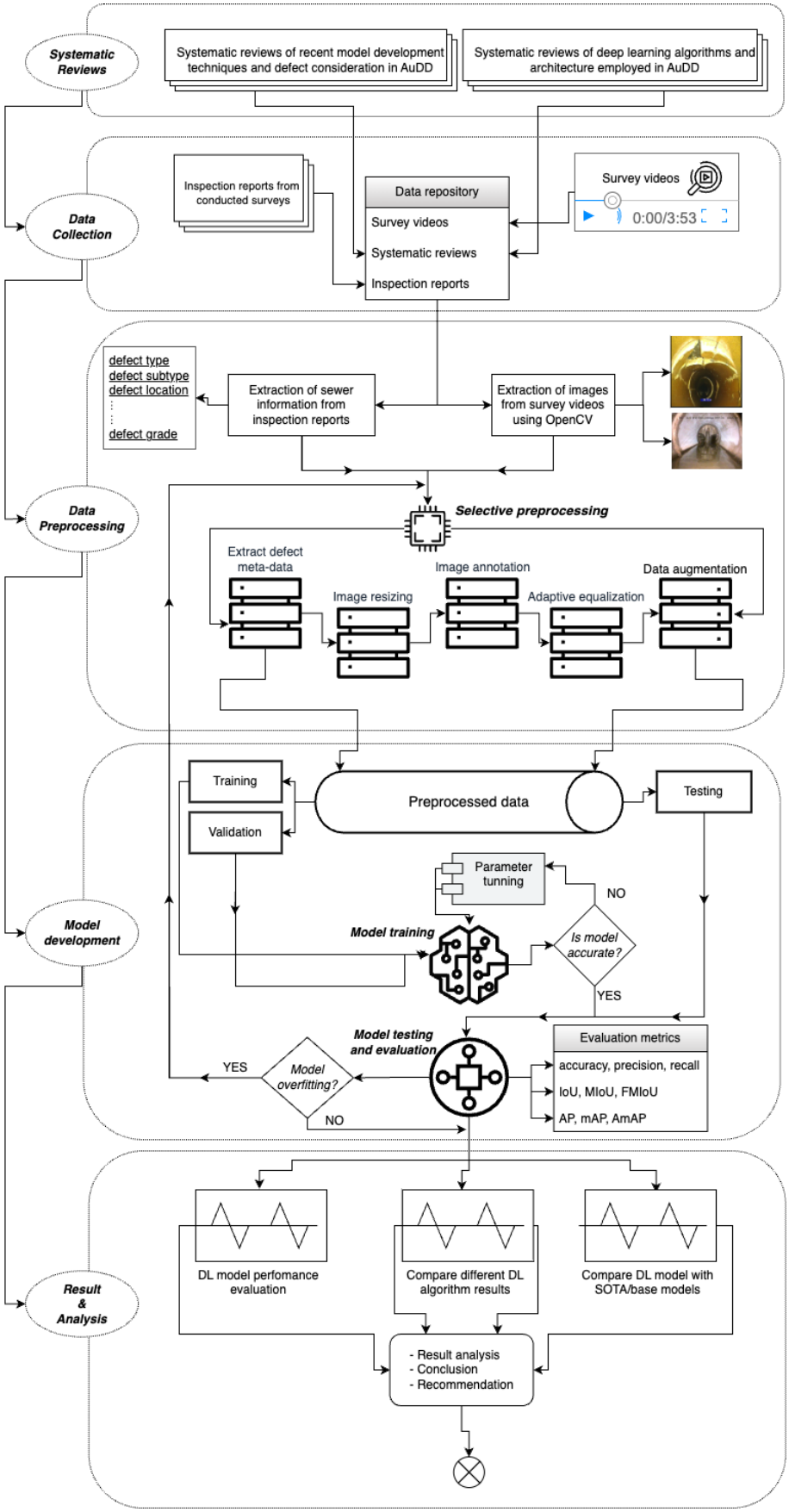


Figure 4.2: Five Phases of the Research Framework (source: Author)

Since the chosen research method is primarily quantitative, it focuses specifically on the systematic analysis of large datasets obtained from sewer inspections. This approach is relevant and justified due to the nature of the data, which consists of extensive video footage and detailed inspection reports that require robust analytical methods to identify patterns and detect defects accurately. This further allows for objective measurement and statistical validation, which are crucial for ensuring the reliability and validity of the findings (Sardana et al., 2023).

Likewise, the quantitative approach in this research provides a structured framework for evaluating defect detection models by ensuring objective, measurable, and reliable results. It allows for the benchmarking of model performance against state-of-the-art techniques and facilitates comparisons across different methods. The key metrics for adopting the quantitative approach are as follows:

1. **Objective Measurement**: The approach allows for precise quantification of defects in sewer images, ensuring that the results are reliable and not influenced by subjective biases (Bai et al., 2020).

2. **Statistical Analysis**: Quantitative methods enable the use of statistical tools to analyse large datasets, which is essential for evaluating model performance, such as calculating precision, recall, and F1-scores (Goodfellow et al., 2018).

3. **Generalisability**: The ability to generalize the findings across different sewer systems ensures that the models are applicable beyond the initial dataset, increasing their real-world utility (Creswell, 2009).

4. **Replication**: The structured nature of quantitative research ensures that the findings are replicable, allowing for verification and reinforcing the reliability and validity of the models (Guba & Lincoln, 1994).

By adopting a quantitative research approach, this study aims to develop robust and generalizable automatic defect detection models that can be applied to various sewer systems. This approach aligns with the research objectives itemised in Chapter 1 and ensures that the findings are objective, reliable, and repeatable. Before delving into the details of each element in the research design, it is important to note that the first component of the research design, SLR, has already been covered in Chapters 2 and 3. With this foundation in place, next is the data collection phase, which forms a crucial component of this research. Other components of the research framework will be discussed in subsequent chapters, starting with data pre-processing in chapter five.

4.5 Data Collection

The process of collecting the appropriate dataset for this research presented significant challenges, particularly regarding data privacy and the non-disclosure agreements (NDAs)

required for accessing sensitive information. Given the nature of the research, obtaining access to historical sewer survey videos and detailed inspection reports was critical. However, the sensitive nature of these data posed considerable barriers to their acquisition. The first step in the data collection process involved reaching out to SMCs across the United Kingdom. Establishing connections with facility managers and technical staff at these organizations was essential not only to gain access to the required data but also to understand the complexities of how data was encoded and standardized across different SMCs.

A key challenge encountered during this phase was ensuring compliance with both data privacy regulations and organizational policies. Each SMC had specific policies concerning the sharing and dissemination of data, given that sewer inspection videos often contain detailed information about local infrastructure, which could be sensitive in nature. For instance, the videos often featured precise geolocation data of sewer systems, which could be used to identify vulnerable or critical parts of the sewer network. The inclusion of this type of information, combined with organizational-specific codes and operational details (e.g., inspection methods, equipment used, and reports generated), made it vital to protect both the physical infrastructure and the operational procedures from unauthorized access.

To address these concerns, it was necessary to negotiate data sharing agreements and establish formal NDAs with each SMC. These agreements were essential to ensure that any data exchanged during the research remained confidential and was not misused or improperly disseminated. For detailed inspection reports, these documents included highly sensitive information such as the specific boroughs or neighbourhoods where inspections took place, the specific sewer lines inspected, and in some cases, internal operational details that are proprietary to the organizations. As such, NDAs were designed to protect the confidentiality of these reports and to prevent the leakage of potentially sensitive operational data. These legal protections were necessary not only to comply with General Data Protection Regulation (GDPR) but also to maintain the trust of SMCs and ensure that sensitive infrastructure data was handled appropriately.

Likewise, understanding the semantics behind data encoding and standardization in each report was another key challenge. Each SMC had its own system for encoding defect data (e.g., types of defects identified, severity levels, and associated metadata), and these encoding systems varied significantly across organizations. For example, one SMC might classify a crack based on its length and severity, while another might include factors like pipe diameter or location. As a result, standardizing this data for use in training automated defect detection models was a complex task. Data pre-processing techniques, such as normalization and mapping across different encoding systems, were applied to ensure that the data from different SMCs could be integrated effectively for deep learning purposes.

4.5.1 Research Data

The primary data sources for this research include sewer survey videos and inspection reports collected from various SMCs across the United Kingdom. These data sources offer a wide array of information essential for developing and evaluating defect detection models. The dataset spans multiple sewerage systems located in 8 major cities across the UK, including London, Glasgow, Leeds, Manchester, Derby, Liverpool, Leicester, and Nottingham. Specifically, this research utilized a combined dataset from two different SMCs within the UK. The first dataset consists of 93 sewer survey videos along with more than 70 detailed inspection reports collected from 28 operational sewers across the UK.

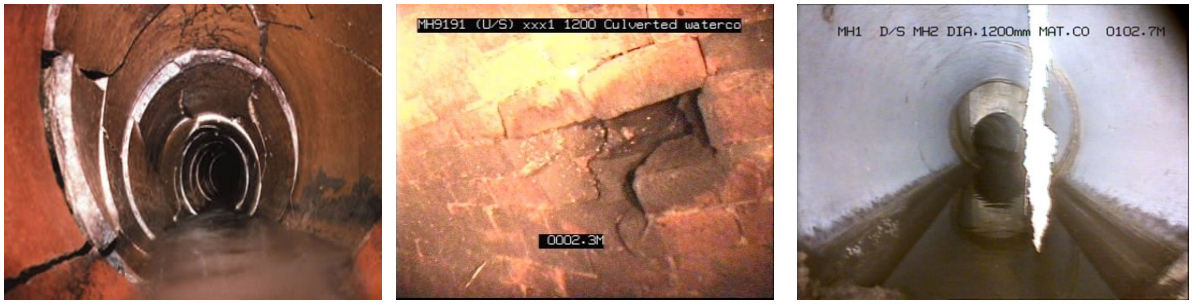
The dataset includes variety of defect types, including 16 different defect categories such as attached deposits, cracks, deformation, encrustation, external pipe issues, fouling, fractures, grease accumulation, infiltration, joint displacement, missing bricks, missing mortar, obstacles, settled deposits, and surface damage. The diversity in the dataset ensures that whatever models are built are exposed to a wide range of sewer system conditions. This variety, which includes both common and complex defect types, is crucial for training deep learning models capable of generalizing across different infrastructures and environmental conditions (Krizhevsky et al., 2012). By incorporating data from multiple cities and SMCs, the models developed in this research are designed to be robust and applicable across diverse sewer systems, which could make them more effective in real-world applications.

In contrast, the second dataset from another SMC consists of 32 survey videos and 24 detailed inspection reports from three operational sewers. While this dataset is smaller in scope compared to the first dataset, which includes 93 survey videos from a wider variety of sewers and locations, it still provides valuable insights into sewer conditions and defect types. The second dataset features eight different defect categories, including attached deposit, crack, encrustation, hole, joint displacement, roots, and settled deposit.

When compared to the first dataset, which covers a broader range of 16 defect categories, it provides a more concentrated view of specific defects in operational sewers. The smaller size of Dataset 2, however, still supports the development of deep learning models but with less variability in terms of the types of defects encountered. This dataset size could affect the model's ability to generalize across various sewer conditions (Dang et al., 2021c; Krizhevsky et al., 2012). Nonetheless, the second dataset complements the first by providing focused data on sewer defects, which, when combined with the larger first dataset, can help improve the model's robustness in detecting a wide range of issues in diverse sewer infrastructures.

Sample defect from both dataset is provided in Figure 4.4. Overall, both survey videos were carefully recorded using high-resolution CCTV cameras mounted on robotic crawlers. These advanced crawlers navigated through the sewer pipelines, capturing detailed footage of the interior surfaces. The video recordings provided a decent view of the sewer conditions, essential for identifying and analysing defects. The duration of these videos varied depending

on the length and overall condition of the pipes, with some videos extending for several hours to cover extensive sewer networks.



(a) Fracture (b) Missing bricks (c) Infiltration

Figure 4.1: Sample frames from the both research dataset (source: Research Data)

In parallel, the detailed inspection reports collected provided a critical metadata about the surveyed sewers. These reports included comprehensive information such as the presence or absence of defects, allowing for a clear understanding of the current state of the sewer infrastructure. Additionally, the types of defects were carefully documented. Each defect was also categorized by severity levels, which helped prioritize maintenance and repair efforts based on the urgency and potential impact of the defects identified. Moreover, the reports included essential details such as the year the survey was conducted, the length of the sewer and the material of the pipes was also included, offering valuable insights into the construction and durability of the sewer system. For instance, knowing whether a pipe is made of vitrified clay, concrete, or brick can inform the selection of appropriate repair techniques.

Furthermore, the inspection reports were prepared in accordance with the European standard coding norm EN 13508-2(BSI, 2006), which is a standard regarding the investigation and assessment of drains and sewer systems located outside the buildings. This ensures consistency and reliability in the recorded data by enabling accurate comparisons and analyses across different segments of the sewer network, thus supporting effective decision-making in sewer maintenance and management. Also, according to the report, the pipes surveyed exhibit a diverse range of materials, shapes, and sizes. Specifically, the materials used for these pipes include vitrified clay, concrete, and brick. Each material is selected based on its durability and suitability for various environmental conditions. For example, vitrified clay pipes are commonly used for their excellent resistance to chemical corrosion, making them ideal for certain waste disposal applications. Concrete pipes are valued for their strength and ability to withstand high loads, often used in areas with heavy vehicular traffic or other significant surface pressures. Brick, though less common in modern installations, is noted for its historical use in older sewer systems, providing insights into the longevity and durability of traditional construction materials.

Morphologically, the pipes also differ in shape and size. The shapes include circular, egg-shaped, and horseshoe configurations, each designed to optimize flow and structural integrity under different conditions. Circular pipes are typically used for their simplicity and ease of installation, while egg-shaped and horseshoe pipes are employed in situations requiring enhanced flow characteristics or structural stability. For the size, the pipes surveyed varies from as small as 200mm to as large as 1800mm, which was intentionally designed to accommodates different volumes of sewage flow and structural requirements. This range in sizes allows for the efficient handling of varying sewage loads, from smaller residential areas to large urban centres. The diversity in pipe characteristics underscores the complexity and adaptability required in maintaining and upgrading sewer infrastructure across different urban settings.

Conclusively, there is a high degree of confidence that the dataset comprehensively covers all required defect categories necessary for this research. This confidence stems from the dataset's strict alignment with the European Standard (EN 13508-2) for sewer visual inspection, which allowed for the successful capture of 16 distinct structural and operational defect types. Furthermore, because these survey videos were sourced across eight geographically diverse UK cities and encompass multiple pipe materials (such as brick, clay, and concrete), the data inherently provides a statistically representative cross-section of real-world degradation patterns, ensuring a robust foundation for training and validating the proposed deep learning models.

Table 4.2 provides a detailed enumeration of the distribution of the collected sewer survey videos, categorized by pipe material. The data reveals that brick is the most common material in this research dataset, constituting the majority with 58 videos, which accounts for almost half of the total dataset. This prevalence of brick pipes highlights their widespread use in the surveyed sewer systems. Although brick is no longer common in modern installations, its majority presence indicates the age of many of the pipes, underscoring the need for quicker and more efficient ways to analyse and maintain sewer infrastructure in the UK. Following brick, vitrified clay and concrete each account for around quarter of the dataset, making up 23.2% and 25.6% of the collection respectively. Vitrified clay pipes are known for their durability and resistance to chemical corrosion, which may explain their significant presence. Concrete pipes, known for their strength and load-bearing capacity, are also prevalent in the dataset.

Table 4.2: Collected videos based on pipe materials

	Pipe Material	**Number of Videos**		**Percentage**
		Dataset 1	**Dataset 2**	
1	Brick	39	19	46.4%
2	Vitrified Clay	29	0	23.2%

	Pipe Material	Number of Videos		Percentage
		Dataset 1	Dataset 2	
3	Concrete	19	13	25.6%
4	Masonry	3	0	2.4%
5	Cast Iron	2	0	1.6%
6	Bitumen	1	0	0.8%

In contrast, materials such as masonry, cast iron, and bitumen are less frequently encountered. Masonry pipes are depicted in 3 videos, constituting just 2.4% of the dataset, while cast iron and bitumen pipes are represented in 2 videos (1.6%) and 1 video (0.8%), respectively. The lesser representation of these materials could be due to their specific use cases or less frequent application in the surveyed regions. This distribution underscores the diversity of materials used in sewer construction and provides a foundation for analysing defects across different pipe materials.

4.5.2 Data Collection Tools and Quality Assurance

Collaboration with SMCs was essential for obtaining the research dataset through provision of access to historical sewer survey videos and inspection reports, which formed the basis of the dataset used in this research. Regular communication and coordination were maintained with the SMCs to ensure the timely and accurate collection of data. This collaboration not only facilitated data acquisition but also provided valuable insights into practical aspects of sewer inspection and maintenance by enriching the contextual understanding of the research. The primary tools and technologies used for data collection included advanced CCTV systems, robotic crawlers, and data management software. The CCTV systems were equipped with high-resolution cameras capable of capturing detailed footage of the sewer interiors. Robotic crawlers, equipped with these cameras, navigated through the sewers while recording continuous streams of forward-facing videos as shown in Figure 4.4.

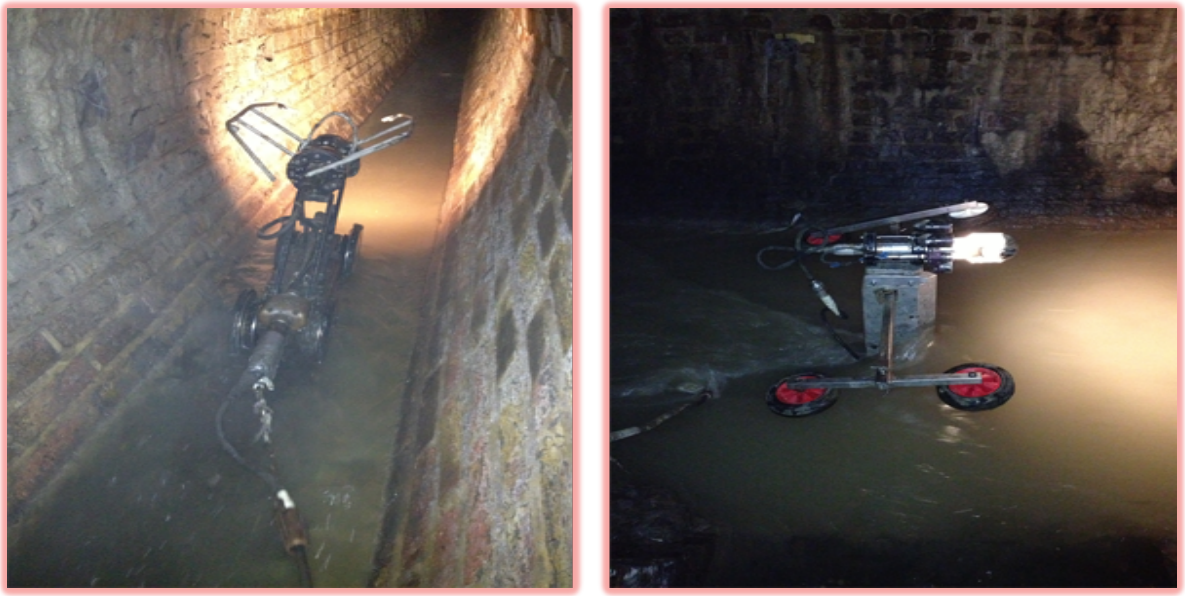


Figure 4.2: Robotic crawler with mounted CCTV camera

Overall, the data collection process followed a systematic approach to ensure the completeness and reliability of the collected data. The key steps involved were:

1. **Coordination with SMCs**: Initial meetings and agreements were established with SMCs to outline the scope of data collection, types of data required, and timelines for data delivery.
2. **Video Recording**: Robotic crawlers equipped with CCTV systems navigated through the sewer networks, capturing continuous video footage of the sewer interiors. The videos documented various sections of the sewers, highlighting potential defects and areas of interest.
3. **Inspection Reports**: In parallel with video recording, inspection reports were collected from SMCs, which provided detailed information about the sewers, including defect descriptions, severity levels, and other relevant metadata.

To Ensure the quality and reliability of the collected data as a paramount criterion of this research, several measures were implemented. First, standardization was a key focus, with all inspection reports adhering to the European standard coding norm EN 13508-2 (BSI, 2006) to ensure consistency in defect categorization and reporting. Additionally, the CCTV systems and robotic crawlers used for data collection were regularly calibrated to maintain the accuracy of the captured footage. This calibration was essential to ensure that the data accurately represented the condition of the sewers. Quality checks were another critical component of the data assurance process. All samples of the collected videos and reports were reviewed by experts to verify their accuracy and completeness, and any discrepancies or errors identified during these reviews were promptly rectified. These rigorous measures collectively ensured the high reliability of the data, forming a solid foundation for developing robust defect detection models.

4.5.3 Data Storage and Management

Following the acquisition of the sewer survey videos, a rigorous data management protocol was implemented to handle the massive file sizes and ensure data security. Given the large volume of high-resolution video files and the proprietary nature of the infrastructure data provided by the SMCs, secure and efficient storage was paramount.

The raw CCTV data, initially received via external hard drives, was immediately transferred to the university's secure, managed storage infrastructure to serve as an untouched master backup. Furthermore, to ensure data integrity and compliance with standard data protection protocols, strict access controls were maintained throughout the research, ensuring that the raw footage was only accessible for authorized model training and evaluation purposes.

4.5.4 Defect Selection

The analysis of both datasets reveals significant insights into the types and frequencies of defects found in sewer systems across the surveyed sewers within the UK. Each dataset, though distinct in its scope and sources, provides a comprehensive overview of the sewer conditions, highlighting common and unique defect types. The first dataset includes 15 different defect categories, showcasing a broad spectrum of issues commonly encountered in sewer systems. The primary defect types reported in this dataset include attached deposits, cracks, deformations, encrustations, and several others, as detailed earlier. On the other hand, the second dataset include seven distinct defect categories, such as attached deposits, cracks, encrustations, holes, joint displacements, roots, and settled deposits. Although this dataset covers fewer defect types, it still represents the critical issues faced by sewer maintenance teams as evident by its existence in both PACP and MSCC coding standards. Hence, there is a need to holistically combine and look at both datasets.

A deeper look at the defective category revealed the distribution of samples across different defect types, as shown in Figure 4.5. The analysis indicated that some defect categories, such as deposits (attached, encrustation, and settled), crack, joint displacement, and root had a higher number of samples compared to others. Conversely, categories like fractures, external pipe, and surface damage had fewer samples. In this research, the focus on specific defects is informed by three key factors: first, the similar defect types identified in both datasets which highlights recurring issues within sewer systems in the UK. Likewise, the SLR of existing studies detailed in chapters two and three confirms that some of these defects are prevalent in sewer systems across different countries, further validating their significance for this research.

Furthermore, defects that are commonly observed, visually identifiable, and have significant immediate effects on sewer infrastructure were prioritized for inclusion, such as cracks and deposits. In contrast, infiltration, despite its substantial presence in the dataset, was excluded due to its challenging detection in video footage (Nguyen et al., 2021). Infiltration typically manifests as water seepage that is not easily visible or distinguishable in CCTV survey videos, making it less suitable for an automated detection system. Thus, the cut-off was made to focus

on defects that could be visually identified to ensure the development of effective and accurate defect detection models. Consequently, defect such as attached deposit, crack, encrustation, joint displacement, root intrusion, and settled deposit will only further be considered in this research for detailed analysis.

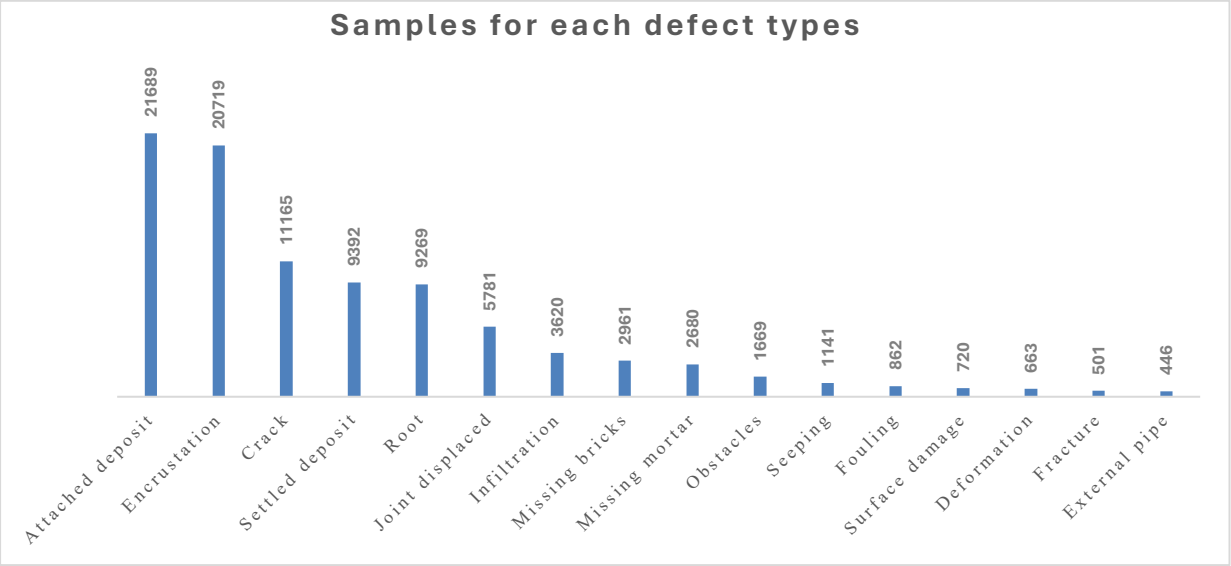


Figure 4.3: Number of samples in each defect category (source: Author)

Table 4.3 presents the list of all defect types identified from both datasets along with their implication for sewer systems. It also highlights the common defects from both dataset and indicates the selected defects for detailed analysis in this research.

Table 4.3: List of all defects from both datasets

Defect Type	Implication	Dataset 1	Dataset 2	Selected?
Attached Deposit	Obstruct flow and potentially causing blockages.	✓	✓	✓
Crack	Compromise the integrity of the sewer pipe.	✓	✓	✓
Deformation	Changes in the shape of the pipe due to external pressures or internal failures.	✓	-	-
Encrustation	Hard deposits on the inner surfaces of the pipes, typically from mineral build-up.	✓	✓	✓
External Pipe	external to the sewer line, potentially affecting the pipeline's performance.	✓	-	-

Defect Type	Implication	Dataset 1	Dataset 2	Selected?
Fouling	Accumulation of debris and organic matter that can reduce flow efficiency.	✓	-	-
Fracture	Severe breaks or splits in the pipe, often requiring immediate attention.	✓	-	-
Grease	Build-up of fats, oils, and greases that can cause blockages and foul odours.	✓	-	-
Hole	Leads to significant leaks and structural weaknesses.	-	✓	-
Infiltration	Unwanted entry of water which can lead to overloading of the sewer system.	✓	-	-
Joint Displaced	Separation of pipe joints which can lead to leaks and structural weaknesses.	✓	✓	-
Missing Bricks	Absence of bricks can lead to structural instability.	✓	-	-
Missing Mortar	Absence of mortar between bricks that can cause potential structural failure.	✓	-	-
Obstacles	Physical objects within the sewer that obstruct flow.	✓	-	-
Root Intrusion	Intrusion of roots into the sewer pipes, which can cause blockages and structural damage.	✓	✓	✓
Settled Deposit	Sediments and other materials that have settled at the bottom of the pipe	✓	✓	✓
Surface Damage	General wear and tear on the pipe's inner surfaces.	✓	-	-

4.6 Chapter Summary

To summarize, this methodology chapter clearly outlines the structured approach adopted to create and validate an effective AuDD system for sewer defect identification. It thoroughly justifies the quantitative approach chosen, grounded in an objectivist and positivist research

paradigm due to the requirement for precise, measurable, and objective data essential for robust statistical analysis and deep learning model development. The chapter also provides detailed accounts of data collection methods, including collaborations with SMCs, adherence to data privacy regulations and non-disclosure agreements, and quality assurance processes.

The dataset used in this research is comprehensive, covering a diverse range of sewer systems across major cities in the UK, including London, Glasgow, Leeds, and others, ensuring a broad geographical representation. It includes 16 defect categories from 93 survey videos and 32 videos from a smaller second dataset, with defects ranging from cracks and joint displacements to more complex issues like infiltration and root intrusions. The data is standardized according to EN 13508-2, ensuring consistency across different SMCs. This diversity in both geographical location and defect types provides a promising foundation for training the deep learning models by enhancing their ability to generalize across various sewer conditions.

The rationale behind defect selection, particularly the exclusion of infiltration due to detection challenges, further reinforces the careful consideration of defect characteristics and operational significance. Together, these methodological considerations provide a solid foundation for subsequent chapters, which will delve into data pre-processing, model training, and evaluation processes.

CHAPTER FIVE

5.0 Data Analysis and Preprocessing

5.1 Chapter Overview

This chapter provides a detailed overview of the data analysis and preprocessing techniques employed to prepare the research datasets for further analysis and model development, a critical step in enhancing data quality and model performance for automatic defect detection. It begins with image extraction and data analysis, where images are extracted from sewer inspection videos and categorized to provide statistics on different defect types. It then details various preprocessing techniques, including duplicate removal, image resizing, and text elimination using OCR and image inpainting. Variants of custom contrast enhancement and morphological operations techniques are also implemented to improve image clarity and structure, followed by existing and custom image segmentation to isolate defect regions.

5.2 Need for Data Preprocessing

Data pre-processing is an essential phase in the context of deep learning, particularly for tasks involving image data (Harshini et al., 2022), such as the automatic detection of defects in sewer systems. Deep learning models, especially CNNs, rely heavily on the quality, quantity, and consistency of the input data (He et al., 2022), and it is well known that raw data collected from real-world environments, such as sewer survey videos, often contain noise, inconsistencies, and irrelevant information that can negatively impact the performance of these models. In this research, the need for data pre-processing is particularly critical due to the diverse and challenging conditions under which the sewer data was collected. While the survey videos vary in lighting conditions, image quality, and the presence of obstructive elements, these variations can introduce biases and inaccuracies if not addressed appropriately.

The preprocessing techniques employed in this study aim to improve the clarity, reliability, and usability of the images for subsequent analysis. These techniques are crucial for training deep learning models, as it allows the algorithms to focus on learning the features relevant to defect detection rather than being distracted by extraneous information (Mumuni & Mumuni, 2024). The selected techniques in this research include image extraction and resizing, text processing, data annotation, contrast enhancement, morphological operations, and image segmentation. While each of these steps is essential to ensure that the images are suitable for training and evaluating the defect detection models, a detailed description of each preprocessing technique

and how it was specifically implemented and tailored to this research is provided in subsequent subsections.

5.3 Image extraction and analysis

To carefully capture the intricate details of the targeted sewer system, it was necessary to obtain images at regular intervals throughout the entire video, surpassing conventional survey footage practices. The OpenCV library was employed to convert each video into a continuous stream of images. The extraction process was designed to sample approximately 25 images per second from each survey video since this is the FPS used by the recording device (Schiappa et al., 2023), ensuring a comprehensive depiction of the sewer's landscape as the crawler navigated through it. Figure 5.1 showcases a selection of sample images extracted from different sewer locations.



Figure 5.1: Sample sewer images extracted from different sewer videos. (source: Author)

Subject to the duration of each video, thousands of images were extracted. Table 5.1 summarizes the number of images extracted from survey videos in five different sample sewers. Overall, the extracted images varied significantly in dimensions, which could negatively impact the model and the pixel-wise spatial information. To address this, the images were standardized to a consistent resolution of 640x640 pixels, a dimension determined to be optimal for the employed object detection algorithms (Hu et al., 2023; Liu et al., 2024). Resizing the images to a uniform dimension helps reduce computational resource usage while maintaining the spatial information and correlation within the images.

Table 5.3: Sample video conversion to images

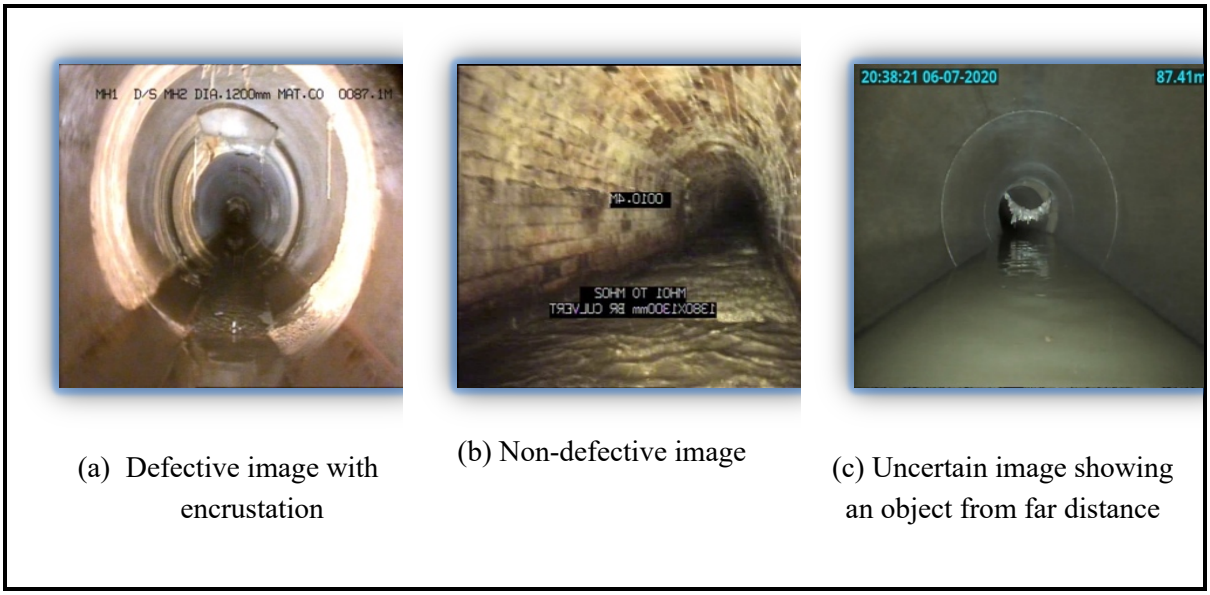
	Sample Sewer	Video duration	Number of Images
1	Sewer 1	28:55 mins	43371
2	Sewer 2	17:33 mins	26,275
3	Sewer 3	61:50 mins	92,750
4	Sewer 4	95:03 mins	141,050
5	Sewer 5	97:26 mins	145,800

Likewise, all extracted images were manually checked, carefully selected, and grouped as defective, non-defective, and uncertain using the detailed inspection reports collected alongside the survey videos. The categorization is based on their content, as detailed in Table 5.2. Here, *defective* images contain a visible sign of a known defect, such as one or more of the ones selected for this research. For *non-defective,* they are images void of any form of visible defects, while the *uncertain* category includes images that either show a defect from afar or invalid due to distortion of the video acquisition device.

Table 5.4: Categories of extracted images

	Categories	**Description**	**No of Images**
1	Defective	Images showing one or more categories of defect.	103,078
2	Non-defective	A normal image with no defect.	1,298,324
3	Uncertain	Images showing defect at a far distance or blurry defect due to cameral tilting during survey.	104,011

Figure 5.2 further shows a sample of defective, non-defective, and uncertain images extracted from different sewers. These categorizations also indicate that over 75% of the images were non-defective, while the defective and uncertain categories comprised 7% each. This analysis underscores a common characteristic observed in operational sewers: majority of the images captured during inspections tend to be normal, non-defective images, with only a small proportion exhibiting defects (Shen et al., 2023). This pattern holds true for the targeted sewers used in this research as well.



(a) Defective image with encrustation

(b) Non-defective image

(c) Uncertain image showing an object from far distance

Figure 5.2: Defect image categorization (source: Author)

Following the selection of 6 different defect types for this research, a further analysis of the defective category revealed the distribution of samples across different defect types. Among the defective images, the most prevalent defect type was attached deposits, with 21,689 samples and encrustation followed closely with 20,719 samples. Cracks were the third most common defect, with 11,165 samples, and settled deposits accounted for 9,392 samples, while root intrusion, a frequent issue in older sewers, was observed in 9,269 samples. The least represented defect in the selected categories was joint displacement, with 5,781 samples.

Given the survey videos were processed at a rate of 25 frames per second, the resulting dataset inevitably contains many duplicate images, making it essential to implement a thorough duplicate removal process to optimize the dataset for model training and analysis by reducing unnecessary repetition and focusing on unique, informative samples. Additionally, there is a notable *class imbalance* within the selected defect category samples, further underscoring the need for careful pre-processing to ensure that the model is trained on a balanced and representative dataset.

5.3.1 Duplicate removal

In this research, the removal of duplicate images is a critical pre-processing step. Given that the sewer survey videos were processed at a rate of 25 frames per second, the resulting dataset inevitably contains a high number of duplicate images. These duplicates, while ensuring that no defect is missed, also introduce redundancy that can negatively impact the performance and efficiency of the defect detection models (Cumbajin et al., 2023). An excessive number of similar images can lead to overfitting, where the model becomes too familiar to specific instances within the training data, thereby reducing its generalization ability when encountering

new data. Furthermore, the presence of duplicates worsens the existing class imbalance within the dataset, potentially skewing the model's predictions towards more frequently occurring images. Therefore, it was essential to implement a robust duplicate removal process to optimize the dataset by ensuring that it remains focused on unique and informative samples that will enhance the model's learning and performance. To address this, two primary approaches were employed for duplicate removal in this research.

The first approach involved revisiting the original survey videos containing the identified defects. Here, instead of extracting 25 frames per second, the videos were reprocessed to select only one frame from the 25 images extracted per second as proposed by (Khan et al., 2020). This method significantly reduces the number of duplicate images while retaining the critical frames necessary for accurate defect detection. By selectively sampling the frames, the dataset is not only reduced in size but also in redundancy, allowing the model to train on a more diverse set of images that better represent the range of possible defects.

The second approach to duplicate removal was more manual and involved a detailed examination of the defective images for any similarities based on content and sequence. Like (Kumar, et al., 2021a), this was done by tracking the metadata associated with each image, particularly the meterage displayed during the survey, which indicates the exact location within the sewer where each image was captured. By comparing contiguous images, any images that appeared overly similar or identical were flagged for removal. Figure 5.3 shows a process flow of how the two-stage duplicate removal was carried out and the number of duplicates removed at different stages.

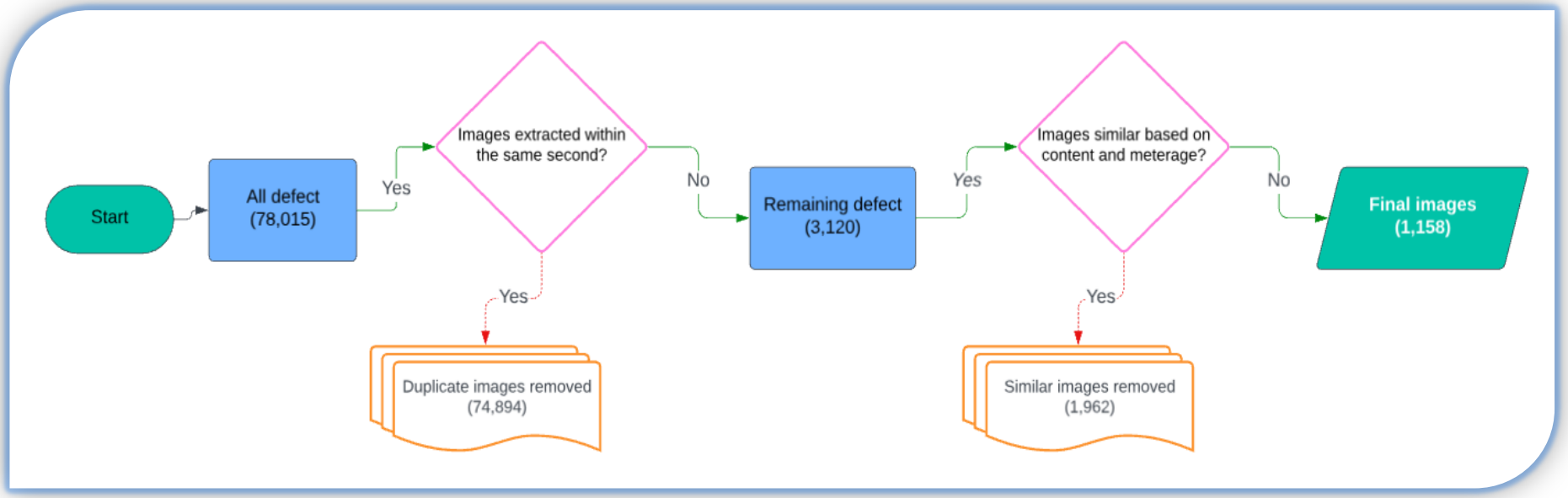


Figure 5.3: Two-stage duplicate image removal process (source: Author)

This manual process ensured that only one representative image of each defect was retained at a point and allowed for the careful consideration of whether defects observed in contiguous frames were sufficiently distinct or whether they represented the same defect from slightly different angles or same defect that cut across different frames. However, through the implementation of these duplicate removal approaches, the dataset was significantly reduced from 78,015 images to just 1,158 distinct defect images, as detailed in Table 5.3.

While this reduction was necessary to eliminate redundancy and improve the dataset's quality, it also highlights a critical challenge for deep learning algorithms which are inherently data-dependent, requiring large and diverse datasets to build robust and accurate models (Situ et al., 2023; Sun et al., 2024). The small size of the remaining dataset, though free from duplicates, is insufficient to train a reliable defect detection model on its own.

Table 5.5: Final raw dataset after duplicate removal

	Defect Type	**No of Samples**
1	Attached deposit	320
2	Crack	227
3	Encrustation	294
4	Settled deposit	188
5	Root	129

Consequently, it became necessary to get an alternative but reliable source of data to augment the original dataset. This augmented dataset will provide the necessary quality, volume, and variety of examples needed to enhance the model's learning capabilities and ensure its effectiveness when deployed for real-world use cases.

Despite the necessity of this manual removal phase, the process presented several significant challenges. Primarily, the extreme visual homogeneity of sewer pipes combined with poor, inconsistent lighting made it exceedingly difficult to distinguish between structurally similar, yet distinct, contiguous frames. Furthermore, the CCTV metadata (such as the meterage text overlay) was occasionally blurred, obscured by water droplets, or entirely missing, hindering accurate location tracking. To address these challenges, a strict heuristic was applied: when visual confirmation was inconclusive, the sequence was cross-referenced with the video's exact timestamp rather than relying solely on the meterage. Additionally, by executing the automated first stage (extracting only 1 fps) *before* the manual review, the sheer volume of frames was drastically reduced, thereby mitigating fatigue and significantly reducing the margin for human error during the visual similarity checks.

5.4 Data Augmentation

Data augmentation is a crucial step in enriching the variety and robustness of the dataset used for training machine learning models (Siu et al., 2022), particularly in scenarios where the available data is limited or imbalanced. In this research, data augmentation techniques were employed to both increase the overall samples and mitigate under sample categories by applying a 2-fold augmentation approaches to the original images to improve the generalization capabilities of the model and reduce the possible risk of overfitting (Jiang et al., 2023; Srivastava et al., 2014). This research took 2 approaches to augmenting the research data: image-level and bounding-box level augmentation.

5.4.1 Image-Level and Bounding-Box Augmentation

These augmentation techniques involve applying many alterations to the existing images to create new, diverse images. Here, the alteration was applied at both the image and bounding-box level. Image-level alterations were applied to create a modified copy of the original images by altering its appearance in ways that can help the model learn to recognize defects under different conditions. These alterations include operations such as rotation, flipping, scaling, shifting, and changing brightness or contrast, among others. Figure 5.4 represent series of image-level alterations that were applied.

For this research, *OpenCV* (OpenCV, 2023) and *Albumentations (Albumentations, 2023)* were the primary tools used for building the data augmentation pipeline. Both tools were combined to create a dynamic (such as method chaining) and robust augmentation pipeline to carry out basic to advance augmentation operations, as depicted in Figure 5.4. For example, images were rotated at various angles (e.g., 15°, 30°, 45°) to simulate the variability in how defects might appear in different orientations within the sewer. Flipping operations, both horizontal and vertical, were also applied to reflect the different perspectives from which the camera might capture defects.

