https://doi.org/10.20965/jaciii.2024.p0103

Review:

Burnt-in Text Recognition from Medical Imaging Modalities: Existing Machine Learning Practices

Efosa Osagie†, Wei Ji, and Na Helian

Department of Computer Science, University of Hertfordshire
College Lane Campus, Hatfield, Hertfordshire AL10 9AB, United Kingdom
E-mail: {e.osagie, w.1.ji, n.helian}@herts.ac.uk
†Corresponding author

[Received February 2, 2023; accepted August 17, 2023]

In recent times, medical imaging has become a significant component of clinical diagnosis and examinations to detect and evaluate various medical conditions. The interpretation of these medical examinations and the patient's demographics are usually textual data, which is burned in on the pixel content of medical imaging modalities (MIM). Example of these MIM includes ultrasound and X-ray imaging. As artificial intelligence advances for medical applications, there is a high demand for the accessibility of these burned-in textual data for various needs. This article aims to review the significance of burned-in textual data recognition in MIM and recent research regarding the machine learning approach, challenges, and open issues for further investigation on this application. The review describes the significant problems in this study area as low resolution and background interference of textual data. Finally, the review suggests applying more advanced deep learning ensemble algorithms as possible solutions.

Keywords: medical image character recognition, OCR challenges, burned-in text, medical imaging, medical image processing

1. Introduction

Medical imaging modalities (MIM) are used to probe the human body parts for clinical examinations, and the interpretation of the resulting examinations is burned into the images in the form of textual data, including the patient's demographics. Still, these textual data come with various constraints due to the unique nature of these images. These include background interference, low contrast, distortion, and low resolution of MIM. The optical character recognition (OCR) technique has been applied to enhance visual interpretation and allow automated or semi-automated text recognition and extraction from MIM [1]. In combination with OCR, sophisticated and extensive image processing techniques are also used to eliminate manual textual data recognition [2]. Most processes involve characterisation, multiple image transformations, and other methods to extract the textual data of interest from MIM [1]. Therefore, an investigation of the problem of low resolution and background interference in MIM in current work is highly needed. This review aims to discover the practices currently in research to recognise the low-resolution burned-in text information from MIM by using machine learning (ML) techniques. The challenges and open issues in using ML and deep learning (DL) techniques for recognising burned-textual data on MIM will be presented.

The rest of this review is structured as follows: Section 2 states the significance of textual data recognition in MIM, Section 3 presents detailed background literature on various ML-based techniques, and Section 4 discusses the current challenges and open issues. Finally, Section 5 presents the conclusion, which includes recommendations. A generalised workflow in this article is shown below in **Fig. 1**.

2. Textual Data Recognition from MIM

This section highlights the significance of recognising burned-in textual data from MIM and briefly explains the improvements it can bring to medical informatics. **Fig. 2** shows an X-ray image with magnified areas, including the burned-in textual data.

These textual data are automatically burned-in on the pixel content of the MIM and stored during the imaging acquisition process by acquisition machines.

These textual data can be valuable information with the help of ML and statistical techniques available, which can aid clinical decision-making, patient engagement, and treatment recommendation by leveraging the textual data existing on these MIM [3]. There has been a growing need to integrate heterogeneous data from multiple sources, including burned-in textual data on MIM, in making critical decisions in diagnosis, prognosis, and patient treatment plans [4].

It is highly relevant to recognise burned-in textual data in MIM and integrate it into electronic health record (EHR), which contains non-imaging patient health data (such as blood test results) based on clinical history, vital signs, medications, procedures, and laboratory exam-