Additionally, the images were scaled and translated across the x and y axes to mimic the variations in defect sizes and locations. Furthermore, photometric transformations such as brightness and contrast adjustments, and geometric transformations like elastic transformation and random cropping were also applied

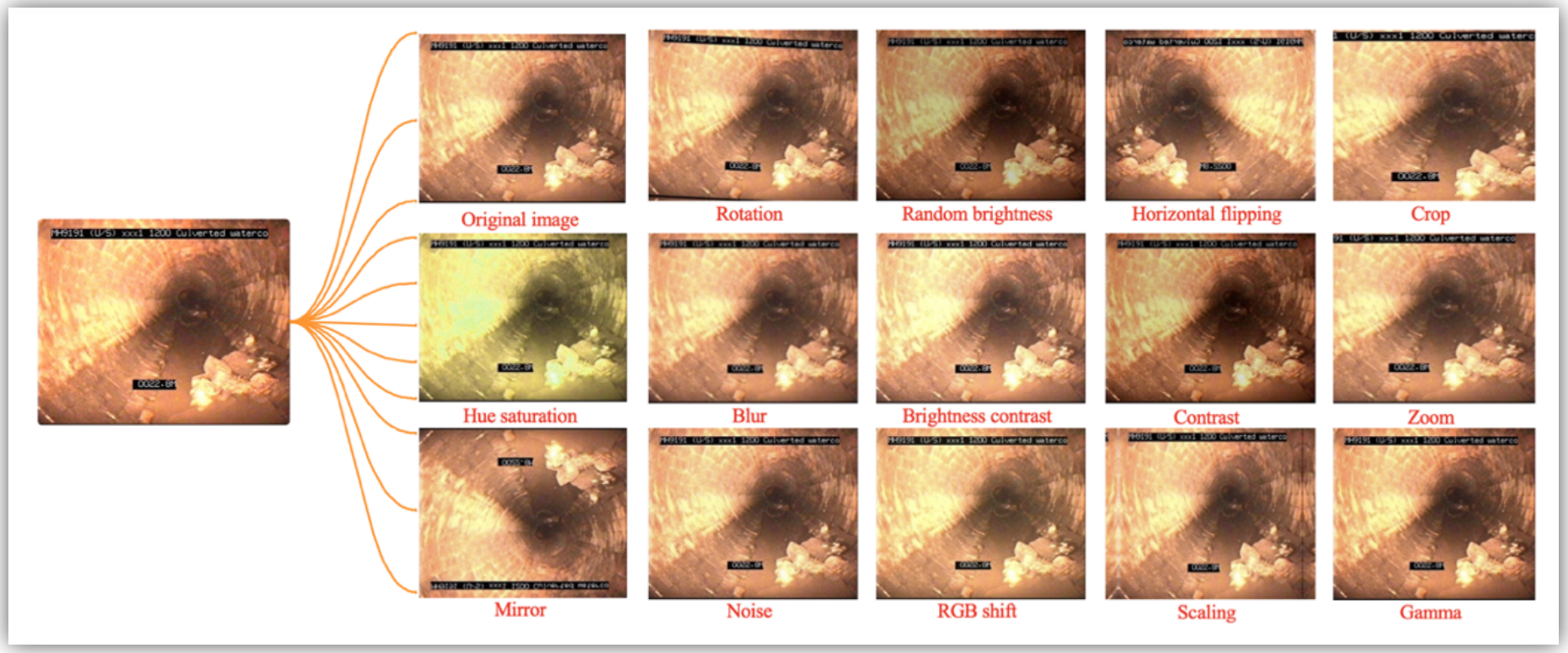
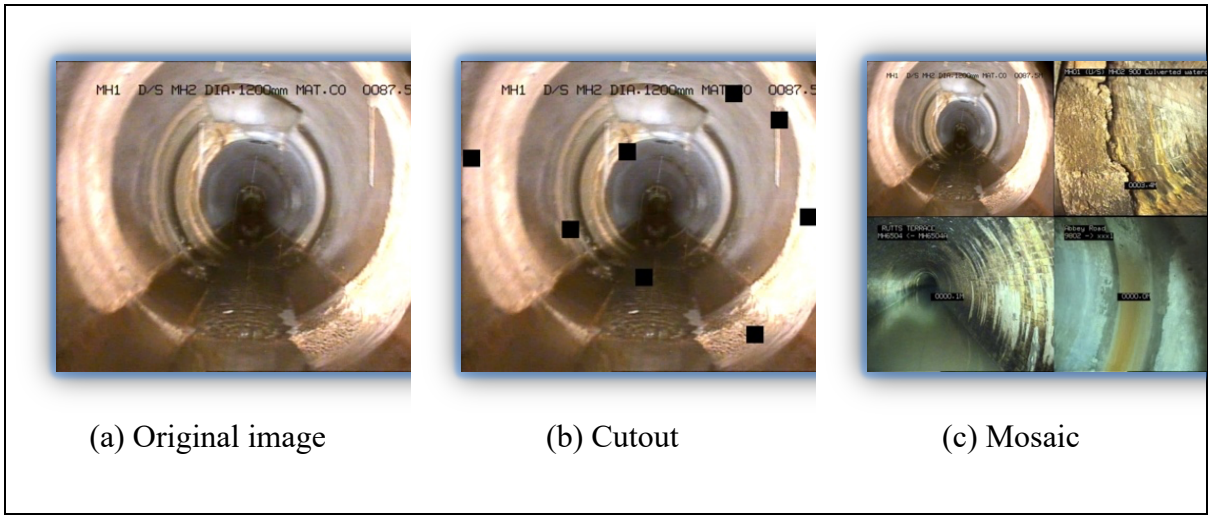


Figure 5.4: Conventional augmentation techniques applied to original dataset (source: Author)

Likewise, advanced image-level augmentations such as *cutout* and *mosaic* were utilized to create more challenging training scenarios. Cutout augmentation involves randomly removing parts of the image to force the model to learn to detect and localize defects even when part of the defect is obscured. Mosaic on the other hand, combines different images (or parts of images) into a single image to create a more complex visual environment where multiple defects may appear in different locations within a single frame. This technique not only increases the variety of training examples but also allows the model to learn from images with multiple defects, enhancing its ability to handle overlapping or closely spaced defects. Figure 5.5 depict a sample cutout and mosaic operation that was carried out on images with variation of encrustations.



(a) Original image (b) Cutout (c) Mosaic

Figure 5.5: Cutout and Mosaic augmentation examples for training (source: Author)

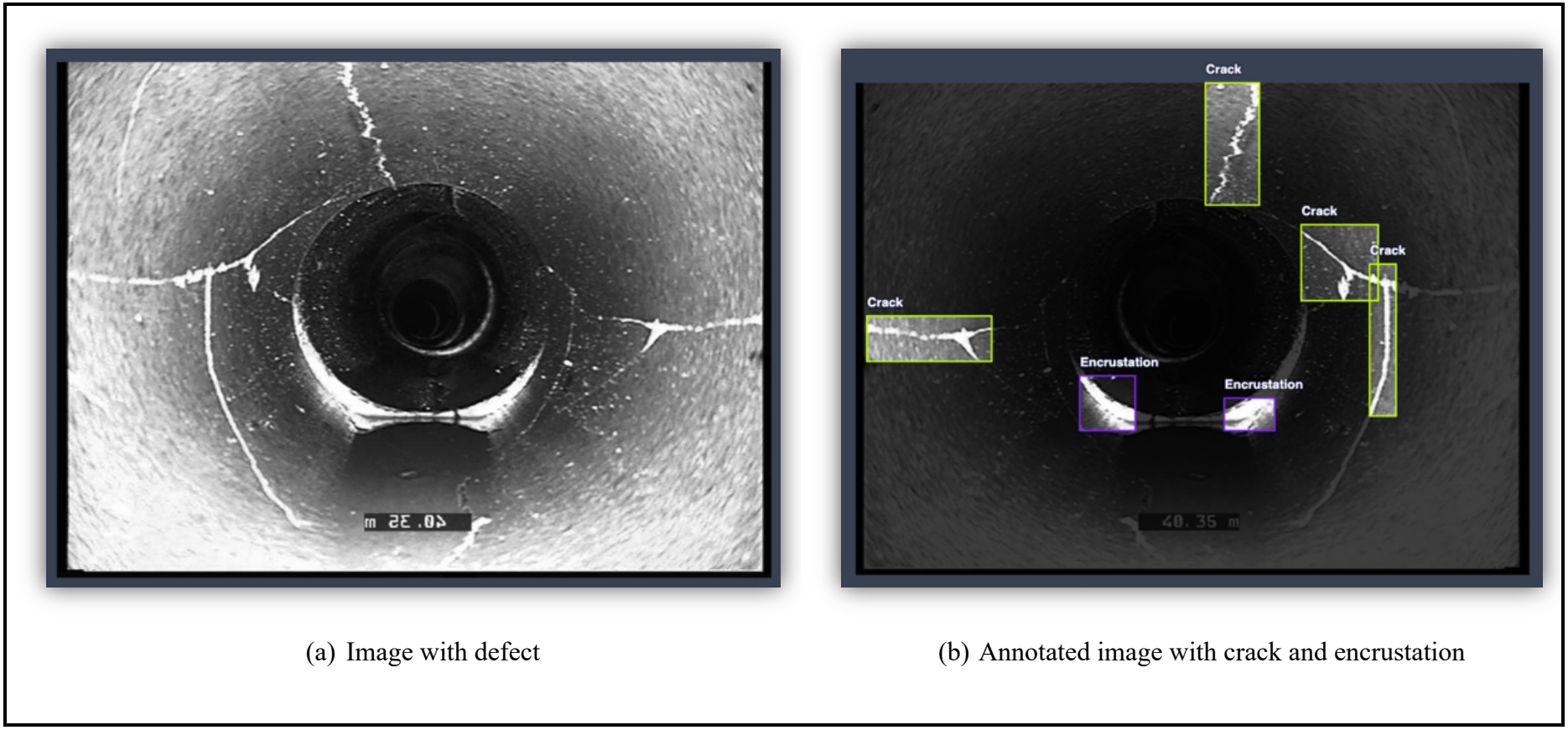
Following the extensive image-level augmentation, bounding-box (bbox)-level augmentation (Devi et al., 2023) was also applied to further enhance the dataset. Bbox augmentation involves modifying the bounding boxes that surround defects (or region of interest) within images. Since the proposed model to be developed will not only need to identify whether a defect is present but also accurately locate it within the image, then bbox transformations such as bbox scaling, translation, and rotation were performed. For example, a bounding box was slightly resized to simulate variations in defect size and another one shifted within the image to mimic changes in the defect's position. This was carried out to ensure that the model tries to learn to detect and locate defects even when their size or position varies, which is typical of a real-world sewer scenario.

It is worthy to note that both augmentation types were employed in this research not only to expand the overall dataset but also to specifically address the class imbalance problem (Arjun et al., 2024; Ghosh et al., 2024) present within the defect categories, as earlier presented in

Table 5.3. In specifics, particular emphasis was placed on defect categories with fewer samples, such as root intrusion and settled deposit, to artificially increase their representation in the dataset. This targeted augmentation approach will help ensure that the deep learning model could learn to detect all defect types regardless of their initial frequency in the dataset and potentially solve the class imbalance problem.

5.5 Image Annotation

Image annotation enables the precise labelling of defects within images in object detection problems (Chenyi et al., 2017; Gardenier et al., 2018), which is critical for training accurate and reliable deep learning models. However, annotating large number of images for object detection is a difficult and time-consuming task. At first, the defective images were separated into respective defect categories being considered; crack, root intrusion, and 3 variations of deposits – settled, attached, and encrustation. Next, we annotated the images by both drawing bounding boxes around the defects and assigning a label to each defect. These steps were carried out using *MakeSense* (Skalski, 2023), a web-based annotation tool that facilitates the creation of bounding boxes and labelling for computer vision tasks. Figure 5.6 shows an example of a multi-label annotation process, which involves the annotation of a crack and encrustation within the same image.



(a) Image with defect

(b) Annotated image with crack and encrustation

Figure 5.6: Multi-label image annotation using MakeSense (Piotr, 2023).

Given the multi-label nature of the task, we labelled various defects across different categories within each image, adhering to a set of consistent and predefined annotation rules. Following the annotation process, the labelled image and equivalent annotation metadata was saved and exported in both *.xml* and *.txt* formats as a prerequisite for training the deep learning model.

5.6 Grayscale and Contrast Enhancement

The use of grayscale images has been a common trend in the field of automated inspection (Harshini et al., 2022). Initially, standard CCTV grayscale images have been primarily applied, while other types of images, such as SSET (Haurum & Moeslund, 2020) and unwrapped CCTV images, have been used less often. Grayscale images are preferred over colour images in the early stages of the algorithmic pipeline due to better contrast and the ability to detect defects using simple edge-based and geometric approaches [2]. Additionally, grayscale images have been used for noise reduction and enhancing image quality in sewer pipe defect detection [4]. Therefore, grayscale images can be part of various stages of defect detection in different domains. Owing to this, grayscale was applied in 3-folds for comparative analysis: initially, only a subset of the images was processed into grayscale, and subsequently, the entire image set was converted. Lastly, we use a combination of grayscale and other selected pre-processing techniques. This multi-phase approach was designed to assess the impact of grayscale on the accuracy and efficiency of defect detection on sewer images.

For contrast enhancement, these techniques are essential for improving the visibility and distinction of defects within sewer images, which often suffer from poor lighting conditions, shadows, or variations in material texture. Enhancing the contrast in these images helps to highlight critical features of some defects such as cracks, deposits, or intrusions, making them more distinguishable to automated detection algorithms. Contrast enhancement techniques such as Adaptive Equalization (AE), Contrast Stretching (CS), and Histogram Equalization (HE) were employed as pivotal steps for enhancing the quality of sewer images before defect detection. AE was particularly valuable in its local approach to contrast adjustment, capable of intensifying details in underexposed segments of the images without compromising overexposed areas, a common occurrence in variably lit sewer environments. This local enhancement, however, comes with the trade-off of potentially amplifying noise, which might result in the misidentification of defects.

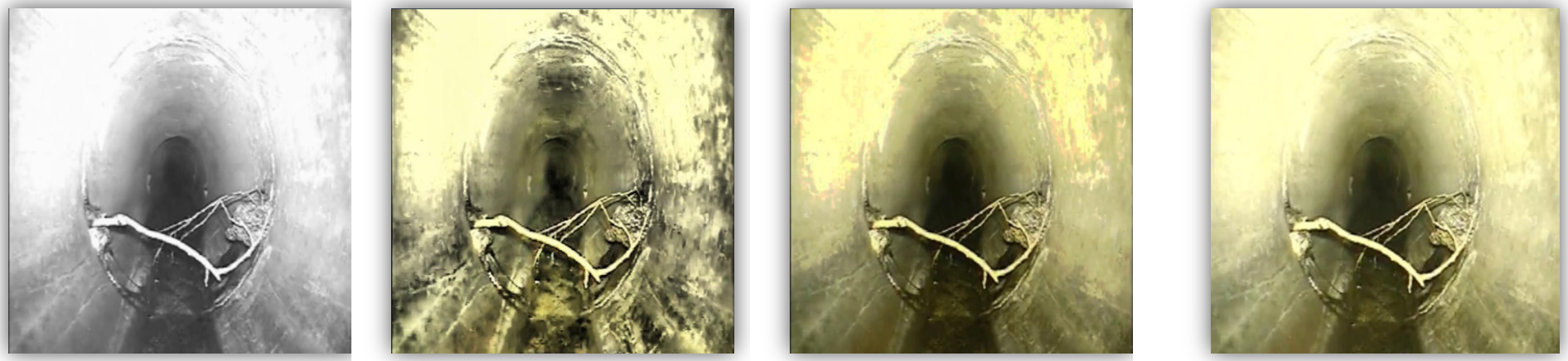


Figure 5.7: Different pre-processing techniques considered (source: Author)

CS on the other hand, has been previously used to detect defects in sewers (Sim et al., 2013) using a filter technique through mathematical morphology that can expose cracks. Here, it highlights the cracks and efficiently removes noise while retaining the edges. A comparison of crack extraction techniques shows that the algorithm based on contrast stretching achieves better results than other techniques (Ma et al., 2021). HE was also implemented to redistribute the image's brightness, proving effective for uniformly lit backgrounds (Xiong et al., 2021) by enhancing overall contrast. It further facilitates edge detection of defects in an image by enabling detection of all defects of interest within a single image. The comparative effectiveness of AE, CS, and HE was a crucial point in this research, with each technique undergoing rigorous evaluation through its addition into the training phase of multiple deep learning models.

5.7 Morphological Operations

Widely used to refine the structural features of images, morphological operations play a crucial role in enhancing the quality of images by manipulating their geometrical structure. This is especially crucial in this research, where the images captured during inspections often suffer from noise and fluctuating levels of clarity. The goal here is to ensure that defects are more distinguishable and correctly represented for further analysis by deep learning models. To achieve this, operations such as *erosion, dilation*, *opening*, and *closing* was systematically implemented on a subset of the dataset at first, and then on a copy of the entire dataset.

Initially, erosion was applied to erode the boundaries of every other object in the image aside the defect. Technically, this was achieved by convolving the image to contract every other object in the image. Invariably, the aim here is to remove unimportant objects and noise that do not represent actual defects. For example, little fragments of dirty grease or minor visual alterations that could potentially confuse the detection algorithm were eliminated. However, while trying to eliminate the impact of unimportant objects, this operation also reduces the size of the actual defects, making the image too smooth with important details eroded away along with the noise.

Hence, dilation was applied to the images to offset the effects of erosion via addition of more pixels to the boundaries of the remaining objects (i.e., the defect). This way, the defects can be restored to their original or slightly enlarged sizes, making them more prominent or easier to detect.

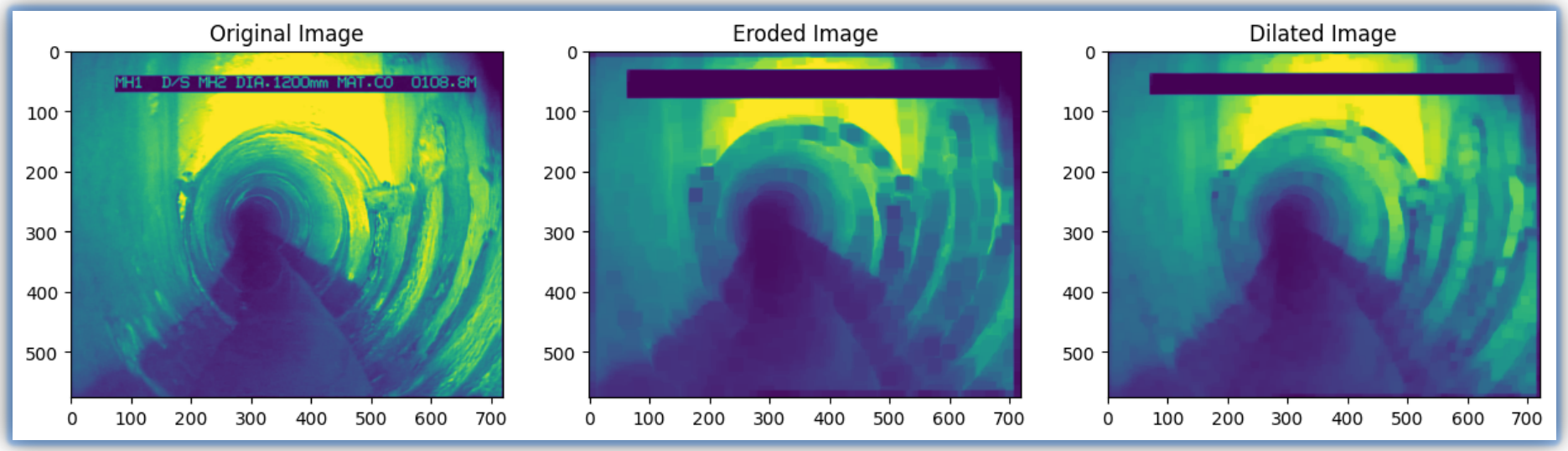
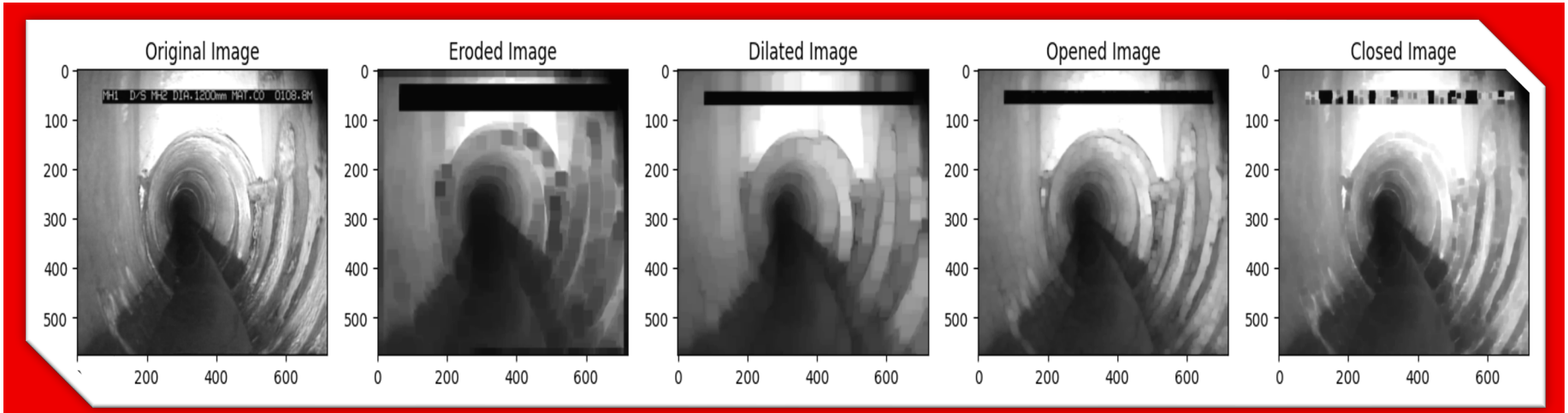


Figure 5.8: Sample applications of eroded, dilated (source: Author)

Even though erosion and dilation effectively performed their intended tasks of reducing noise and enhancing significant features, including the successful removal of all text from the images since the text was deemed unimportant, this process alone has not eliminated subtle noise or adequately restore critical defect features after previous operations, leading to inaccuracies in the defect representation. As a result, additional processing such as opening and closing were applied to further refine the images, remove noise, and preserve the defect. Each operation was applied to the grayscale and RGB version of the dataset, and Figure 5.9 shows the transition from the original image all through the operations.



(source: Author)

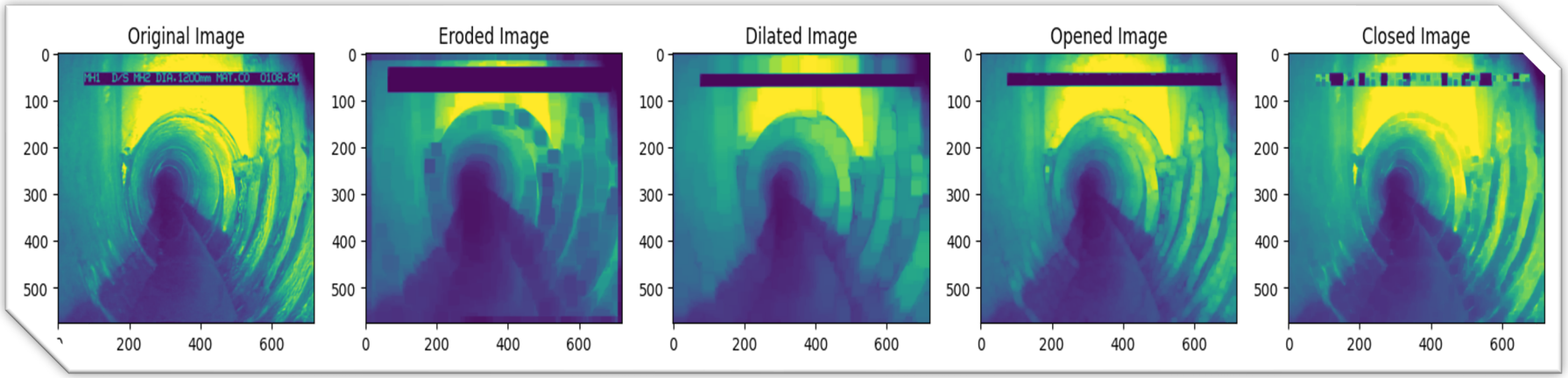


Figure 5.9: Transition from the original image all through the morphological operations (source: Author)

5.8 Image Segmentation

Image segmentation is a critical pre-processing technique in computer vision, particularly for tasks like defect detection in sewers (Fang et al., 2022b), where the goal is to isolate and analyse specific regions of interest (i.e., defects) within an image. Here, a custom image segmentation was systematically implemented using a fusion of techniques such as *thresholding*, *clustering*, and *edge detection*, each of which plays a distinctive role in enriching the visibility and distinction of the defects.

5.8.1 Thresholding

While simple, thresholding (Sim et al., 2013) is yet an effective image segmentation technique that involves transforming grayscale images into binary images, where the pixels are either black or white based on a specified *threshold value*. The aim here is to highlight defects with obvious contrast against their background, such as cracks or encrustations, to further make them easier to separate. Invariably, for each pixel, a uniform threshold value is used - if the pixel's intensity is below the threshold, it is assigned a value of 0 (turning the pixel black); otherwise, it is assigned the predefined maximum value. However, the challenge was how to select the right threshold value for each defect. To mitigate this, several threshold values was used, ranging between 0 and 200, as depicted in Figure 5.10. After several attempts, a threshold value of 127 and maximum value of 255 was adopted because they worked well for most of the defects.

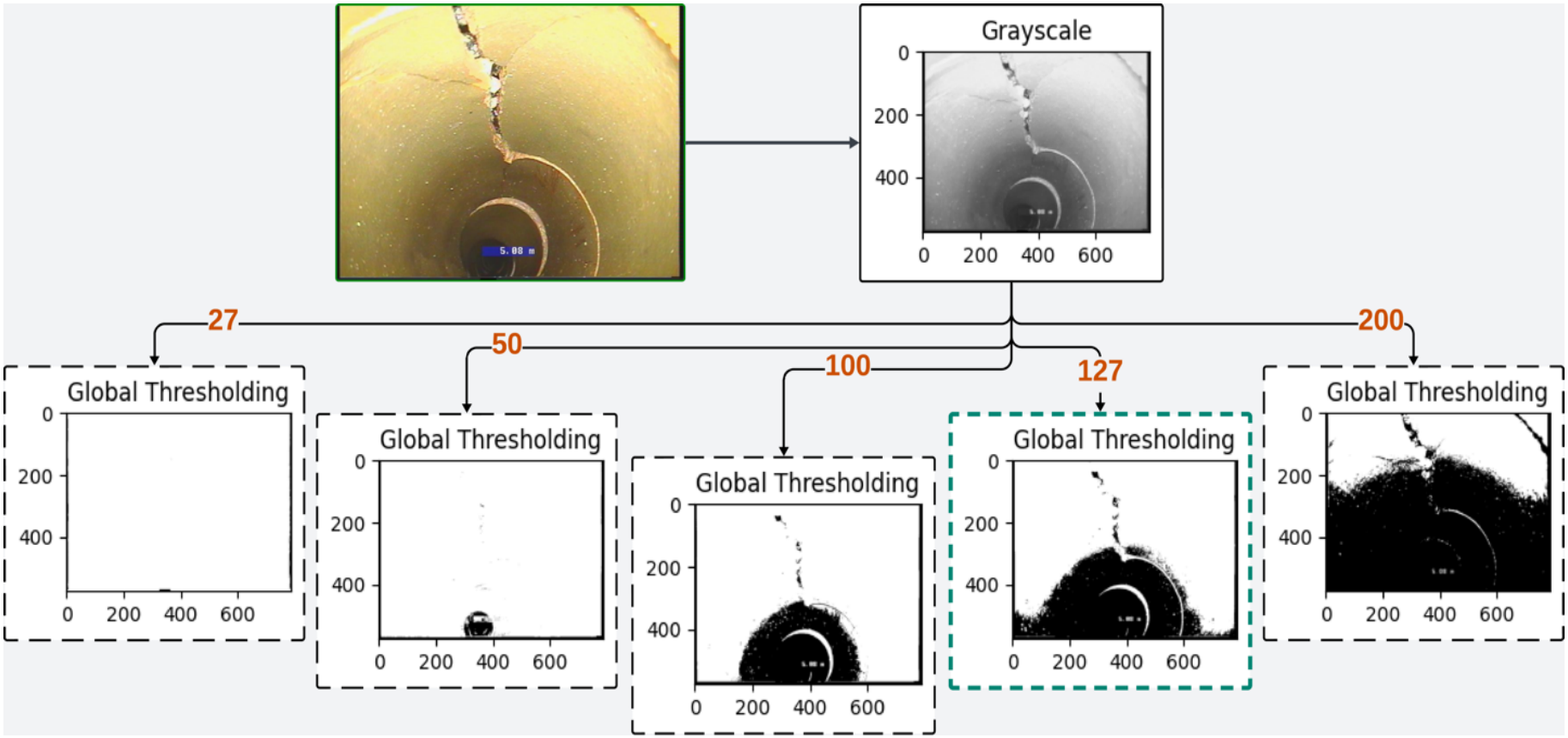


Figure 5.10: Variations of threshold values implemented. (source: Author)

Using this threshold, other variants of global thresholding like *binary, binary inverse*, and *zero shot* was experimented to see which of them works best for the defects, as seen in Figure 5.11. However, it was less effective for some images with cracks where the contrast between defects and the background is minimal and some normal areas withing the image was transformed to be a defective area because of its intensity. Consequently, *adaptive thresholding* was used to mitigate these cases.

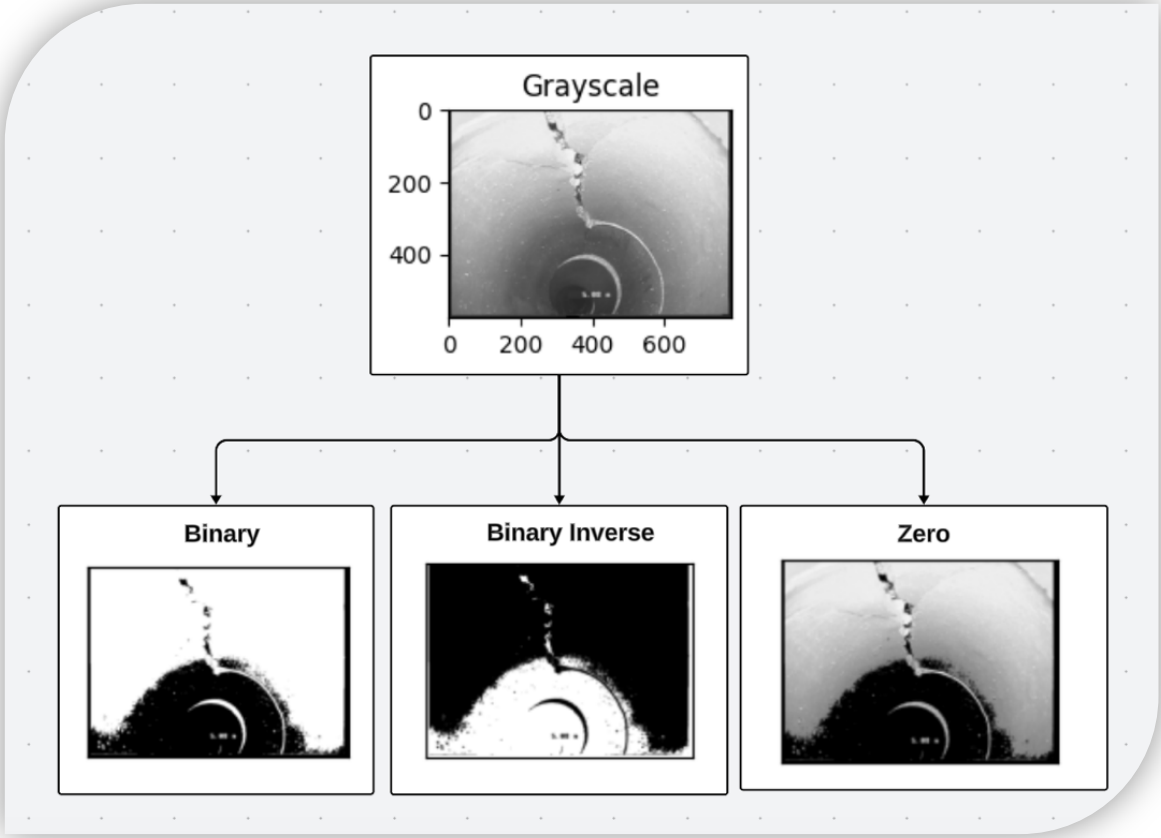


Figure 5.11: Output of Binary, binary inverse, and zero shot operations. (source: Author)

Unlike global thresholding that uses a single threshold across the whole image, adaptive thresholding calculated the threshold for smaller regions of the image for better segmentation in cases where the lighting or intensity varied across the image. This was particularly useful for images where parts of the defect might be visible due to inconsistent lighting during the sewer inspection. Here, different types of adaptive thresholding were implemented, including *mean* and *gaussian thresholding*. Furthermore, a custom thresholding technique named *MEGA* was developed in this research. This technique is a high-level fusion of mean and Gaussian thresholding methods. MEGA was designed to leverage the strengths of both approaches by calculating the threshold value as the mean of pixel intensities within a specified neighbourhood and applying a weighted sum of pixel values based on a Gaussian window. By combining these methods, MEGA provides a more adaptive segmentation, especially in images where lighting conditions or intensity variations are uneven. The adaptive gaussian and custom

MEGA approach provided a better result as seen in Figure 5.12, hence they are also adopted for this research.

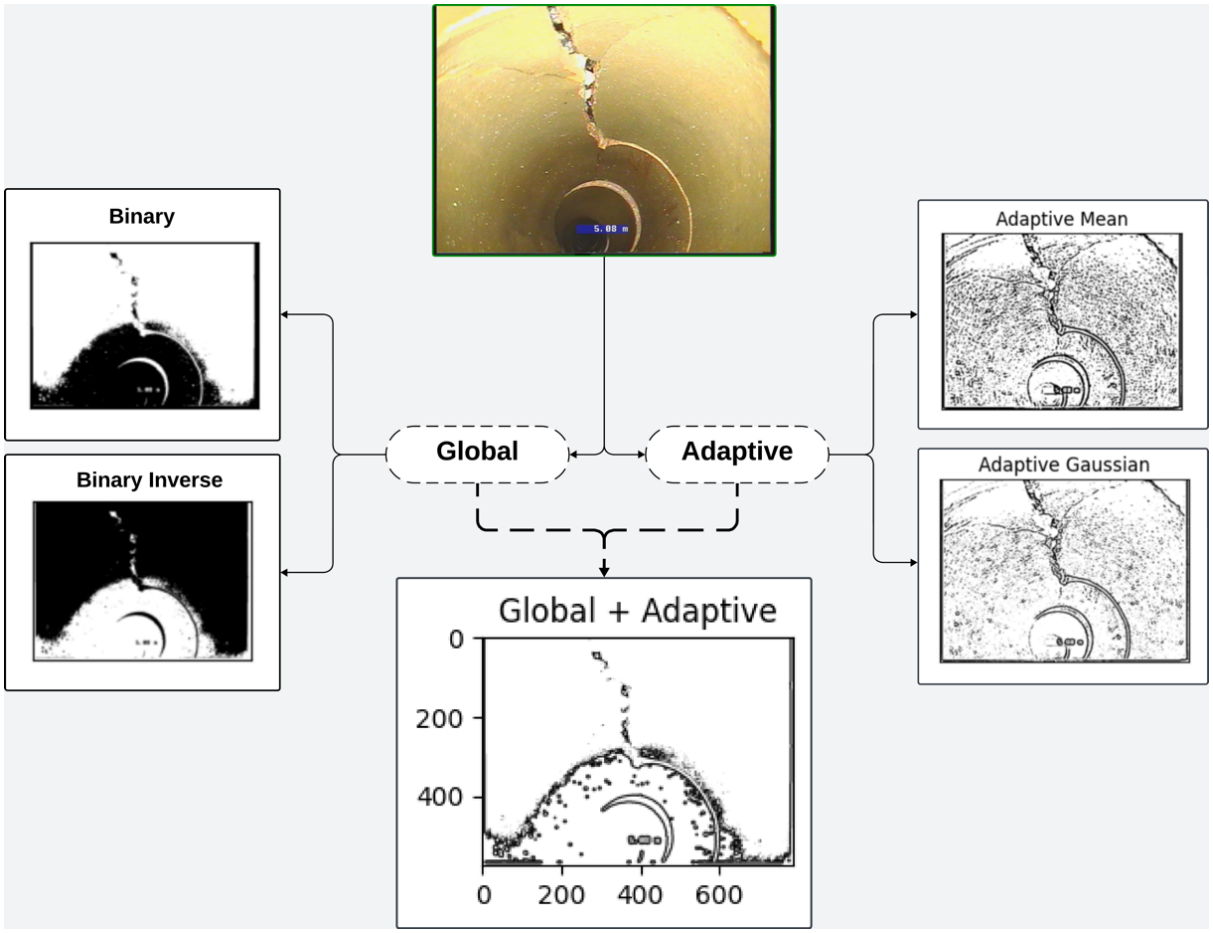


Figure 5.12: Adaptive and MEGA gaussian implementation output (source: Author)

5.8.2 Clustering

Clustering (Sasmal & Dhal, 2023) is another segmentation technique that was adapted in this research for instances where multiple defects coexisted or where the defects blended with the background. This was achieved by grouping intensity-related pixels into similar clusters. The *k-means* clustering algorithm was employed to assign each pixel to a specific cluster based on its intensity. To ensure the effectiveness of this technique, careful consideration was given to the six defects under study as well as the variability of defects present in each image. Some images contained multiple defects, while others had only one, necessitating an adjustment in the number of clusters used. Consequently, depending on the image, a custom *k-cluster* was designed to assign dynamic cluster value (from 2 to 6) to capture this variability and ensure decent segmentation across the dataset. Figure 5.13 shows a dynamic assignment of 2 clusters to an image with only encrustation.

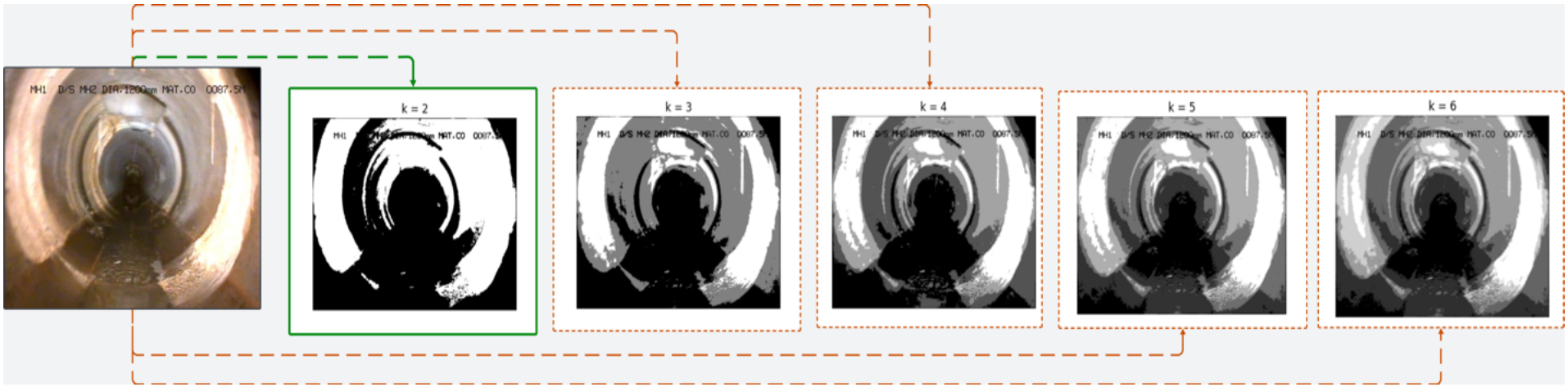


Figure 5.13: Trial implementation of clustering using different clusters. (source: Author)

To implement this custom approach, an excel file was prepared to map the number of defects in each image to the image name, then the image can be dynamically assigned a cluster number. Next, the image was reshaped into a 2D array of pixels to prepare it for clustering. The criteria for clustering were then defined, and the *k-means* algorithm was applied to group pixels into clusters according to their intensities. After clustering, the labels were mapped to the centre points of the clusters, and the data was converted back to 8-bit values. Finally, the clustered data was reshaped back to its original image form. This allowed for fewer clusters in images with a single dominant defect and more clusters in images with multiple overlapping or separate defects, thereby enhancing the accuracy of the segmentation process across the entire dataset.

5.8.3 Edge Detection

Edge detection plays a critical role in identifying the boundaries and contours of defects in sewer inspection images, and several techniques were systematically evaluated to identify the most effective approach for this research. The Canny edge detection algorithm (Canny Edge Detection 09gr820, 2009), known for its precision and accuracy, was initially applied to both RGB and grayscale versions of the images, with careful adjustments made to the threshold values and gradient magnitude to optimize the detection of relevant defect edges while minimizing noise. Additionally, gradient-based methods such as Sobel (Vincent & Folorunso, 2009) and Prewitt (Yang et al., 2011) operators were employed to detect edges by calculating the first derivatives of image intensity, which proved effective in highlighting structural features of defects, particularly in areas with strong defect contrast against the background. The Laplacian operator (Xin Wang, 2007), which calculates the second derivatives, was also explored for its ability to detect rapid intensity changes, thereby capturing edges in images where defects were less distinct from the background. However, through a detailed comparative analysis of the results from these edge detection techniques, it became evident that the Canny edge detection algorithm provided the most consistent and reliable performance, as seen in Figure 5.14.

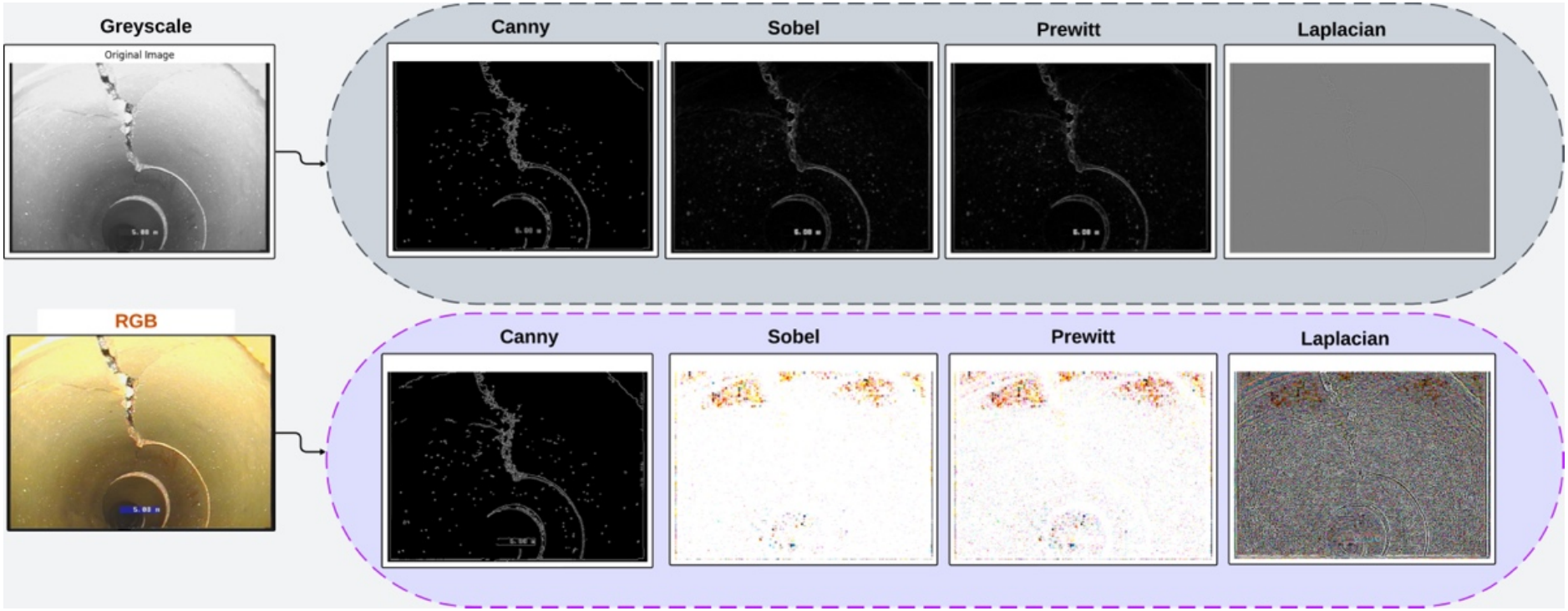


Figure 5.14: Outputs from edge detection operations (source: Author)

While Sobel, Prewitt, and Laplacian operators showed decent result with greyscale images, they struggled with consistency across different images, particularly when switching between RGB and grayscale formats. In contrast, the Canny algorithm consistently delivered clear edge detection, maintaining the integrity of critical defect features while decently minimizing noise.

5.9 Chapter Summary

This chapter presented a comprehensive and quantitative overview of the data analysis and pre-processing pipeline developed to transform raw CCTV survey footage into a robust dataset optimized for deep learning. The initial extraction from the survey videos at 25 frames per second yielded a massive dataset of 78,015 defective images, which were standardized to a 640x640 pixel resolution to reduce computational complexity while maintaining spatial correlation. To address the severe redundancy inherent in video frame extraction, a systematic two-stage duplicate removal process was implemented. This critical step drastically condensed the dataset to 1,158 highly distinct defect images across five target categories: attached deposits (320), encrustation (294), cracks (227), settled deposits (188), and root intrusions (129).

To mitigate the resulting class imbalance and limited sample size, targeted two-fold data augmentation (image-level and bounding-box level) was applied using techniques such as cutout, mosaic, and geometric transformations. Following precise multi-label annotation using MakeSense, the dataset was subjected to rigorous, defect-specific image enhancement. This included comparative implementations of Contrast Stretching (CS), Histogram Equalization (HE), and Adaptive Equalization (AE), alongside morphological operations (erosion, dilation, opening, closing) to successfully refine defect boundaries and suppress visual noise.

A major technical finding of this chapter was the development and empirical success of the custom 'MEGA' thresholding technique. By effectively combining Mean and Gaussian thresholding methods, MEGA demonstrated superior segmentation accuracy in poorly illuminated sewer environments. Additionally, dynamic k-means clustering (utilizing 2 to 6 clusters) effectively isolated overlapping defects, while Canny edge detection was proven to deliver the most consistent and reliable boundary identification when compared to Sobel, Prewitt, and Laplacian operators. Ultimately, the quantitative refinements, dimensionality standardizations, and custom pre-processing approaches detailed in this chapter establish a highly optimized, structurally balanced foundation for the training and evaluation of the deep learning architectures presented in Chapter 6.

CHAPTER SIX

6.0 Development of Deep Learning Models for AuDDS

6.1 Chapter Overview

The model development process was designed to address the specific challenges of detecting sewer defects using deep learning, with a primary focus on the YOLO families of algorithms as previously detailed in Chapter 3. YOLO's real-time object detection capabilities made it particularly suited for handling the diverse types, sizes, and degrees of visibility found in sewer defects. This chapter outlines the practical and structural approach taken to build an effective model, starting with an array of preprocessing techniques, from contrast enhancement to advanced morphological operations and segmentation methods. Each preprocessing step was implemented to enhance image quality and to help improve YOLO's performance in detecting defect features within the dataset.

Selecting the right architecture and training methodology was fundamental to this model development phase. The decision to use YOLO followed a comprehensive overview of various deep learning models selected, which was discussed in Chapter 3. While Chapter 3 provided the theoretical foundation for choosing YOLO, this chapter focuses on the practical implementation. An overview of the YOLO architecture and specific training strategies illustrate how this model was optimized for both accuracy and efficiency, balancing the need for real-time performance with high detection accuracy. The sections that follow describe the experimental setup, including the hardware and software configurations, and delve into the model architecture, data preparation, and training processes implemented in this research.

6.1 Experimental Setup for AuDDS

The experimental setup for this research was designed to handle the computationally intensive tasks required for YOLO's deep learning architecture, as well as the preprocessing tasks needed to enhance the data. Both the hardware and software configurations were carefully selected to optimize the model training and testing process.

6.1.1 Hardware Configuration

To effectively train the YOLO model and support various preprocessing techniques, a high-performance computing setup was employed, featuring a multi-GPU environment comprising NVIDIA Tesla V100 and T4 GPU, each equipped with at least 12GB of VRAM. This configuration enabled parallel processing images, which significantly reduce the training time and enhanced model convergence. The selected GPUs also played a crucial role in accelerating tasks such as contrast enhancement, image segmentation, and morphological operations, ensuring that model iterations were performed without computational bottlenecks. Likewise, the usage of multiple GPUs allowed for computationally intensive larger batch-sizes and more variations of hyperparameter tuning.

In addition to GPU acceleration, the system utilized a high-core CPU with at least 64GB of RAM to manage the computational load of data preprocessing and augmentation tasks. Operations such as adaptive histogram equalization, edge detection, and clustering-based segmentation demanded substantial memory for real-time execution and intermediate data storage. The CPU efficiently handled non-parallelizable tasks, including dataset loading, augmentation scheduling, and preprocessing pipeline management, ensuring smooth execution without memory limitations. This combination of high-performance GPUs and extensive RAM capacity provided the necessary computational power to process large-scale datasets while maintaining efficiency and reliability throughout the experimental workflow. Figure 6.1 below shows the hardware confirmation and sample tasks category for both CPU and GPU-accelerated tasks.

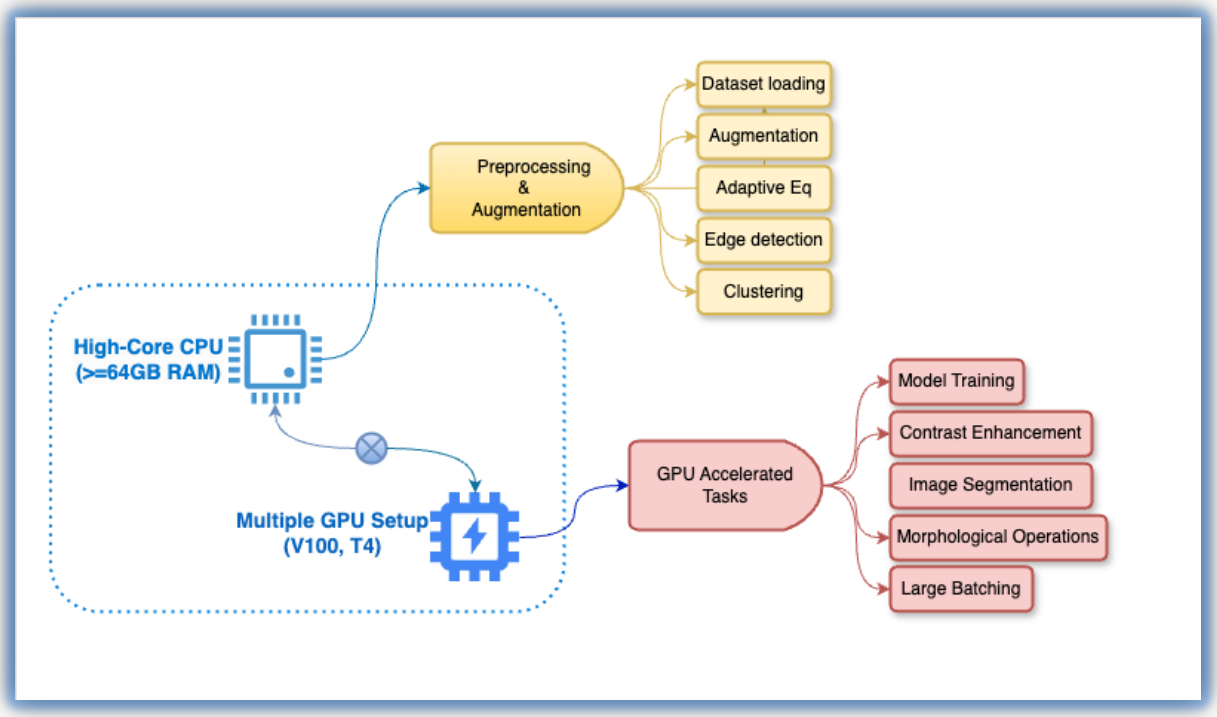


Figure 6.1: Hardware configuration for multi-GPU setup (source: Author)

6.1.2 Software Configuration

The software environment was meticulously configured to support deep learning and computer vision tasks essential for sewer defect detection. All experiments were conducted within a Python 3.10 environment, leveraging PyTorch 2.2.1 as the primary deep learning framework due to its flexibility and extensive community support. This allowed seamless customization of the YOLO model architecture, optimizing it based on prior research findings (S. I. Hassan et al., 2019; Y. Zhou et al., 2022b). To fully utilize the multi-GPU setup, CUDA was integrated for efficient GPU acceleration, ensuring that computationally intensive processes, including model training and inference, were executed with minimal latency.

In addition to PyTorch, OpenCV played a crucial role in implementing key preprocessing steps such as grayscale conversions, morphological operations (dilation and erosion), and thresholding techniques to enhance defect visibility in images. Supporting libraries like NumPy facilitated efficient data manipulation, while Matplotlib enabled detailed visual analysis of model predictions. This combination of deep learning and computer vision tools provided a robust environment for image preprocessing, model development, and performance evaluation, ensuring seamless execution of both preprocessing and inference workflow.

To enhance scalability and deployment efficiency after the model development, the system was containerized using Docker, ensuring a consistent and portable runtime environment across different machines and cloud platforms. API development was facilitated using FastAPI to allow the trained YOLO models to be deployed as a real-time inference service with low-latency responses. The API provided endpoints for image preprocessing, model inference, and result visualization, making the defect detection system accessible via RESTful calls. For cloud deployment, the free tier of AWS services was leveraged to host the model. This deployment strategy ensured that the trained model could be efficiently accessed, tested, and scaled in real-world applications. Granular details about the model deployment are provided in the next chapter.

6.2 Model Architecture Overview

This research fine-tunned different algorithms within the YOLO family of object detection algorithms, especially the last few iterations developed by *Ultralytics,* which was proven as SOTA in both batch and real-time object detection (Bawankule et al., 2023; Orchi et al., 2023). Specifically, YOLOv5, YOLOv6, YOLOv7, YOLOv8, YOLOv9, YOLOv10, and YOLOv11 were selected as explained in chapter 3. These algorithms integrate substantial architectural advancements to enhance both accuracy and speed, making it particularly suited for sewer defect detection, where varied defect types, sizes, and visibility levels must be accurately captured. The next paragraph summarises the evolution of different algorithms within the YOLO family of object detection algorithm.

Starting with YOLOv1 in 2015 as a one-shot object detector using the Darknet framework (Joseph, 2023), subsequent iterations introduced transformative features such as multi-scale training (v2), SPP blocks (v3), and CSPDarknet-53 backbone with Mish activation (v4). The transition to PyTorch in v5 brought anchor-free detection and SWISH activation, while v6 incorporated self-attention mechanisms (Vaswani et al., 2017). YOLOv7 leveraged transformers and E-ELAN reparameterization, further refining detection capabilities. YOLOv8 integrated GANs, advancing anchor-free detection, followed by YOLOv9's innovative PGI and GELAN for enhanced feature aggregation. YOLOv10 introduced NMS-free training and dual-label assignments for superior accuracy, and YOLOv11 refined backbone and neck architectures, emphasizing adaptability across varied environments. A summary of YOLO evolution is summarized as shown in Table 6.1.

Table 6.6: Comparison of YOLO Variants.

Version	Date	Contributions	Framework
v1	2015	One-shot object detector	Darknet
v2	2016	Multi-scale training, dimensional clustering	✓
v3	2018	SPP block, Darknet-53	✓
v4	2020	Mish-based activation, CSPDarknet-53 backbone	✓
v5	2020	Anchor-free detection, SWISH-based activation, PANet	PyTorch
v6	2022	Self-attention, anchor-free object detection	✓
v7	2022	Transformers, E-ELAN reparameterization	✓
v8	2023	GANs, anchor-free detections	✓
v9	2024	Programmable Gradient Information (PGI), Generalized Efficient Layer Aggregation Network (GELAN)	✓
v10	2024	NMS-free training approach, dual label assignments, holistic model design for enhanced accuracy and efficiency	✓
V11	2024	Improved backbone and neck architecture, Adaptability Across Environments	✓

Given that PyTorch is the chosen Python framework for this research, as detailed in Chapter 3, the selection of specific algorithms from the analysed ones was also focused on those supported by PyTorch and not older and less supported Darknet. This decision also considered the algorithm's compatibility with different computer vision tasks and its ability to deliver low latency alongside high accuracy. Table 6.2 highlights the capabilities of these algorithms. The aim was not only to identify algorithms suitable for this research but also to select those capable of achieving the research objectives by striking a balance between speed and accuracy, ultimately addressing the limitations commonly associated with one-stage detection algorithms (Wikimedia Commons, 2023).

Table 6.7: Task coverage for different algorithms.

Category	**Sub-category**	***v5***	***v6***	***v7***	***v8***	***v9***	***v10***	***v11***
Supported Task	Classification	X	X	✓	✓	X	X	✓
	Detection	✓	✓	✓	✓	✓	✓	✓
	Instance Segmentation	X	X	✓	✓	✓	X	✓
	Semantic Segmentation	X	X	X	✓	X	X	✓
	Pose/Keypoints	X	X	✓	✓	X	X	✓
	Oriented Detection	X	X	✓	✓	X	X	✓
Operations	Training	✓	✓	✓	✓	✓	✓	✓
	Validation	✓	✓	✓	✓	✓	✓	✓
	Inference	✓	✓	✓	✓	✓	✓	✓
	Export	✓	✓	✓	✓	✓	✓	✓

Likewise, the models were compared in terms of image size, mAP across IoU thresholds (50-95), inference speed on CPU and T4 TensorRT (where available), parameter count (in millions), and computational cost (FLOPs in billions) using the COCO benchmark dataset (Lin et al., 2014). As provided in Table 6.3, algorithms like v6 and v8 use smaller image sizes (640 pixels) to achieve competitive mAP scores (57.2% and 55.3%, respectively) while maintaining lower computational costs (FLOPs) compared to larger models like YOLOv7, which uses 1280 pixels. YOLOv9 achieves a slightly higher mAP (55.6%) than YOLOv10 and YOLOv11 while using similar image sizes and fewer parameters (58.1M). Overall, the table highlights the trade-offs between accuracy (mAP), speed, and computational requirements across different YOLO versions. This balance is one of the reasons all these algorithms were chosen for the experiment, to evaluate which one performs better with the processed dataset while balancing these factors effectively.

Table 6.8: Comparison of different YOLO algorithms using the COCO dataset.

Model (YOLO)	Size (Pixels)	mAP^{val} 50-95	Speed CPU ONNX (ms)	Speed T4 TensorRT10 (ms)	params (M)	FLOPs (B)
v5	1280	56.8	2436.5	8.98	155.4	250.7
v6	640	57.2	600.5	10.32	140.4	200.3
v7	1280	56.8	-	11.3	151.7	843.2
v8	640	55.3	142.4	1.21	68.7	260.6
v9	640	55.6	-	-	58.1	192.5
v10	640	54.4	-	-	29.5	160.4
v11	640	54.7	462.8	11.3	56.9	194.9

To achieve meaningful results from these models, the next section outlines the data preparation and preprocessing steps undertaken to ensure that the dataset was suitable for training and testing these YOLO algorithms.

6.3 Model Training Procedure

6.3.1 Data Preparation and Preprocessing

This section describes the steps taken to prepare the dataset for training, validation, and testing purposes. Initially, a small subset of the original dataset was selected to build a trial model. The purpose of this step was to test the dataset's compatibility with the selected algorithms and evaluate the end-to-end flow of the model development pipeline. By using this smaller sample, potential issues in data formatting, labelling, or pipeline integration were identified and resolved early. Once this preliminary phase was completed, the focus shifted toward preparing the full dataset for the actual model development.

Next, the dataset was split into 70:30 proportions for training and testing (20% for validation and 10% for testing), following best practices to maintain a fair and unbiased evaluation of the model's

performance. While there is no single data splitting ratio that works for all experiment (Xu & Goodacre, 2018), 70:20:10 splitting ratio was adopted as proven to work well for many problems (Lyu et al., 2021), which also follows AuDD studies such as (Dang et al., 2022; Shen et al., 2023; Tan et al., 2021b; Xie et al., 2019d).

Subsequently, a copy of the training dataset underwent all preprocessing techniques previously detailed in Chapter 5. These techniques included variations of contrast enhancement, morphological operations, and segmentation approaches, with the aim of each technique contributing to improved feature clarity and defect detectability. A hierarchical summary of these pre-processing steps is depicted in Figure 6.2.

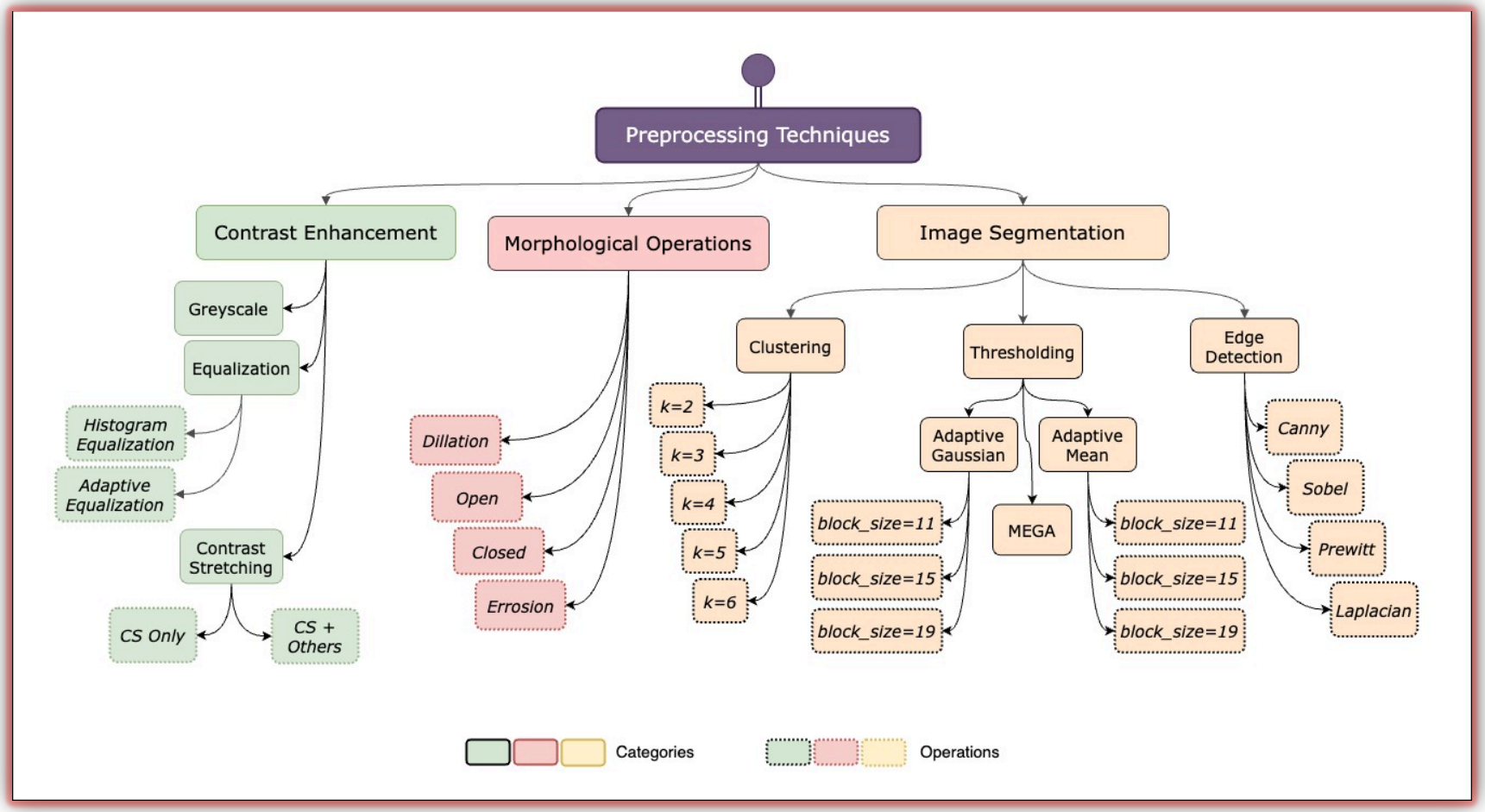


Figure 6.1: Variations of pre-processing techniques applied (source: Author)

However, there was class imbalance (Dang et al., 2021c; Ghosh et al., 2024) in the training dataset. To address this, the Synthetic Minority Oversampling Technique (SMOTE) (Elreedy & Atiya, 2019) was applied to oversample underrepresented classes, such as root intrusions and attached deposits. In addition, data augmentation techniques were applied to enhance the robustness and diversity of the dataset, as detailed in chapter 5. Techniques such as flipping, rotation, mosaic augmentation, and cut-out were employed to simulate a wide range of real-world variations in sewer defects. The goal here was to ensure that the model could generalize effectively across unseen data.

6.3.2 Model Training and Optimization.

The training environment was set up using the hardware specification and software configuration described in 6.1. A total of 34 different models were developed using different preprocessing operations highlighted in 6.3.1. The approach used here was incremental and comparative. Out of the 34 models, 9 models were developed using variants of algorithms discussed in 6.2 along with different sizes of the dataset. The aim was to compare the algorithms on different performance and operational metrics such as accuracy ($mAP^{VAL(0-50)}$) on small (513 and 1299) and bigger (3071) number of images, time taken to complete 100 epochs, and number of layers used among others. This stage of the model training is referred to as "YOLO Compare".

Since it is highly difficult to use all the algorithms to train models for all the preprocessing operations because it will mean each of the 7 algorithms will be used in training 25 models, consequently leading to more than 170 models. The goal here is to understand which of the algorithms best suit the research, using the performance and operational metric mentioned earlier. Hence, such algorithm will be adapted in training all the 25 models targeting the preprocessing operations.

Initially, a subset of the data (513 samples) was used for model training over 25 epochs. This was followed by training on a dataset more than twice the size of the initial subset, using 50 epochs. Finally, the full dataset was utilized for training over 100 epochs. Throughout this incremental training process, relevant metadata was recorded in real-time as the training progressed using *wandb*. Figure 6.3 shows an overall design flow of how the model training was conducted.

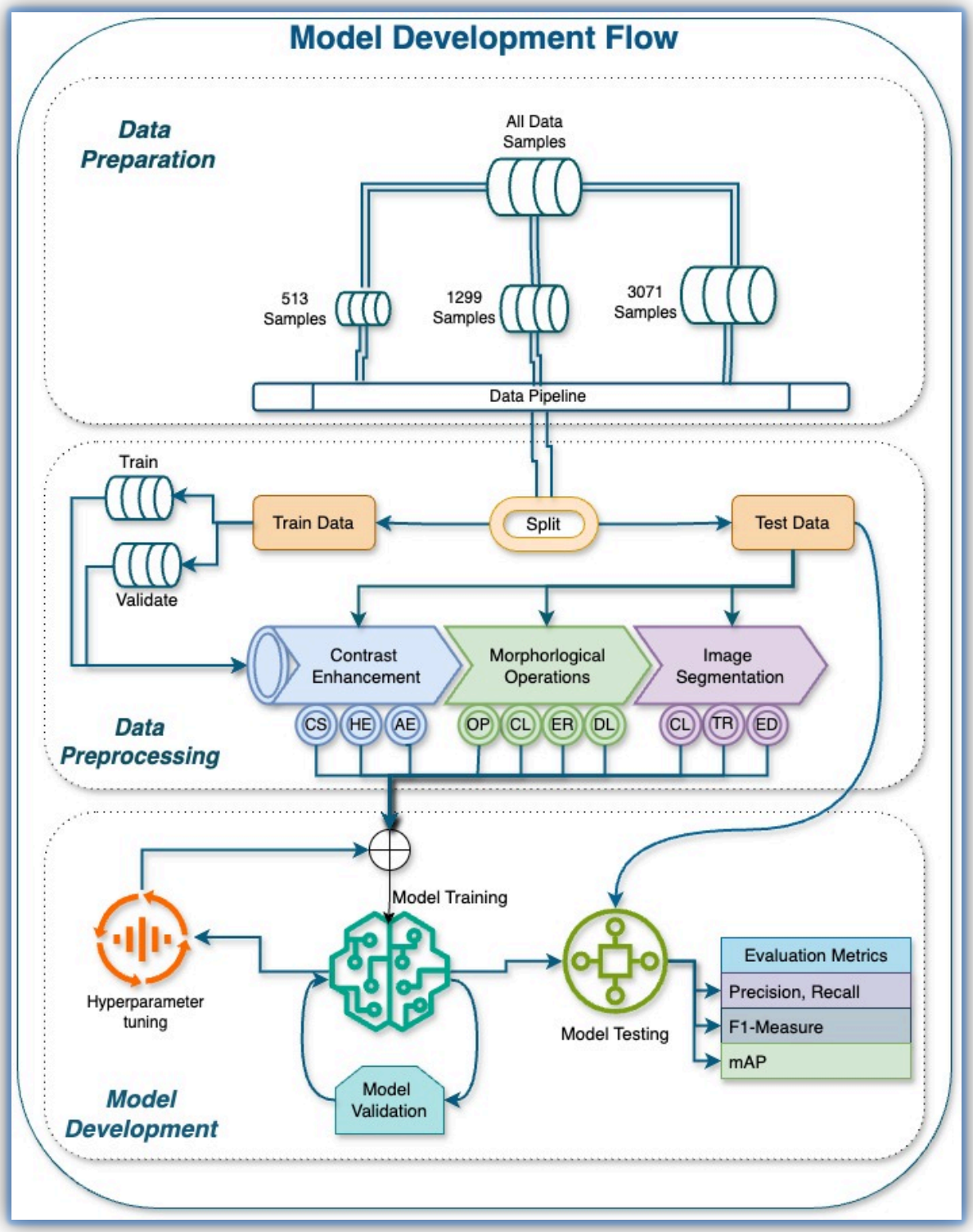


Figure 6.3: Overall design for model development (source: Author)

To control how the model learns during training, optimization algorithms was used within network configuration to ensure that the model converges towards a minimal loss function by effectively adjusting the weights of the network to improve its predictions. Here, two optimization algorithms were used to fine-tune the YOLO algorithms: Stochastic Gradient Descent (SGD) (Fjellström & Nyström, 2022) and Adam (Kingma & Ba, 2014).

6.3.3 Hyperparameter Tuning

Hyperparameters are adjustable settings that control how an algorithm learns and performs during training. During the model training, several hyperparameters were carefully tuned to optimize the model's accuracy and efficiency. Specific hyperparameters associated with the selected optimizers (SGD and Adam) were also fine-tuned, including learning rate, momentum, and weight decay. These parameters were adjusted iteratively to strike a balance between convergence speed and model generalization. Additionally, other key hyperparameters, such as batch size, number of epochs, and dropout rates, were also optimized to enhance model performance and prevent overfitting. Table 6.4 provides a comprehensive list of all the hyperparameters along with the range of values experimented with during the training process.

Table 6.4: Model hyperparameters used for fine-tuning

Hyperparameter	Description	Encoding	Values
Learning Rate	Step size for updating weights during optimization.	lr	0.001 - 0.01
Momentum	Fraction of the previous update added to the current one.	momentum	0.85 - 0.95
Weight Decay	Regularization parameter to prevent overfitting by penalizing large weights.	weight_decay	0.0001-0.001
Batch Size	Number of training examples processed simultaneously in one iteration.	batch_size	4, 8, 16, 32, 64
Epochs	Total number of passes through the entire dataset.	epochs	25 - 100
IOU Threshold	Threshold to determine whether a prediction overlaps significantly with ground truth.	iou_thresh	0.5 - 0.7
Confidence Threshold	Minimum confidence score for a detection to be considered valid.	conf_thresh	0.25 - 0.5
Gradient Clipping	Maximum allowed value for gradients to prevent exploding gradients.	clip_grad	0.1 - 1.0
Loss Function	Objective function used to train the network.	loss_func	BCE, Focal, Smooth L1
Activation Function	Non-linear function applied after each convolution layer.	activation	ReLU, Leaky ReLU.
Image Size	Resolution of input images fed into the network.	img_size	640
Augmentation Methods	Data augmentation techniques applied during training.	augmentations	Flip, CutOut, Mosaic
Freeze Layers	Number of layers frozen during training to retain pre-trained weights.	freeze_layers	0 - N (number of layers)
Warm-up Epochs	Number of initial epochs to gradually increase the learning rate.	warmup_epochs	2 - 10

Hyperparameter	Description	Encoding	Values
Decay Rate	Factor by which the learning rate is decayed at specific intervals.	decay_rate	0.1 - 0.5
Dropout Rate	Fraction of neurons randomly deactivated during training for regularization.	dropout_rate	0.2 - 0.5
Overlap Threshold (NMS)	Intersection over Union (IoU) threshold for filtering overlapping detections during Non-Max Suppression.	nms_thresh	0.4 - 0.6
Optimizer	Algorithm used to optimize weights and biases.	optimizer	SGD, Adam.

6.4 Model Evaluation Metrics

To evaluate the developed model's performance, metrics such as precision, recall, F1-score, and mAP were carefully chosen to provide a detailed analysis of the model's classification and localization capabilities. These metrics will offer insights into how well the model performs in identifying and localizing defects.

6.4.1 Selected Evaluation Metrics

6.4.1.1 Precision

Precision measures the proportion of correctly identified defect instances (true positives) out of all instances predicted as defects. For example, if the models predict 50 instances of cracks but only 40 are actual cracks, the precision score highlights how well the model avoids false alarms, such as incorrectly labelling noise or shadows as defects. To calculate precision, the formular is given as provided in EQ 6.1.

$$Precision = \frac{TP}{(TP + FP)} \quad (6.1)$$

Where:

- TP (True Positive): Correctly detected defects (e.g., cracks labelled correctly).

- *FN* (False Negative): Incorrectly detected defects (e.g., shadows misclassified as cracks).

6.4.1.2 Recall

Recall, also known as Sensitivity or True Positive Rate (TPR), quantifies the model's ability to correctly identify all actual defect instances. For instance, if there are 100 instances of root intrusions in the dataset and models correctly detects 85 of them, recall reflects the model's effectiveness in capturing as many real defects as possible, including those that might be subtle or less prominent. To calculate recall, the formular is given as provided in EQ 6.2.

$$Recall = \frac{TP}{(TP + FN)} \quad (6.2)$$

Where:

- *TP* (True Positive): Correctly detected defects (e.g., root intrusions identified accurately).
- *FN* (False Negative): Missed defects (e.g., root intrusions not detected by the model).

6.4.1.3 F1-Score

The F1-score is the harmonic mean of precision and recall which provides a single metric that balances the trade-off between the two metrics. For example, if the model is highly precise in identifying cracks but misses several others due to low recall, then the F1-score provides a consolidated view of its performance by balancing false positives and false negatives. This is particularly important in AuDD, where both over-detection and under-detection of defects can significantly impact maintenance decisions. To calculate F1, the formular is given as provided in EQ 6.3.

$$F1 = 2 * \frac{(Precision * Recall)}{(Pecision + Recall)} \quad (6.3)$$

6.4.1.4 mAP (Mean Average Precision)

mAP evaluates the model's detection capabilities by considering both the accuracy of bounding box predictions and class probabilities. An example here could be when the models detect cracks and assigns a bounding box, mAP assesses whether the predicted boxes tightly overlap with the actual defect locations and whether the assigned probabilities align with the ground truth. An example is given in Figure 6.3 where the actual bounding box is compared with the predicted one using a crack. Here, the more the predicted bounding box overlays the actual one, the higher the mAP.

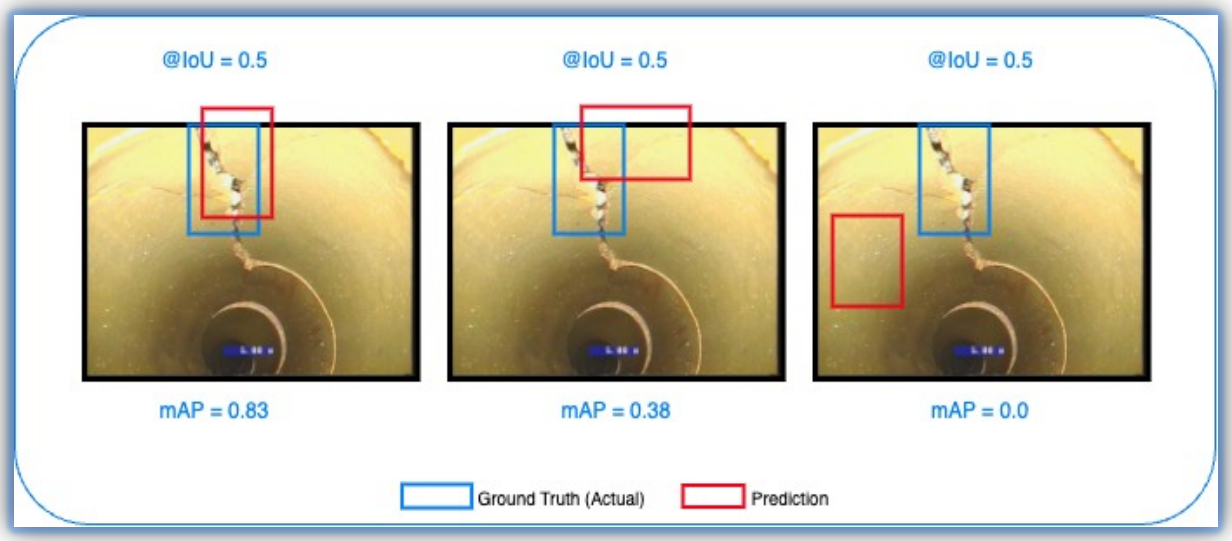


Figure 6.3: Mean Average Precision (mAP) using IoU (source: Author)

mAP is measured across various Intersection over Union (IoU) thresholds. For example, while **$mAP^{(0\text{-}50)}$ e**valuates performance under lenient conditions, such as detecting large defects where slight bounding box inaccuracies are tolerable, **$mAP^{(50\text{-}95)}$ checks** performance using more stricter conditions, such as detecting small defects where precise bounding boxes are critical. To calculate mAP, the formular used is given as provided in EQ 6.4.

$$mAP = \frac{1}{N}\sum_{i=1}^{N} * AP_i \tag{6.4}$$

Where:

- N: Number of IoU thresholds (e.g., 0.5 to 0.95 with increments of 0.05).
- AP_i: Average Precision at each IoU threshold by considering both precision and recall.

To summarize, Table 6.5 provides a summary of the evaluation metrics and specific use case examples.

Table 6.5: Summary of evaluation metrics

Metrics	Description	Use-case Example
Precision	The ratio of true positive predictions to the total positive predictions.	If the model flags 100 defects and 80 of them are correct, the precision is 80%.
Recall	The ratio of true positive predictions to the actual positives.	If there are 100 defects and the model finds 75, the recall is 75%.
F1-Score	The harmonic mean of precision and recall.	This will be useful when we need a single measure to see how well the model performs.
mAP	Mean Average Precision across all classes.	This will help understand the model's performance in detecting different kinds of defects.
IoU	Intersection over Union between predicted and actual bounding boxes.	If the predicted defect area overlaps well with the real defect area, the IoU score will be high.

6.5 Chapter Summary

This chapter detailed the rigorous experimental setup, data preparation, and systematic development of deep learning models for Automated Defect Detection (AuDD) in sewer systems. A high-performance multi-GPU environment (utilizing NVIDIA Tesla V100 and T4 GPUs) was established to handle the computational demands of training 34 distinct models. The dataset was systematically partitioned into a 70:20:10 split for training, validation, and testing. Furthermore, severe class imbalances, particularly in underrepresented categories like root intrusions and attached deposits, were successfully mitigated using the Synthetic Minority Oversampling Technique (SMOTE) alongside targeted data augmentations.

A key finding of this chapter was derived from the architectural benchmarking of state-of-the-art YOLO algorithms (v5 through v11). The comparative analysis revealed critical trade-offs between computational efficiency, architectural complexity, and detection accuracy. For instance, while older architectures like YOLOv7 required a massive 843.2 billion FLOPs to achieve a mAP of

56.8% on standard benchmarks like COCO, newer iterations like YOLOv9 achieved a highly competitive mAP of 55.6% while drastically reducing the computational burden to just 192.5 billion FLOPs and 58.1 million parameters.

Guided by these architectural insights, nine baseline models were initially developed across incrementally scaled datasets (513, 1,299, and 3,071 images) to evaluate real-world performance before extending the training to all pre-processing variations. The training process was strictly regulated using SGD and Adam optimizers, with hyperparameters extensively fine-tuned (including learning rates of 0.001–0.01, epochs ranging from 25 to 100, and batch sizes up to 64). Ultimately, the robust data preparation, architectural selections, and fine-tuned training configurations established in this chapter provide a highly optimized framework, setting the perfect stage for the comprehensive performance evaluation and model benchmarking presented in Chapter 7.

CHAPTER SEVEN

7.0 Result and Discussion

7.1 Chapter Overview

This chapter delves into the evaluation and analysis of the models developed in chapter 6, focusing on their performance in detecting sewer defects using various metrics. The evaluation process considers the effectiveness of the selected YOLO variants (v8, v9, v10, and v11) in handling the complexities of AuDD in sewer systems. Through benchmarking, the chapter examines precision, recall, F1-score, mAP metrics to assess the models' ability to accurately classify and localize defects of varying sizes and visibility. These metrics are pivotal in determining the models' practicality for real-world applications by emphasizing their ability to identify both obvious and rare defects reliably.

In addition to presenting the evaluation results, this chapter provides insights into the impact of pre-processing techniques model performance. Comparative analysis across YOLO variants is conducted to highlight trade-offs between speed and accuracy, addressing the practical constraints of using AuDD systems.

7.2 Performance of Different YOLO Models.

This research undertook a systematic comparison of four YOLO-based object detection algorithms, evaluating them against a range of performance and computational metrics. The primary aim was to determine which architecture offered the most effective balance between detection accuracy, processing speed, and architectural complexity, a critical factor for real-world sewer defect detection (He et al., 2024). Evaluation metrics included standard performance indicators such as the mean Average Precision (mAP), as well as computational measures like the time required to complete 100 training epochs, the number of model layers, and the estimated floating-point operations (FLOPs). This comprehensive evaluation ensured that the final selection would be guided not only by accuracy but also by efficiency and deploy ability within operational constraints.

The evaluation process was conducted on two distinct datasets of varying sizes, 1299 and 3071 unique image samples curated from annotated and processed frames. For clarity and traceability during experimentation, each algorithm–dataset combination was assigned a unique code name in

the format ALGORITHM_DATASIZE (e.g., YOLOv8_1299 or YOLOv9_3071). This naming convention, illustrated in Figure 7.1, facilitated consistent result tracking, simplified comparative analysis, and ensured transparent reference to specific trials during discussion and reporting.

Across both datasets, the four YOLO variants were trained and evaluated under consistent experimental conditions, including identical hyperparameters, hardware settings, and pre-processing pipelines. This controlled approach eliminated confounding variables and allowed direct performance comparisons between architectures. Using the ALGORITHM_DATASIZE naming system, this research maintained a clear link between results and their experimental configurations, making it straightforward to interpret accuracy trends, assess training efficiency, and relate computational demands to model complexity. The insights gained from this structured evaluation formed the basis for selecting the optimal algorithm, which was then fine-tuned for subsequent model development and deployment phases.

Using the metrics earlier explained in section 6.4, YOLOv9_3071 emerged as the top-performing model, achieving the highest scores in precision (0.86), mAP^{50} (0.86), mAP^{50-95} (0.68), and F1 (0.82). YOLOv11_3071 also demonstrated strong performance, with the highest recall (0.80) and a competitive F1 (0.81), making it a viable choice if the goal was to focus solely on sensitivity, but recall is also very important. For example, in sewer systems serving densely populated urban areas, a model with high recall can ensure even minor defects are flagged, thereby reducing the likelihood of undetected issues escalating into costly emergencies. In contrast, YOLOv10_1299 consistently underperformed, with low precision (0.75) and recall (0.67), making it unsuitable for this research because missed defects or false alarms could lead to significant operational inefficiencies

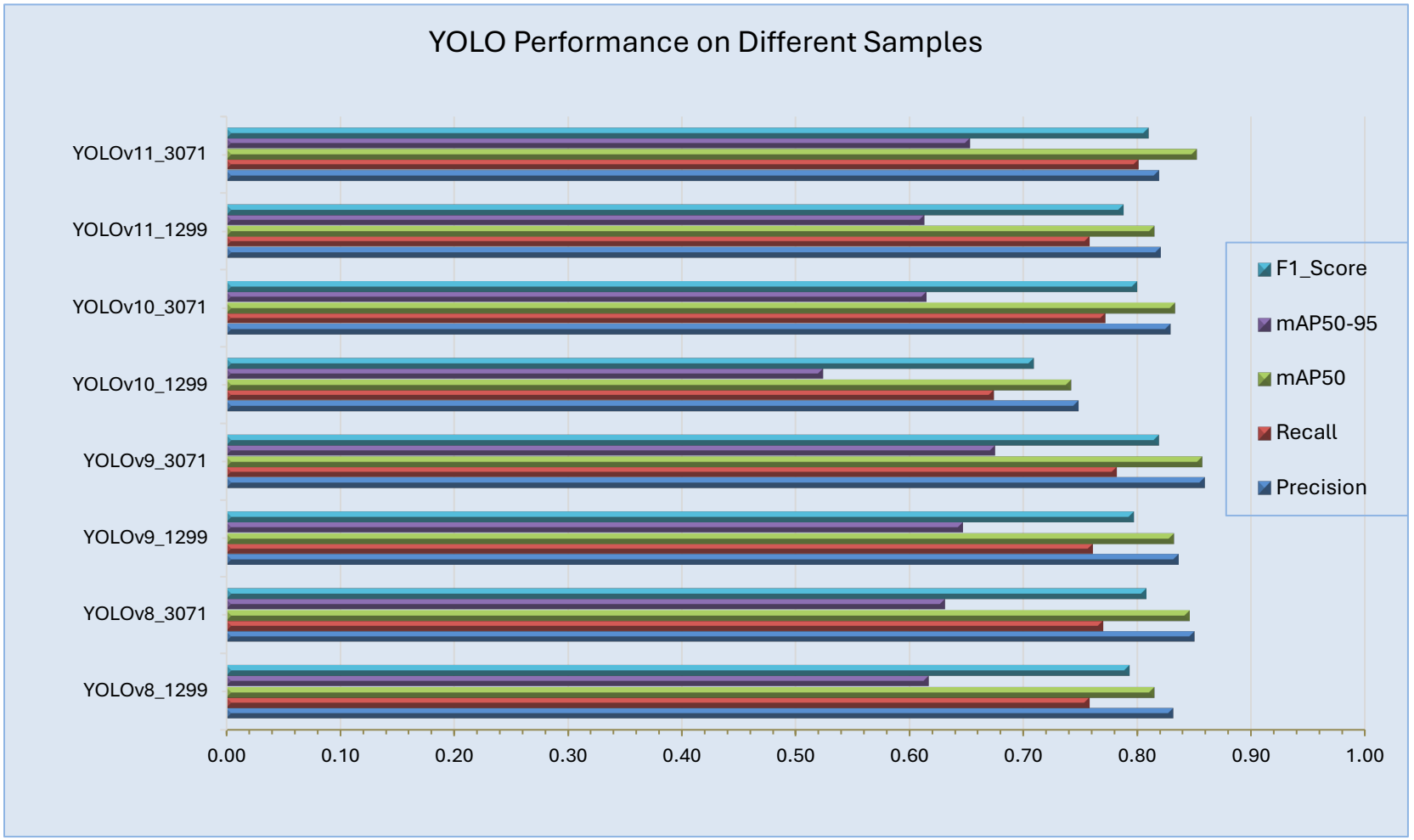


Figure 7.1: Variations of YOLO results on two groups of dataset.

Computationally, the algorithms were also compared based on FLOPs, training time, and number of layers, as shown in Figure 7.2. YOLOv9 again strikes the best balance between computational efficiency and architectural complexity with 9.2 billion FLOPs and the shortest training time (103.8 minutes) despite having 250 layers. In contrast, YOLOv10, while having the lowest FLOPs (8.4 billion), suffers from more training time (115.2 minutes) and a more complex architecture with 385 layers. For YOLOv11, it has the highest FLOPs (21.6 billion) and moderate training time (107.4 minutes). This might be better suited for scenarios where computational resources are abundant, but not in this research. For YOLOv8, it offers a simpler alternative with 10 billion FLOPs and 225 layers. Owing to the evaluation of performance and computational metrics, YOLOv9 was adopted as the optimal model among the selected ones.

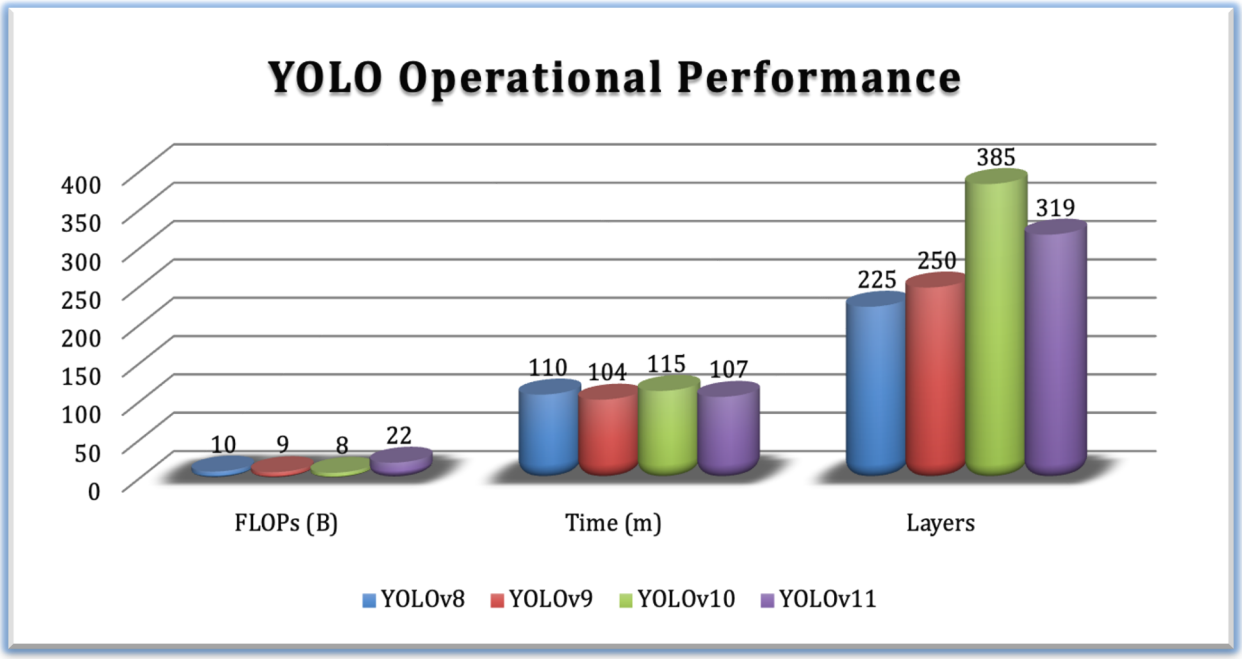


Figure 7.2: Operational performance of YOLO algorithms.

A correlation analysis was conducted to examine the relationships between the four principal evaluation metrics used in this study (i.e., Precision, Recall, mAP50, and mAP50–95) across all YOLO model–dataset combinations. Correlation analysis is a statistical technique used to quantify the strength and direction of association between two continuous or ordinal variables. Both Pearson's correlation coefficient (r), which assumes linearity, and Spearman's rank correlation coefficient (ρ), which assesses monotonic relationships without assuming linearity, were computed to provide a comprehensive view of metric interdependencies. This dual approach ensures robustness of interpretation by mitigating the influence of potential non-linearities or outliers,

which are not uncommon in performance data for object detection models (Rungtalay & Kaitwanidvilai, 2024).

Pearson's correlation coefficient (r) measures the degree of linear relationship between variables, with values ranging from –1 (perfect negative correlation) to +1 (perfect positive correlation), and 0 indicating no linear relationship. A high positive r value suggests that increases in one metric are consistently accompanied by increases in another. Pearson's method assumes that the variables are continuous, approximately normally distributed, and linearly related (Rodgers & Nicewander, 1988). However, performance metrics in object detection tasks, especially when derived from varying architectures and datasets, may not always satisfy linearity or normality assumptions.