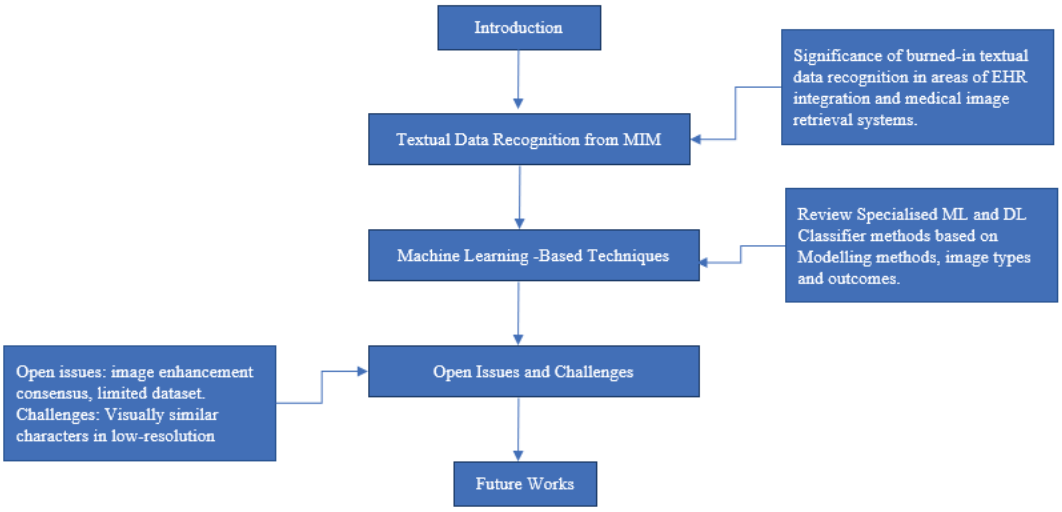

Fig. 1. A generalised workflow of the article.

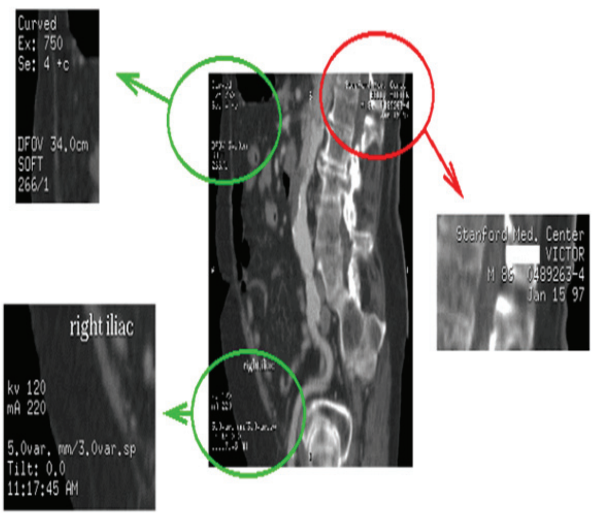

Fig. 2. Sample X-ray image [2].

inations [5]. This will enable health workers to interpret imaging examination findings faster for the appropriate patient due to automatically recognising burned-in textual data using suitable methods. This will lead to higher accuracy in clinical diagnosis, comprehensive organisation and decision of treatment plans, and patient treatment outcomes [6].

In this digital age, the high volume of medical imaging examinations has increased health workers' workload. This increased workload can lead to burn-out, excess fatigue, and an increased error rate [7], especially in manually recognising burned-in text in MIM, which is usually in low resolutions and has background interference. Hence, it is vital to recognise this burned-in text for extraction and fusion with EHR to get these benefits. Historically, ML has emerged in this field with increasing algorithms being developed. These ML algorithms can understand patterns in an image's pixel content and achieve an outstanding classification rate, which is required in recognising these burned-in textual data.

Clinical data entry for MIM is a significant and challenging task that health workers face daily [8], and the significance of an efficient medical image character recognition system would significantly improve the speed, accuracy, and management of medical data entry systems. An efficient clinical data entry ensures patient data are reviewed regularly and queries are fired continuously and quickly addressed. The manual data entry process, especially for MIM, remains time-consuming even with available data capture systems. To get increased accuracy in the clinical data entry process, there is a need to develop a highly efficient and automated burned-in textual data recognition system.

Another significant need is to recognise burned-in textual data to assist health workers who may be visually impaired and even those who are not. Vision is essential to see objects fully and for dark adaptation, contrast sensitivity, image background interference, balance, and colour perception. Some health workers may have a visual impairment, especially in a constrained situation. Therefore, a technique is highly required to recognise the textual data on MIM. Even for healthy vision, there is a need to automate the process and eliminate human differences in extracted textual information (such as knowledge, experiences, and image content interpretation). These recognised burned-in textual data can be converted to easily readable and accessible forms, such as plain text, for rapid integration into medical decision-making systems.

Several methods have been proposed to recognise and apply these burned-in textual data to improve clinical decision-making, reduce data entry inaccuracies and human differences in clinical data entry in text-based medical image retrieval systems. These include using open-source and commercial OCR alone or with various image pre-processing techniques. The recent approach has been to combine image pre-processing techniques and ML models. The following section will present findings on the ML approach in this field.

3. ML-Based Techniques

Designing and implementing an ML model that accurately identifies burned-in textual data on MIM has been challenging over the years. Several works have proposed different combinations of image pre-processing techniques and ML models to recognise burned-in textual data on low-resolution MIM either at the word level with a pre-determined wordlist or the character level. These varied MIM are usually unstandardised images, making conventional OCR methods unreliable because of MIM variety in low contrast, distortion, low resolution, and background interference. The main problems in recognising burned-in textual data are tackling the unstandardised image problems with critical emphasis on low resolution and background interference. This is because the low contrast, distorted text lines, skew and background noise result in poor OCR outcomes. In addition, the low resolution makes characters more zigged and merge with the image's background, which leads to erroneous recognition of characters. This area needs to be investigated when considering high recognition rates from open-source and commercial OCRs on the standard printed text and unconstrained text. Different past authors have applied ML algorithms that have been successful in other classification tasks in medical image processing, such as computerised tomography (CT) image classification to detect lesion classification, X-ray classification to diagnose pneumonia and classifying magnetic resonance (MR) brain images of patients for mild cognitive impairment. Most widely used algorithms for these general classification tasks in MIM include; random forest, gradient boosting classifiers [9], support vector machine, support vector regression, naive Bayes, k-nearest neighbours, and decision tree algorithms [10]. These ML algorithms classify which parts of the human body, presented by the medical image, are infected by the disease using various feature extraction and selection techniques. The poor outcomes of these ML algorithms in burned-in text recognition are due to the problems outlined above, requiring further in-depth research in this area.

A detailed analysis by Newhauser et al. [11] showed that the most popular OCR (Tesseract) was specially designed for recognising text on office documents with character size scanned at a resolution of 300–400 dpi. These documents have a pixel dimension of 1,700 × 2,000 pixels. Therefore, a text with an 8-pt font size under this resolution of 300–400 dpi will be about 22 pixels and can be easily recognised using popular OCRs. A text of 22 pixels means that each character takes up 22 pixels on the image from the top of the character to its bottom. In contrast, a regular CT image has a pixel dimension of 512 × 512 pixels, and a burned-in text of 8-pt font will be approximately 9 pixels [11]. The resolution in MIM is significantly less than what the popular OCR can recognise. The accuracy of these popular OCRs drops rapidly for a text with an 8-pt font size, resolution less than 300 dpi, and fewer pixels than 22 pixels per character [11].

A study by Menasalvas and Gonzalo-Martin [12] on the analysis of non-structured text on MIM also indicated that there would be a need to develop new algorithms and methodologies that can take full advantage of the burned-in textual data contained in these MIM. Menasalvas and Gonzalo-Martin [12] stated that a significant problem in this area lies in the variety of these MIM in background content and low resolution, hindering the applicability of conventional OCR solutions. Hence, designing an ML-based OCR with a high recognition accuracy for any input MIM from any organisation, country, and lexicon is a technically challenging task. In the general domain of OCR under difficult conditions (such as text in natural scenes and degraded hand-held camera-captured document images), several authors have proposed various solutions such as structure extraction by graph spectral decomposition and component selection criterion [13], multiple commercial OCRs with majority logic [14], reinforcement learning, and multiple recognisers [15]. Good results were achieved by these works, such as an 82.3% recognition rate for decorated characters [13], a 98.83% recognition rate for printed Japanese characters [14], and a 90.1% recognition rate for Chinese characters with unique character shapes [15]. However, these solutions are unsuitable for MIM because of their much lower resolution and background interference.