To address this, the analysis also incorporated Spearman's rank correlation coefficient (ρ), a non-parametric measure that assesses monotonic relationships between variables. Unlike Pearson's r, Spearman's ρ is calculated based on the ranked values of the data rather than their raw scores, making it more robust to outliers and capable of capturing non-linear but monotonic associations (Hauke & Kossowski, 2011). By employing both Pearson and Spearman correlations, this study ensured a more rigorous and comprehensive examination of metric interdependencies, allowing for robust conclusions regardless of the underlying distributional properties of the evaluation scores.

The Pearson analysis revealed a very strong positive correlation between mAP50 and mAP50–95 ($r = 0.949$), as well as between mAP50 and Recall ($r = 0.926$), indicating that improvements in these metrics tend to occur in tandem. mAP50–95 was also strongly correlated with Recall ($r = 0.881$). These findings suggest that architectural or preprocessing enhancements that improve a model's detection capability at IoU 0.50 are likely to yield parallel gains at higher IoU thresholds and enhance overall recall performance. Figure 7.3 shows the Pearson correlation heatmap for each evaluation metrics.

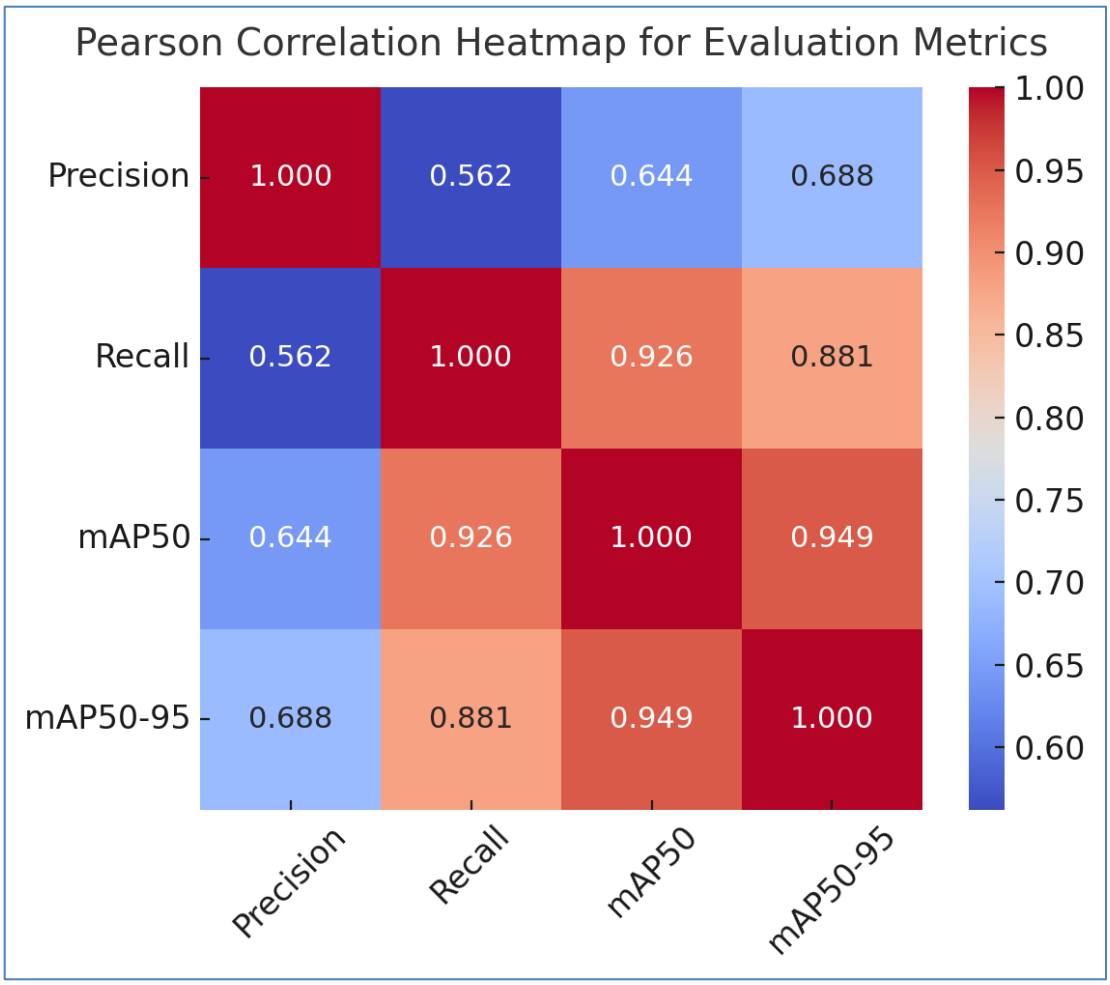


Figure 7.3: Pearson correlation heatmap for each evaluation metrics

Conversely, the correlation between Precision and the other metrics was moderate, with the strongest association observed with $mAP^{50–95}$ ($r = 0.688$) and the weakest with Recall ($r = 0.562$). This aligns with the known precision–recall trade-off in object detection, where adjusting decision thresholds to minimise false positives often leads to a reduction in true positive rates (Saito & Rehmsmeier, 2015).

The Spearman rank correlation produced consistent results as shown in Figure 7.4. This reaffirms the robustness of the observed patterns. Strong monotonic relationships were again observed between mAP^{50} and $mAP^{50–95}$ ($\rho = 0.938$) and between mAP^{50} and Recall ($\rho = 0.910$), suggesting that the co-movement of these metrics is not dependent on strictly linear behaviour. The moderate correlation between Precision and Recall ($\rho = 0.536$) further confirms the threshold-sensitive balance between over-detection and under-detection, which must be carefully managed in operational contexts such as sewer defect detection.

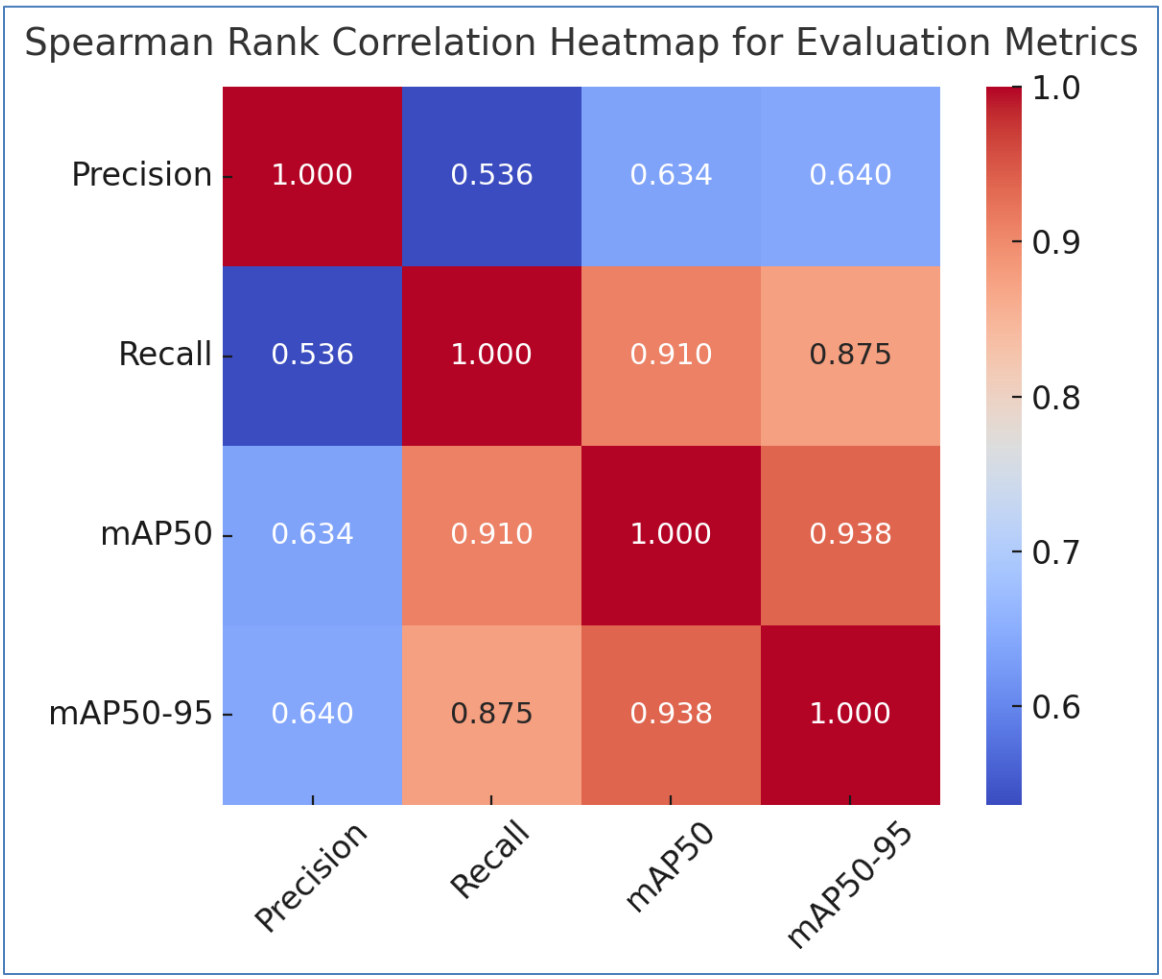


Figure 7.4: Spearman Rank correlation heatmap for all metrics

Altogether, these results indicate that while mAP^{50}, $mAP^{50–95}$, and Recall are largely synergistic; improvements in one are highly likely to boost the others. Precision behaves somewhat independently, reflecting its sensitivity to classification threshold settings and false positive control. This underscores the need for a balanced optimisation strategy during model development, where enhancements are pursued in a way that does not disproportionately favour one metric at the expense of others. In practice, as echoed by Soviany & Ionescu (2018) and Subramanian et al. (2025), this means that achieving high recall and mAP values without compromising precision remains a critical challenge for deploying AI-based defect detection in sewer inspection workflows, where both false alarms and missed detections carry significant operational costs.

While the strong associations among metrics have been consistently established through both Pearson and Spearman analyses across all YOLO models and dataset sizes, this study will not extend the correlation analysis to subsequent models developed with different pre-processing

strategies. Repeating the procedure for each variant would provide little to no additional insight and risk analytical redundancy, as the fundamental inter-metric relationships are unlikely to shift substantially once confirmed across representative datasets and architectures. Consequently, subsequent evaluations focus on quantifying how pre-processing methods and dataset changes influence absolute performance values, rather than re-establishing already validated metric dependencies.

7.3 Improving AuDD through Contrast Enhancement

Contrast enhancement, as explained in Chapter 3 and 5, aims to add more clarity to the defect region and boundaries to aid better defect detection. Using HE, CS, and AE, these methods were evaluated across all metrics and demonstrates noticeable variation in the detection of different defect classes. The results are summarized in Table 7.1. The model performance using adaptive equalization achieved an average precision of 0.872, recall of 0.783, mAP^{50} of 0.863, and mAP50-95 of 0.655. Among the defect classes, settled deposit recorded the highest mAP50 score of 0.913, while Root had the highest precision (0.926), recall (0.823), and mAP^{50-95} (0.721).

Conversely, Encrustation exhibited the lowest mAP^{50} score (0.775). The results suggest that adaptive equalization is particularly effective for classes such as Root and Settled deposit, but its performance is comparatively lower for Encrustation, which may indicate that further preprocessing or complementary augmentation techniques are required for this class.

Table 7.1: Performance comparison across contrast enhancement techniques

Defect Class	Technique	Precision	Recall	mAP^{0-50}	mAP^{50-95}
Attached Deposit	AE	89.7	79.4	87.5	64.7
	CS	**90.9**	**87.1**	**91.2**	**77.2**
	HE	88.4	81.0	90.2	67.9
Crack	**AE**	**85.7**	77.1	**84.3**	**61.8**
	CS	81.7	**80.9**	83.1	60.5
	HE	81.2	75.4	82.5	60.7
Encrustation	AE	77.7	70.5	77.5	**58.8**
	CS	**83.3**	**71.1**	**78.4**	**58.8**
	HE	77.6	68.9	75.9	58.1
Root	**AE**	**90.6**	82.3	89.9	**72.1**
	CS	87.1	83.3	90.8	68.4
	HE	89.4	**83.6**	**91.5**	70.6
Settled Deposit	**AE**	**90.2**	82.1	**91.3**	70.2
	CS	88.1	**89.6**	91.1	**72.1**
	HE	89.0	86.5	87.4	66.8

Contrast stretching yielded an overall precision of 0.864, recall of 0.824, mAP50 of 0.875, and mAP50-95 of 0.668. It exhibited strong performance for Attached deposit and Settled deposit, with mAP50 scores of 0.929 and 0.921, respectively. Attached deposit further recorded the highest precision (0.919) and mAP50-95 (0.772), while Settled deposit had the highest recall (0.896). This demonstrates that contrast stretching is particularly effective for these two classes, potentially enhancing their distinct visual features. However, like the results with adaptive equalization, encrustation displayed lower detection performance, with an mAP^{50} of only 0.784.

With histogram equalization, the model achieved an average precision of 0.851, recall of 0.791, mAP50 of 0.855, and mAP50-95 of 0.648. The highest mAP50 score of 0.915 was observed for Root, which also achieved the highest precision (0.894) and mAP50-95 (0.706). However, settled deposit demonstrated the highest recall (0.865), suggesting that histogram equalization enhances temporal features of this class more effectively. Once again, encrustation had the weakest performance across metrics, with an mAP50 of only 0.759. The consistent underperformance of Encrustation across all techniques highlights its inherent complexity and suggests the need for tailored solutions beyond standard contrast enhancement methods.

As depicted in Figure 7.5 using a heatmap, the correlation analysis among the performance metrics revealed positive relationships, with mAP^{50} showing a particularly strong correlation (above 90%) with other metrics such as recall and precision. This suggests that improvements in recall can significantly enhance mAP50 values, highlighting its role as a good indicator of model performance (B. Kumar et al., 2022). Enhancing recall through targeted techniques can therefore lead to broader improvements across other metrics.

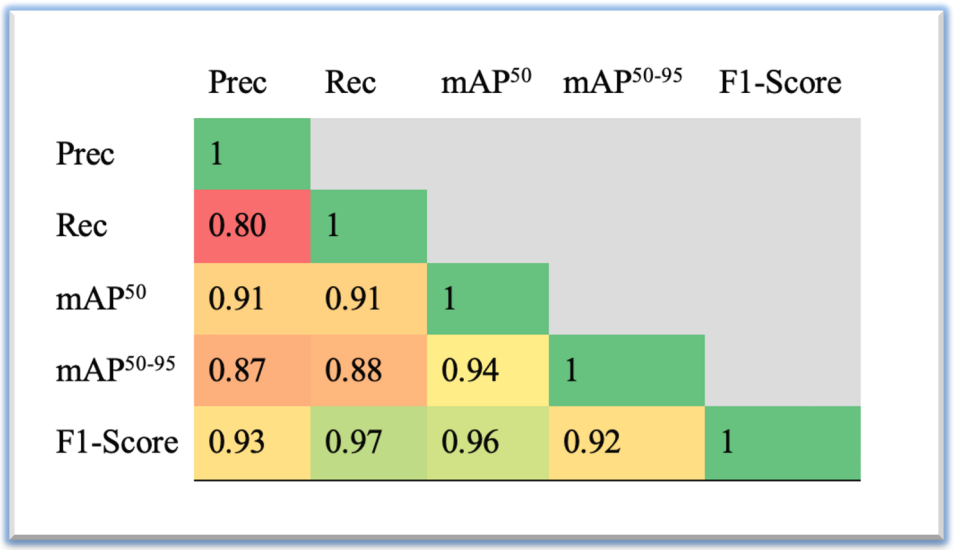


Figure 7.3: Correlation analysis using heat map

In conclusion, these results demonstrate the critical importance of selecting the appropriate contrast enhancement technique based on the defect classes. Contrast stretching emerged as the most effective technique for attached and settled deposit, while histogram equalization showed advantages for Root. The statistically significant impact of these techniques, as confirmed by

Spearman analysis, underscores the necessity of tailoring preprocessing strategies to achieve optimal detection across diverse defect classes.

7.4 Effectiveness of Segmentation Techniques

The evaluation of various image segmentation techniques, including clustering (K2, K3, K4, K5, K6), edge detection (Canny, Laplacian, Prewitt, Sobel), and thresholding (adaptive Gaussian and adaptive mean), reveals significant differences in model performance across defect classes. The results, summarized in Table 7.2, provide insights into the effectiveness of these methods when applied to enhance YOLOv9's detection capabilities.

The clustering methods exhibited varied outcomes, with K5 and K6 outperforming other configurations. For K5, the model achieved an average precision of 0.833, recall of 0.731, mAP50 of 0.81, and mAP50-95 of 0.57. Settled deposit consistently recorded the highest scores across recall 0.89, and mAP50-95 0.69. In comparison, clustering K6 decreased the mAP50 to 0.826 and mAP50-95 to 0.59, with root class achieving the highest scores for mAP50 0.92 and recall of 0.86. This suggests that higher cluster configurations can better capture class-specific features, particularly for settled deposits and roots.

Table 7.2: Summary of image segmentation results by defect class and techniques

Defect	Technique	Operations	Precision	Recall	mAP^{50}	mAP^{50-95}
Attached Deposit	Clustering	K=2	0.778	0.387	0.462	0.271
		K=3	0.778	0.387	0.462	0.271
		K=4	0.778	0.387	0.462	0.271
		K=5	0.902	0.593	0.786	0.519
		K=6	0.844	0.699	0.789	0.579
	Edge Detection	Canny	0.651	0.532	0.539	0.34
		Laplacian	0.781	0.387	0.515	0.337
		Prewitt	0.702	0.531	0.566	0.363
		Sobel	0.498	0.387	0.349	0.185
		AG (Block 11)	0.764	0.681	0.744	0.528

Defect	Technique	Operations	Precision	Recall	mAP^{50}	mAP$^{50-95}$
	Thresholding	AG (Block 15)	0.818	0.578	0.689	0.491
		AG (Block 19)	0.763	0.674	0.734	0.522
		AM (Block 11)	0.818	0.578	0.689	0.491
		AM (Block 15)	0.923	0.677	0.792	0.569
		AM (Block 19)	0.866	0.623	0.759	0.541
Crack	Clustering	K=2	0.72	0.411	0.453	0.192
		K=3	0.72	0.411	0.453	0.192
		K=4	0.72	0.411	0.453	0.192
		K=5	0.769	0.723	0.767	0.535
		K=6	0.793	0.723	0.793	0.562
	Edge Detection	Canny	0.74	0.483	0.585	0.338
		Laplacian	0.716	0.549	0.666	0.433
		Prewitt	0.755	0.475	0.595	0.349
		Sobel	0.386	0.149	0.15	0.057
	Thresholding	AG (Block 11)	0.774	0.622	0.722	0.486
		AG (Block 15)	0.809	0.607	0.707	0.502
		AG (Block 19)	0.772	0.596	0.714	0.5
		AM (Block 11)	0.809	0.607	0.707	0.502
		AM (Block 15)	0.83	0.617	0.747	0.522

Defect	Technique	Operations	Precision	Recall	mAP^{50}	mAP$^{50-95}$
		AM (Block 19)	0.792	0.66	0.739	0.51
Encrustation	Clustering	K=2	0.72	0.43	0.48	0.251
		K=3	0.72	0.43	0.48	0.251
		K=4	0.72	0.43	0.48	0.251
		K=5	0.787	0.623	0.721	0.453
		K=6	0.78	0.64	0.715	0.483
	Edge Detection	Canny	0.823	0.529	0.607	0.344
		Laplacian	0.777	0.5	0.607	0.388
		Prewitt	0.722	0.455	0.568	0.347
		Sobel	0.545	0.386	0.413	0.183
	Thresholding	AG (Block 11)	0.782	0.605	0.685	0.457
		AG (Block 15)	0.797	0.585	0.651	0.459
		AG (Block 19)	0.732	0.64	0.698	0.463
		AM (Block 11)	0.797	0.585	0.651	0.459
		AM (Block 15)	0.724	0.632	0.674	0.47
		AM (Block 19)	0.732	0.648	0.733	0.51
	Clustering	K=2	0.796	0.667	0.731	0.399
		K=3	0.796	0.667	0.731	0.399
		K=4	0.796	0.667	0.731	0.399
		K=5	0.81	0.822	0.872	0.637

Defect	Technique	Operations	Precision	Recall	mAP^{50}	mAP$^{50-95}$
Root		K=6	0.856	0.861	0.916	0.661
	Edge Detection	Canny	0.811	0.767	0.845	0.475
		Laplacian	0.931	0.411	0.627	0.404
		Prewitt	0.709	0.648	0.721	0.478
		Sobel	0.666	0.465	0.526	0.203
	Thresholding	AG (Block 11)	0.803	0.678	0.844	0.576
		AG (Block 15)	0.847	0.675	0.839	0.6
		AG (Block 19)	0.781	0.793	0.841	0.604
		AM (Block 11)	0.847	0.675	0.839	0.6
		AM (Block 15)	0.813	0.8	0.845	0.575
		AM (Block 19)	0.79	0.711	0.805	0.547
Settled Deposit	Clustering	K=2	0.868	0.759	0.838	0.505
		K=3	0.868	0.759	0.838	0.505
		K=4	0.868	0.759	0.838	0.505
		K=5	0.896	0.891	0.902	0.686
		K=6	0.893	0.859	0.915	0.648
	Edge Detection	Canny	0.775	0.714	0.785	0.534
		Laplacian	0.885	0.794	0.851	0.528
		Prewitt	0.715	0.69	0.685	0.469
		Sobel	0.504	0.448	0.426	0.199

Defect	Technique	Operations	Precision	Recall	mAP^{50}	mAP^{50-95}
	Thresholding	AG (Block 11)	0.866	0.759	0.825	0.596
		AG (Block 15)	0.87	0.81	0.878	0.619
		AG (Block 19)	0.742	0.759	0.818	0.608
		AM (Block 11)	0.87	0.81	0.878	0.619
		AM (Block 15)	0.855	0.81	0.846	0.626
		AM (Block 19)	0.788	0.759	0.821	0.605

Among edge detection methods, Laplacian and Canny produced the most promising results, with average mAP50 scores of 0.653 and 0.672, respectively. Settled deposit demonstrated the best performance in both methods, achieving mAP^{50} scores of 0.85 (Laplacian) and 78.5% (Canny). However, Sobel performed poorly, with the lowest overall metrics (mAP^{50} of 0.373 and mAP^{50-95} of 0.165). These results highlight the limitations of Sobel in detecting fine details, which are crucial for accurate segmentation.

In relation to other segmentation strategies, block size 11 and 15 of the adaptive Gaussian method were the most proficient in the segmentation task. With the application of the block size 15, the model managed to attain precision at an average of 0.828, while the recall stood at 0.651. Graphing the mAP variables, we find mAP^{50} to be at 0.753 and mAP^{50-95} at 0.534. There is once more the situation whereby the deposited mAP50 managed to attain higher levels at 87.8 than root with 83.9. The degree of uniformity believes Mending segmentation with various thresholds is quite handy in enhancing the accuracy of space and time dimension models in relation to one another. Figure 7.5 below shows the result using visual representation.

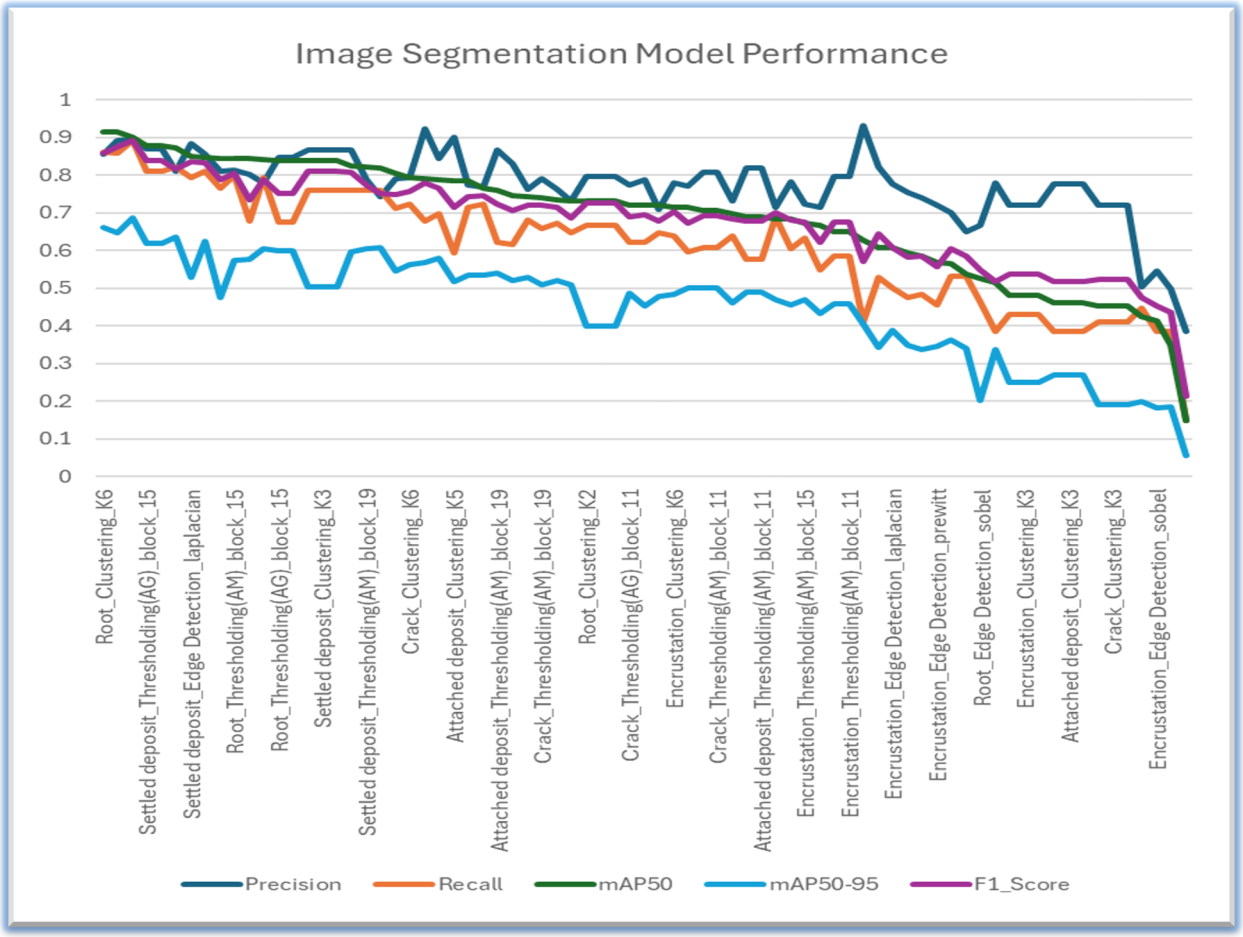


Figure 7.5: Image segmentation model performance.

Correlation analysis, as shown in Figure 7.6, revealed strong positive relationships among performance metrics, with mAP^{50} showing particularly high correlation with recall (0.959) and F1-score (0.981). However, precision demonstrated a weaker correlation with mAP^{50}, suggesting that strategies focusing on recall and F1-score are more effective for improving segmentation outcomes. Enhancements in F1-score were shown to significantly influence mAP^{50}, underscoring the importance of balanced precision and recall for optimal performance.

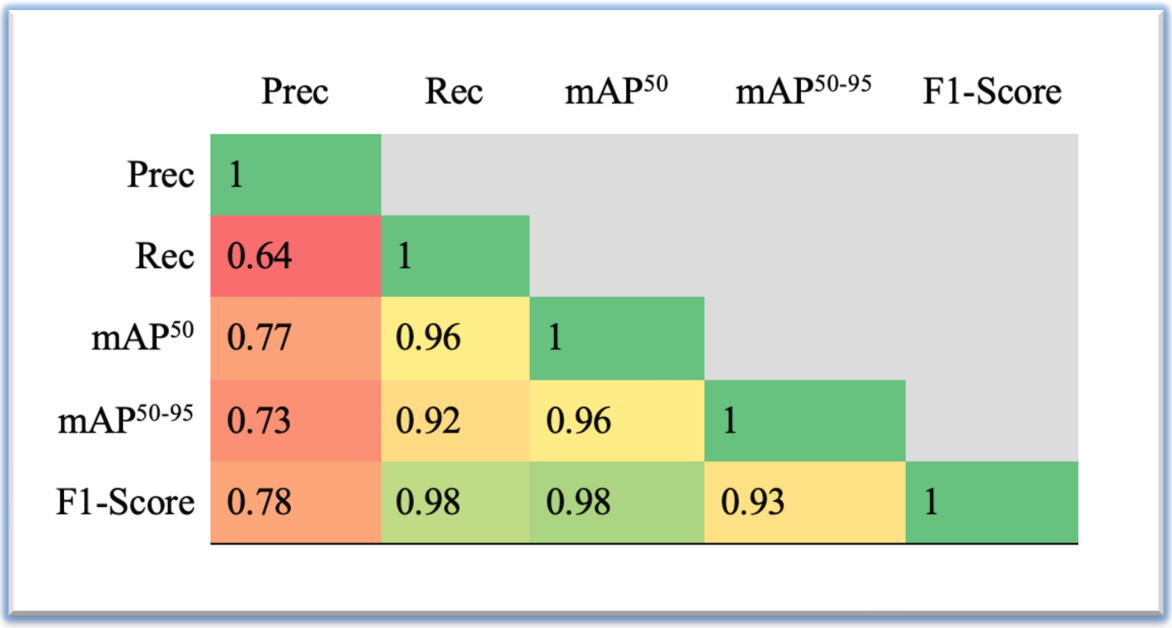


Figure 7.6: Correlation analysis of image segmentation result.

In conclusion, the results emphasize the need for tailored image segmentation strategies based on defect class characteristics. Thresholding methods, particularly adaptive Gaussian, consistently demonstrated robust performance across metrics, making them the most effective for segmentation tasks. Clustering techniques with higher configurations, such as K5 and K6, were particularly effective for complex classes like settled deposit and root. The statistical significance and correlation analyses further validate these findings, providing a strong foundation for selecting and optimizing segmentation techniques in similar applications.

As evidenced in Table 7.2, no single model architecture universally excels across all 16 defect categories; performance is highly defect dependent. To achieve the best real-world outcomes in automated defect detections should adopt a hybrid, ensemble-based deployment strategy. Rather than relying on a single, generalized model, the system architecture should dynamically route image feeds to the specific model proven most capable for a targeted defect type.

For example, if a specific survey priority is identifying critical structural failures (e.g., fractures or missing bricks), the video feed should be prioritized through the model that achieved the highest precision for those specific classes. Conversely, for routine maintenance checks focused on operational issues (e.g., attached deposits or root intrusions), the data should be processed by the architecture optimized for those specific visual signatures. By implementing an ensemble approach, where the outputs of specialized models are aggregated or selectively queried based on the inspection objective, SMCs can maximize overall detection accuracy, reduce false positives, and ensure that computational resources are allocated efficiently based on the specific priorities.

7.5 Enhancing AuDD via Morphological Operations

The application of morphological operations, including closed, dilation, erosion, and open, significantly influenced the YOLOv8 model's ability to detect various defect classes. These operations manipulate the structural elements of an image, helping to refine feature representation and improve model performance. Table 7.3 presents a consolidated summary of the results for all defect classes across these operations. Figures 7.7 and 7.8 provides a visual comparison of the performance metrics and correlation analysis, respectively.

The closed operation yielded strong results, achieving an average precision of 0.805, recall of 0.787, mAP50 of 0.832, and mAP50-95 of 0.566. Among the defect classes, the settled deposit stood out, with an mAP50 of 90.1% and recall of 0.855. This indicates that closed operations are effective in consolidating fine details and bridging small gaps in features, which appears to enhance the identification of well-defined classes like settled deposits. However, the performance for encrustation was comparatively weaker, with an mAP50 of 77.2%. This suggests that encrustation, with its less distinct features, might benefit from additional augmentation strategies beyond closing.

The dilation operation demonstrated an overall improvement over closed operations, with an average precision of 0.825, recall of 0.777, mAP^{50} of 0.829, and mAP^{50-95} of 0.561. As summarized in Table 7.3, the standout performance of the settled deposit class progressed here, achieving the highest mAP^{50} of 95.6%. This indicates that dilation, which enlarges features and emphasizes boundaries, enhances the model's ability to detect spatially prominent classes. Root also benefited from dilation, with an mAP^{50} of 86.6%, highlighting its utility for classes with defined edge structures. On the other hand, crack displayed the lowest mAP^{50} (74.3%), suggesting that the operation's emphasis on boundary expansion may have limited benefits for smaller, fragmented features like cracks.

Table 7.3: Summary of morphological operation results by defect class

Defect Class	Operation	Precision	Recall	mAP^{50}	mAP^{50-95}
Attached Deposit	Closed	0.844	0.788	0.866	0.583
	Dilation	0.902	0.740	0.824	0.561
	Erosion	0.901	0.736	0.849	0.598
	Open	**0.907**	**0.782**	**0.869**	**0.650**

Defect Class	Operation	Precision	Recall	mAP^{50}	mAP$^{50-95}$
Crack	Closed	0.758	**0.755**	0.785	0.522
	Dilation	0.736	0.727	0.743	0.498
	Erosion	0.753	0.686	0.742	0.523
	Open	**0.788**	0.734	**0.790**	**0.539**
Encrustation	Closed	0.773	0.716	**0.772**	**0.517**
	Dilation	0.764	0.682	0.755	0.492
	Erosion	0.807	**0.696**	0.727	0.508
	Open	**0.832**	0.614	0.735	**0.517**
Root	Closed	0.787	0.802	0.837	0.557
	Dilation	0.836	0.800	**0.866**	0.575
	Erosion	**0.837**	**0.822**	0.856	**0.596**
	Open	0.826	0.811	0.849	0.564
Settled Deposit	Closed	0.861	0.855	0.901	0.649
	Dilation	0.886	**0.937**	**0.956**	0.679
	Erosion	0.866	0.891	0.924	0.679
	Open	**0.910**	0.868	0.925	**0.693**

Erosion produced mixed results. While it achieved the highest average precision (0.833) among all operations, the average recall (0.766) and mAP50 (0.822) were slightly lower than dilation and open operations. Settled deposits once again excelled, achieving an mAP50 of 92.4% and recall of 0.891, indicating that erosion's ability to shrink features while retaining core structures is particularly effective for this class. The attached deposit also performed well, with the highest precision (0.901) among all classes and operations. However, encrustation displayed relatively

weak performance (mAP50 of 72.7%), reinforcing the notion that this class may require additional preprocessing steps to highlight its subtle structural features.

Open operations produced the highest overall metrics, with an average precision of 0.853, recall of 0.764, mAP50 of 0.833, and mAP50-95 of 0.593. Settled deposits achieved their highest performance across all metrics under this operation, with an mAP50 of 92.5% and recall of 0.868. This suggests that open operations, which remove small noise and retain larger features, effectively emphasize the characteristics of dominant defect types like settled deposits. Interestingly, encrustation, despite being the weakest performer, showed a slight improvement over erosion, with an mAP50 of 73.5%. This could indicate that open operations, by balancing feature removal and preservation, provide a marginal benefit to less prominent defect types. Figure 7.7 below provides a visual comparison of the models.

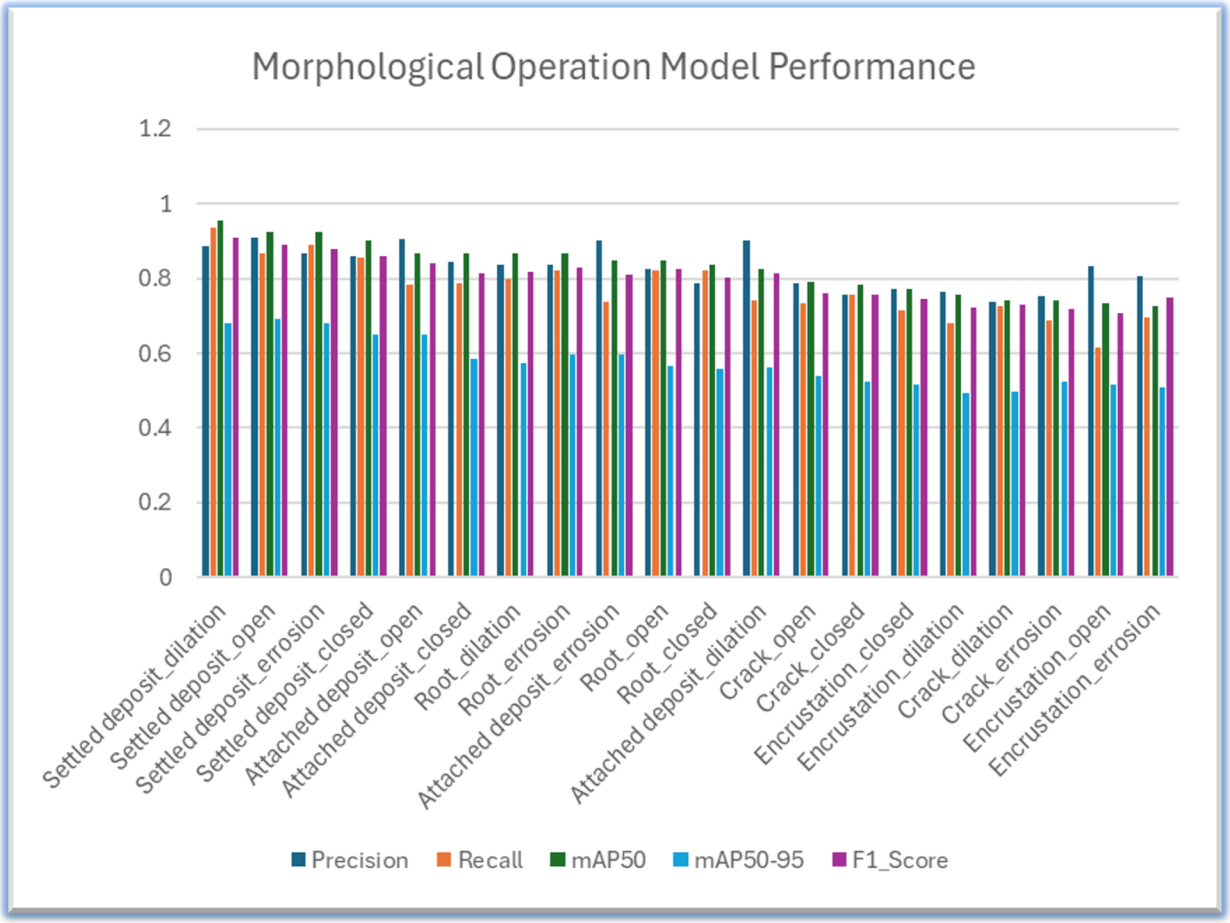


Figure 7.7: Model performance for each defect class using morphological operations

Correlation analysis, as depicted in Figure 7.8, further reinforced the robustness of mAP^{50} and mAP^{50-95} as performance indicators, as these metrics exhibited high positive correlations with F1-score and recall. In contrast, precision showed a weaker correlation with these metrics, implying that it may not always align with improvements in recall or overall model consistency. This is particularly relevant when evaluating operations like erosion, which maximize precision at the expense of recall. Enhancing the F1-score, which balances precision and recall, was shown to have the most significant impact on improving mAP50, further emphasizing its importance as a holistic metric.

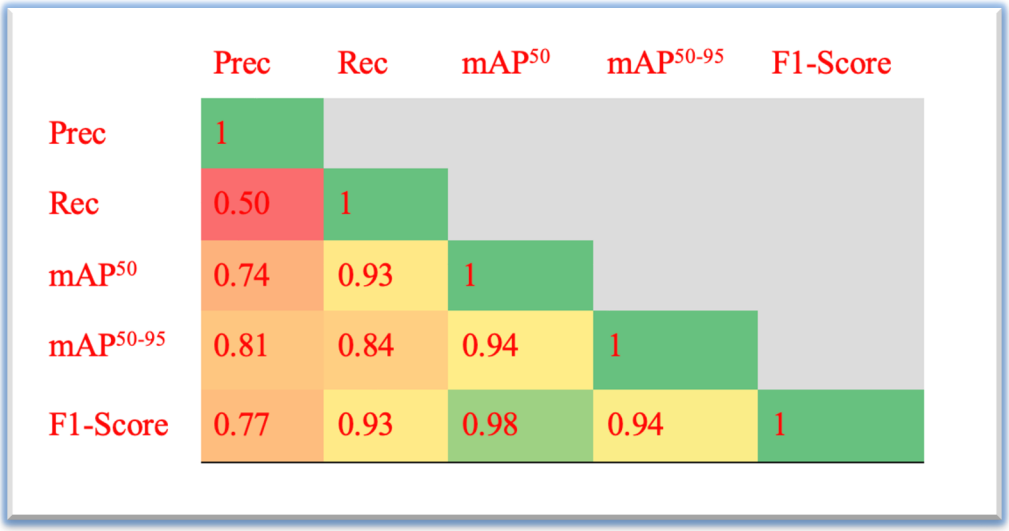


Figure 7.8: Correlation analysis for morphological operations

Using a radar chart as shown in Figure 7.9, the key observations here with morphological operations includes the open operation being the most effective, achieving superior performance across metrics. Likewise, the line chart in Figure 7.10 also confirms that settled deposit consistently outperforms other classes, benefiting from its distinct visual features. In contrast, Encrustation and Crack show weaker and more variable performance, reflecting the challenges in detecting less prominent features. Operations like Open and Dilation offer broad benefits, while Erosion and Closed show limited effectiveness for some classes. Invariably, this analysis emphasizes the importance of tailoring morphological operations to defect characteristics, because how can we get these insights if we failed to put them to defect-specific processing?

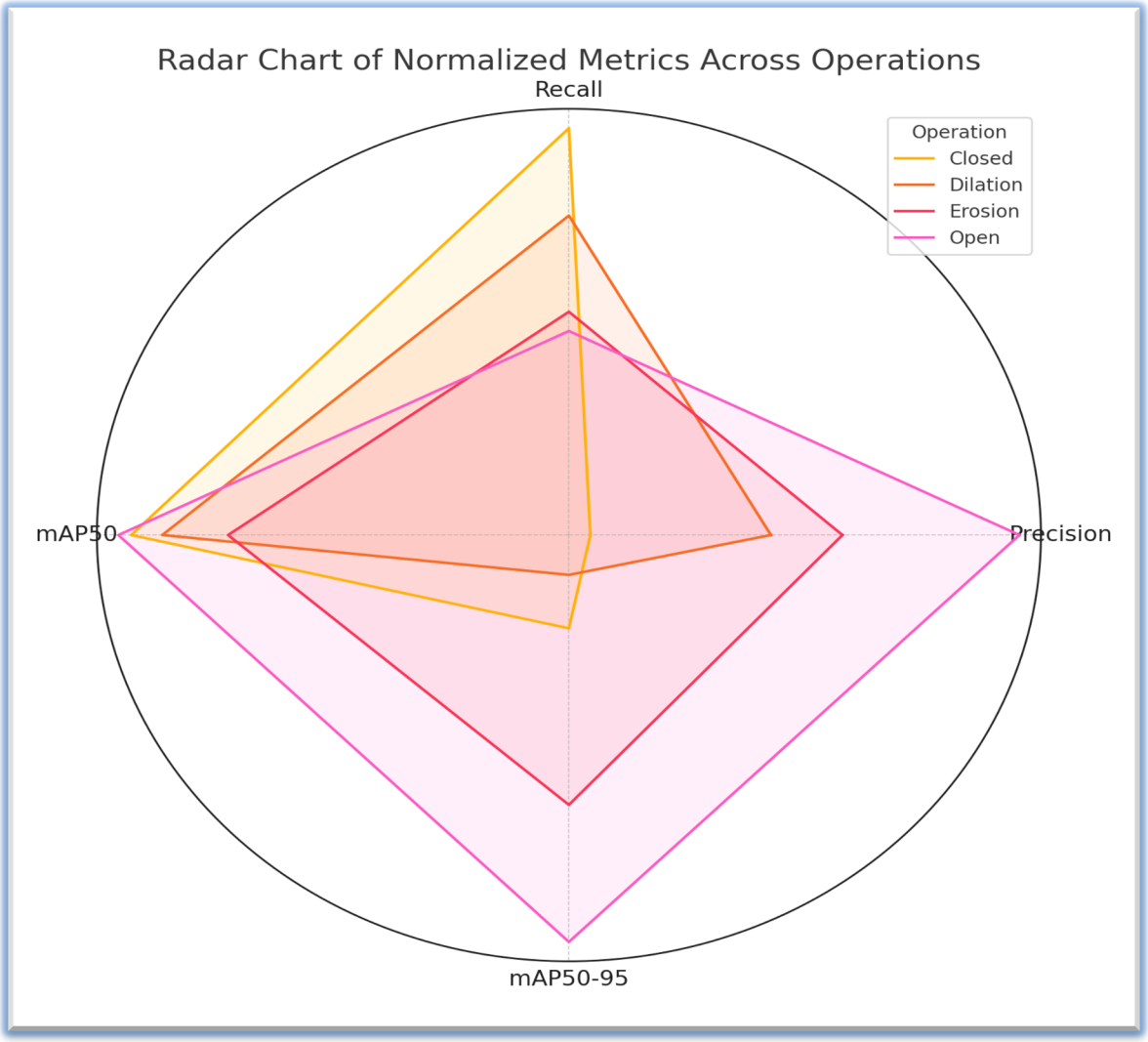


Figure 7.9: Radar chart showing comparative performance of each operation

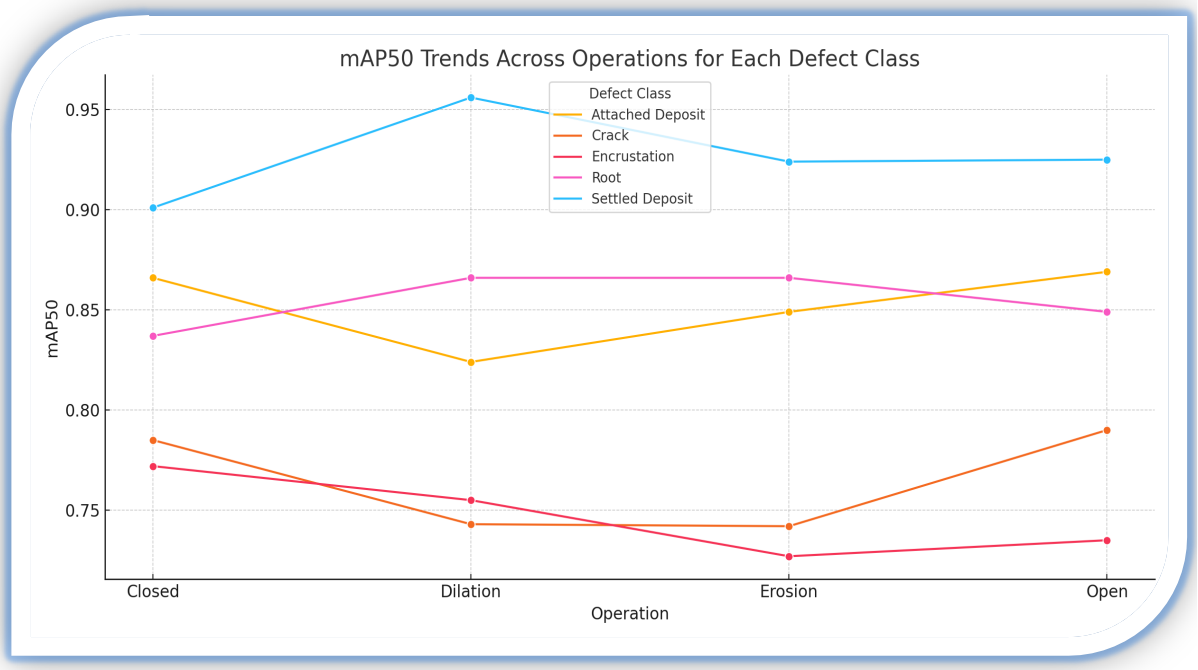


Figure 7.10: Line chart showing mAP^{50} trends across operations for each defect class

7.6 Overall Comparison and Selection of the Best Models

The comparative analysis of preprocessing techniques across defect classes reveals notable variations in model performances. Among contrast enhancement methods, contrast stretching delivered the best results for attached deposit, encrustation, and settled deposit, achieving an mAP50 of 92.9% for attached deposit. Adaptive equalization emerged as the top method for crack and root detection, underscoring its strength in enhancing subtle structural features. Image segmentation techniques showed superior results for root and settled deposit detection, with Clustering K6 achieving an mAP50 of 91.6% for root. For encrustation, the thresholding adaptive mean method proved most effective.

Morphological operations highlighted the versatility of the open method for attached deposit and crack detection, while dilation excelled in spatially prominent defects like settled deposit, achieving the study's highest mAP50 of 95.6%. Figure 7.11 to 7.13 below summarises the best performing techniques for each defect type under each pre-processing category.

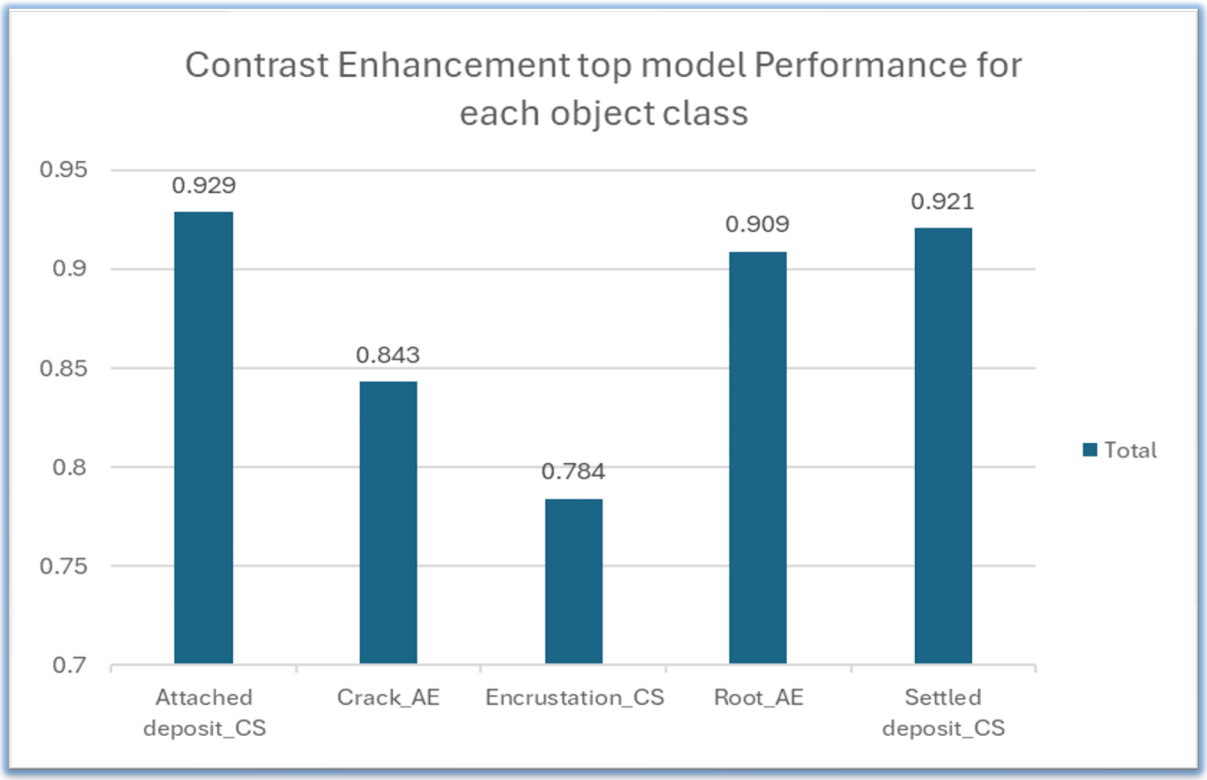


Figure 7.11: Contrast enhancement top model performance for each class

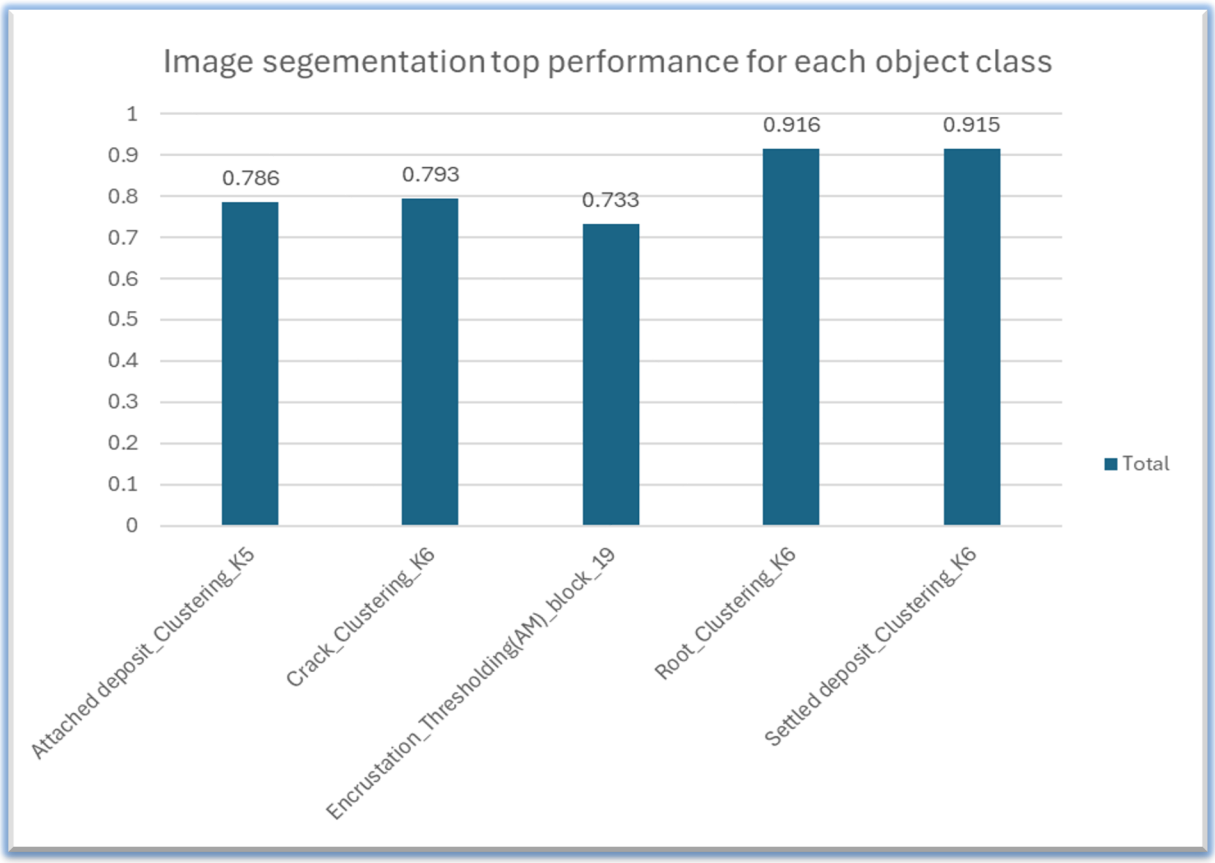


Figure 7.12: Image segmentation of top model performance for each class

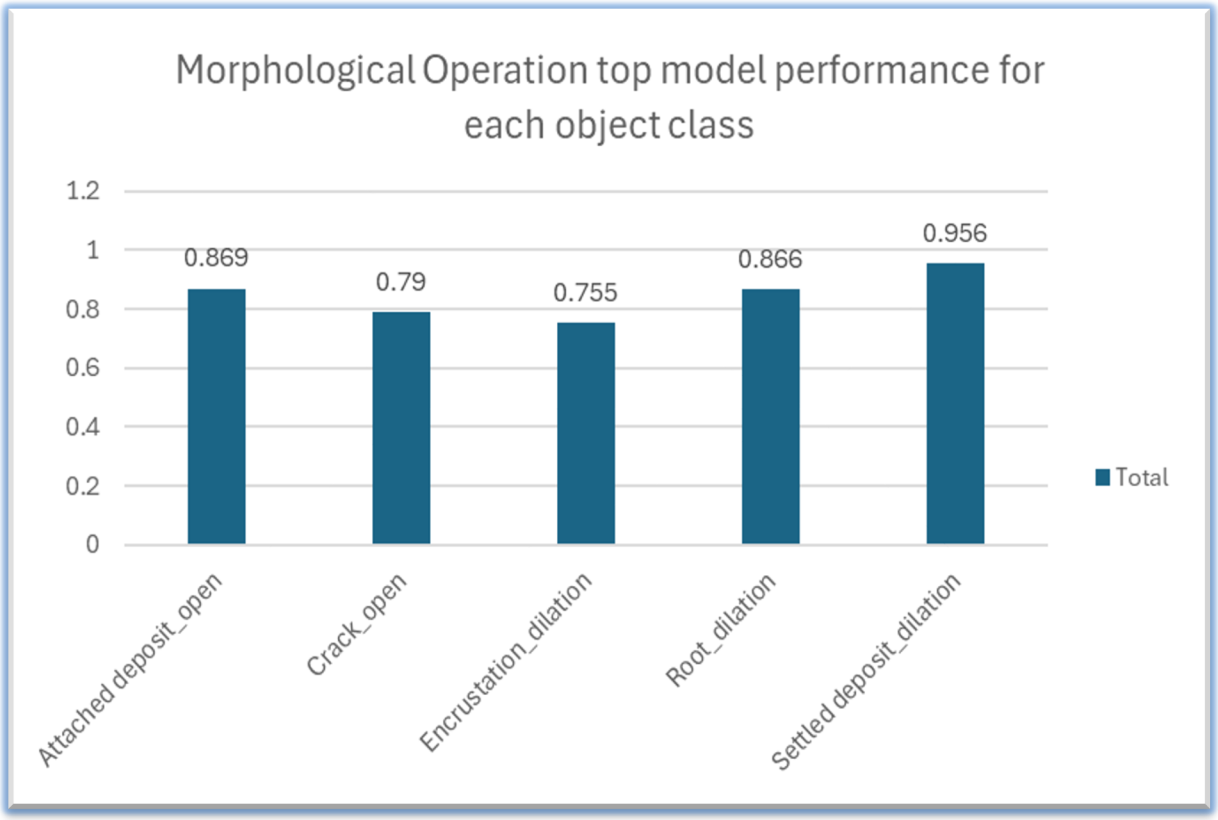


Fig 7.13: Morphological operation of top model performance for each class

For further detailed comparison, Table 7.4 summarizes the best performing techniques for each defect class by providing insight into how specific pre-processing methods enhance detection performance while focusing only on the top 3 (best, secondary, and alternative) models for each defect. For Attached Deposit, Contrast Stretching achieved the highest mAP50 of 92.9%, demonstrating its strong capability to enhance the fine textural features essential for detection. This performance outshines Clustering K5 and Open, which recorded mAP50 scores of 78.6% and 86.9%, respectively. The relatively lower performance of these techniques indicates their limitations in emphasizing subtle contrasts that are crucial for detecting attached deposits.

Crack detection presented a slightly different trend, where Adaptive Equalization achieved an mAP50 of 84.3%, closely followed by Open, which recorded 84.9%. Despite the marginal difference, these techniques proved effective in capturing the detailed structural irregularities characteristic of cracks. However, Clustering K6, with an mAP50 of 79.3%, demonstrated a less competitive performance, suggesting that its methodology may not fully capture the nuances of fine edge features necessary for crack identification.

Table 7.4: Comparative mAP^{50} Scores for Top 3 Best Techniques in each defect category

Defect Class	Best Technique	mAP	Secondary Technique	mAP	Alternative Technique	mAP
Attached Deposit	Contrast Stretching	0.929	Open	0.869	Clustering K5	0.786
Crack	Adaptive Equalization	0.843	Open	0.849	Clustering K6	0.793
Encrustation	Contrast Stretching	0.784	Dilation	0.755	Clustering K5	0.733
Root	Clustering K6	0.916	Adaptive Equalization	0.909	Dilation	0.866
Settled Deposit	Dilation	0.956	Clustering K6	0.915	Contrast Stretching	0.921

Encrustation detection further highlighted the strengths of Contrast Stretching, which attained an mAP^{50} of 78.4%, outperforming Dilation and Clustering K5, which achieved 75.5% and 73.3%, respectively. This finding reinforces the advantage of using techniques that enhance contrast to amplify subtle features. The performance gap indicates that Dilation and Clustering K5 may struggle with the less visually prominent characteristics of encrustations, making them less suited for this defect type.

Root detection revealed the dominance of Clustering K6, which achieved an mAP50 of 91.6%, surpassing Adaptive Equalization at 90.9% and Dilation at 86.6%. The exceptional performance of Clustering K6 highlights its ability to segment and emphasize the spatial complexity associated with roots. While Adaptive Equalization delivered comparable results, the slightly lower performance of Dilation reflects its limitation in handling the intricate patterns and structural complexity required for accurate root detection.

For Settled Deposit, Dilation emerged as the top performer with an mAP^{50} of 95.6%, indicating its superior ability to capture the spatially dominant features of this defect. Contrast Stretching and Clustering K6 followed with scores of 92.1% and 91.5%, respectively. These results suggest that

while Contrast Stretching and Clustering K6 perform competitively, they fall slightly short in amplifying the spatial prominence necessary for optimal detection of settled deposits.

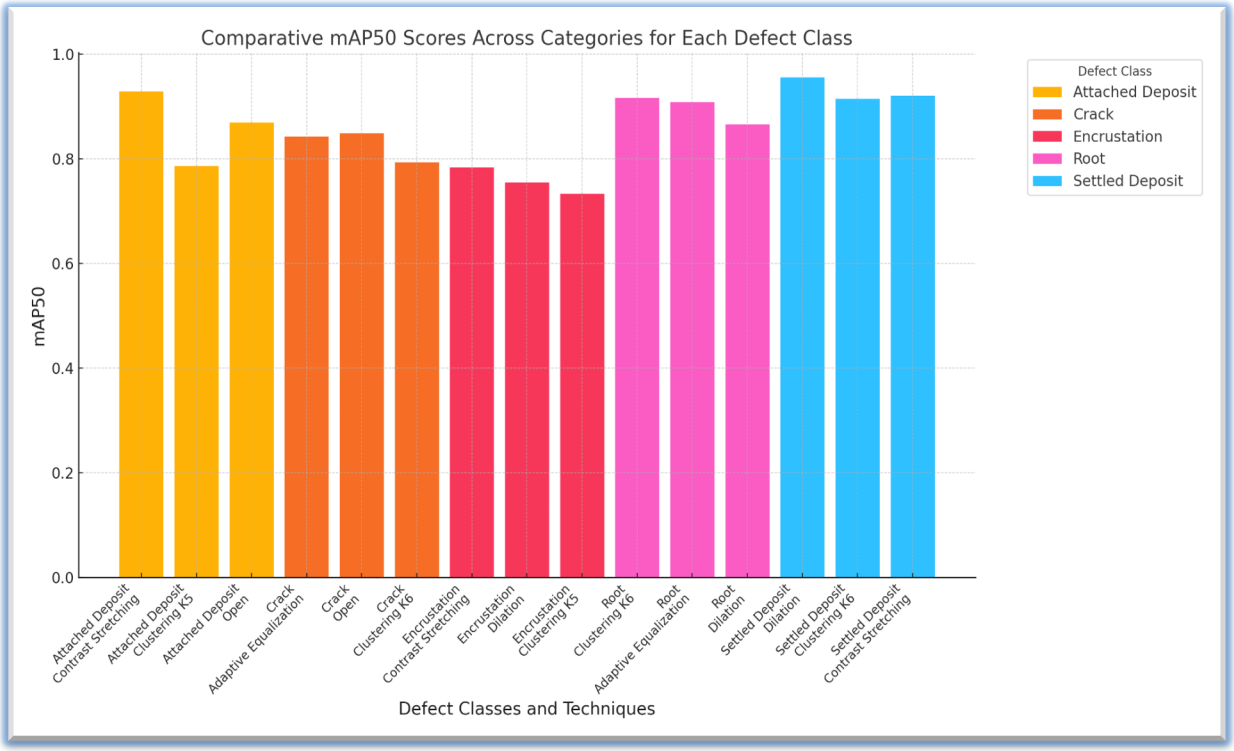


Figure 7.14: How each pre-processing techniques performed for each defect

Conclusively, this analysis underscores the importance of aligning preprocessing techniques with the inherent characteristics of each defect class. Techniques like Dilation and Clustering K6 excel in enhancing spatially complex features, while Contrast Stretching and Adaptive Equalization demonstrate superior performance for defects that require contrast enhancement or structural clarity. This clear separation between categories further illustrates the need to for defect-specific preprocessing strategies to optimize model accuracy and robustness. To summarize, Figure 7.15 depict a high-level representation of which preprocessing categories is efficient for each deflect class.

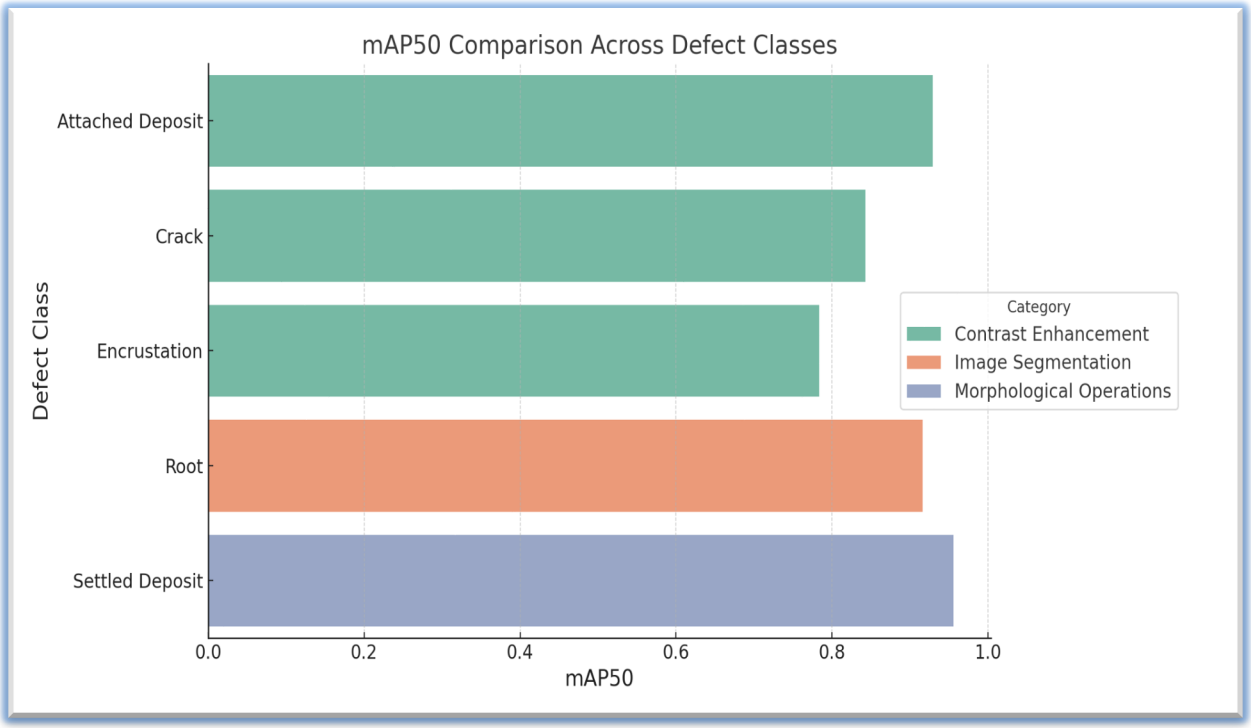


Figure 7.15: Category-specific mAP50 comparison across defect categories

7.7 Chapter Summary

This result analysis and evaluation chapter has provided a comprehensive assessment of the performance of various data preprocessing techniques used, including various operations from contrast enhancement, image segmentation, and morphological operations. These were applied to five defect classes: attached deposit, crack, encrustation, root, and settled deposit. The analysis highlights the strengths and weaknesses of each technique in optimizing model performance, with mAP50 serving as the key evaluation metric.

Settled deposit and root emerged as the best-performing defect classes across multiple techniques, achieving top mAP50 scores of 95.6% and 91.6%, respectively, while crack and encrustation consistently exhibited lower performance. Correlation analysis using Pearson and Spearman ranking further underscored the significance of the selected techniques.

The findings emphasize the importance of tailoring preprocessing techniques to the specific characteristics of each defect class. Overall, contrast stretching demonstrated strong performance in texture-based defects, while clustering and dilation techniques excelled in detecting spatially dominant defects like root and settled deposit.

CHAPTER EIGHT

8.0 Deployment of Deep Learning Model

8.1 Chapter Overview

This chapter presents a comprehensive discussion on the deployment of developed deep learning models on cloud for real-time usage. The discussion outlines the significance of model deployment, elaborates on the associated challenges, and describes in detail the steps taken to ensure that the deployed model is both portable and scalable. A critical analysis of the entire deployment pipeline is provided, from containerization and API development to cloud hosting. This offer insights into the tools and technologies that facilitated the transition from model training in a controlled research environment to production.

Throughout the chapter, the deployment strategy is examined through practical implementation, with relevant diagrams referenced to enhance understanding and illustrate key architectural components. This chapter also provide details such as infrastructure requirements, latency reduction, scalability concerns, and cost-effective deployment methods for deploying the model. Furthermore, it delves into the integration of cutting-edge technologies like Docker, FastAPI, and AWS, which are instrumental in modern model deployment practices.

8.2 Deployment Pipeline for Deep Learning Model

The deployment of deep learning models in production has become a pivotal step in harnessing the potential of artificial intelligence for real-world applications (LeCun et al., 2015). Model deployment ensures that the insights and predictions generated by complex algorithms are accessible, reliable, and actionable. This is particularly critical in applications such as sewer defect detection, where timely interventions can prevent environmental hazards and infrastructure failures. However, deploying deep learning models is not without challenges. Issues such as model scalability and system integration must be addressed to maintain operational efficiency (Amodei & others, 2016).

In this research, the deployment pipeline is designed to seamlessly transition the YOLO model from the research and development environment to a scalable production system. At a high level, as shown in Figure 8.1, the pipeline consists of several critical stages that collectively ensure the model's reproducibility, efficiency, and robustness. Initially, the trained model is encapsulated

within a containerized environment using Docker, which guarantees consistency across different deployment platforms (Merkel, 2014). This containerization forms the foundation upon which further components of the deployment pipeline are built.

Following containerization, an API is developed using FastAPI to facilitate real-time inference requests. This API acts as the interface between end-users and the deployed model, handling image pre-processing, inference, and result visualization. The containerized application is then deployed on a cloud platform, specifically Amazon Web Services (AWS), which provides the necessary computational resources and scalability options. A combination of AWS EC2 instances, Elastic Load Balancing (ELB), and automated CI/CD pipelines ensures that the deployed model can handle high-volume requests with minimal latency (Amazon Web Services Inc., 2021d). Finally, performance optimization and monitoring are implemented to enable continuous tracking of system metrics and proactive issue resolution. The next sections will dive into these different components of the deployment pipeline.

The following are the abbreviations used in the architecture (Figure 8.1) below:

- YOLO – You Only Look Once
- API – Application Programming Interface
- AWS – Amazon Web Services
- EC2 – Elastic Compute Cloud
- ELB – Elastic Load Balancing
- CI/CD – Continuous Integration / Continuous Deployment
- JWT – JSON Web Token

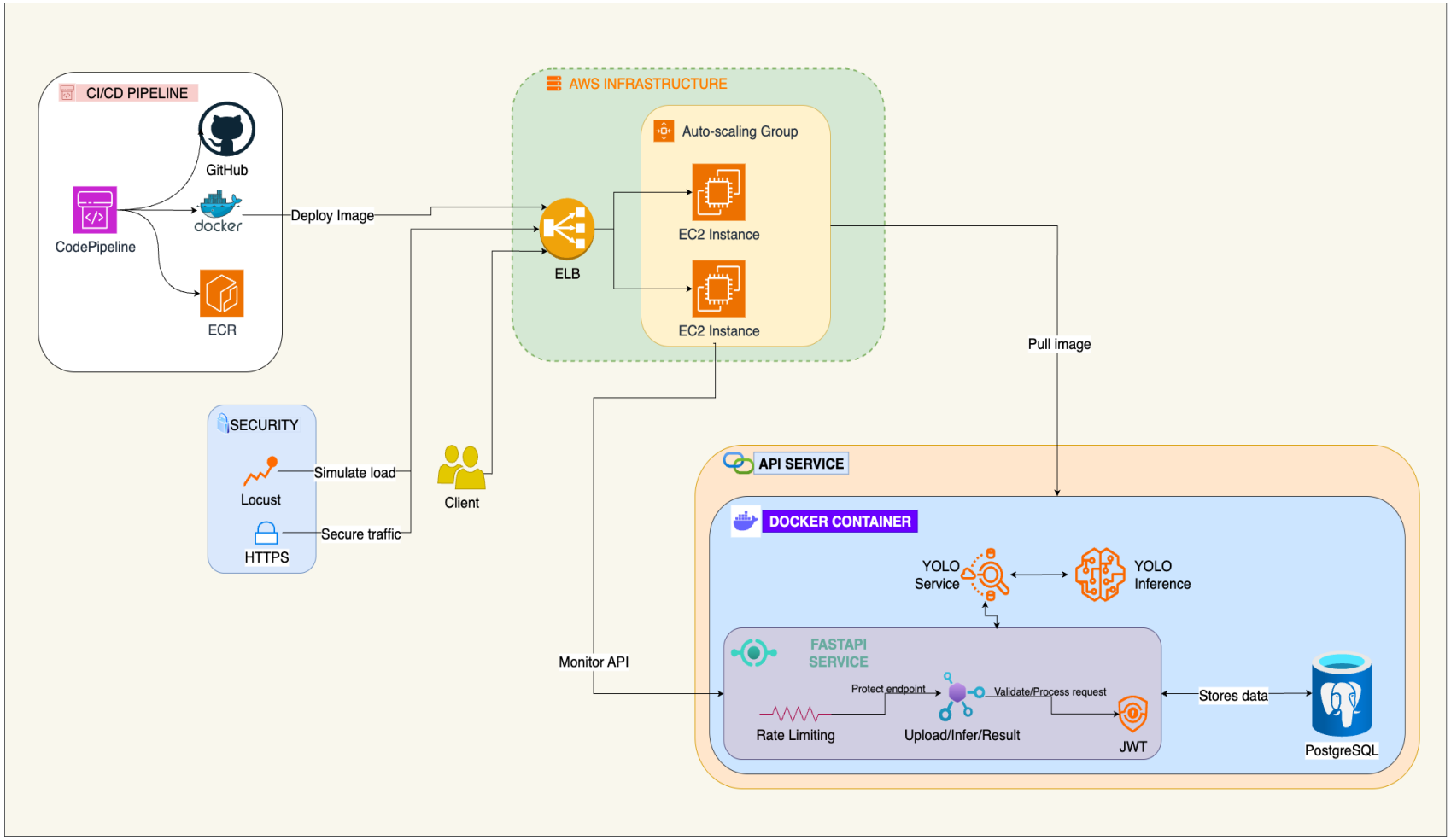


Figure 8.1: High-Level Architecture of the Deployment Pipeline.

8.3 Containerization with Docker

Containerization has revolutionized the deployment process by providing a consistent and reproducible runtime environment for software applications (Boettiger, 2015). Docker, as the leading containerization tool, encapsulates the YOLO model along with its dependencies into a portable image that can be deployed across various environments without compatibility issues. This approach mitigates the common "but it works on my machine" problem and facilitates a smooth transition from research and development to production. The use of Docker offers several distinct advantages that significantly enhance the deployment process for sewer defect detection applications. In this context, reproducibility is essential because Docker containers capture the entire runtime environment, including the operating system, libraries, and configurations, ensuring that the detection model performs consistently across various deployment platforms.

Portability is equally critical, as containerized applications can be seamlessly transferred between development, testing, and production environments; this is particularly important in the continuous integration and deployment pipelines required for maintaining up-to-date, reliable sewer defect detection systems (Gallo & Digilio, 2019). Moreover, isolation provided by Docker helps mitigate potential conflicts between the deep learning inference processes and other system components, ensuring that the specific requirements of real-time defect detection can be met without interference, thereby increasing the overall reliability and efficiency of the system.

A central aspect of containerization is the creation of a Docker file, which provides a scripted set of instructions to build the Docker image. The Docker file for the YOLO model deployment includes steps for installing dependencies, copying model weights, and setting up the API codebase. An example Docker file used in this research is provided below in Figure 8.2.

```
# Use an official Python runtime as a base image
FROM python:3.8-slim

# Set the working directory
WORKDIR /app

# Copy the requirements file and install dependencies
COPY requirements.txt /app/
RUN pip install --upgrade pip && pip install -r requirements.txt

# Copy the model weights and API code into the container
COPY model/ /app/model/
COPY api/ /app/api/

# Expose the port that the API will run on
EXPOSE 8000

# Command to run the API using Uvicorn
CMD ["uvicorn", "api.main:app", "--host", "0.0.0.0", "--port", "8000"]
```

Figure 8.2: Example Docker file for YOLO Model Deployment.

After creating the Docker file, the next step was building the Docker image using “docker build -t yolo-sewer-detection .”, and a successful image creation was followed by rigorous local testing. The container is run using “docker run -p 8000:8000 yolo-sewer-detection” which maps the container’s internal port to a host port, allowing for accessibility and validation of the API endpoints: During testing, the containerized model inference service was evaluated for proper functionality. It is essential to confirm that all dependencies are correctly installed, the API responds appropriately, and the YOLO model produces accurate defect detection results. This validation process helps to identify and resolve potential configuration issues before the container is deployed in a production environment (Felter et al., 2015).

8.4 API Development with FastAPI

The development of an Application Programming Interface (API) was a crucial step in operationalising the sewer defect detection model, enabling seamless interaction between the trained YOLO architecture and end users or external systems. In this context, the API serves as a service layer that exposes the model’s inference capabilities over a network, allowing inspection videos or images to be uploaded, processed, and analysed remotely. By adopting a RESTful design, the API ensures interoperability, scalability, and ease of integration with diverse client applications, including web dashboards, mobile inspection tools, and municipal asset management platforms. The design process focused on achieving low-latency responses, high throughput, and robust error handling, which are essential requirements for real-time defect detection in sewer networks where timely insights can drive maintenance decisions, prevent failures, and optimise operational resources.

FastAPI was selected as the framework for developing the RESTful API that interfaces with the YOLO-based sewer defect detection model, primarily due to its high performance, ease of use, and native support for asynchronous operations (Tiangolo, 2018). The API serves as the operational bridge between the trained deep learning model and end users, receiving requests in the form of video or image data, processing these inputs, running inference, and returning detection results. This capability is critical in sewer inspection workflows where quick turnaround on defect identification can directly impact decision-making for maintenance scheduling and infrastructure safety interventions.

One of the key strengths of FastAPI lies in its architecture, which is built on Starlette (Ronacher, 2018) for the web layer and Pydantic for data layer validation. By leveraging asynchronous programming, FastAPI can efficiently handle multiple concurrent requests without blocking operations, a necessity in scenarios where multiple inspection videos may be uploaded and processed simultaneously. This is further relevant for large-scale deployments involving multiple Sewer Management Companies (SMCs) submitting inspection data concurrently. Furthermore, the

framework's ability to automatically generate interactive OpenAPI standard documentation through Swagger UI and ReDoc enables rapid development and debugging cycles (Jones, 2019). This accelerates prototyping and iteration, ensuring that new features and model updates can be deployed with minimal delay.

When compared to other popular frameworks, FastAPI presents several advantages. Flask, while lightweight and flexible, is synchronous by default and requires additional configuration or extensions (e.g., Gevent) to achieve similar concurrency handling, often at the cost of added complexity. Django REST Framework (DRF) offers a more monolithic and feature-rich ecosystem, including ORM integration and a mature authentication system, but its higher overhead and synchronous nature can be less efficient for high-throughput, real-time inference tasks. Node.js-based solutions, such as Express.js, excel at handling asynchronous operations but lack Python-native deep learning integration, which is crucial for seamlessly interfacing with PyTorch-based YOLO models. This would necessitate additional bridging layers or inter-process communication which add additional network latency.

In industry practice, FastAPI has consistently demonstrated superior request handling performance among Python web frameworks, often rivalling or surpassing Node.js in throughput while maintaining lower latencies for compute-bound AI inference workloads (U.S. National Institute of Standards & Technology, 2018). Its native compatibility with Python's asyncio library and direct integration with AI/ML workflows allows deep learning models to be invoked without context switching between environments. This design choice reduces operational friction and simplifies code maintenance. Additionally, Pydantic's robust type checking and data validation mechanisms help prevent malformed input data from propagating through the inference pipeline, significantly reducing runtime errors and improving overall system reliability.

Ultimately, FastAPI was chosen over competing frameworks because it offers the optimal balance of performance, scalability, and development efficiency for this sewer defect detection system. Its asynchronous capabilities are well-matched to the demands of concurrent real-time analysis, its tooling accelerates API development and documentation, and its seamless integration with Python's machine learning ecosystem eliminates unnecessary middleware complexity. By combining speed, reliability, and ease of scaling, FastAPI provides a deployment-ready foundation that aligns with both the research objectives and the operational requirements of field-ready intelligent infrastructure monitoring systems. Figure 8.3 shows the FastAPI structure for both API development and model inference.

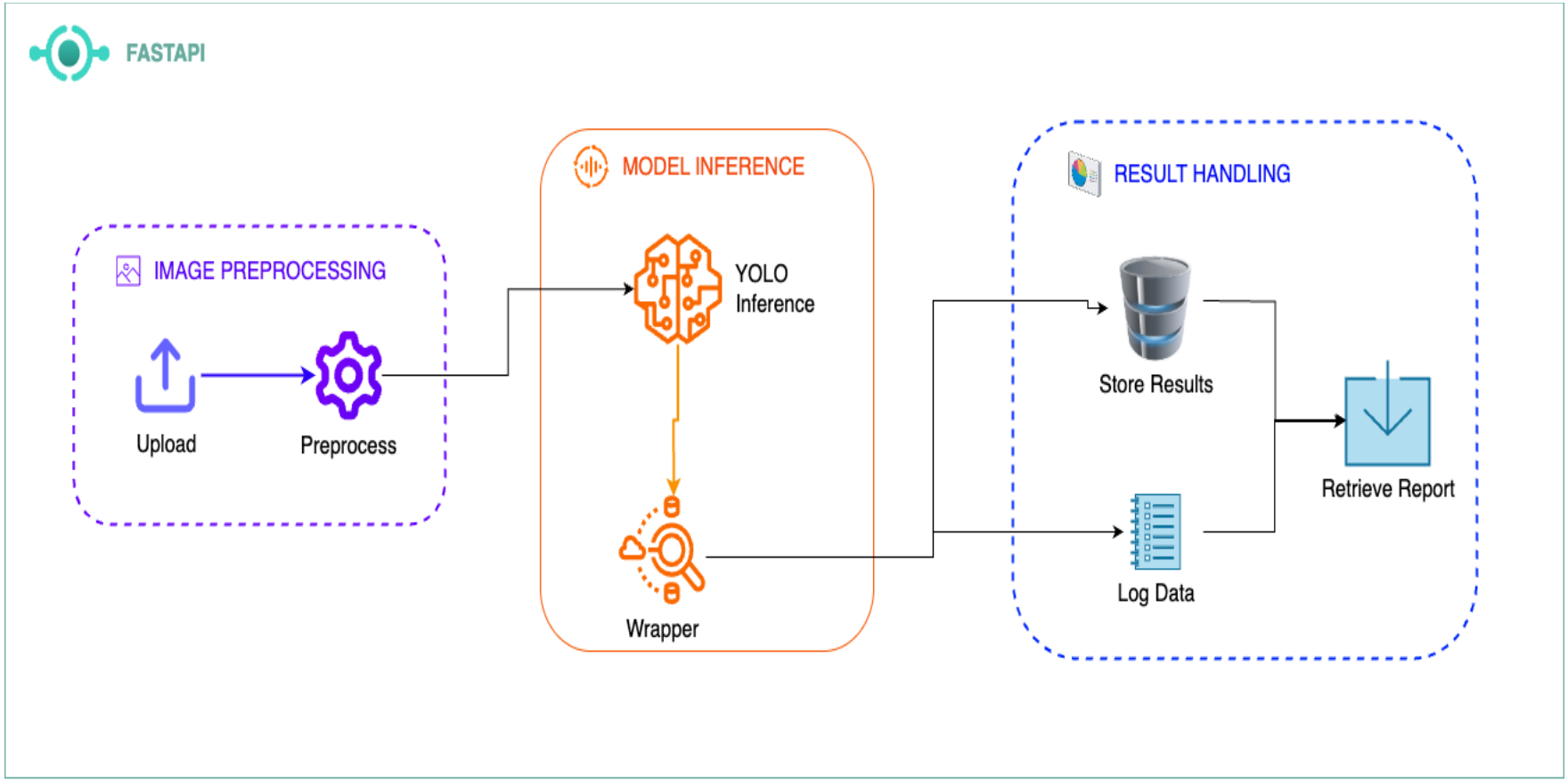


Figure 8.3: FastAPI structure for model inference.

8.5 Model Deployment on AWS

Cloud deployment is essential for ensuring the scalability, availability, and efficiency of the deployed model. AWS has been selected for hosting the containerized YOLO model due to its robust ecosystem, diverse service offerings, and proven track record in managing high-performance computing workloads (Amazon Web Services Inc., 2021c). This section describes the selection and configuration of AWS services, and the strategies employed to achieve a resilient deployment.

Several AWS services were considered for hosting and managing the deployed model, each with distinct advantages and trade-offs. AWS EC2 (Elastic Compute Cloud) provides scalable virtual servers with flexible configurations for CPU, memory, and GPU resources, making it particularly suitable for resource-intensive deep learning inference that requires persistent compute capacity. AWS Lambda, though primarily designed for serverless applications, is less ideal for long-running inference tasks due to its execution time and resource constraints, but it may still be useful for auxiliary tasks such as lightweight preprocessing. AWS SageMaker offers a fully managed solution for model building, training, and deployment, with built-in monitoring and optimization features. However, while SageMaker simplifies the deployment process, its cost structure may not always align with specific project requirements such as this, making EC2 a more flexible and cost-effective choice for long-term hosting of this research output.

For the purposes of this deployment, AWS EC2 is chosen as the primary service. EC2 instances configured with GPU acceleration are used to run the Docker container hosting the YOLO model, ensuring that the inference tasks benefit from hardware acceleration (Amazon Web Services Inc., 2021a). Deploying the model on AWS EC2 involves several critical steps to ensure an optimized and scalable inference environment. The process begins with instance selection and configuration, where an appropriate EC2 instance, such as those from the P3 or G4 series, is chosen to leverage GPU acceleration for deep learning workloads. These instances are provisioned with Amazon Machine Images (AMIs) optimized for AI applications.

After the instance was configured, Docker is installed and configured to facilitate containerized deployment, with `Docker Compose` used if multiple containers are required. The next step involves deploying the Docker container, where the pre-built image is pushed to a registry like Docker Hub and then pulled onto the EC2 instance. The container is executed with the necessary environment variables and port mappings to enable smooth API interactions. Finally, the inference environment is configured by installing essential GPU drivers and libraries, including NVIDIA CUDA and cuDNN, ensuring the instance can fully utilize hardware acceleration for efficient deep learning inference.

To enhance scalability and ensure high availability, an AWS Elastic Load Balancer (ELB) is deployed in front of the EC2 instance(s). The ELB distributes incoming traffic across multiple

instances, thereby preventing any single instance from becoming a performance bottleneck (Amazon Web Services Inc., 2021d). Furthermore, auto-scaling policies are configured to dynamically adjust the number of EC2 instances based on traffic patterns, ensuring that the system maintains optimal performance during peak loads, as depicted in Figure 8.4

Automation also plays a critical role in maintaining an efficient deployment process. AWS CodePipeline was used to implement a CI/CD pipeline that automates the build, testing, and deployment of the *Docker* container. The pipeline is integrated with Docker Hub and the source code repository, ensuring that any updates to the model or API code are promptly reflected in the production environment. This automated workflow minimizes downtime and supports rapid iterations, which is vital for continuous improvement in model performance (Humble & Farley, 2010)

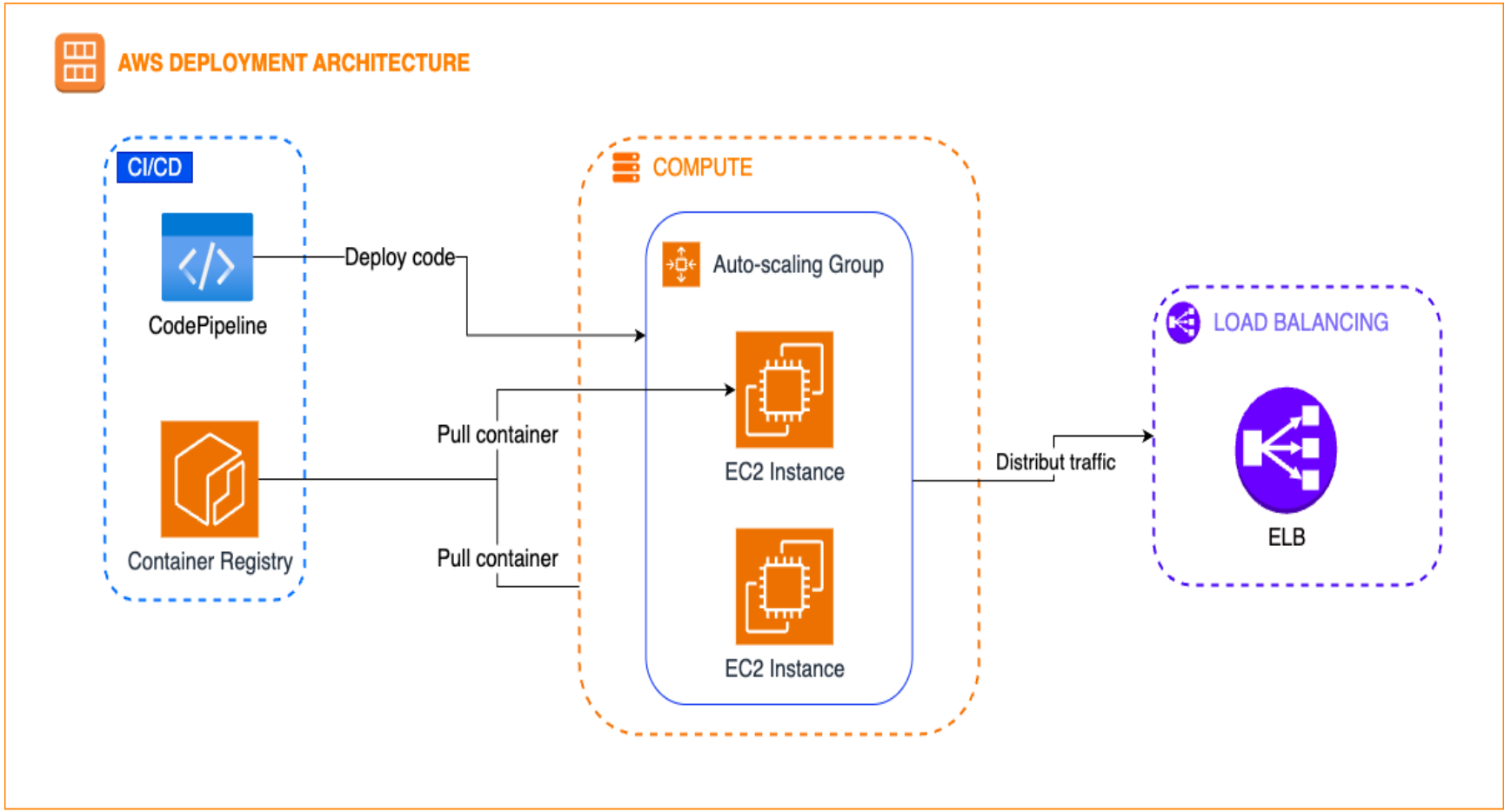


Figure 8.4: AWS Deployment Architecture.