Regarding ML approach to recognising burned-in textual data, different authors have proposed combinations of various image-filtering algorithms with traditional ML models [16]. More recently, convolutional neural networks (CNN) have been employed in solving multiple difficulties experienced in using open-source OCR [17]. The challenge in the ML approach is the need for a high-performance classifier that can distinguish similar characters with low resolution in MIM, even in the presence of background interference. **Table 1** below provides an overview of previous works on the use of ML techniques for the modelling of different solutions to recognise these burned-textual data and presents some main reference papers in the recent literature along with the authors' names, methods used, the dataset sources, the applied evaluation metric and the outcomes of the works. The ML approach usually requires a dataset collected from a public or private source [26]. The collected dataset can be divided into training and validation datasets. The training dataset is used to train the ML model. The validation dataset is used to estimate the trained model's performance while tuning the model's hyperparameters.

Table 1 shows that these past studies encountered similar challenges in recognising burned-in textual data on MIM, which are (a) the problem of background interference and (b) the problem of low resolution. The problem of background interference in MIM occurs mainly as having a grey background, fuzzy font, overlapping text, and inconsistent image quality (too-bright or too-dark image). The problem of low resolution occurs at a resolution of 72–150 dpi for varied modalities such as ultrasound, CT, and others.

The studies by [2, 11, 16, 18, 21] all followed similar methodologies to recognise the burned-in textual data

Table 1. Overview of previous works in the use of ML in recognising burned-in text in MIM.

Works	Image pre-processing technique	Modelling technique	Dataset	Evaluation metrics	Outcomes
Wang [2]	Daubechies wavelet's image transformation	Open-source OCR	100 medical images were collected from a public source and Stanford medical centre.	Character recognition rate (CRR)	The results cannot be generalised due to the small validation dataset used and the problem of low resolution.
Tsui and Chan [18]	Regional thresholding and morphology for character segmentation	Tesseract OCR with weighted similarity	189 ultrasound images from 6 volunteers. Simulated images were produced from 660 previously anonymised medical images.	CRR	CRR of 99.5% but relied on a pre-determined dictionary and a human-assisted revision for error correction.
Newhauser et al. [11]	Threshold-redaction algorithm	Tesseract OCR	NIH-funded studies from 13 patients for cancer treatment.	CER	50% CER achieved, but poor results on X-ray images.
Monteiro et al. [19]	Total-variation denoising, adaptive bilateral filtering, and binary thresholding	Restricted Boltzmann machine and random forest classifier	Training data was from a public character dataset and validated on a 60 ultrasound image collected from a Portuguese medical centre.	False positive rate, false negative rate, F1-score, precision, and recall	Could not recognise certain small font sizes and types in low-resolution MIM.
Ma and Wang [20]	Using local features such as edge density	AdaBoost classifier	100 medical images with text-ultrasound, MR, CT, X-ray. The size was between 300 × 600 pixels and 800 × 1,200 pixels.	Computational cost, precision, and recall	The precision was 74%, and the recall was 77%—difficulty in recognising varied fonts in low-resolution images.
Reul et al. [21]	An expectation-driven method by using prior knowledge of the position and appearance of the textual data in the image	Open-source OCR	22,500 ultrasound images were collected from an investigation of 26 peripheral nerves, and 225 measurements are performed on at least 100 subjects.	CER and word error rate (WER)	A user-assisted revision method with a low error rate of 0.06%. Poor generalisation with complex processes.
Monteiro et al. [22]	Total-variation de-noising, adaptive bilateral filtering, and binary threshold	CNN	Training data was from a public character dataset and validated on privately collected 500 ultrasound images.	CNN model's precision, recall, and F1-score	It depended on complex processes and could not be applied to varied MIM with low resolution.
Silva et al. [23]	Adaptive bilateral filtering and total-variation denoising	CNN	400 high-resolution varied medical images were collected from a private clinic facility.	CER and WER	The model could not recognise certain font types and similar characters.
Vcelak et al. [16]	Binarisation for image transformation	Tesseract OCR	15,334 images for training and 70,191 for validation were collected from the university hospital in the Czech Republic.	Weighted average recall and inverse recall, Cohen's kappa coefficient, and False positive rate (FPR)	FPR of 1.81%–4.00% requires a pre-determined dictionary.
Xu et al. [24]	Image blending	CRNN	2,500 images from the Medpix cardiac atlas (MRI) database.	CRNN model's precision, recall, and F1-measure	Could not recognise similar characters in low-resolution MIM.
Antunes et al. [25]	Template matching from pre-existing MIM metadata	Open-source OCR	Several hundreds of ultrasound images (exact quantity not mentioned).	CRR	Poor performance in complex backgrounds with overlapping text data.