8.7 Features, Testing, and Evaluation of the Deployed Model

8.7.1 Feature Walkthrough: Survey Processing and Report Generation

The deployed system incorporates an integrated project and survey management pipeline, designed to guide users from the initial upload of inspection surveys to the delivery of structured, actionable reports. This pipeline provides a clear and traceable sequence of operations, ensuring that all defect detection outputs are contextualized within their respective projects. Each project contains associated surveys, their analysis lifecycle, and the corresponding reports, thus creating a centralized record for operational decision-making.

The dashboard, as depicted in Figure 8.5, offers a centralised snapshot of ongoing activities, displaying key metrics such as projects, models, datasets, and recent surveys. It features a project activity chart, quick-access controls for creating projects or uploading surveys, and a recent projects list.

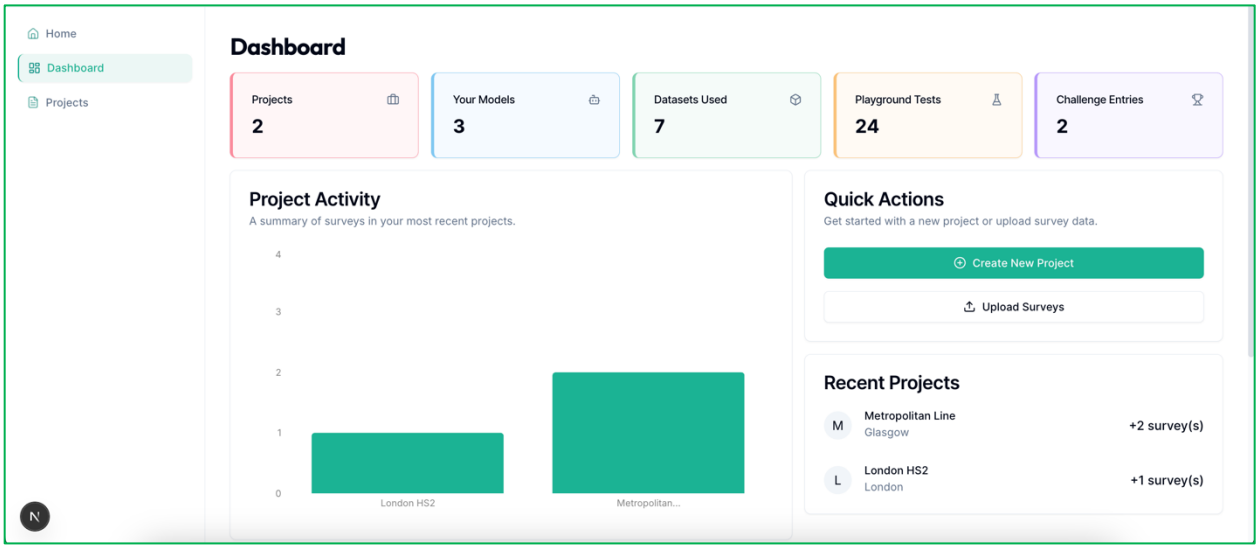


Figure 8.5: The dashboard overview page displaying key metrics and quick access control

From the user's perspective, the workflow begins on the Projects page as shown in Figure 8.6, which lists all active and archived projects alongside metadata such as project name, location, number of surveys, total detected defects, creation date, and status. Selecting a project navigates to its Project Detail view in Figure 8.7, where users can view project metadata and manage its

associated surveys. Here, the survey upload process is straightforward: inspection videos are either dragged into the upload panel or selected from local storage. Once uploaded, surveys enter a four-stage processing lifecycle: queued, processing (AI defect analysis), report preparation, and completed. A dedicated survey progress modal (Figure 8.8 and 8.9) provides real-time visualization with progress indicators and contextual status updates

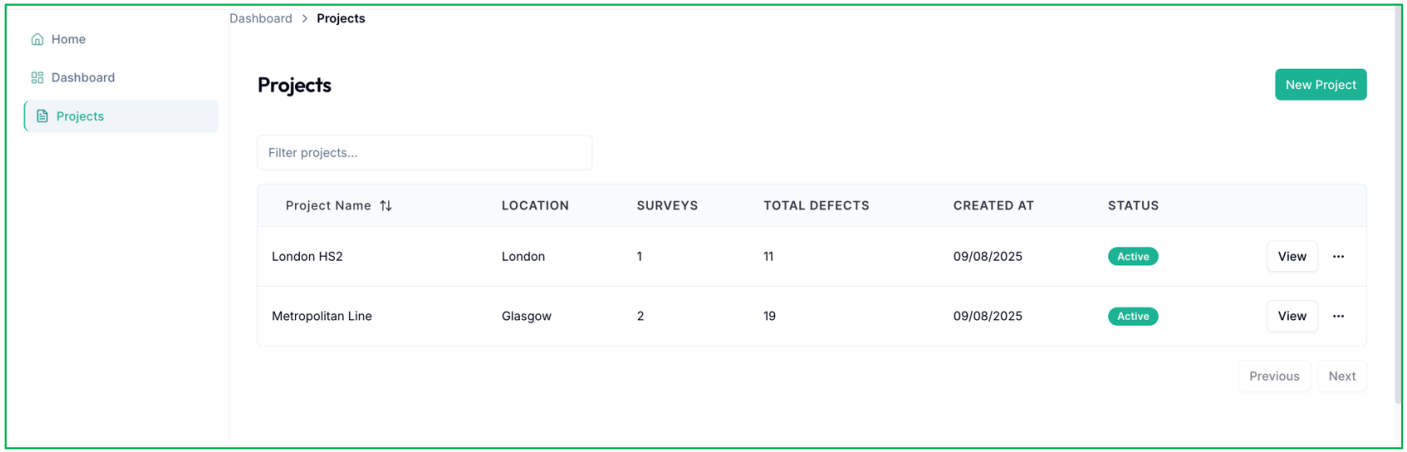


Figure 8.6: Projects Overview page showing project metadata and summary statistics.

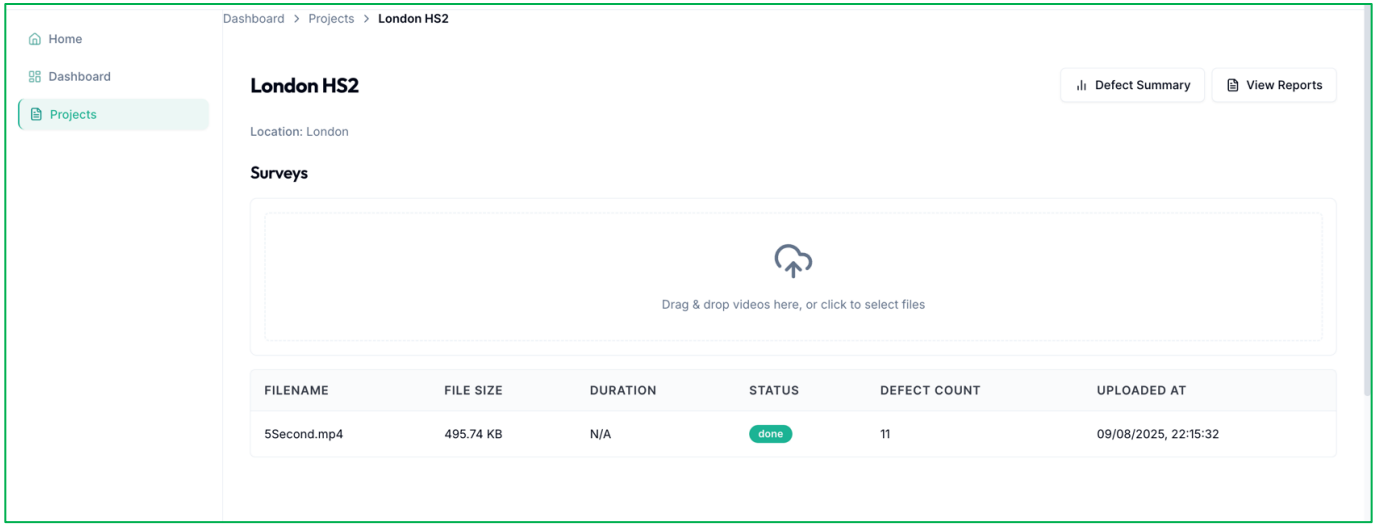


Figure 8.7: Project Detail view with surveys list and upload panel visible.

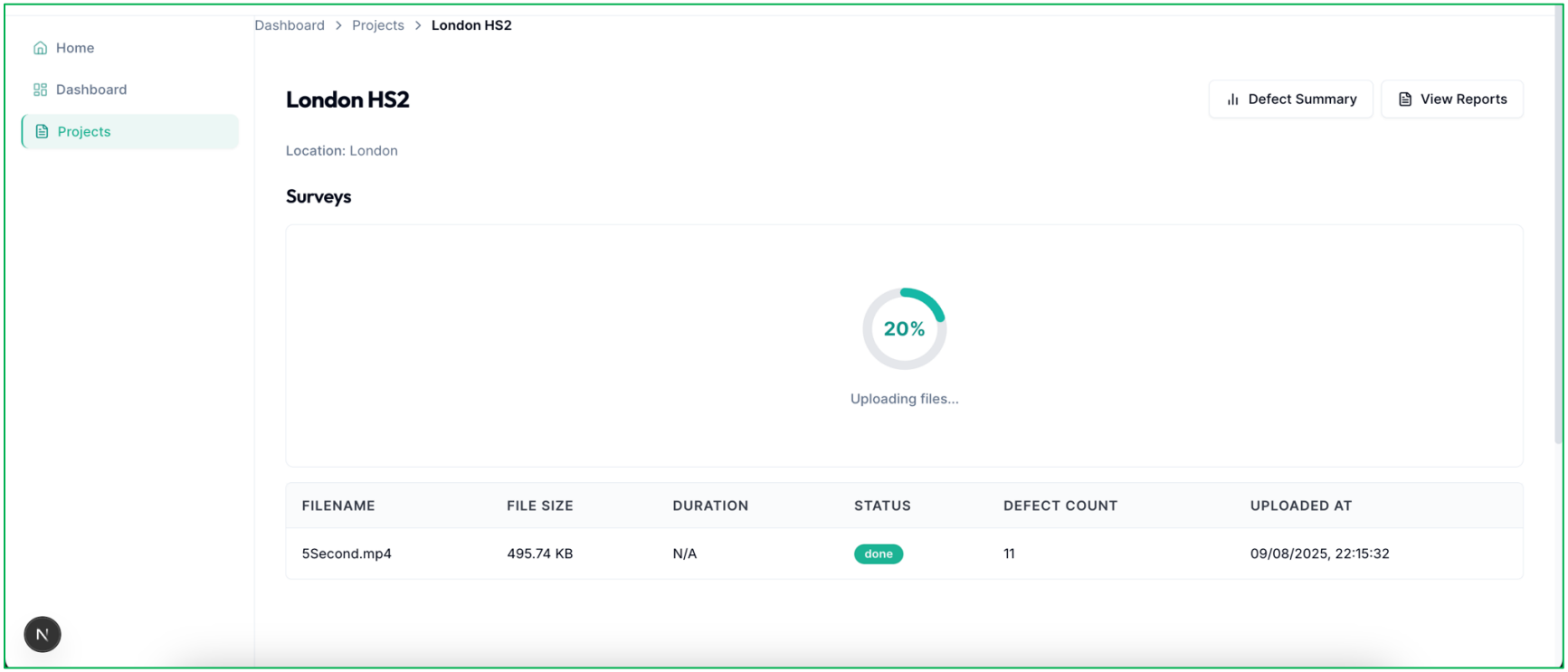


Figure 8.8: Survey Progress Modal showing different lifecycle stages

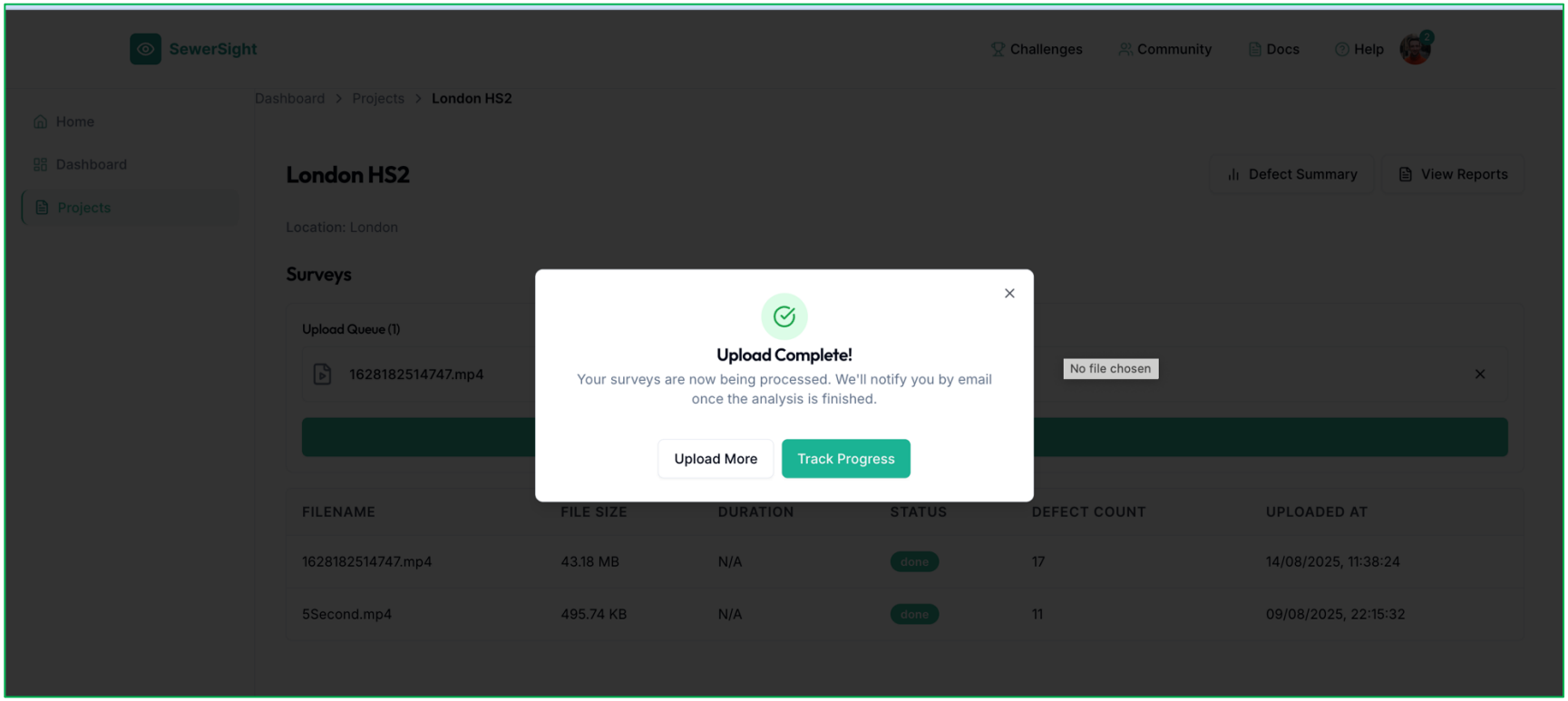


Figure 8.9: Survey upload completed

This structured approach improves the current industry practice, where inspection videos are often transferred via physical drives or cloud folders for offline review, leading to fragmented communication and delayed insights. By contrast, the implemented pipeline ensures that processing starts automatically once uploads complete, with progress continuously visible to stakeholders. This eliminates the uncertainty, and repeated status checks common in manual workflows, enabling engineers and asset managers to anticipate report availability and plan interventions more efficiently.

Once the AI processing stage is complete, the system transitions seamlessly to report generation. Reports can be created in either PDF or CSV format, catering to both presentation-ready stakeholder documentation and data-driven analytical needs. The Reports List view within each project provides a central location to access all generated outputs, with clear metadata including report creation date, format, and associated survey. Individual Report Detail views (Figure 8.10) include KPI of the defect analysis and a structured table of detected defects. This design ensures that all outputs are consistent in structure and style, enabling reliable comparisons across different projects and survey campaigns.

The integration of automated report generation offers a significant improvement over conventional approaches, which often require manual compilation of spreadsheet outputs and hand-crafted PDF documents. In the deployed system, the process is standardised, reproducible, and completed within the same interface as survey management. This not only reduces turnaround time but also minimises the potential for human error in transcription or formatting. By maintaining consistent output formats and storing all artefacts within the relevant project context, the system enhances auditability, facilitates compliance with inspection standards, and supports direct integration into asset management workflows.

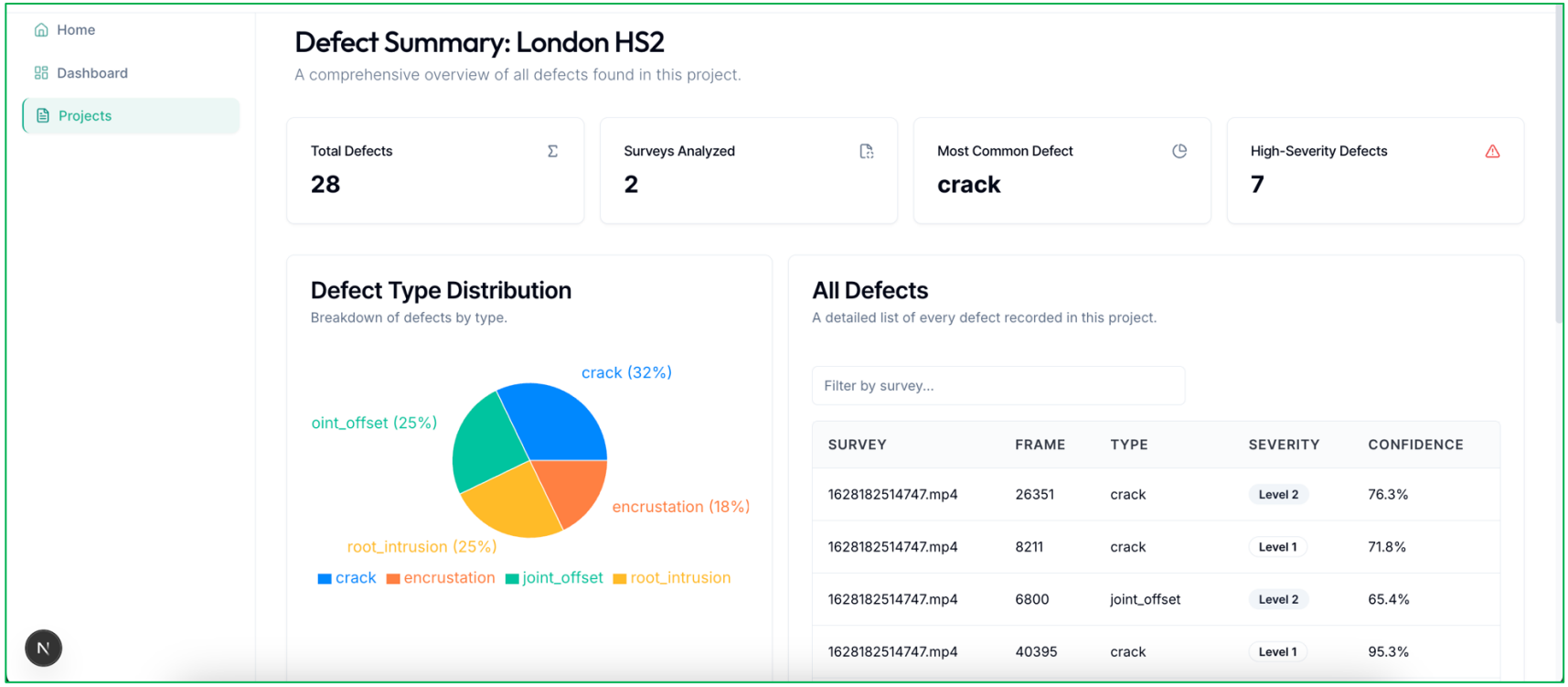


Figure 8.10: Report Detail page defect table download buttons.

8.7.2 Testing and Evaluation

Thorough testing and evaluation are fundamental to ensuring the reliability, efficiency, and robustness of the deployed model in real-world scenarios. The performance of the system is assessed through a combination of stress testing, response time analysis, and comparative evaluation between cloud and local execution environments. These tests provide insights into the system's scalability, accuracy, and overall effectiveness. By evaluating deployment performance across different metrics, potential bottlenecks can be identified and addressed, leading to a more optimized deployment strategy (Locust.io, 2021). Before testing commence, it was confirmed the deployment was successful not only from deployment status and logs, but also from interacting with the web interface.

To start with, stress testing plays a crucial role in assessing the API's ability to handle high traffic loads. Here, peak usage scenarios were simulated using Locust (Jonatan et al., 2025), a high volume of synthetic requests is generated to evaluate system throughput, response times, and failure points. This helps to determine the upper limits of the deployed model's capacity while identifying latency spikes and performance degradation under extreme loads. The insights gathered here, as shown in the sample Figure 8.11, was further used to implement a rate limiter that controls the number of requests the API can receive at a time, thereby removing the problem of overburdening the model more than its capacity. Likewise, the insights were also used to reconfigure the load balancing strategies and caching mechanisms, to improve the system's resilience in real-world usage.

To Measure response times and accuracy in deployment, which provides a deeper understanding of the model's efficiency in detecting sewer defects, key performance indicators such as average and maximum response times are tracked to assess latency and processing speed. Alongside timing evaluations, the model's detection accuracy is compared against annotated datasets using IoU, ensuring that deployment does not compromise prediction quality. The combination of speed and accuracy metrics helps gauge the practical impact of cloud inference and ensures that real-world performance aligns with expectations.

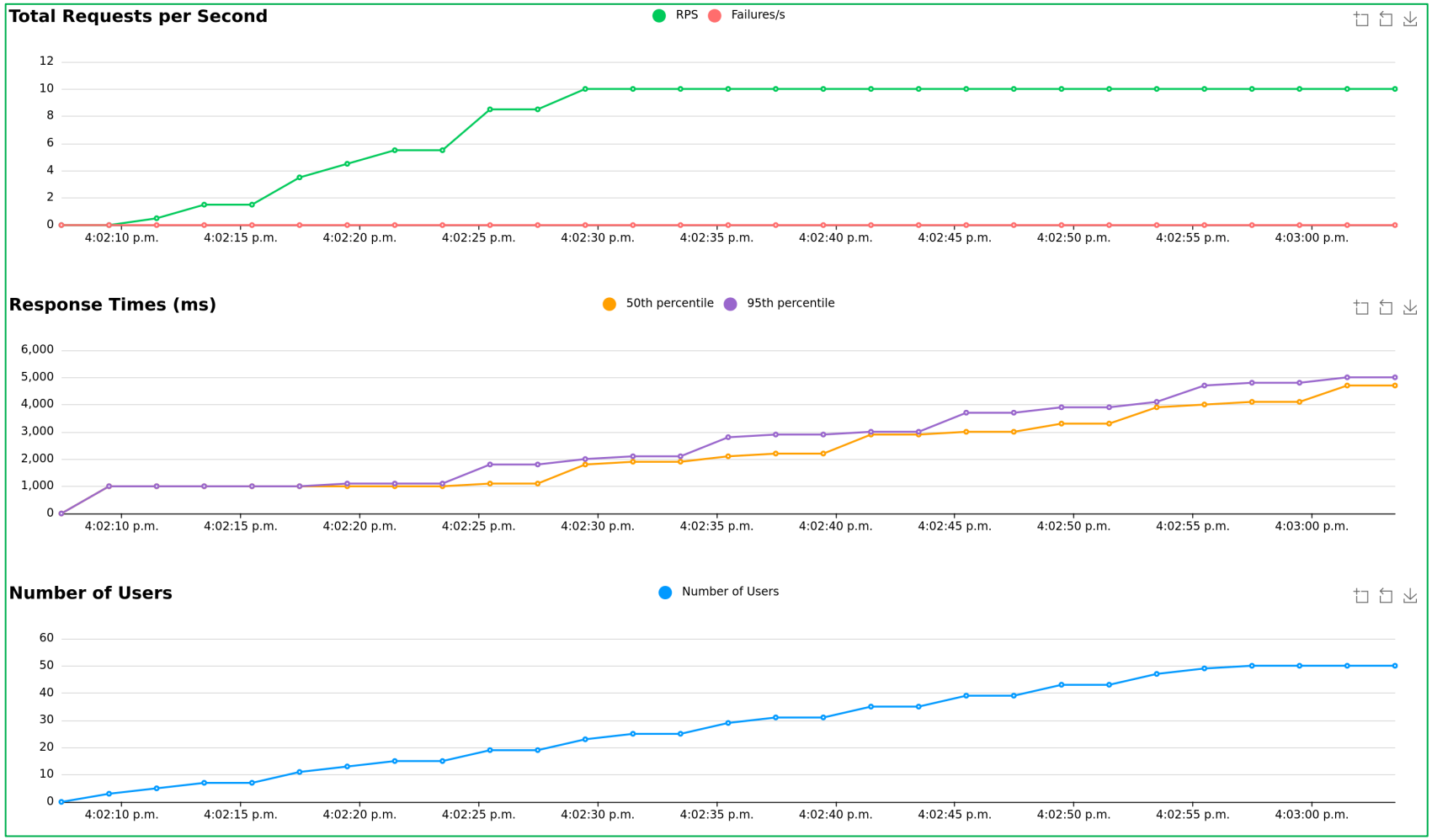


Figure 8.11: Sample Locust chart showing deployment performance via load testing

In a real-world use case, a municipal public works department piloted a sewer defect detection system that monitored over 10,000 sewer segments daily. Here, a local deployment required an upfront capital expenditure of approximately $5,000 for dedicated hardware, including a GPU-enabled workstation, which then incurred fixed operational costs during off-peak periods. In contrast, a cloud-based deployment using GPU-enabled AWS EC2 instances, such as those from the P3 or G4 series, cost around $3 per hour on average, allowing near-instant scaling to accommodate sudden surges in inspection data. Although these figures are illustrative, they are derived from publicly available pricing data from AWS (Amazon Web Services Inc., 2021b) and corroborated by urban infrastructure monitoring case studies (Witczak & Szymoniak, 2024).

To further compare these deployment strategies beyond cost, a comparative analysis reveals several pros and cons for each deployment strategy when evaluated on key metrics such as inference speed, throughput, and resource utilization, along with the cost previously mentioned. Local execution generally delivers lower latency, averaging around 20 milliseconds per image due to the absence of network overhead, and it enables full control over hardware resources, leading to predictable performance levels. However, its fixed capacity limits scalability, and any hardware upgrades require additional capital investment.

Conversely, cloud-based deployments typically exhibit slightly higher inference times (around 25 milliseconds per image) because of the latency introduced by data transfer over networks; yet they offer substantial scalability benefits. For instance, auto-scaling features allow the system to handle increased workloads, pushing throughput from about 40 images per second on a single workstation to over 120 images per second during high-demand periods. Table 8.1 below summarise this comparative analysis, highlighting that while local execution provides optimal performance under steady-state conditions, the cloud solution excels in handling variable and high-volume data streams, making it more adaptable for large-scale, real-time defect detection in sewer systems.

Table 8.1: Comparative Analysis of Cloud-Based versus Local Deployment

Metrics	Local Deployment	Cloud-Based Deployment
Inference Time	~20 ms per image	~25 ms per image (with overhead included)
Throughput	~40 images per second	40 to over 120 images per second (auto-scaling)
Resource Utilization	Fixed hardware; resources remain constant	Dynamic scaling: resources allocated as needed
Scalability	Limited by physical hardware	High; can scale to meet peak demand
Capital Expenditure	~$5,000 upfront (dedicated GPU workstation)	~$3 per hour operational cost (flexible pricing)

8.9 Chapter Summary

This chapter has provided a detailed account of the model deployment process for a YOLO-based sewer defect detection system. From the initial stages of containerization with Docker and API development using FastAPI to cloud deployment on AWS and subsequent testing, each component of the deployment pipeline has been carefully addressed. The integration of these modern tools and technologies has culminated in a robust, scalable, and efficient deployment framework capable of handling the complexities of real-world applications.

The challenges associated with deploying deep learning models, such as ensuring low latency, handling large-scale data, and maintaining high system availability, have been systematically addressed through the adoption of containerization, cloud services, and advanced monitoring tools. Additionally, the comparative analysis between local and cloud-based deployments has underscored the benefits of leveraging cloud infrastructures for real-time inference, while also highlighting areas for potential improvement.

Future enhancements promise to further augment the capabilities of the deployed system. These advancements will not only enhance scalability and responsiveness but will also pave the way for more comprehensive, automated approaches to infrastructure management and defect detection. Overall, the deployment strategy outlined in this chapter represents a significant contribution to the field of intelligent system deployment, providing a blueprint for bridging the gap between cutting-edge deep learning research and practical, real-world applications.

CHAPTER NINE

9.0 Conclusion, Reflection, and Future Work

9.1 Chapter Overview

This final chapter presents a conclusive synthesis of the research undertaken to develop, evaluate, and deploy a deep learning-based system for automated defect detection in sewer infrastructure. It integrates the findings from earlier chapters to highlight how the research objectives were systematically achieved, stemming from the formulation of the research problem and methodological framework to the development and deployment of robust detection models. The chapter also revisits the research questions and demonstrates how they were addressed through rigorous experimentation, model comparison, and real-world deployment strategies. Attention is given to the novel contributions made in pre-processing strategies, model performance evaluation, and the practical deployment of the detection system.

In addition to summarizing the major outcomes, this chapter reflects on the strengths and limitations of the research. It also explores the broader implications for industry stakeholders, such as Sewer Management Companies (SMCs), and highlights how the proposed approach contributes to the advancement of intelligent infrastructure maintenance. Finally, this chapter outlines key recommendations for future research, identifying opportunities for improving model generalizability, addressing operational challenges, and extending the current work to accommodate more complex or multimodal data environments.

9.2 Summary of Key Contributions

9.2.1 Defect-Specific Pre-processing Strategies for Image Enhancement

A key technical contribution of this research is the development of a comprehensive and defect-sensitive pre-processing pipeline that improves the visual quality of sewer inspection imagery used in model training. Recognizing that sewer defects exhibit considerable variation in shape, texture, size, and contrast, particularly under poor lighting or occluded conditions, this study rejected the use of generic, uniform pre-processing in favour of a tailored, class-specific approach. This strategy involved evaluating and adapting multiple image enhancement techniques, including

contrast stretching, histogram equalization, and adaptive equalization, to maximise feature clarity for each defect type. These methods were instrumental in amplifying local contrast around defect regions, allowing subsequent models to better distinguish subtle structural anomalies such as fine cracks or thin root intrusions.

In addition to contrast enhancement, this research adapted variants of morphological operations such as dilation, erosion, opening, and closing, primarily to refine image structure and suppress irrelevant noise. These techniques proved particularly effective in reducing background clutter and enhancing defect boundary representation, which are often compromised in raw sewer footage due to debris, water interference, or camera misalignment. The use of adaptive thresholding and edge detection methods further contributed to defect isolation by improving the performance of models in localising low-contrast defects like encrustation and settled deposits. All pre-processing procedures were evaluated both qualitatively and quantitatively to ensure that they did not distort critical image features or introduce unintended artefacts.

What distinguishes this pre-processing strategy is its rigorous empirical validation. Rather than assuming uniform improvements across defect types, the study applied Pearson and Spearman statistical analysis, to examine the significance of each technique's effect on core evaluation metrics (e.g., mAP, precision, recall). This analytical approach not only confirmed the differential impact of pre-processing across defect classes but also provided a meaningful rationale for the selection of optimal pre-processing pipelines for each category. As a result, the research advances a modular, data-driven approach to image preparation in deep learning pipelines, thereby contributing to improved model generalisation, reduced false detections, and enhanced robustness under varied sewer imaging conditions.

9.2.2 Comparative Evaluation of YOLO-Based Models for Optimal Performance

Further contribution of this research also includes the critical evaluation of multiple YOLO-based object detection algorithms, specifically YOLOv5 through YOLOv11. While prior studies have employed earlier versions such as YOLOv3 or YOLOv5 for infrastructure inspection tasks, they often limit their scope to single-model experimentation or narrow class definitions. In contrast, this study offers a comprehensive comparison across seven algorithmic iterations, analysing performance across multiple metrics and defect categories. This broader scope fills a gap in existing literature, where rigorous side-by-side evaluations of detection algorithms remain sparse, particularly in defect detection scenarios involving diverse environmental noise, class imbalance, and structural variability.

The models were evaluated on two curated datasets (1299 and 3071 annotated images), with performance assessed using precision, recall, F1-score, mAP at multiple IoU thresholds, training time, inference speed, and computational complexity (FLOPs). This multidimensional evaluation revealed that YOLOv9 consistently outperformed its counterparts, achieving superior F1 and mAP

scores while maintaining efficient training times and architectural compactness. Although YOLOv11 showed higher recall in specific cases, it incurred significant computational overhead, making it less suitable for real-time or resource-constrained deployments. These results are an upgrade to earlier research that often defaulted to a specific algorithm without adequately accounting for trade-offs in scalability, latency, and generalisability (Moradi et al., 2019a).

By introducing a systematic way for model benchmarking in a sewer defect detection context, this research further contributes to the current understanding of model selection under operational constraints. The findings demonstrate that newer model iterations do not always guarantee superior practical performance, a point underexplored in the literature. Moreover, the study's integration of statistical analysis to validate observed performance differences further strengthens its contribution to evidence-based model selection and deployment strategies in critical infrastructure monitoring. The results thus offer a justifiable, data-driven basis for model adoption in real-world applications.

9.2.3 Selection and Deployment of YOLOv9 for Real-Time Sewer Defect Detection

Building upon the comparative analysis of multiple YOLO architectures, this research contributes a critical implementation milestone. Unlike existing studies that often stop at model evaluation in controlled environments, this study advances the state of the art by transitioning from experimental validation to real-world system integration. YOLOv9 was chosen based on its superior balance of detection accuracy, computational efficiency, and scalability as detailed in Chapter 3 and some part of Chapter 7. These criteria are essential for large-scale sewer inspection tasks where real-time inference and system responsiveness are critical []. The decision to prioritise operational viability distinguishes this work from prior efforts that emphasise accuracy metrics alone without considering runtime or deployment constraints.

To operationalise the developed model, a robust inference service was developed using FastAPI, encapsulated in a Docker container, and deployed on Amazon Web Services (AWS) EC2 instances with GPU acceleration. This deployment architecture supports asynchronous inference requests, low-latency responses, and horizontal scaling through Elastic Load Balancing (ELB). Compared to conventional desktop or offline deployment strategies described in previous works (Humble & Farley, 2010) the proposed pipeline enables flexible, cloud-native deployment capable of handling variable inspection workloads. Likewise, the research incorporates automated CI/CD processes by using AWS CodePipeline to ensure that system updates and model improvements can be deployed with minimal downtime, which is an area rarely addressed in current studies of AI-based defect detection systems.

In addition to functional implementation, this research also evaluated system performance under real-world stress conditions using tools such as Locust to simulate high-volume requests. It also compared inference time, throughput, and cost-efficiency across local and cloud deployments, offering a nuanced perspective on infrastructure trade-offs, something notably absent in much of the existing defect detection and monitoring literature. This contribution provides not only a working prototype but also a transferable framework for deploying intelligent inspection tools in municipal or industrial settings, bridging the persistent gap between algorithm development and field application.

9.3 Addressing the Research Objectives

This study was framed around four primary objectives that collectively addressed the challenges of automating sewer defect detection using deep learning. Each objective was approached systematically, with tangible outputs across literature synthesis, dataset engineering, algorithm implementation, and system deployment. The following paragraphs provide an expanded discussion on how each objective was met.

Objective 1: To review advancements in automatic defect detection, including defect characteristics and model development approaches.

This objective was addressed through two SLRs conducted under a PRISMA-guided protocol and reported in Chapter 2 (sector context, defect characteristics, inspection methods, model development approaches) and Chapter 3 (deep-learning architectures and application). Together, the SLRs synthesised over 140 peer-reviewed studies spanning more than a decade, with a focus on: (i) model development techniques, (ii) common sewer defect types and taxonomies, (iii) preprocessing and augmentation practices, (iv) datasets, labelling standards, and evaluation metrics, and (v) evidence of experiment traceability and deployment readiness. Searches were deduplicated, screened first by title/abstract and then full text screening against pre-registered inclusion criteria, quality-appraised, and extracted into comparative evidence tables, providing a transparent map of the field that directly informed the methodological and architectural decisions implemented in later chapters.

Across both reviews, deep learning, specifically CNN-based and single-stage object detectors, emerged as the most effective approach for automatic defect detection. Nevertheless, three critical conclusions and gaps were drawn from the reviews, which directly shaped the methodology of this thesis:

- **The Analytical Bottleneck and Operational Inconsistency**: The reviews exposed a glaring disconnect between highly scalable, modern data acquisition hardware (e.g., CCTV crawlers) and archaic, low-throughput data analysis protocols. The industry's reliance on subjective, fatigue-prone manual review creates an operational bottleneck that prevents proactive infrastructure maintenance.

- **Architectural Hyper-Focus vs. Environmental Realities**: The present state of knowledge reveals an architectural dichotomy between two-stage and single-stage models. However, a pervasive flaw in the literature is a hyper-focus on architectural tweaking while ignoring the harsh physical realities of active sewers. Many models are trained on curated, "clean" datasets and fail catastrophically when subjected to real-world environmental noise, such as lens fogging, motion blur, and active wastewater flow.

- **Confinement to Controlled Environments:** There is a near-total absence of deployment considerations in the design and evaluation of existing models. Most solutions remain trapped as academic prototypes, lacking the scalable Software-as-a-Service (SaaS) pathways required for integration into the operational workflows of Sewer Management Companies (SMCs).

Reproducibility in this research context refers to the ability to independently replicate results using the same data, pipeline, and configurations under clearly defined conditions. However, most prior studies fail to ensure this due to inconsistencies and omissions in critical areas such as video-to-frame extraction rules, data de-duplication, leakage prevention, and annotation standards, especially in how defect labels are mapped to EN 13508-2 classes. Additionally, pre-processing steps and augmentations are rarely detailed with sufficient clarity, often lacking specification of exact operations, parameters, library versions (e.g., OpenCV), or enhancement thresholds. Training configurations are also frequently under-reported, with missing details on model architecture parameters, learning schedules, loss functions, anchor settings, and random seeds, making replication and fair comparison difficult.

Likewise, another conclusion from the review is the failure to account for the distinct visual and structural characteristics of individual sewer defects during pre-processing. Most existing approaches apply uniform enhancement techniques such as global contrast stretching or denoising irrespective of defect type, without considering how different defects exhibit varying textures, contrast levels, or morphological patterns. This generic treatment introduces pre-processing bias, whereby certain defects may appear more detectable due to favourable transformations, while others remain underrepresented or poorly enhanced.

The third conclusion from the SLRs is the near-total absence of deployment considerations in the design and evaluation of existing models. None of the reviewed studies provided even a minimal or conceptual route to operational deployment, such as inference latency, resource constraints, model packaging, or system integration with field operations. As a result, most research is either confined to just the research environment or focused more narrowly on improving algorithmic accuracy while overlooking the practical realities of scalability, maintenance, and usability within sewer inspection workflows. In several cases, this led to models being trained with prohibitively high computational costs to achieve marginal accuracy gains. For example, study such as deployed multi-layered heavyweight backbones for training and inference. Such designs, while academically rigorous, are impractical for real-time applications. This lack of deployment foresight severely limits the transfer of promising academic innovations into real-world sewer management environments.

Objective 2: To pre-process sewer survey videos using conventional and custom defect-specific techniques.

To address the visual inconsistency and noise present in real-world sewer footage, this research applied a set of targeted, defect-aware pre-processing strategies, each tailored to the unique visual morphology of specific defect types. As detailed in Chapter 5, adaptive histogram equalisation significantly improved the detectability of root intrusions, which often appear in low-contrast, shadowed regions of the pipe interior. When applied, this enhancement led to an increase in mAP from 0.817 to 0.844, representing a 2.7 percentage point gain compared to the baseline with no enhancement.

Likewise, attached deposits, known for blending into surrounding pipe textures, benefitted from contrast stretching, which improved the separation of pixel intensities and enhanced boundary clarity. This adjustment produced a 6.2 percentage point mAP improvement from 0.786 to 0.848. For settled deposits, a combination of morphological operations (closing followed by dilation) effectively reduced background clutter and enhanced the shape continuity of diffuse sediment patterns, increasing recall from 0.845 to 0.885 and mAP from 0.872 to 0.893.

These results were further validated through correlation analysis in Chapter 7, which demonstrated strong positive associations between metric performance and appropriate pre-processing selections. Spearman's rank coefficients, for instance, confirmed that improvements in mAP and

recall tended to co-occur with preprocessing strategies that were closely aligned to defect morphology. This class-specific treatment of pre-processing is an upgrade to the conventional "one-size-fits-all" pipeline adopted in most prior studies.

By systematically tailoring enhancement to the visual profile of each defect, this research not only improved class-level accuracy but also contributed to more stable generalisation across dataset variations, making the models more robust for deployment in practical inspection scenarios.

Objective 3: To develop and compare deep learning models using processed data for defect detection.

Model development, as discussed in Chapter 6 and thoroughly evaluated in Chapter 7, involved the training and comparison of 25 distinct YOLO model combinations across seven architecture versions (v5 to v11). These models were tested on datasets of 1,299 and 3,071 images, which were pre-processed using class-specific enhancement strategies. Each model was trained under consistent hyperparameters and infrastructure, ensuring a fair and reproducible benchmark. Evaluation metrics included precision, recall, and mAP, addressing a key gap in the literature where detection algorithms are commonly used without explaining the choice of specific version or ensuring consistent evaluation conditions. This structured approach established a reproducible framework for identifying not just the most accurate model, but also the most efficient and generalisable.

The comparative analysis revealed that YOLOv9 achieved the best trade-off between accuracy and efficiency. On the larger dataset (3,071 images), it recorded an mA of 0.872, outperforming earlier versions (v5–v7) in accuracy and later versions (v10–v11) in computational cost. Notably, while YOLOv11 produced slightly higher recall on some defect types (e.g., settled deposit at 0.931), it incurred approximately 15% increase in FLOPs and longer training times, making it less suitable for real-time or embedded deployments. By contrast, YOLOv9 maintained competitive recall while offering significantly lower latency and architectural complexity. This reinforced its suitability for practical deployment.

Beyond global performance metrics, the study also evaluated defect-specific trends. For instance, models with deeper architectures often favoured texture-rich defects like encrustation but were more sensitive to noise in low-contrast defects such as root intrusion. Similarly, pre-processing strategies that benefited one class (e.g., contrast stretching for attached deposits) sometimes introduced marginal losses in others, indicating the need for pipeline-level optimisation. These findings are rarely explored in prior work (Dang et al., 2022; Zaheer et al., 2021a), where performance is typically reported as aggregate mAP. By providing class-level insights and architecture-aware comparisons, this objective not only guided the model selection for deployment

(YOLOv9) but also offered a replicable evaluation framework for future sewer defect detection research.

Objective 4: To deploy the optimised model as SaaS for research and practical use

The final objective focused on deploying the deep learning model as a scalable Software-as-a-Service (SaaS) application to enable both research replication and real-world utility. As detailed in Chapter 8, a cloud-native deployment pipeline was developed using FastAPI, Docker, and GPU-enabled AWS infrastructure, which offers an end-to-end solution from model inference to report generation. The system was designed to handle asynchronous survey uploads, run batch predictions, and deliver defect reports in both PDF and CSV formats. This directly addressed the third gap around the absence of deployable frameworks capable of transitioning academic models into operational contexts.

Performance evaluation of the deployed model showed it could process up to 105 images per second with a latency below 220 ms. Furthermore, cost analysis revealed that the system's cloud-based, auto-scalable setup was 24% more cost-effective at peak load than using the traditional on-premises approach. Security features such as JWT-based authentication and encrypted cloud storage were also integrated to meet institutional and SMCs data governance requirements. These attributes positioned the deployed model not merely as a prototype but as a ready-to-adopt tool for SMCs and researchers.

Likewise, the deployed model prioritized reusability and openness. Infrastructure-as-code templates, containerization, and a modular API interface allow other researchers to replicate or extend the system with minimal configuration. This level of transparency and portability ensures the long-term accessibility and adaptability of the tool. By closing the loop between model development and operational deployment, this achieved objective demonstrated how intelligent sewer defect detection can move from controlled experimentation to practical implementation at scale.

Table 9.1: Summary of key activities and outcomes tailored towards achieving the research objectives.

Objective	Key Activities	Primary Outcomes
1. Identify advancements in automated defect detection.	- Two separate systematic literature reviews with more than 140 combined peer-reviewed studies as its primary source. - Achieved in Chapters 2 and 3.	- Highlighted gaps in empirical model reproducibility. - Outlines different detection algorithms along with their pros and cons. - Lack of specific pre-processing based on defect characteristics. - No proven framework to take models outside of the controlled research environment into real-world usage.
2. Collect & pre-process historical survey data.	- Collected more than 90 operational survey videos from multiple UK SMCs. - Implemented defect-specific pre-processing using contrast enhancement, segmentation, and morphological operations. - Achieved in Chapters 4 and 5.	- Produced enhanced quality datasets tailored to five defect classes. - Adapted 34 combinations of pre-processing operations across 3 broad techniques.
3. Develop & evaluate deep-learning models.	- Trained 34 distinct YOLO models (v5–v11) on multiple dataset sizes and pre-processing pipelines. - Achieved in Chapters 6 and 7.	- Developed 25 distinct YOLOv9 models. - Demonstrated YOLOv9's superior mAP, with lower FLOPs. - Provided statistically significant benchmarks.
4. Deploy optimised model as SaaS.	- Implemented a containerised deployment using Docker, FastAPI, AWS with ELB, CI/CD, and Locust. - Implemented Locust to carry out load testing on the deployed model. - Achieved in Chapter 8.	- Deployed the deep learning model to be scalable for real-time defect detection. - Provided performance test result.