while focusing on the problem of background interference. The authors of these studies mainly proposed image transformation techniques to improve the background contrast and fed the enhanced image to an open-source or commercial OCR. Wang [2] applied the Daubechies wavelet image processing, and Tsui and Chan [18] performed morphological operations. Several studies [11, 16, 21] implemented multiple thresholding techniques to improve the background. These studies focused on increasing the image's local contrast, that is, the contrast between burned-in textual data and background pixels, to make it easier to recognise the characters.

Several studies [22, 24, 27] focused on the low-resolution problem in MIM and proposed various solutions to recognise burned-in textual data on these MIM. The work by Antunes et al. [25], after proposing a character template solution to recognise burned-in textual data focusing on the MIM's low resolution of 352 dpi, concluded that the quality (resolution) of generated character dataset (before recognition) is a principal factor that determines the character recognition accuracy. Monteiro et al. [22] performed recognition on ultrasound images with a 6-layer CNN, achieving a recognition rate of 89.2% on 500 processed images. The low-resolution problem was suggested as the reason why the CNN-based solution could not recognise certain font types and characters. Xu et al. [24] went further to propose a solution for the low-resolution problem by using a convolutional recurrent neural network (CRNN) combining scale variant features during training. Though Xu et al. [24] achieved a recall of 65%, a precision of 70% and an F1-measure of 67% in cardiac MR imaging (MRI), Xu et al. [24] concluded that the solution was not transferable to other types of MIM. The system could not distinguish similar characters in low-resolution MIM. The study [24] did not evaluate the character recognition rate of MIM but provided only the model's performance metrics.

This section shows that there is a need to explore further research in ML and DL to improve state-of-the-art recognition accuracy. Such research focuses on designing a specialised ML or DL model to recognise these characters as accurately as possible under these problematic conditions of low resolution, background interference and noise corruption [28]. The background interference occurs on varied MIM because of the lightning and image acquisition process [29]. The low resolution is standard on varied MIM because of the limited storage of the acquisition machines, leading to reduced image quality [30]. Hence, this paper suggests that a prompt understanding of MIM's background interference and low resolution is required to design an optimal classifier effectively. This understanding will enable research into the design of creating a specific classifier for each modality with a critical focus on the problem of low resolution and background interference.

4. Open Issues and Challenges

Various works have been done using traditional image pre-processing techniques and open-source OCR, and more recently, using ML algorithms in recognising this burned-in text. Yet, some challenges and issues still need to be solved. These challenges and issues exist because of the constraints of acquiring MIM in different difficult conditions, resulting in low resolution and background interference problems due to hardware limits [31]. The open issues discussed in this section include (a) no consensus to measure image enhancement and (b) a limited medical imaging dataset. The significant challenge identified from the different existing ML approaches from an extensive literature review is (a) the discrimination of visually similar characters in low-resolution MIM with background interference.