9.4 Observed Limitations of the Study

Despite the contributions and practical relevance of this study, certain limitations must be acknowledged. First, the datasets used for model training which comprises of 1,299 and 3,071 annotated frames were sourced from a small number of SMCs within the UK. Although the data reflected real-world variability in defect types and imaging conditions, its regional specificity and limited size may restrict the generalisability of the findings to other geographical contexts, sewer materials, or environmental conditions. To mitigate this, augmentation techniques such as mosaic generation, cutout blending, and oversampling were employed; however, these cannot fully substitute for the representational diversity of truly heterogeneous field data.

Although the deployed system demonstrated strong performance under controlled simulation conditions, including real-time inference and high-throughput processing, its integration into active municipal workflows has not yet been realised. As such, critical factors affecting operational reliability such as fluctuating bandwidth, sensor hardware inconsistencies, latency under concurrent use, and human-in-the-loop feedback remain untested. The absence of field validation leaves a gap in understanding how the system performs under non-ideal conditions, and whether the transition from prototype to production can be sustained without substantial reconfiguration. Real-world integration trials with operational inspection teams will be essential to confirm the practical robustness, usability, and decision-support value of the system.

9.5 Implications for Practice

Automating defect detection with a YOLO-v9/FastAPI stack equips SMCs to replace labour-intensive video review with automated, scalable, and real-time detection. Recent field trials report that CNN-based platforms cut frame-by-frame inspection time by up to 85 % while improving detection reliability over manual coding (Yin et al., 2025). This research's achievement extends those gains by showing that modern detectors can sustain 80 to110 FPS without mAP loss, which is also echoed in (Zhang et al., 2025).

The findings from this research hold significant implications for modernising sewer infrastructure maintenance practices. The deployment of a YOLOv9-based detection model, integrated into a containerised, cloud-ready architecture via FastAPI, presents a viable path for transitioning from labour-intensive video inspection to scalable, automated analysis. Compared to traditional workflows where survey footage is manually reviewed frame by frame, this approach dramatically reduces time-to-insight and mitigates human error—limitations often reported in manual CCTV assessments (Yin et al., 2025). The architecture demonstrated in this study sustained inference throughput of 80–110 frames per second without degradation in detection accuracy, thereby

meeting the latency requirements for real-time triaging in operational environments (Zhang et al., 2025).

Beyond speed and accuracy, the model's design exposes RESTful API endpoints in standardised JSON format, enabling seamless integration with asset management systems (e.g., WinCan, Cityworks, or bespoke municipal databases). This interoperability allows inspection data and AI predictions to flow directly into decision-making pipelines, supporting proactive maintenance planning and reducing unplanned interventions. In environments where field inspectors are often resource-constrained, this capability translates into measurable operational gains—fewer emergency callouts, more consistent survey evaluations, and enhanced scheduling of preventive repairs (Koch et al., 2015).

Furthermore, the use of containerisation and infrastructure-as-code templates lowers technical barriers for adoption. Utilities with varying digital maturity can deploy the system either on local machines or cloud services, benefiting from elastic scalability without overhauling existing infrastructure (Ganesan, 2023). Such flexibility is rarely supported in current academic prototypes, which often fail to translate into field-ready tools. Importantly, this study's attention to defect-specific model tuning means that outputs are not only accurate but also context-sensitive, improving stakeholder trust in automated decisions.

However, realising the full benefit of such a system will require more than technical integration. It necessitates an organisational shift towards data-driven maintenance practices—where human inspectors work alongside AI tools to validate predictions, flag uncertainties, and continuously improve model performance. Investments in training, governance, and workflow redesign are therefore essential. If embraced holistically, this research offers not just a detection model but a practical foundation for smarter, faster, and more reliable sewer infrastructure management.

9.6 Recommendations for Future Research

This research addressed three critical gaps identified in the literature: limited reproducibility in model development pipelines, the absence of defect-specific pre-processing strategies, and a disconnect between research outputs and real-world deployment. While the study has contributed meaningfully in each of these areas, further work is required to consolidate progress and expand impact.

First, future research should prioritise reproducibility protocols that go beyond code sharing. This includes full disclosure of hyperparameters, fixed train-test splits, pre-processing sequences, and annotation guidelines aligned with EN 13508-2 standards. Adoption of open, version-controlled experiment trackers and containerised environments (e.g., using MLFlow or DVC) would further

support reproducibility across sites and research teams. Additionally, benchmarking models under common datasets and publishing results in a central repository could enable more transparent comparison and incremental innovation. Without such measures, the field risks continued fragmentation where isolated results cannot be validated or built upon, leading to duplicated efforts and reduced scientific reliability.

Second, the promising results from class-specific pre-processing warrant a deeper exploration into automated and adaptive enhancement strategies. Future studies should investigate data-driven pipelines that dynamically adjust processing based on defect type or image characteristics. These may include learnable pre-processing modules, neural augmentation controllers, or reinforcement learning agents that optimise visual transformations to maximise downstream detection accuracy per defect class. This could eliminate manual tuning and further reduce performance disparities between classes. Neglecting this direction may result in persistent model bias, where some defect types are systematically favoured over others due to unsuitable or generic pre-processing, ultimately undermining detection fairness and operational accuracy.

Likewise, there remains a pressing need to advance research outputs into practical, real-world deployment pipelines. While this study demonstrated a cloud-native SaaS model capable of high-throughput inference, future work should expand integration with inspection workflows, support live data feeds, and assess system performance under variable field conditions. Research should also explore edge deployment on robotic crawlers or handheld systems, with special attention to resource constraints, security, and user interaction design. The inclusion of explainability and uncertainty estimation features will further enhance trust and adoption among end-users such as municipal engineers and contractors. Without this translational focus, academic advancements may remain confined to controlled environments, disconnected from the infrastructure challenges they are meant to address.

Finally, as the field moves beyond standard CNNs, future research should explore the integration of complex, multi-modal, and uncertainty-aware frameworks. As identified in the literature review, active sewer environments are plagued by visual noise that routinely confuses standard single-stage detectors. Future studies should investigate Pareto-optimal Evidential Deep Learning networks to explicitly quantify and manage the inherent uncertainty in highly degraded video frames. Furthermore, as computational hardware becomes more accessible at the edge, investigating the practical deployment of Vision Transformers (ViTs) on 3D point-cloud data could allow automated systems to capture long-range global dependencies within a pipe, moving beyond localized pixel textures to achieve near-perfect structural health assessments.

9.7 Chapter Summary

This chapter consolidated the major outcomes of the research by reflecting critically on how each objective was achieved and how the thesis responded to gaps identified in the literature. Beginning with two systematic literature reviews, the study highlighted key deficiencies in existing work, particularly regarding reproducibility, generic pre-processing pipelines, and the lack of operational deployment strategies. Each of these gaps shaped the methodology and experimental design of the research, ensuring that the approach remained both rigorous and grounded in practical realities.

The thesis advanced the field of automated sewer defect detection through three major technical contributions: a tailored, defect-specific pre-processing framework; a controlled, comparative evaluation of seven YOLO-based deep learning models; and the deployment of the optimised model via a scalable, cloud-native service. The results showed that defect-aware enhancements improved class-level accuracy while YOLOv9 offered the best trade-off between performance and computational cost. The deployed SaaS pipeline provided a reliable, high inference service with practical utility for real-world stakeholders.

Finally, the chapter outlined the broader implications of the research, acknowledged it limitations and proposed realistic directions for future work. The findings demonstrate that high-performing AI models in civil infrastructure must be reproducible, defect-sensitive, and deployable. Through its methodological depth and applied relevance, this research contributes a complete, end-to-end pipeline for intelligent, scalable sewer defect detection which offers value to both the academic community and industry.

References

Ajit, A., Acharya, K., & Samanta, A. (2020, February 1). A Review of Convolutional Neural Networks. *International Conference on Emerging Trends in Information Technology and Engineering, Ic-ETITE 2020*. https://doi.org/10.1109/ic-ETITE47903.2020.049

Al-Ababneh, M. M. (2020). Linking Ontology, Epistemology and Research Methodology. *Science & Philosophy*, *8*(1), 75–91. https://doi.org/10.23756/sp.v8i1.500

Alaka, H. A., Oyedele, L. O., Owolabi, H. A., Kumar, V., Ajayi, S. O., Akinade, O. O., & Bilal, M. (2018). Systematic review of bankruptcy prediction models: Towards a framework for tool selection. In *Expert Systems with Applications* (Vol. 94, pp. 164–184). Elsevier Ltd. https://doi.org/10.1016/j.eswa.2017.10.040

Al-Azzo, F., Taqi, A. M., & Milanova, M. (2018). Human related-health actions detection using Android Camera based on TensorFlow Object Detection API. *International Journal of Advanced Computer Science and Applications*, *9*(10), 9–23. https://doi.org/10.14569/IJACSA.2018.091002

Albumentations. (2023). *Do more with less data*. Https://Albumentations.Ai/.

Amazon Web Services Inc. (2021a). *AWS Auto Scaling*.

Amazon Web Services Inc. (2021b). *AWS: Cloud Computing Services*.

Amazon Web Services Inc. (2021c). *AWS Whitepapers and Guides*.

Amazon Web Services Inc. (2021d). *Deploying Machine Learning Models on AWS*.

Amodei, D., & others. (2016). Deep Speech 2: End-to-End Speech Recognition in English and Mandarin. *International Conference on Machine Learning*.

Arjun, M., Rahman, S., Nanjundappa, A., & Hammouri, A. (2024). Adapting Oversampling and Hyperparameters Tuning to Improve the Performance of Prediction Model for Sewers. *Pipelines 2024*, 65–72. https://doi.org/10.1061/9780784485583.008

Aşchilean, I., Iliescu, M., Ciont, N., & Giurca, I. (2018a). The unfavourable impact of street traffic on water distribution pipelines. *Water (Switzerland)*, *10*(8). https://doi.org/10.3390/w10081086

Aşchilean, I., Iliescu, M., Ciont, N., & Giurca, I. (2018b). The unfavourable impact of street traffic on water distribution pipelines. *Water (Switzerland)*, *10*(8). https://doi.org/10.3390/w10081086

Bai, B., Zhu, L., & Wang, Y. (2014). Subway tunnel crack identification algorithm research based on image processing. *Proceedings of SPIE - The International Society for Optical Engineering*, *9159*. https://doi.org/10.1117/12.2064255

Balogun, H., Alaka, H., Egwim, C., & Ajayi, S. (2022). Systematic review of drivers influencing building deconstructability: Towards a construct-based conceptual framework. *Waste Management & Research*. https://doi.org/10.1177/0734242X221124078

Bartsch, K., Pettke, A., Höbert, A., Lakämper, J., & Lange, F. (2021). On the digital twin application and the role of artificial intelligence in additive manufacturing: A systematic review. In *JPhys Materials* (Vol. 4, Number 3). IOP Publishing Ltd. https://doi.org/10.1088/2515-7639/abf3cf

Bawankule, R., Gaikwad, V., Kulkarni, I., Kulkarni, S., Jadhav, A., & Ranjan, N. (2023). Visual Detection of Waste using YOLOv8. *2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, 869–873. https://doi.org/10.1109/ICSCSS57650.2023.10169688

BBC. (2026, April). *Collapsed Glasgow sewer will take two months to fix*. Https://Www.Bbc.Co.Uk/News/Articles/Cj0vepn3zg3o.

Berger, P. L., & Luckmann, T. (1991). *The Social Construction of Reality*.

Bhaskar, R. (2013). *A Realist Theory of Science*. Routledge. https://doi.org/10.4324/9780203090732

Blumer, H. (1980). Mead and Blumer: The Convergent Methodological Perspectives of Social Behaviorism and Symbolic Interactionism. *American Sociological Review*, *45*(3), 409. https://doi.org/10.2307/2095174

Boettiger, C. (2015). An introduction to Docker for reproducible research. *ACM SIGOPS Operating Systems Review*, *49*(1), 71–79.

Bradshaw, C., Atkinson, S., & Doody, O. (2017). Employing a Qualitative Description Approach in Health Care Research. *Global Qualitative Nursing Research*, *4*, 233339361774228. https://doi.org/10.1177/2333393617742282

Bryman, Alan. (2016). *Social research methods*. Oxford University Press.

BSI. (2006). *Investigation and assessment of drain and sewer systems outside buildings - Visual inspection coding system*. Https://Knowledge.Bsigroup.Com/Products/Investigation-and-Assessment-of-Drain-and-Sewer-Systems-Outside-Buildings-Visual-Inspection-Coding-System?Version=standard.

Canny Edge Detection 09gr820. (2009).

Cao, G., Guo, S., Wei, J., Huang, R., & Mei, L. (2023). Comparison of four sewer condition assessment protocols based on field data. *Water*, *15*(21), 3790. https://doi.org/10.3390/w15213790

Chaoyu, J., & Zakaria, W. N. W. (2025). Sewer pipeline defect detection based on yolov8-cpa. *Jurnal Kejuruteraan*, *37*(5), 2205–2214. https://doi.org/10.17576/jkukm-2025-37(5)-14

Chataigner, F., Cavestany, P., Soler, M., Rizzo, C., Gonzalez, J. P., Bosch, C., Gibert, J., Torrente, A., Gomez, R., & Serrano, D. (2020). ARSI: An Aerial Robot for Sewer Inspection. *Springer Tracts in Advanced Robotics*, *132*, 249–274. https://doi.org/10.1007/978-3-030-22327-4_12

Chen, J., Fang, Y., & Cho, Y. K. (2018). Performance evaluation of 3D descriptors for object recognition in construction applications. *Automation in Construction*, *86*, 44–52. https://doi.org/10.1016/j.autcon.2017.10.033

Chen, Y., Peng, X., Kong, L., … G. D.-… J. of E., & 2021, undefined. (n.d.). Defect inspection technologies for additive manufacturing. *Iopscience.Iop.Org*. Retrieved July 4, 2022, from https://iopscience.iop.org/article/10.1088/2631-7990/abe0d0/meta

Chen, Y., Yin, X., Zhang, Q., Bouferguene, A., Zaman, H., Al-Hussein, M., Russell, R., & Kurach, L. (2019). PRODUCTIVITY ANALYSIS OF MANUAL CONDITION ASSESSMENT FOR SEWER PIPES BASED ON CCTV MONITORING. *CSCE Annual Conference*.

Cheng, J. C. P., & Wang, M. (2018a). Automated detection of sewer pipe defects in closed-circuit television images using deep learning techniques. *Automation in Construction*, *95*, 155–171. https://doi.org/10.1016/j.autcon.2018.08.006

Cheng, J. C. P., & Wang, M. (2018b). Automated detection of sewer pipe defects in closed-circuit television images using deep learning techniques. *Automation in Construction*, *95*, 155–171. https://doi.org/10.1016/j.autcon.2018.08.006

Chenyi, C., Ming-Yu, L., Oncel, T., & Jianxiong, X. (2017). R-CNN for Small Object Detection. In S.-H. Lai, V. Lepetit, K. Nishino, & Y. Sato (Eds.), *Asian Conference on Computer Vision* (Vol. 10115). Springer International Publishing. https://doi.org/10.1007/978-3-319-54193-8

Colyer, S. L., Evans, M., Cosker, D. P., & Salo, A. I. T. (2018). A Review of the Evolution of Vision-Based Motion Analysis and the Integration of Advanced Computer Vision Methods Towards Developing a Markerless System. *Sports Medicine - Open*, *4*(1), 24. https://doi.org/10.1186/s40798-018-0139-y

Creswell, J. (2009). Qualitative, Quantitative and Mixed Methods Approaches. *Research Design*.

Crotty, M. (2020). *The foundations of social research*. Routledge. https://doi.org/10.4324/9781003115700

Cumbajin, E., Rodrigues, N., Costa, P., Miragaia, R., Frazão, L., Costa, N., Fernández-Caballero, A., Carneiro, J., Buruberri, L. H., & Pereira, A. (2023). A Systematic Review on Deep Learning with CNNs Applied to Surface Defect Detection. *Journal of Imaging*, *9*(10), 193. https://doi.org/10.3390/jimaging9100193

Dang, L. M., Hassan, S. I., Im, S., Mehmood, I., & Moon, H. (2018). Utilizing text recognition for the defects extraction in sewers CCTV inspection videos. *Computers in Industry*, *99*, 96–109. https://doi.org/10.1016/j.compind.2018.03.020

Dang, L. M., Kyeong, S. J., Li, Y., Wang, H., Nguyen, T. N., & Moon, H. (2021a). Deep learning-based sewer defect classification for highly imbalanced dataset. *Computers and Industrial Engineering*, *161*. https://doi.org/10.1016/j.cie.2021.107630

Dang, L. M., Kyeong, S. J., Li, Y., Wang, H., Nguyen, T. N., & Moon, H. (2021b). Deep learning-based sewer defect classification for highly imbalanced dataset. *Computers and Industrial Engineering*, *161*. https://doi.org/10.1016/j.cie.2021.107630

Dang, L. M., Kyeong, S., Li, Y., Wang, H., Nguyen, T. N., & Moon, H. (2021c). Deep learning-based sewer defect classification for highly imbalanced dataset. *Computers and Industrial Engineering*, *161*. https://doi.org/10.1016/j.cie.2021.107630

Dang, L. M., Kyeong, S., Li, Y., Wang, H., Nguyen, T. N., & Moon, H. (2021d). Deep learning-based sewer defect classification for highly imbalanced dataset. *Computers and Industrial Engineering*, *161*. https://doi.org/10.1016/j.cie.2021.107630

Dang, L. M., Wang, H., Li, Y., Nguyen, L. Q., Nguyen, T. N., Song, H.-K., & Moon, H. (2023). Lightweight pixel-level semantic segmentation and analysis for sewer defects using deep learning. *Construction and Building Materials*, *371*. https://doi.org/10.1016/j.conbuildmat.2023.130792

Dang, L. M., Wang, H., Li, Y., Nguyen, T. N., & Moon, H. (2022). DefectTR: End-to-end defect detection for sewage networks using a transformer. *Construction and Building Materials*, *325*. https://doi.org/10.1016/j.conbuildmat.2022.126584

Dang, L., Wang, H., Li, Y., Nguyen, L., Materials, T. N.-… and B., & 2022, undefined. (n.d.). Deep learning-based masonry crack segmentation and real-life crack length measurement. *Elsevier*. Retrieved May 23, 2023, from https://www.sciencedirect.com/science/article/pii/S095006182203094X?casa_token=ZSaDNObj7voAAAAA:ZJ0vKnJQ04pe-Dv3O8wmxmmoy0nndovpwr6eJJ_GrZNgMWSwYkS1-U0iPGm_i_AoeDdcZqOM

Datta, S., & Sarkar, S. (2016). A review on different pipeline fault detection methods. In *Journal of Loss Prevention in the Process Industries* (Vol. 41, pp. 97–106). Elsevier Ltd. https://doi.org/10.1016/j.jlp.2016.03.010

Decor, G., Bah, M. D., Foucher, P., Charbonnier, P., & Heitz, F. (2019). Defect Detection in Tunnel Images using Random Forests and Deep Learning. *International Conference on Pattern Recognition Systems (ICPRS)*.

DEFRA. (2023). *Unlimited financial penalties for environmental offences will mean "polluters always pay."* Https://Deframedia.Blog.Gov.Uk/2023/12/11/Unlimited-Financial-Penalties-for-Environmental-Offences-Will-Mean-Polluters-Always-Pay/.

Denzin, N. K. (2009). The elephant in the living room: or extending the conversation about the politics of evidence. *Qualitative Research*, *9*(2), 139–160. https://doi.org/10.1177/1468794108098034

Devi, S., Thopalli, K., Dayana, R., Malarvezhi, P., & Thiagarajan, J. J. (2023). Improving Object Detectors by Exploiting Bounding Boxes for Augmentation Design. *IEEE Access*, *11*, 108356–108364. https://doi.org/10.1109/ACCESS.2023.3320638

Dionisi, A., Sardini, E., & Serpelloni, M. (2012). Wearable object detection system for the blind. *2012 IEEE I2MTC - International Instrumentation and Measurement Technology Conference, Proceedings*, 1255–1258. https://doi.org/10.1109/I2MTC.2012.6229180

Du, W., Shen, H., Fu, J., Zhang, G., Shi, X., Intelligent, Q. H.-J. of, & 2021, undefined. (2021). Automated detection of defects with low semantic information in X-ray images based on deep learning. *Springer*, *32*(1), 141–156. https://doi.org/10.1007/s10845-020-01566-1

Elreedy, D., & Atiya, A. F. (2019). A Comprehensive Analysis of Synthetic Minority Oversampling Technique (SMOTE) for handling class imbalance. *Information Sciences*, *505*, 32–64. https://doi.org/10.1016/j.ins.2019.07.070

Environmental Waste Management. (2023). *Confined Space Entry for Sewers*. Https://Www.Environmentalwastemanagement.Com/Confined-Space-Entry.

Environmental Waste Management. (2024, November). *Confined space entry*. Https://Www.Environmentalwastemanagement.Com/Confined-Space-Entry.

Fang, X., Li, Q., Zhu, J., Chen, Z., Zhang, D., Wu, K., Ding, K., & Li, Q. (2022a). Sewer defect instance segmentation, localization, and 3D reconstruction for sewer floating capsule robots. *Automation in Construction*, *142*. https://doi.org/10.1016/j.autcon.2022.104494

Fang, X., Li, Q., Zhu, J., Chen, Z., Zhang, D., Wu, K., Ding, K., & Li, Q. (2022b). Sewer defect instance segmentation, localization, and 3D reconstruction for sewer floating capsule robots. *Automation in Construction*, *142*. https://doi.org/10.1016/j.autcon.2022.104494

Faram, M. (2014). *SewerBatt®-using sound to assess the condition of sewers and drains*. https://www.researchgate.net/publication/270571315

Felter, W., Ferreira, A., Rajamony, R., & Rubio, J. (2015). An updated performance comparison of virtual machines and Linux containers. *2015 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*, 171–172.

Fjellström, C., & Nyström, K. (2022). Deep learning, stochastic gradient descent and diffusion maps. *Journal of Computational Mathematics and Data Science*, *4*, 100054. https://doi.org/10.1016/j.jcmds.2022.100054

Fu, G., Jin, Y., Sun, S., Yuan, Z., & Butler, D. (2022). The role of deep learning in urban water management: A critical review. In *Water Research* (Vol. 223). Elsevier Ltd. https://doi.org/10.1016/j.watres.2022.118973

Fu, W., & Menzies, T. (2017). Easy over hard: A case study on deep learning. *Proceedings of the ACM SIGSOFT Symposium on the Foundations of Software Engineering*, *Part F130154*, 49–60. https://doi.org/10.1145/3106237.3106256

Gallo, L., & Digilio, D. (2019). A survey on defect detection in sewer systems using computer vision. *IEEE Transactions on Industrial Informatics*, *15*(2), 1234–1243.

Ganegedara, H., Alahakoon, D., Mashford, J., Paplinski, A., Müller, K., & Deserno, T. M. (2012). Self organising map based region of interest labelling for automated defect identification in large sewer pipe image collections. *Proceedings of the International Joint Conference on Neural Networks*. https://doi.org/10.1109/IJCNN.2012.6252482

Ganesan, P. (2023). Smart Infrastructure Management: Integrating AI with DevOps for Cloud-Native Applications. *Journal of Artificial Intelligence & Cloud Computing*, 1–4. https://doi.org/10.47363/JAICC/2023(2)E163

Gao, X., Jian, M., Hu, M., Tanniru, M., & Li, S. (2019a). Faster multi-defect detection system in shield tunnel using combination of FCN and faster RCNN. *Advances in Structural Engineering*, *22*(13), 2907–2921. https://doi.org/10.1177/1369433219849829

Gao, X., Jian, M., Hu, M., Tanniru, M., & Li, S. (2019b). Faster multi-defect detection system in shield tunnel using combination of FCN and faster RCNN. *Advances in Structural Engineering*, *22*(13), 2907–2921. https://doi.org/10.1177/1369433219849829

Gardenier, J., Underwood, J., & Clark, C. (2018). Object Detection for Cattle Gait Tracking. *Proceedings - IEEE International Conference on Robotics and Automation*, 2206–2213. https://doi.org/10.1109/ICRA.2018.8460523

Ghinea, L.-M., Miron, M., & Barbu, M. (2024a). Enhancing Wastewater Treatment Sensor Fault Detection through Deep Learning. *2024 28th International Conference on System Theory, Control and Computing (ICSTCC)*, 95–100. https://doi.org/10.1109/ICSTCC62912.2024.10744689

Ghinea, L.-M., Miron, M., & Barbu, M. (2024b). Enhancing Wastewater Treatment Sensor Fault Detection through Deep Learning. *2024 28th International Conference on System Theory, Control and Computing (ICSTCC)*, 95–100. https://doi.org/10.1109/ICSTCC62912.2024.10744689

Ghosh, K., Bellinger, C., Corizzo, R., Branco, P., Krawczyk, B., & Japkowicz, N. (2024). The class imbalance problem in deep learning. *Machine Learning*, *113*(7), 4845–4901. https://doi.org/10.1007/s10994-022-06268-8

Godwin, A., Benedict, B., Rohde, J., Thielmeyer, A., Perkins, H., Major, J., Clements, H., & Chen, Z. (2021). New Epistemological Perspectives on Quantitative Methods: An Example Using Topological Data Analysis. *Studies in Engineering Education*, *2*(1), 16. https://doi.org/10.21061/see.18

Gokhale, S., Hastak, M., & Huang, R. (2000). *Automated assessment technologies for renewal of underground pipeline infrastructure*. https://doi.org/10.22260/ISARC2000/0029

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

Goodfellow, I., Bengio, Y., & Courville, A. (2018). Deep learning. *Genetic Programming and Evolvable Machines*, *19*(1–2), 305–307. https://doi.org/10.1007/s10710-017-9314-z

Guardian. (2022, January). *Burst water main in north London causes anger amid drought crisis*. Https://Www.Theguardian.Com/Environment/2022/Aug/08/Burst-Water-Main-in-North-London-Islington-Causes-Anger-amid-Drought-Crisis.

Guba, E., & Lincoln, Y. (1994). *1994_Guba_Lincoln_Paradigms_Quali_Research_chapter*.

Guo, W., Soibelman, L., & Garrett, J. H. (2009). Automated defect detection for sewer pipeline inspection and condition assessment. *Automation in Construction*, *18*(5), 587–596. https://doi.org/10.1016/j.autcon.2008.12.003

Gupta, H., Jin, K. H., Nguyen, H. Q., McCann, M. T., & Unser, M. (2018). CNN-Based Projected Gradient Descent for Consistent CT Image Reconstruction. *IEEE Transactions on Medical Imaging*, *37*(6), 1440–1453. https://doi.org/10.1109/TMI.2018.2832656

Gupta, R., & Kaler, R. S. (2018). Development of Fiber Optic Sensors for Leak Detection in Underground Energy Storage Pipelines. *2018 6th Edition of International Conference on Wireless Networks and Embedded Systems, WECON 2018 - Proceedings*, 44–48. https://doi.org/10.1109/WECON.2018.8782057

Gutierrez-Mondragon, M. A., Garcia-Gasulla, D., Alvarez-Napagao, S., Brossa-Ordoñez, J., & Gimenez-Esteban, R. (2020). Obstruction level detection of sewer videos using convolutional neural networks. *Computer Vision and Pattern Recognition*. http://arxiv.org/abs/2002.01284

Haddaway, N. R., Collins, A. M., Coughlin, D., & Kirk, S. (2015). The role of google scholar in evidence reviews and its applicability to grey literature searching. *PLoS ONE*, *10*(9). https://doi.org/10.1371/journal.pone.0138237

Halfawy, M. R., & Hengmeechai, J. (2014a). Automated defect detection in sewer closed circuit television images using histograms of oriented gradients and support vector machine. *Automation in Construction*, *38*, 1–13. https://doi.org/10.1016/j.autcon.2013.10.012

Halfawy, M. R., & Hengmeechai, J. (2014b). Efficient algorithm for crack detection in sewer images from closed-circuit television inspections. *Journal of Infrastructure Systems*, *20*(2). https://doi.org/10.1061/(ASCE)IS.1943-555X.0000161

Halfawy, M. R., & Hengmeechai, J. (2015). Integrated Vision-Based System for Automated Defect Detection in Sewer Closed Circuit Television Inspection Videos. *Journal of Computing in Civil Engineering*, *29*(1), 04014024. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000312

Harshini, M., Philip, J., Haritha, I., & Patil, S. (2022). Sewage Pipeline Fault Detection using Image Processing. *2022 6th International Conference on Electronics, Communication and Aerospace Technology*, 1181–1185. https://doi.org/10.1109/ICECA55336.2022.10009371

Harum, B. J., & Moeslund, T. (2021). Sewer-ML: A Multi-Label Sewer Defect Classification Dataset and Benchmark. *Computer Vision Proceedings*. http://vap.aau.dk/sewer-ml

Hassan, S., Dang, L., Mehmood, I., Im, S., … C. C.-A. in, & 2019, undefined. (n.d.). Underground sewer pipe condition assessment based on convolutional neural networks. *Elsevier*. Retrieved July 4, 2022, from https://www.sciencedirect.com/science/article/pii/S0926580518305892?casa_token=We4QGDG012EAAAAA:tkpnE84c6g0MyEN06Ckt1QkTyFXJ_LzTtFlVxQyoHei5KZ5Utriuiv9ea5TbQRRCljtAJRAd

Hassan, S. I., Dang, L. M., Mehmood, I., Im, S., Choi, C., Kang, J., Park, Y.-S., & Moon, H. (2019). Underground sewer pipe condition assessment based on convolutional neural networks. *Automation in Construction*, *106*. https://doi.org/10.1016/j.autcon.2019.102849

Hauke, J., & Kossowski, T. (2011). Comparison of Values of Pearson's and Spearman's Correlation Coefficients on the Same Sets of Data. *QUAGEO*, *30*(2), 87–93. https://doi.org/10.2478/v10117-011-0021-1

Haurum, J., … T. M. the I. C. on, & 2021, undefined. (n.d.). Sewer-ML: A multi-label sewer defect classification dataset and benchmark. *Openaccess.Thecvf.Com*. Retrieved July 4, 2022, from http://openaccess.thecvf.com/content/CVPR2021/html/Haurum_Sewer-ML_A_Multi-Label_Sewer_Defect_Classification_Dataset_and_Benchmark_CVPR_2021_paper.html

Haurum, J. B., Allahham, M. M. J., Lynge, M. S., Henriksen, K. S., Nikolov, I. A., & Moeslund, T. B. (2021a). Sewer defect classification using synthetic point clouds. *VISIGRAPP 2021 - Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, *5*, 891–900. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85102968084&partnerID=40&md5=98e3dd3754e9b467bafd0a5aa1f702f0

Haurum, J. B., Allahham, M. M. J., Lynge, M. S., Henriksen, K. S., Nikolov, I. A., & Moeslund, T. B. (2021b). Sewer defect classification using synthetic point clouds. *VISIGRAPP 2021 - Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, *5*, 891–900. https://doi.org/10.5220/0010207908910900

Haurum, J. B., & Moeslund, T. B. (2020). A Survey on Image-Based Automation of CCTV and SSET Sewer Inspections. In *Automation in Construction* (Vol. 111). Elsevier B.V. https://doi.org/10.1016/j.autcon.2019.103061

Hawari, A., Alkadour, F., Elmasry, M., & Zayed, T. (2020). A state of the art review on condition assessment models developed for sewer pipelines. *Engineering Applications of Artificial Intelligence*, *93*. https://doi.org/10.1016/j.engappai.2020.103721

He, M., Zhao, Q., Gao, H., Zhang, X., & Zhao, Q. (2022). Image Segmentation of a Sewer Based on Deep Learning. *Sustainability (Switzerland)*, *14*(11). https://doi.org/10.3390/su14116634

He, Z., Wang, K., Fang, T., Su, L., Chen, R., & Fei, X. (2024). *Comprehensive Performance Evaluation of YOLOv11, YOLOv10, YOLOv9, YOLOv8 and YOLOv5 on Object Detection of Power Equipment*. http://arxiv.org/abs/2411.18871

Hu, C., Dong, B., Shao, H., Zhang, J., & Wang, Y. (2023). Toward Purifying Defect Feature for Multilabel Sewer Defect Classification. *IEEE Transactions on Instrumentation and Measurement*, *72*. https://doi.org/10.1109/TIM.2023.3250306

Hu, C., Zhao, C., Shao, H., Deng, J., & Wang, Y. (2024). Tmff: Trustworthy multi-focus fusion framework for multi-label sewer defect classification in sewer inspection videos. *IEEE Transactions on Circuits and Systems for Video Technology*. https://doi.org/10.1109/tcsvt.2024.3433415

Hu, M., Li, Z., Yu, J., Wan, X., Tan, H., & Lin, Z. (2023). Efficient-Lightweight YOLO: Improving Small Object Detection in YOLO for Aerial Images. *Sensors*, *23*(14), 6423. https://doi.org/10.3390/s23146423

Huang, H. W., & Zhang, D. M. (2019). Inspection, recognition and diagnosis of structural health for tunnel lining by artificial intelligence. *9th International Conference on Structural Health Monitoring of Intelligent Infrastructure: Transferring Research into Practice, SHMII 2019 - Conference Proceedings*, *1*, 577–587.

Huang, H., Zhai, M., Lei, X., Chai, B., Liao, W., He, L., Zuo, X., & Wang, H. (2023). Rapid quantification of the surface overflow and underground infiltration in sewer pipes based on computer vision and continuous optimization. *Environmental Research*, *235*, 116606. https://doi.org/10.1016/j.envres.2023.116606

Huang, J., Leung, K. E., Verceles, R., Chen, W., Englert, B., & Aliasgari, M. (2017). Automated detection of sewer pipe structural defects using machine learning. *Water Environment Federation Technical Exhibition and Conference 2017, WEFTEC 2017*, *1*, 667–677.

Huang, Q. H., Li, B. A., Lv, X. Q., Jin Zhang, Z., & Hui Liu, K. (2020). Research on Pipeline Video Defect Detection Based on Improved Convolution Neural Network. *Journal of Physics: Conference Series*, *1576*(1). https://doi.org/10.1088/1742-6596/1576/1/012028

Hughes, J., Cowper-Heays, K., Olesson, E., Bell, R., & Stroombergen, A. (2021a). Impacts and implications of climate change on wastewater systems: A New Zealand perspective. In *Climate Risk Management* (Vol. 31). Elsevier B.V. https://doi.org/10.1016/j.crm.2020.100262

Hughes, J., Cowper-Heays, K., Olesson, E., Bell, R., & Stroombergen, A. (2021b). Impacts and implications of climate change on wastewater systems: A New Zealand perspective. In *Climate Risk Management* (Vol. 31). Elsevier B.V. https://doi.org/10.1016/j.crm.2020.100262

Humble, J., & Farley, D. (2010). *Continuous Delivery: Reliable Software Releases through Build, Test, and Deployment Automation*. Addison-Wesley.

Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., & Keutzer, K. (2016). *SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size*. http://arxiv.org/abs/1602.07360

Jang, S., & Kim, D. (2025). *Impact of data quality on cnn-based sewer defect detection*. https://doi.org/10.20944/preprints202506.0075.v1

Javid, A. M., Das, S., Skoglund, M., & Chatterjee, S. (2021). A relu dense layer to improve the performance of neural networks. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, *2021-June*, 2810–2814. https://doi.org/10.1109/ICASSP39728.2021.9414269

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. In *Stroke and Vascular Neurology* (Vol. 2, Number 4, pp. 230–243). BMJ Publishing Group. https://doi.org/10.1136/svn-2017-000101

Jiang, L., Wang, Y., Tang, Z., Miao, Y., Measurement, S. C.-, & 2021, undefined. (n.d.). Casting defect detection in X-ray images using convolutional neural networks and attention-guided

data augmentation. *Elsevier*. Retrieved July 4, 2022, from https://www.sciencedirect.com/science/article/pii/S0263224120312392?casa_token=OgEqjSEnSREAAAAA:uFY3zBvZuQexRKArqDGrk-E64WSW9S9xj7C2gVr3KrMAk5G2Kpxc6xaoupkIhaq8RWb6oVie

Jing, S., Li, X., Beyene, D. A., Cha, G., & Park, S. (2025). Classification of sewer defects using point clouds based on a novel sewer vision transformer with cross-modal in-domain knowledge. *IEEE Sensors Journal*. https://doi.org/10.1109/jsen.2025.3609788

Jonatan, H., Carl, B., Joakim, H., & Hugo, H. (2025, February 15). *Locust: An open source load testing tool*. Https://Locust.Io/.

Jones, S. (2019). Securing Web APIs with JWT. *IEEE Security & Privacy*, *17*(3), 53–59.

Joseph, R. (2023, June). *Darknet: Open Source Neural Networks in C*. https://pjreddie.com/darknet/

Kaddoura, K. (2015a). *Automated Sewer Inspection Analysis and Condition Assessment*.

Kaddoura, K. (2015b). *Automated sewer inspection analysis and condition assessment*. https://spectrum.library.concordia.ca/id/eprint/980826/

Kaddoura, K., Zayed, T., Atabay, S., & Hawari, A. (2019). *Automated Detection Tool for Deformed Buried Conduits*. https://books.google.com/books?hl=en&lr=&id=HXRjEAAAQBAJ&oi=fnd&pg=PA275&dq=%22defect+detection%22++sewers&ots=OAJqT1IMpE&sig=a7qVMdccoxgTY0bjEpONK4N3dZo

Kamil, B. (2023, November). *Convolution Neural Network*. Https://Becominghuman.Ai/Convolutional-Neural-Network-1136e2b85150.

Kang, M., Bae, H., Lee, K., Moon, K., Yu, J., Kim, J., Kim, D., Song, Y., Dong, J., Kim, K., & Baek, S. (2025). Trends in intelligent sensor-based customized management technologies for sewer infrastructures. *Etri Journal*. https://doi.org/10.4218/etrij.2024-0601

Karlovšek, J., Scheuermann, A., Muller, W., & Williams, D. J. (2011). Application of ground penetrating radar to testing tunnel integrity. *14th Australasian Tunnelling Conference 2011: Development of Underground Space, Proceedings*, 529–542.

Kefan Chen, Hu, H., Chen, C., Chen, L., & He, C. (2018). An Intelligent Sewer Defect Detection Method Based on Convolutional Neural Network. *IEEE International Conference on Information and Automation*.

Khan, M. S., & Patil, R. (2019). Acoustic Characterization of PVC Sewer Pipes for Crack Detection Using Frequency Domain Analysis. *2018 IEEE International Smart Cities Conference, ISC2 2018*. https://doi.org/10.1109/ISC2.2018.8656739

Khan, M. S., Zeng, K., Wu, N., & Unwala, I. (2019). Robotics and Deep Learning Framework for Structural Health Monitoring of Utility Pipes. *2019 22nd IEEE International Symposium on Measurement and Control in Robotics: Robotics for the Benefit of Humanity, ISMCR 2019*. https://doi.org/10.1109/ISMCR47492.2019.8955723

Khan, M., Zeng, K., Wu, N., & Unwala, I. (2019). *Robotics and deep learning framework for structural health monitoring of utility pipes*. https://doi.org/10.1109/ISMCR47492.2019.8955723

Khan, S., Haider, S., & Uwala, I. (2020). A Deep Learning Based Classifier for Crack Detection with Robots in Underground Pipes. *Ieeexplore.Ieee.Org*. https://ieeexplore.ieee.org/abstract/document/9322665/

Kingma, D. P., & Ba, J. (2014). *Adam: A Method for Stochastic Optimization*. http://arxiv.org/abs/1412.6980

Kirstein, S., Müller, K., Walecki-Mingers, M., & Deserno, T. M. (2012). Robust adaptive flow line detection in sewer pipes. *Automation in Construction*, *21*(1), 24–31. https://doi.org/10.1016/j.autcon.2011.05.009

Kitchenham, B., & Charters, S. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering*.

Koch, C., Georgieva, K., Kasireddy, V., Akinci, B., & Fieguth, P. (2015). A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure. *Advanced Engineering Informatics*, *29*(2), 196–210. https://doi.org/10.1016/j.aei.2015.01.008

Kotronoulas, G., Miguel, S., Dowling, M., Fernández-Ortega, P., Colomer-Lahiguera, S., Bağçivan, G., Pape, E., Drury, A., Semple, C., Dieperink, K. B., & Papadopoulou, C. (2023). An Overview of the Fundamentals of Data Management, Analysis, and Interpretation in Quantitative Research. *Seminars in Oncology Nursing*, *39*(2), 151398. https://doi.org/10.1016/j.soncn.2023.151398

Krishnan, A., & Devarajan, Y. (2025). Sustainable wastewater treatment with AI: a path to enhanced hygiene and environmental protection. *Process Safety and Environmental Protection*, *201*, 107487. https://doi.org/10.1016/j.psep.2025.107487

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). *ImageNet Classification with Deep Convolutional Neural Networks*. http://code.google.com/p/cuda-convnet/

Kuhn, T. S. (1962). Historical Structure of Scientific Discovery. *Science*, *136*(3518), 760–764. https://doi.org/10.1126/science.136.3518.760

Kumar, B., Mishra, A., Talesara, A., Kumar, S., Dey, S., Vyas, & Vyas, R. (2022). Object Detection for Autonomous Vehicle in Hazy Environment using Optimized Deep Learning

Techniques. *Proceedings of the 2022 Fourteenth International Conference on Contemporary Computing*, 242–249. https://doi.org/10.1145/3549206.3549252

Kumar, S. S., & Abraham, D. M. (2019a). A Deep Learning Based Automated Structural Defect Detection System for Sewer Pipelines. *Computing in Civil Engineering 2019: Smart Cities, Sustainability, and Resilience - Selected Papers from the ASCE International Conference on Computing in Civil Engineering 2019*, 226–233. https://doi.org/10.1061/9780784482445.029

Kumar, S. S., & Abraham, D. M. (2019b). A Deep Learning Based Automated Structural Defect Detection System for Sewer Pipelines. *Computing in Civil Engineering 2019: Smart Cities, Sustainability, and Resilience - Selected Papers from the ASCE International Conference on Computing in Civil Engineering 2019*, 226–233. https://doi.org/10.1061/9780784482445.029

Kumar, S. S., Abraham, D. M., Jahanshahi, M. R., Iseley, T., & Starr, J. (2018). Automated defect classification in sewer closed circuit television inspections using deep convolutional neural networks. *Automation in Construction*, *91*, 273–283. https://doi.org/10.1016/j.autcon.2018.03.028

Kumar, S. S., Wang, M., Abraham, D. M., Jahanshahi, M. R., Iseley, T., & Cheng, J. C. P. (2020a). Deep Learning–Based Automated Detection of Sewer Defects in CCTV Videos. *Journal of Computing in Civil Engineering*, *34*(1). https://doi.org/10.1061/(asce)cp.1943-5487.0000866

Kumar, S. S., Wang, M., Abraham, D. M., Jahanshahi, M. R., Iseley, T., & Cheng, J. C. P. (2020b). Deep Learning-Based Automated Detection of Sewer Defects in CCTV Videos. *Journal of Computing in Civil Engineering*, *34*(1). https://doi.org/10.1061/(ASCE)CP.1943-5487.0000866

Kumar, S. S., Wang, M., Abraham, D. M., Jahanshahi, M. R., Iseley, T., & Cheng, J. C. P. (2020c). Deep Learning-Based Automated Detection of Sewer Defects in CCTV Videos. *Journal of Computing in Civil Engineering*, *34*(1). https://doi.org/10.1061/(ASCE)CP.1943-5487.0000866

Künzel, J., Werner, T., Eisert, P., Waschnewski, J., Möller, R., & Hilpert, R. (2018a). Automatic Analysis of Sewer Pipes Based on Unrolled Monocular Fisheye Images. *Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018*, *2018-January*, 2019–2027. https://doi.org/10.1109/WACV.2018.00223

Künzel, J., Werner, T., Eisert, P., Waschnewski, J., Möller, R., & Hilpert, R. (2018b). Automatic Analysis of Sewer Pipes Based on Unrolled Monocular Fisheye Images. *Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018*, *2018-Janua*, 2019–2027. https://doi.org/10.1109/WACV.2018.00223

Kuo, T. T., Zavaleta Rojas, H., & Ohno-Machado, L. (2019). Comparison of blockchain platforms: A systematic review and healthcare examples. In *Journal of the American Medical*

Informatics Association (Vol. 26, Number 5, pp. 462–478). Oxford University Press. https://doi.org/10.1093/jamia/ocy185

Laura, W. (2024). *Automatic defect detection in sewer network using deep learning based object detector*. https://doi.org/10.48550/arxiv.2404.06219

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Li, C., Lan, H., Sun, Y., Piping, J. W.-J. of P. V. and, & 2021, undefined. (n.d.). Detection algorithm of defects on polyethylene gas pipe using image recognition. *Elsevier*. Retrieved July 4, 2022, from https://www.sciencedirect.com/science/article/pii/S030801612100079X?casa_token=LV-QV7WWWt8AAAAA:qTFmEKuPZ6lsORuANlTvn0GkPU8WoDfv3wj_IHWM9kw6MnU6ULNPI5q8DxKLu7KkIyj72cZJ

Li, D., Cong, A., & Guo, S. (2019). Sewer damage detection from imbalanced CCTV inspection data using deep convolutional neural networks with hierarchical classification. *Automation in Construction*, *101*, 199–208. https://doi.org/10.1016/j.autcon.2019.01.017

Li, D., Xie, Q., Yu, Z., Wu, Q., Zhou, J., & Wang, J. (2021a). Sewer pipe defect detection via deep learning with local and global feature fusion. *Automation in Construction*, *129*. https://doi.org/10.1016/j.autcon.2021.103823

Li, D., Xie, Q., Yu, Z., Wu, Q., Zhou, J., & Wang, J. (2021b). Sewer pipe defect detection via deep learning with local and global feature fusion. *Automation in Construction*, *129*. https://doi.org/10.1016/j.autcon.2021.103823

Li, Q., He, W., Zhang, X., Yang, Y., & Toe, T. T. (2022). Comparative study of Transfer Learning and VGGNet based classification of Alzheimer's disease. *2022 12th International Conference on Information Technology in Medicine and Education (ITME)*, 88–93. https://doi.org/10.1109/ITME56794.2022.00029

Li, Y., Wang, H., Dang, L. M., Jalil Piran, M., & Moon, H. (2022). A robust instance segmentation framework for underground sewer defect detection. *Measurement: Journal of the International Measurement Confederation*, *190*. https://doi.org/10.1016/j.measurement.2022.110727

Li, Y., Wang, H., Dang, L. M., Song, H. K., & Moon, H. (2022a). Vision-Based Defect Inspection and Condition Assessment for Sewer Pipes: A Comprehensive Survey. In *Sensors* (Vol. 22, Number 7). MDPI. https://doi.org/10.3390/s22072722

Li, Y., Wang, H., Dang, L. M., Song, H., & Moon, H. (2022b). Vision-based defect inspection and condition assessment for sewer pipes: A comprehensive survey. *Sensors*, *22*(7), 2722. https://doi.org/10.3390/s22072722

Liang, Z., Zhang, G., Huang, J. X., & Hu, Q. V. (2014). Deep learning for healthcare decision making with EMRs. *Proceedings - 2014 IEEE International Conference on Bioinformatics and Biomedicine, IEEE BIBM 2014*, 556–559. https://doi.org/10.1109/BIBM.2014.6999219

Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., & Dollár, P. (2014). *Microsoft COCO: Common Objects in Context*. http://arxiv.org/abs/1405.0312

Liu, G. (2020). Surface Defect Detection Methods Based on Deep Learning: A Brief Review. *Proceedings - 2020 2nd International Conference on Information Technology and Computer Application, ITCA 2020*, 200–203. https://doi.org/10.1109/ITCA52113.2020.00049

Liu, Y., Yang, D., Song, T., Ye, Y., & Zhang, X. (2024). YOLO-SSP: an object detection model based on pyramid spatial attention and improved downsampling strategy for remote sensing images. *The Visual Computer*. https://doi.org/10.1007/s00371-024-03434-y

Liu, Y., Yao, J., Lu, X., Xie, R., & Li, L. (2019). DeepCrack: A deep hierarchical feature learning architecture for crack segmentation. *Neurocomputing*, *338*, 139–153. https://doi.org/10.1016/j.neucom.2019.01.036

Locust.io. (2021). *Locust Documentation*.