There is yet to be a consensus to measure the image enhancement on these MIM to get a high recognition performance from the OCR due to the unique nature of the images. This is partly because there is yet to be a commonly accepted metric to measure the level of image enhancement, though some researchers have proposed the peak-signal-to-noise-ratio (PSNR). Michalak and Okarma [32] used the PSNR to measure the success of a proposed image pre-processing methodology using local image entropy for an OCR in text recognition on illuminated document images. Their study suggested a more helpful approach would be the application of metrics calculated for identified characters based on the Levenshtein distance, known as the character error rate (CER), instead of individual pixels. Bieniecki et al. [33] used the CER to evaluate the image's pre-processing methods for text recognition in distorted document images using open-source ABBYY FineReader. Nomura et al. [34] showed that the CER of the resulting models is commonly used as a measure of the image pre-processing success rate. Nomura et al. [34] concluded on the average CER metric from a quantitative evaluation on a test dataset of 1,194 degraded word images to show the essentiality and effectiveness of their proposed image pre-processing method to increase the character recognition rate. Nomura et al. [34] applied a modified Otsu global thresholding technique, which reduced computational requirements and improved the CER in degraded digital word images on an open-source OCR system. Past researchers have applied these generalised metrics in evaluating their overall pipeline by mainly using the CER to measure the level of image enhancement without directly measuring the image enhancement stage. There is a need for a benchmark metric to determine the success of the image enhancement used as different image filtering algorithms have been applied from past works.

The issue of limited medical imaging datasets is of serious concern in applying ML to medical imaging, including the OCR for burned-in textual data recognition, because ML requires a lot of training data to enable optimal tuning of parameters by the learning algorithm. DL algorithms for image classification require large datasets to produce good results and perform poorly with limited datasets [35]. Privacy protection requirements and accessibility greatly hinder the availability of MIM. The resultant effect has led to limited MIM available for the implementation and validation of pipelines, leading to slow progress in the field. Health centres housing MIM usually follow a regional regulatory framework such as the Health Insurance Portability and Accountability Act of 1996 in the United States, as a mandate on the privacy protection of patient's medical records, which ensures that a guarantee is given on the confidentiality of data during storage and transmission via any secured or unsecured means [36]. Due to the adapted regulatory framework, collecting a sufficiently large-scale, balanced MIM dataset is difficult [37]. Several medical image classification competitions have been organised in recent years, motivated by the need to provide more datasets to the ML community to investigate novel ML algorithms on medical images. A notable competition is the Grand Challenge for biomedical imaging competition organised by the Medical Image Computing and Computer Assisted Intervention in 2007 [29]. This medical imaging competition uses annotated datasets to ensure that a uniform validation protocol is available for all participants [30]. Such annotated datasets cannot be used in MIM research regarding textual data identification.

With the existing problems of low resolution and background interference, discriminating visually similar characters is a significant challenge in various ML approaches to recognising burned-in textual data in MIM. A poorly defined character due to low image resolution, background interference, and small fuzzy font sizes can often distort the geometric shape of the character [38]. Human vision sometimes misinterprets visually similar characters, especially when these characters stand alone. This challenge has guided research in improving the recognition rate of similar characters such as "0" and "O," and "5" and "S." Inkeaw et al. [39] similarly identified this challenge in OCRs and suggested a classifier-based approach may be the solution to improve the recognition rate of these visually similar characters. There is a need to develop a complex classifier using ML and DL techniques to adequately learn discriminative features of visually similar characters existing on MIM. Monteiro et al. [22] and

Vcelak et al. [16] identified and attempted to solve this significant challenge using DL techniques by applying a 6-layer CNN but could not recognise certain visually similar characters.

5. Future Outlook

In this review, we have concentrated on research regarding recognising burned-in textual data, which has low resolution and background interference. All these conditions have made it challenging for conventional OCRs, image pre-processing, and ML approaches to recognise these textual data. As a result, an accurate medical image character recognition system will be considered as a significant milestone in medical informatics. This will improve healthcare delivery by accessing and using these identified textual data in decision-making systems. This will also apply to scene text recognition for low-resolution images with noisy backgrounds.