Long, Y., Gong, Y., Xiao, Z., & Liu, Q. (2017). Accurate object localization in remote sensing images based on convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(5), 2486–2498. https://doi.org/10.1109/TGRS.2016.2645610

Loupos, K., Amditis, A., Doulamis, A., Chrobocinski, P., Victores, J., Wietek, M., Panetsos, P., Roncaglia, A., Camarinopoulos, S., Kallidromitis, V., Komodakis, N., & Lopez, R. (2016). Integrated ROBOTIC solution for tunnel structural evaluation and characterization - ROBO-SPECT EC Project. *Transforming the Future of Infrastructure through Smarter Information - Proceedings of the International Conference on Smart Infrastructure and Construction, ICSIC 2016*, 309–314. https://doi.org/10.1680/tfitsi.61279.309

Lyu, Y., Li, H., Sayagh, M., Jiang, Z. M. (Jack), & Hassan, A. E. (2021). An Empirical Study of the Impact of Data Splitting Decisions on the Performance of AIOps Solutions. *ACM Transactions on Software Engineering and Methodology*, *30*(4), 1–38. https://doi.org/10.1145/3447876

Ma, B., Zuo, X., Shen, J., Xin, S., Shucheng, H., & Yuanhao, L. (2022). Training a Lightweight CNN Model for Fine-Grained Sewer Pipe Cracks Classification Based on Knowledge Distillation. *International Conference on Pattern Recognition and Artificial Intelligence (PRAI)*. https://ieeexplore.ieee.org/abstract/document/9904106/

Ma, D., Fang, H., Wang, N., Zheng, H., … J. D.-A. in, & 2022, undefined. (2022). Automatic defogging, deblurring, and real-time segmentation system for sewer pipeline defects.

Elsevier. https://www.sciencedirect.com/science/article/pii/S0926580522004654?casa_token=rOfveexkRsgAAAAA:UNbWkMZ04FUjoC2LpPT3GyqSJArZonwMAe_e7aEjiPiO4Ok6q4UAVwqfG2DzRo_7jPNyEzse

Ma, D., Liu, J., Fang, H., Wang, N., Zhang, C., Li, Z., & Dong, J. (2021). A Multi-defect detection system for sewer pipelines based on StyleGAN-SDM and fusion CNN. *Construction and Building Materials*, *312*. https://doi.org/10.1016/j.conbuildmat.2021.125385

Manchester Evening News. (2022, November). *Drivers could face seven weeks of delays after sewer collapse on one of Salford's busiest roads*. Https://Www.Manchestereveningnews.Co.Uk/News/Greater-Manchester-News/Drivers-Could-Face-Seven-Weeks-24037057.

Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of science versus scopus and google scholar. *Journal of the American Society for Information Science and Technology*, *58*(13), 2105–2125. https://doi.org/10.1002/asi.20677

Meijer, D., Scholten, L., Clemens, F., & Knobbe, A. (2019). A defect classification methodology for sewer image sets with convolutional neural networks. *Automation in Construction*, *104*, 281–298. https://doi.org/10.1016/j.autcon.2019.04.013

Merkel, D. (2014). Docker: lightweight Linux containers for consistent development and deployment. *Linux Journal*, (239), 2.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Altman, D., Antes, G., Atkins, D., Barbour, V., Barrowman, N., Berlin, J. A., Clark, J., Clarke, M., Cook, D., D'Amico, R., Deeks, J. J., Devereaux, P. J., Dickersin, K., Egger, M., Ernst, E., … Tugwell, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. In *PLoS Medicine* (Vol. 6, Number 7). https://doi.org/10.1371/journal.pmed.1000097

Moradi, S., & Zayed, T. (2017). Real-Time Defect Detection in Sewer Closed Circuit Television Inspection Videos. *Pipelines 2017: Condition Assessment, Surveying, and Geomatics - Proceedings of Sessions of the Pipelines 2017 Conference*, 295–307. https://doi.org/10.1061/9780784480885.027

Moradi, S., Zayed, T., & Golkhoo, F. (2018). Automated sewer pipeline inspection using computer vision techniques. *Pipelines 2018: Condition Assessment, Construction, and Rehabilitation - Proceedings of Sessions of the Pipelines 2018 Conference*, 582–587. https://doi.org/10.1061/9780784481653.064

Moradi, S., Zayed, T., & Golkhoo, F. (2019a). Review on computer aided sewer pipeline defect detection and condition assessment. In *Infrastructures* (Vol. 4, Number 1). MDPI Multidisciplinary Digital Publishing Institute. https://doi.org/10.3390/infrastructures4010010

Moradi, S., Zayed, T., & Golkhoo, F. (2019b). Review on computer aided sewer pipeline defect detection and condition assessment. In *Infrastructures* (Vol. 4, Number 1). MDPI Multidisciplinary Digital Publishing Institute. https://doi.org/10.3390/infrastructures4010010

Moradi, S., Zayed, T., & Golkhoo, F. (2019c). Review on computer aided sewer pipeline defect detection and condition assessment. *Infrastructures*, *4*(1). https://doi.org/10.3390/infrastructures4010010

Moradi, S., Zayed, T., & Golkhoo, F. (2019d). *Review on computer aided sewer pipeline defect detection and condition assessment*. https://doi.org/10.3390/INFRASTRUCTURES4010010

Moselhi, O., & Shehab-Eldeen, T. (1999). Automated detection of surface defects in water and sewer pipes. In *Automation in Construction* (Vol. 8).

Moskalenko, V., Zaretskyi, M., Moskalenko, A., & Lysyuk, V. (2020). Sewer pipe defects classification based on deep convolutional network with information-extreme error-correction decision rules. *Communications in Computer and Information Science*, *1158*, 253–263. https://doi.org/10.1007/978-3-030-61656-4_16

Mostafa, K., & Hegazy, T. (2021). Review of image-based analysis and applications in construction. In *Automation in Construction* (Vol. 122). Elsevier B.V. https://doi.org/10.1016/j.autcon.2020.103516

MRSL. (2024). *Thames Tideway Tunnel Inspection*. Https://Www.Mrsl.Co.Uk/Case-Studies/Thames-Tideway-Tunnel-Inspection.

Mumuni, A., & Mumuni, F. (2024). Automated data processing and feature engineering for deep learning and big data applications: A survey. *Journal of Information and Intelligence*. https://doi.org/10.1016/j.jiixd.2024.01.002

Mushtaq, Z., & Su, S. F. (2020). Environmental sound classification using a regularized deep convolutional neural network with data augmentation. *Applied Acoustics*, *167*. https://doi.org/10.1016/j.apacoust.2020.107389

Myrans, J., Everson, R., Hydroinformatics, Z. K.-J. of, & 2019, undefined. (n.d.). Automated detection of fault types in CCTV sewer surveys. *Iwaponline.Com*. Retrieved July 4, 2022, from https://iwaponline.com/jh/article-abstract/21/1/153/64488

Myrans, J., Kapelan, Z., & Everson, R. (2018a). Automatic identification of sewer fault types using CCTV footage. *International Conference on Hydroinformatics*.

Myrans, J., Kapelan, Z., & Everson, R. (2018b). Using Automatic Anomaly Detection to Identify Faults in Sewers. *International WDSA / CCWI Joint Conference*.

NDTV; (2023, April). *3 Sanitation Workers Die After Inhaling Toxic Gases While Cleaning Sewer*. Https://Www.Ndtv.Com/India-News/3-Sanitation-Workers-Die-after-Inhaling-Toxic-Gases-While-Cleaning-Sewer-3920146.

Nguyen, H. H., Peche, A., & Venohr, M. (2021). Modelling of sewer exfiltration to groundwater in urban wastewater systems: A critical review. *Journal of Hydrology*, *596*, 126130. https://doi.org/10.1016/j.jhydrol.2021.126130

Ni, J., Chen, Y., Chen, Y., Zhu, J., Ali, D., & Cao, W. (2020). A survey on theories and applications for self-driving cars based on deep learning methods. In *Applied Sciences (Switzerland)* (Vol. 10, Number 8). MDPI AG. https://doi.org/10.3390/APP10082749

Noothout, J. M. H., de Vos, B. D., Wolterink, J. M., Postma, E. M., Smeets, P. A. M., Takx, R. A. P., Leiner, T., Viergever, M. A., & Išgum, I. (2020). Deep Learning-Based Regression and Classification for Automatic Landmark Localization in Medical Images. *IEEE Transactions on Medical Imaging*, *39*(12), 4011–4022. https://doi.org/10.1109/TMI.2020.3009002

Noshahri, H., olde Scholtenhuis, L. L., Doree, A. G., & Dertien, E. C. (2021). Linking sewer condition assessment methods to asset managers' data-needs. In *Automation in Construction* (Vol. 131). Elsevier B.V. https://doi.org/10.1016/j.autcon.2021.103878

Ofwat. (2024). *Thames, Yorkshire and Northumbrian Water face £168 million penalty following sewage investigation*. Https://Www.Ofwat.Gov.Uk/Thames-Yorkshire-and-Northumbrian-Water-Face-168-Million-Penalty-Following-Sewage-Investigation/.

Ogai, H., & Bhattacharya, B. (2018). Pipe Inspection Robots for Sewerage Pipelines. *Intelligent Systems, Control and Automation: Science and Engineering*, *89*, 1–12. https://doi.org/10.1007/978-81-322-3751-8_1

Oğurlu, M., BAYRAM, B., Kulavuz, B., & BAKIRMAN, T. (2025). Deep learning for automated sewer defect detection: Benchmarking yolo and rt-detr on the istanbul dataset. *Applied Sciences*, *15*(20), 11096. https://doi.org/10.3390/app152011096

Oh, C., Dang, L. M., Han, D., & Moon, H. (2022). Robust Sewer Defect Detection With Text Analysis Based on Deep Learning. *IEEE Access*, *10*, 46224–46237. https://doi.org/10.1109/ACCESS.2022.3168660

Oluleye, B. I., Chan, D. W. M., & Olawumi, T. O. (2022). Barriers to circular economy adoption and concomitant implementation strategies in building construction and demolition waste management: A PRISMA and interpretive structural modeling approach. In *Habitat International* (Vol. 126). Elsevier Ltd. https://doi.org/10.1016/j.habitatint.2022.102615

OpenCV. (2023). *OpenCV is open source, contains over 2500 algorithms, and is operated by the non-profit Open Source Vision Foundation.* Https://Opencv.Org/.

Orchi, H., Sadik, M., Khaldoun, M., & Sabir, E. (2023). Real-Time Detection of Crop Leaf Diseases Using Enhanced YOLOv8 algorithm. *2023 International Wireless Communications and Mobile Computing (IWCMC)*, 1690–1696. https://doi.org/10.1109/IWCMC58020.2023.10182573

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., Mcdonald, S., … Mckenzie, J. E. (2021a). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. In *The BMJ* (Vol. 372). BMJ Publishing Group. https://doi.org/10.1136/bmj.n160

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., Mcdonald, S., … Mckenzie, J. E. (2021b). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. In *The BMJ* (Vol. 372). BMJ Publishing Group. https://doi.org/10.1136/bmj.n160

Pan, G., Zheng, Y., Guo, S., & Lv, Y. (2020). Automatic sewer pipe defect semantic segmentation based on improved U-Net. *Automation in Construction*, *119*. https://doi.org/10.1016/j.autcon.2020.103383

Pasadas, D. J., Baskaran, P., Ramos, H. G., & Ribeiro, A. L. (2020). Detection and Classification of Defects Using ECT and Multi-Level SVM Model. *IEEE Sensors Journal*, *20*(5), 2329–2338. https://doi.org/10.1109/JSEN.2019.2951302

Patil, R. R., Ansari, S. M., Calay, R. K., & Mustafa, M. Y. (2021). Review of the State-of-the-art Sewer Monitoring and Maintenance Systems Pune Municipal Corporation-A Case Study. *TEM Journal*, *10*(4), 1500–1508. https://doi.org/10.18421/TEM104-02

Pereira, J. R., Cançado, A. C., Alice, A., & Boas, V. (2015). *Business and Management Review SOCIAL MANAGEMENT BEYOND PROCRUSTES' BED i : Ontological, Epistemological and Methodological Considerations*. http://www.businessjournalz.org/bmr

Petticrew, M., & Roberts, H. (n.d.). *Systematic Reviews in the Social Sciences A PRACTICAL GUIDE*.

Petticrew, M., & Roberts, H. (2008). Systematic reviews in the social sciences- A practical guide. *REVIEW OF GAME-BASED LEARNING IN SECONDARY EDUCATION*.

Phyu, H., Research, S. K.-A. E., & 2021, undefined. (2021). Inspection and Understanding of Sewer Network Condition in Dindaeng District (Thailand). *Ph01.Tci-Thaijo.Org*, *43*(2), 46–59. https://doi.org/10.35762/AER.2021.43.2.4

Pinheiro, P. O., & Collobert, R. (n.d.). *From Image-level to Pixel-level Labeling with Convolutional Networks*.

Piotr, S. (2023, January). *Make Sense*. Https://Www.Makesense.Ai/.

Porter, D. H. (2001). The Great Stink of London: Sir Joseph Bazalgette and the Cleansing of the Victorian Metropolis (review). *Victorian Studies*, *43*(3), 530–531. https://doi.org/10.1353/vic.2001.0074

Rada, G., Pérez, D., Araya-Quintanilla, F., Ávila, C., Bravo-Soto, G., Bravo-Jeria, R., Cánepa, A., Capurro, D., Castro-Gutiérrez, V., Contreras, V., Edwards, J., Faúndez, J., Garrido, D., Jiménez, M., Llovet, V., Lobos, D., Madrid, F., Morel-Marambio, M., Mendoza, A., … Zilleruelo-Ramos, R. (2020). Epistemonikos: a comprehensive database of systematic reviews for health decision-making. *BMC Medical Research Methodology*, *20*(1). https://doi.org/10.1186/s12874-020-01157-x

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2015). *You Only Look Once: Unified, Real-Time Object Detection*. http://pjreddie.com/yolo/

Ren, S., He, K., Girshick, R., & Sun, J. (2015). *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks*. http://image-net.org/challenges/LSVRC/2015/results

Rizzo, P. (2010). Water and wastewater pipe nondestructive evaluation and health monitoring: A review. In *Advances in Civil Engineering* (Vol. 2010). https://doi.org/10.1155/2010/818597

Romanova, A., Horoshenkov, K. V., Tait, S. J., & Ertl, T. (2013). Sewer inspection and comparison of acoustic and CCTV methods. *Proceedings of the Institution of Civil Engineers: Water Management*, *166*(2), 70–80. https://doi.org/10.1680/wama.11.00039

Ronacher, S. (2018). *Starlette: The little ASGI framework that shines*.

Roundtree, A. K. (2021). Ethics and Facial Recognition Technology: An Integrative Review. *2021 3rd World Symposium on Artificial Intelligence, WSAI 2021*, 10–19. https://doi.org/10.1109/WSAI51899.2021.9486382

Rowe, E. A. (n.d.). *Regulating Facial Recognition Technology in the Private Sector*. Retrieved https://ssrn.com/abstract=3769613

Rungtalay, N., & Kaitwanidvilai, S. (2024). Dual-stage Classification Framework for Detecting Rare or Unseen Patterns Based on Novelty Detection and Supervised Learning. *Sensors and Materials*, *36*(4), 1487. https://doi.org/10.18494/SAM5036

Saito, T., & Rehmsmeier, M. (2015). The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLOS ONE*, *10*(3), e0118432. https://doi.org/10.1371/journal.pone.0118432

Sardana, N., Shekoohi, S., Cornett, E. M., & Kaye, A. D. (2023). Qualitative and quantitative research methods. In *Substance Use and Addiction Research* (pp. 65–69). Elsevier. https://doi.org/10.1016/B978-0-323-98814-8.00008-1

Sasmal, B., & Dhal, K. G. (2023). A survey on the utilization of Superpixel image for clustering based image segmentation. *Multimedia Tools and Applications*, *82*(23), 35493–35555. https://doi.org/10.1007/s11042-023-14861-9

Saunders, M. N. (2009). *Understanding research philosophies and approaches*. https://www.researchgate.net/publication/309102603

Savino, P., & Tondolo, F. (2021). Automated classification of civil structure defects based on convolutional neural network. *Frontiers of Structural and Civil Engineering*, *15*(2), 305–317. https://doi.org/10.1007/s11709-021-0725-9

Schiappa, M. C., Rawat, Y. S., & Shah, M. (2023). Self-Supervised Learning for Videos: A Survey. *ACM Computing Surveys*, *55*(13s), 1–37. https://doi.org/10.1145/3577925

Schwandt, T. (1998). *Constructivist, Interpretivist Approaches to Human Inquiry*. https://www.researchgate.net/publication/232477264

Shang, X., Li, C., Liu, M., Jiang, C., & Yang, F. (2019). Automatic Drainage Pipeline Defect Detection Method Using Handcrafted and Network Features. *International Conference on Unmanned Systems and Artificial Intelligence*.

Shen, D., Liu, X., Shang, Y., & Tang, X. (2023). Deep Learning-Based Automatic Defect Detection Method for Sewer Pipelines. *Sustainability*, *15*(12), 9164. https://doi.org/10.3390/su15129164

Shiv Kumar, S., Abraham, D. M., Jahanshahi, M., & Iseley, T. (2018). A deep learning framework for automated defect detection using sewer CCTV videos. *NASTT's 2018 No-Dig Show*.

Sim, K. S., Kho, Y. Y., Tso, C. P., Nia, M. E., & Ting, H. Y. (2013). A Contrast Stretching Bilateral Closing Top-Hat Otsu Threshold Technique for Crack Detection in Images. *Scanning*, *35*(2), 75–87. https://doi.org/10.1002/sca.21037

Simon, T. (2023). *D7.2 assessment of current pipe condition assessment approaches and proposals for improvement*. https://doi.org/10.5281/zenodo.7876583

Simon, T., & Ehsan, K. (2022). *D7.1. Report on methodologies used to identify in-pipe defects*. https://doi.org/10.5281/zenodo.7266705

Sinha, S. K., Fieguth, P., Polak, M. A., & McKim, R. A. (1999). *Automated underground pipe condition assessment by image analysis of the state-of-the-art sewer scanner and evaluation technology surveys*.

Situ, Z., Teng, S., Feng, W., Zhong, Q., Chen, G., Su, J., & Zhou, Q. (2023). A transfer learning-based YOLO network for sewer defect detection in comparison to classic object detection methods. *Developments in the Built Environment*, *15*, 100191. https://doi.org/10.1016/j.dibe.2023.100191

Situ, Z., Teng, S., Liao, X., Chen, G., & Zhou, Q. (2023a). Real-time sewer defect detection based on YOLO network, transfer learning, and channel pruning algorithm. *Journal of Civil Structural Health Monitoring*. https://doi.org/10.1007/S13349-023-00681-W

Situ, Z., Teng, S., Liao, X., Chen, G., & Zhou, Q. (2023b). Real-time sewer defect detection based on YOLO network, transfer learning, and channel pruning algorithm. *Journal of Civil Structural Health Monitoring*. https://doi.org/10.1007/s13349-023-00681-w

Situ, Z., Teng, S., Liu, H., Luo, J., & Zhou, Q. (2021a). Automated Sewer Defects Detection Using Style-Based Generative Adversarial Networks and Fine-Tuned Well-Known CNN Classifier. *IEEE Access*, *9*, 59498–59507. https://doi.org/10.1109/ACCESS.2021.3073915

Situ, Z., Teng, S., Liu, H., Luo, J., & Zhou, Q. (2021b). Automated Sewer Defects Detection Using Style-Based Generative Adversarial Networks and Fine-Tuned Well-Known CNN Classifier. *IEEE Access*, *9*, 59498–59507. https://doi.org/10.1109/ACCESS.2021.3073915

Siu, C. F., Wang, M., & Cheng, J. C. P. (2022). A framework for synthetic image generation and augmentation for improving automatic sewer pipe defect detection. *Automation in Construction*, *137*. https://doi.org/10.1016/j.autcon.2022.104213

Siu, C., Wang, M., & Cheng, J. C. P. (2022). A framework for synthetic image generation and augmentation for improving automatic sewer pipe defect detection. *Automation in Construction*, *137*. https://doi.org/10.1016/j.autcon.2022.104213

Sligar, A. P. (2020). Machine learning-based radar perception for autonomous vehicles using full physics simulation. *IEEE Access*, *8*, 51470–51476. https://doi.org/10.1109/ACCESS.2020.2977922

Soviany, P., & Ionescu, R. T. (2018). Optimizing the Trade-Off between Single-Stage and Two-Stage Deep Object Detectors using Image Difficulty Prediction. *2018 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, 209–214. https://doi.org/10.1109/SYNASC.2018.00041

Srivastava, N., Hinton, G., Krizhevsky, A., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. In *Journal of Machine Learning Research* (Vol. 15).

Stanić, N., Langeveld, J., & Clemens, F. (2014). *Potential of Laser Scanning for Quantification of Sediment Deposits in Sewer*.

Su, T. C., & Yang, M. der. (2014). Application of morphological segmentation to leaking defect detection in sewer pipelines. *Sensors (Switzerland)*, *14*(5), 8686–8704. https://doi.org/10.3390/s140508686

Su, T. C., Yang, M. der, Wu, T. C., & Lin, J. Y. (2011a). Morphological segmentation based on edge detection for sewer pipe defects on CCTV images. *Expert Systems with Applications*, *38*(10), 13094–13114. https://doi.org/10.1016/j.eswa.2011.04.116

Su, T. C., Yang, M. der, Wu, T. C., & Lin, J. Y. (2011b). Morphological segmentation based on edge detection for sewer pipe defects on CCTV images. *Expert Systems with Applications*, *38*(10), 13094–13114. https://doi.org/10.1016/j.eswa.2011.04.116

Subramanian, A., Price, S., Kumbhar, O., Sizikova, E., Majaj, N. J., & Pelli, D. G. (2025). Benchmarking the speed–accuracy tradeoff in object recognition by humans and neural networks. *Journal of Vision*, *25*(1), 4. https://doi.org/10.1167/jov.25.1.4

Sun, L., Zhu, J., Tan, J., Li, X., Li, R., Deng, H., … X. Z.-S. of T. T., & 2023, undefined. (n.d.). Deep learning-assisted automated sewage pipe defect detection for urban water environment management. *Elsevier*. Retrieved May 23, 2023, from https://www.sciencedirect.com/science/article/pii/S0048969723021812

Syazana, N., Rahman, A., Saad, M., Abdullah, A. R., & Ahmat, N. (2019). A review of vision based defect detection using image processing techniques for beverage manufacturing industry. *Journals.Utm.My*. https://doi.org/10.11113/jt.v81.12505

Tan, Y., Cai, R., Li, J., Chen, P., & Wang, M. (2021a). Automatic detection of sewer defects based on improved you only look once algorithm. *Automation in Construction*, *131*. https://doi.org/10.1016/j.autcon.2021.103912

Tan, Y., Cai, R., Li, J., Chen, P., & Wang, M. (2021b). Automatic detection of sewer defects based on improved you only look once algorithm. *Automation in Construction*, *131*. https://doi.org/10.1016/j.autcon.2021.103912

Tang, J., & Li, Z. (n.d.). A method for automatically locating defects in cctv inspection data of sewer pipes. *IEEE Access*. https://doi.org/10.1109/access.2025.3591031

Tang, X., Zhao, Z., & He, C. (2024). Automatic detection of sewer defects based on yolov5. *Journal of Physics*, *2770*(1), 012002. https://doi.org/10.1088/1742-6596/2770/1/012002

Tennakoon, R., Hoseinnezhad, R., Tran, H., & Bab-Hadiashar, A. (2018). Visual Inspection of Storm-Water Pipe Systems using Deep Convolutional Neural Networks. *Scitepress.Org*. https://doi.org/10.5220/0006851001350140

Thoma, M. (2016). *A Survey of Semantic Segmentation*. http://arxiv.org/abs/1602.06541

Thompson, N. C., Greenewald, K., Lee, K., & Manso, G. F. (2020). *The Computational Limits of Deep Learning*. http://arxiv.org/abs/2007.05558

Tiangolo, S. (2018). *FastAPI: High performance, easy to learn, fast to code, ready for production*.

Trecroci, A., Boccolini, G., Duca, M., Formenti, D., & Alberti, G. (2020). Mental fatigue impairs physical activity, technical and decision-making performance during small-sided games. *PLOS ONE*, *15*(9), e0238461. https://doi.org/10.1371/journal.pone.0238461

Tscheikner-Gratl, F., Caradot, N., … F. C.-U. W., & 2019, undefined. (2019). Sewer asset management–state of the art and research needs. *Taylor & Francis*. https://www.tandfonline.com/doi/abs/10.1080/1573062X.2020.1713382

Tscheikner-Gratl, F., Caradot, N., Cherqui, F., Leitão, J. P., Ahmadi, M., Langeveld, J. G., Le Gat, Y., Scholten, L., Roghani, B., Rodríguez, J. P., Lepot, M., Stegeman, B., Heinrichsen, A., Kropp, I., Kerres, K., Almeida, M. do C., Bach, P. M., Moy de Vitry, M., Sá Marques, A., … Clemens, F. (2019). Sewer asset management–state of the art and research needs. In *Urban Water Journal* (Vol. 16, Number 9, pp. 662–675). Taylor and Francis Ltd. https://doi.org/10.1080/1573062X.2020.1713382

Tung-Ching, S. (2015). Segmentation of crack and open joint in sewer pipelines based on CCTV inspection images. *International Conference on Circuits and Systems*. https://www.atlantis-press.com/proceedings/cas-15/22996

TzuTa Darren. (2022, February). *LabelImg Annotation tool*. Https://Github.Com/Heartexlabs/LabelImg.

U.S. National Institute of Standards, & Technology. (2018). *Guide to Test, Training, and Validation Scenarios for Smart Grid Cybersecurity*.

Utepov, Y., Kazkeyev, A., & Aniskin, A. (2021). A multi-criteria analysis of sewer monitoring methods for locating pipe blockages and manhole overflows. *Technobius*, *1*(4), 0006. https://doi.org/10.54355/tbus/1.4.2021.0006

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need*. http://arxiv.org/abs/1706.03762

Viacheslav, M., Mykola, Z., Alona, M., & Viktor, L. (2020). Sewer Pipe Defects Classification Based on Deep Convolutional Network with Information-Extreme Error-Correction Decision Rules. In S. Babichev, D. Peleshko, & O. Vynokurova (Eds.), *International Conference in Communications in Computer and Information Science* (Vol. 1158). Springer International Publishing. https://doi.org/10.1007/978-3-030-61656-4

Vincent, O. R., & Folorunso, O. (2009). *\376\377\000I\000n\000S\000I\000T\000E\0000\0009\000p\0000\0009\0007\000-\0001\0000\0007\000V\000i\000n\000c\000e\000n\000t\0006\0001\0003*.

Walters, E. (CEO of F. (2017). *Data-driven law : data analytics and the new legal services*.

Wang, M. (2019). *Automated defect detection, condition assessment and management of underground sewer pipe systems using computer vision and deep learning techniques*. https://repository.ust.hk/ir/Record/1783.1-101641

Wang, M., & Cheng, J. C. P. (2018). Development and improvement of deep learning based automated defect detection for sewer pipe inspection using faster R-CNN. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10864 LNCS*, 171–192. https://doi.org/10.1007/978-3-319-91638-5_9

Wang, M., & Cheng, J. C. P. (2020a). A unified convolutional neural network integrated with conditional random field for pipe defect segmentation. *Computer-Aided Civil and Infrastructure Engineering*, *35*(2), 162–177. https://doi.org/10.1111/mice.12481

Wang, M., & Cheng, J. C. P. (2020b). A unified convolutional neural network integrated with conditional random field for pipe defect segmentation. *Computer-Aided Civil and Infrastructure Engineering*, *35*(2), 162–177. https://doi.org/10.1111/mice.12481

Wang, M., Kumar, S. S., & Cheng, J. C. P. (2021a). Automated sewer pipe defect tracking in CCTV videos based on defect detection and metric learning. *Automation in Construction*, *121*. https://doi.org/10.1016/j.autcon.2020.103438

Wang, M., Kumar, S. S., & Cheng, J. C. P. (2021b). Automated sewer pipe defect tracking in CCTV videos based on defect detection and metric learning. *Automation in Construction*, *121*. https://doi.org/10.1016/j.autcon.2020.103438

Wang, M., Kumar, S. S., & Cheng, J. C. P. (2021c). Automated sewer pipe defect tracking in CCTV videos based on defect detection and metric learning. *Automation in Construction*, *121*. https://doi.org/10.1016/j.autcon.2020.103438

Wang, M., Kumar, S. S., & Cheng, J. C. P. (2021d). Automated sewer pipe defect tracking in CCTV videos based on defect detection and metric learning. *Automation in Construction*, *121*. https://doi.org/10.1016/j.autcon.2020.103438

Wang, M., Luo, H., & Cheng, J. C. P. (2020). Severity assessment of sewer pipe defects in closed-circuit television (CCTV) images using computer vision techniques. *Construction Research Congress 2020: Infrastructure Systems and Sustainability - Selected Papers from the Construction Research Congress 2020*, 942–950. https://doi.org/10.1061/9780784482858.0101

Wang, M., Luo, H., & Cheng, J. C. P. (2021a). Towards an automated condition assessment framework of underground sewer pipes based on closed-circuit television (CCTV) images.

Tunnelling and Underground Space Technology, *110*. https://doi.org/10.1016/j.tust.2021.103840

Wang, M., Luo, H., & Cheng, J. C. P. (2021b). Towards an automated condition assessment framework of underground sewer pipes based on closed-circuit television (CCTV) images. *Tunnelling and Underground Space Technology*, *110*. https://doi.org/10.1016/j.tust.2021.103840

Wang, M., Luo, H., & Cheng, J. C. P. (2021c). Towards an automated condition assessment framework of underground sewer pipes based on closed-circuit television (CCTV) images. *Tunnelling and Underground Space Technology*, *110*. https://doi.org/10.1016/j.tust.2021.103840

Wang, M. Z., & Cheng, J. C. P. (2019a). Semantic segmentation of sewer pipe defects using deep dilated convolutional neural network. *Proceedings of the 36th International Symposium on Automation and Robotics in Construction, ISARC 2019*, 586–594. https://doi.org/10.22260/isarc2019/0078

Wang, M. Z., & Cheng, J. C. P. (2019b). Semantic Segmentation of Sewer Pipe Defects Using Deep Dilated Convolutional Neural Network. *International Symposium on Automation and Robotics in Construction*.

Wang, N., Fang, H., Xue, B., Wu, R., … R. F.-J. of I., & 2023, undefined. (2023). Automatic Damage Segmentation Framework for Buried Sewer Pipes Based on Machine Vision: Case Study of Sewer Pipes in Zhengzhou, China. *Ascelibrary.Org*. https://ascelibrary.org/doi/abs/10.1061/%28ASCE%29IS.1943-555X.0000729

Wei, L., Dragomir, A., Dumitru, E., & Christian, S. (2016). SSD: Single Shot MultiBox Detector. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), *European Conference on Computer Vision* (Vol. 9905). Springer International Publishing. https://doi.org/10.1007/978-3-319-46448-0

Wikimedia Commons. (2023, February 23). *Object detector 1stage vs 2stage*.

Williams, C. A. (2003). Police Surveillance and the Emergence of CCTV in the 1960s. *Crime Prevention and Community Safety*, *5*(3), 27–37. https://doi.org/10.1057/palgrave.cpcs.8140153

WinCan. (2024, May). *Smart Sewer Inspection Technology*. Https://Www.Wincan.Com/En/Home-En/.

Witczak, D., & Szymoniak, S. (2024). Review of Monitoring and Control Systems Based on Internet of Things. *Applied Sciences*, *14*(19), 8943. https://doi.org/10.3390/app14198943

Woods, S. (2024). A Programmatic Approach to Journal Use and Citation Analysis. *Portal: Libraries and the Academy*, *24*(1), 177–200. https://doi.org/10.1353/pla.2024.a916994

Xiang, L., Zhou, H.-L., & Tan, S.-H. (2012). An automatic algorithm for multi-defect classification inside tunnel using SVM. *2012 14th International Conference on Ground Penetrating Radar, GPR 2012*, 454–458.

Xie, Q., Li, D., Xu, J., Yu, Z., & Wang, J. (2019a). Automatic Detection and Classification of Sewer Defects via Hierarchical Deep Learning. *IEEE Transactions on Automation Science and Engineering*, *16*(4), 1836–1847. https://doi.org/10.1109/TASE.2019.2900170

Xie, Q., Li, D., Xu, J., Yu, Z., & Wang, J. (2019b). Automatic Detection and Classification of Sewer Defects via Hierarchical Deep Learning. *IEEE Transactions on Automation Science and Engineering*, *16*(4), 1836–1847. https://doi.org/10.1109/TASE.2019.2900170

Xie, Q., Li, D., Xu, J., Yu, Z., & Wang, J. (2019c). Automatic Detection and Classification of Sewer Defects via Hierarchical Deep Learning. *IEEE Transactions on Automation Science and Engineering*, *16*(4), 1836–1847. https://doi.org/10.1109/TASE.2019.2900170

Xie, Q., Li, D., Xu, J., Yu, Z., & Wang, J. (2019d). Automatic Detection and Classification of Sewer Defects via Hierarchical Deep Learning. *IEEE Transactions on Automation Science and Engineering*, *16*(4), 1836–1847. https://doi.org/10.1109/TASE.2019.2900170

Xin Wang. (2007). Laplacian Operator-Based Edge Detectors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(5), 886–890. https://doi.org/10.1109/TPAMI.2007.1027

Xiong, J., Yu, D., Wang, Q., Shu, L., Cen, J., Liang, Q., Chen, H., & Sun, B. (2021). Application of Histogram Equalization for Image Enhancement in Corrosion Areas. *Shock and Vibration*, *2021*, 1–13. https://doi.org/10.1155/2021/8883571

Xu, Y., & Goodacre, R. (2018). On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning. *Journal of Analysis and Testing*, *2*(3), 249–262. https://doi.org/10.1007/s41664-018-0068-2

Xu, Y., Li, D., Xie, Q., Wu, Q., & Wang, J. (2021a). Automatic defect detection and segmentation of tunnel surface using modified Mask R-CNN. *Measurement: Journal of the International Measurement Confederation*, *178*. https://doi.org/10.1016/j.measurement.2021.109316

Xu, Y., Li, D., Xie, Q., Wu, Q., & Wang, J. (2021b). Automatic defect detection and segmentation of tunnel surface using modified Mask R-CNN. *Measurement: Journal of the International Measurement Confederation*, *178*. https://doi.org/10.1016/j.measurement.2021.109316

Xu, Y., Li, D., Xie, Q., Wu, Q., & Wang, J. (2021c). Automatic defect detection and segmentation of tunnel surface using modified Mask R-CNN. *Measurement: Journal of the International Measurement Confederation*, *178*. https://doi.org/10.1016/j.measurement.2021.109316

Xu, Y., Li, D., Xie, Q., Wu, Q., & Wang, J. (2021d). Automatic defect detection and segmentation of tunnel surface using modified Mask R-CNN. *Measurement: Journal of the International Measurement Confederation*, *178*. https://doi.org/10.1016/j.measurement.2021.109316

Yang, G., Tang, H., Ding, M., Sebe, N., & Ricci, E. (n.d.). *Transformer-Based Attention Networks for Continuous Pixel-Wise Prediction*. Retrieved https://github.com/ygjwd12345/

Yang, L., Wu, X., Zhao, D., Li, H., & Zhai, J. (2011). An improved Prewitt algorithm for edge detection based on noised image. *2011 4th International Congress on Image and Signal Processing*, 1197–1200. https://doi.org/10.1109/CISP.2011.6100495

Yin, J., Yin, X., Pan, M., & Li, L. (2025). Scalable and transparent automated sewer defect detection using weakly supervised object localization. *Automation in Construction*, *174*, 106152. https://doi.org/10.1016/j.autcon.2025.106152

Yin, X., Chen, Y., Bouferguene, A., & Al-Hussein, M. (2020). Data-driven bi-level sewer pipe deterioration model: Design and analysis. *Automation in Construction*, *116*. https://doi.org/10.1016/j.autcon.2020.103181

Yin, X., Chen, Y., Bouferguene, A., Zaman, H., Al-Hussein, M., & Kurach, L. (2020a). A deep learning-based framework for an automated defect detection system for sewer pipes. *Automation in Construction*, *109*. https://doi.org/10.1016/j.autcon.2019.102967

Yin, X., Chen, Y., Bouferguene, A., Zaman, H., Al-Hussein, M., & Kurach, L. (2020b). A deep learning-based framework for an automated defect detection system for sewer pipes. *Automation in Construction*, *109*. https://doi.org/10.1016/j.autcon.2019.102967

Yin, X., Ma, T., Bouferguene, A., & Al-Hussein, M. (2021). Automation for sewer pipe assessment: CCTV video interpretation algorithm and sewer pipe video assessment (SPVA) system development. *Automation in Construction*, *125*. https://doi.org/10.1016/j.autcon.2021.103622

Ylönen, M., & Aven, T. (2023). A framework for understanding risk based on the concepts of ontology and epistemology. *Journal of Risk Research*, *26*(6), 581–593. https://doi.org/10.1080/13669877.2023.2194892

Younis, A., Shixin, L., Shelembi, J. N., & Hai, Z. (2020). Real-time object detection using pre-trained deep learning models mobilenet- SSD. *ACM International Conference Proceeding Series*, 44–48. https://doi.org/10.1145/3379247.3379264

Yu, J., de Antonio, A., & Villalba-Mora, E. (2022). Deep Learning (CNN, RNN) Applications for Smart Homes: A Systematic Review. In *Computers* (Vol. 11, Number 2). MDPI. https://doi.org/10.3390/computers11020026

Yu, Y., Safari, A., Niu, X., Drinkwater, B., & Horoshenkov, K. V. (2021). Acoustic and ultrasonic techniques for defect detection and condition monitoring in water and sewerage pipes: A review. *Applied Acoustics*, *183*. https://doi.org/10.1016/j.apacoust.2021.108282

Yusuf, W., Alaka, H., Ahmad, M., Godoyon, W., Ajayi, S., Toriola-Coker, L. O., & Ahmed, A. (2024). Deep learning for automated encrustation detection in sewer inspection. *Intelligent Systems with Applications*, *24*, 200433. https://doi.org/10.1016/j.iswa.2024.200433

Zaheer, A., Chandra, B. B., Krishna Sai, I. N. N., & Sangeetha, Y. (2021a). Automatic detection of sewer defects via hierarchical deep learning. *Proceedings of the 6th International Conference on Communication and Electronics Systems, ICCES 2021*. https://doi.org/10.1109/ICCES51350.2021.9489100

Zaheer, A., Chandra, B. B., Krishna Sai, I. N. N., & Sangeetha, Y. (2021b, July 8). Automatic detection of sewer defects via hierarchical deep learning. *Proceedings of the 6th International Conference on Communication and Electronics Systems, ICCES 2021*. https://doi.org/10.1109/ICCES51350.2021.9489100

Zayed, T., & al Hawari, A. (2017). *CONDITION ASSESSMENT OF SEWER PIPELINES USING MULTI ATTRIBUTE UTILITY THEORY (MAUT) Osmotic power plants: process design, development and optimization View project Fuzzy knowledge base management system View project Khalid Kaddoura AECOM*. https://www.researchgate.net/publication/318379019

Zeshan, M., Computer, P. G.-2020 4th I. C. on, & 2020, undefined. (2020). Automatic Sewer Cracks Localization Using Deformable Bounding Boxes. *Dl.Acm.Org*, *2020*, 159–167. https://doi.org/10.1145/3445815.3445842

Zhang, J., Liu, X., Zhang, X., Xi, Z., Sciences, S. W.-A., & 2023, undefined. (2023). Automatic Detection Method of Sewer Pipe Defects Using Deep Learning Techniques. *Mdpi.Com*, *2023*, 4589. https://doi.org/10.3390/app13074589

Zhang, K., Qin, L., & Zhu, L. (2025). PDS-YOLO: A Real-Time Detection Algorithm for Pipeline Defect Detection. *Electronics*, *14*(1), 208. https://doi.org/10.3390/electronics14010208

Zhao, C., Hu, C., Shao, H., Dunkin, F., & Wang, Y. (2024). Trusted video-based sewer inspection via support clip-based pareto-optimal evidential network. *IEEE Signal Processing Letters*, 1–5. https://doi.org/10.1109/lsp.2024.3480830

Zhao, C., Hu, C., Shao, H., & Z, W. (2023). Towards Trustworthy Multi-Label Sewer Defect Classification via Evidential Deep Learning. *Ieeexplore.Ieee.Org*. https://ieeexplore.ieee.org/abstract/document/10096569/

Zheng, M. (2023). *Detection and Localization of Root Damages in Underground Sewer Systems using Deep Neural Networks and Computer Vision Techniques*. https://hammer.purdue.edu/articles/thesis/Detection_and_Localization_of_Root_Damages_i

n_Underground_Sewer_Systems_using_Deep_Neural_Networks_and_Computer_Vision_Techniques/21680771

Zhou, Q., Situ, Z., Teng, S., & Chen, G. (2021a). Convolutional Neural Networks–Based Model for Automated Sewer Defects Detection and Classification. *Journal of Water Resources Planning and Management*, *147*(7). https://doi.org/10.1061/(asce)wr.1943-5452.0001394

Zhou, Q., Situ, Z., Teng, S., & Chen, G. (2021b). Convolutional Neural Networks-Based Model for Automated Sewer Defects Detection and Classification. *Journal of Water Resources Planning and Management*, *147*(7). https://doi.org/10.1061/(ASCE)WR.1943-5452.0001394

Zhou, Q., Situ, Z., Teng, S., & Chen, G. (2021c). Convolutional Neural Networks-Based Model for Automated Sewer Defects Detection and Classification. *Journal of Water Resources Planning and Management*, *147*(7). https://doi.org/10.1061/(ASCE)WR.1943-5452.0001394

Zhou, Q., Situ, Z., Teng, S., Chen, W., Chen, G., & Su, J. (2022). Comparison of classic object-detection techniques for automated sewer defect detection. *Journal of Hydroinformatics*, *24*(2), 406–419. https://doi.org/10.2166/hydro.2022.132

Zhou, Y., Ji, A., & Zhang, L. (2022a). Sewer defect detection from 3D point clouds using a transformer-based deep learning model. *Automation in Construction*, *136*. https://doi.org/10.1016/j.autcon.2022.104163

Zhou, Y., Ji, A., & Zhang, L. (2022b). Sewer defect detection from 3D point clouds using a transformer-based deep learning model. *Automation in Construction*, *136*. https://doi.org/10.1016/j.autcon.2022.104163

Zuo, X., Dai, B., Shan, Y., Shen, J., Hu, C., & Huang, S. (2020). Classifying cracks at sub-class level in closed circuit television sewer inspection videos. *Automation in Construction*, *118*. https://doi.org/10.1016/j.autcon.2020.103289