Furthermore, the recognised burned-in textual data will be relevant for post-OCR purposes such as anonymisation, sensitive data obfuscation, automatic integration into EHR systems, and developing a controlled search mechanism for MIM in a large medical database using text-based queries [40]. Therefore, there is a need for further research in this area to create an automatic and highly accurate burned-in textual data recognition model. As regards a way forward in solving these challenges and open issues discussed in the previous section, we suggest efforts in the areas of (a) collaboration between medical imaging centres and the ML research community and (b) an advanced DL approach.

A possible approach to manage the issue of the limited dataset and lack of image enhancement measurement consensus is, first of all, to encourage collaboration between medical imaging centres and the ML research community. This collaboration will enable seamless MIM dataset collection and sharing to allow a more massive pattern recognition to be carried out by the researchers. There is a need to share MIM with privacy considerations to researchers and scientists using acceptable ethical standards [38]. This would come with similar efforts in finding optimal image pre-processing standards for each MIM to provide guidelines on steps that can be applied to MIM before feeding onto an ML algorithm. For instance, most of the ML and DL algorithms for OCR cannot be applied directly to the original image to avoid poor performance because a strong and significant representation of the pixel content of these images is highly relevant for the overall success of the algorithm [39]. With more datasets to conduct more research, a consensus can be reached by the ML community for acceptable standard metrics.

In recent years, the advancement of ML has led to the outstanding development of DL models that possess multiple optimisation strategies and multiple layers architectures. These advancements can solve the limitations posed by using single classifiers in past works. Each DL layer can capture patterns and deeper representation and abstraction, especially in image classification tasks [12]. Though these are not new ideas in the area of medical informatics, a more advanced approach to DL techniques, such as the multi-column deep neural networks, weighted majority voting ensemble, stack ensembles, and DL-tree classifier ensembles, have not been researched in literature, to see how accurate these advanced techniques can recognise these burned-in textual data. These methods have been applied in other areas, such as handwriting benchmark recognition with averaged predictions on the benchmark MNIST dataset [41] and CNN to classify into 1,000 class images in the ImageNet dataset [42]. However, these advanced techniques are yet to be exploited to recognise burned-in textual data primarily because of limited training data. A detailed analysis of classifiers' voting techniques in the domain of pattern recognition by Lam and Suen [43] showed that using a combination of classifiers resulted in an outstanding improvement in overall recognition results in the OCR domain, and this was not depending on the nature of the classifiers [44]. A detailed study by Kovács-V [45] on the voting combination strategy on the NIST Special Database for hand-printed characters gave an error rate of 2.59% using three classifiers operated in parallel with a final supervisor classifier.

We consider research into this area and the possibility of designing an ensemble DL solution that can recognise low-resolution burned-in textual data with background interference on MIM. We suggest an ensemble of DL networks as an effective technique that can learn from the limited dataset in the domain of medical imaging. A major reason for suggesting an ensemble of DL networks is to improve the average character prediction accuracy over any single member in the ensemble. The mechanism for the improved character prediction accuracy will be the reduction in the variance component of prediction errors made by the contributing members. In the aspect of no consensus to measure the level of image enhancement, with the multiple image pre-processing methods available, a single metric may not be sufficient, especially in MIM, because each modality has a varying background and requires a different image pre-processing method. Hence, this review supports the CER and WER as comparable measures of image enhancement only in the OCR domain. Conclusively, these suggestions may be the source of the solutions for the open issues and challenges in recognising burned-in textual data on MIM using conventional methods and the ML approach.

Conflict of Interest

The authors have no conflicts of interest to declare relevant to this article's content.

References:

[1] K. Baskar, "A survey on feature selection techniques in medical image processing," Computer Science, 2018. https://www.semanticscholar.org/paper/A-Survey-on-Feature-Selection-Techniques-in-Medical-Baskar/280694439253fc179a5a4157af18f09177af105c [Accessed November 6, 2022]

[2] J. Z. Wang, "Security filtering of medical images using OCR," Proc.

of the 3rd All-Russian Scientific Conf., pp. 118-122, 2001.

[3] T. Davenport and R. Kalakota, "The potential for artificial intelligence in healthcare," Future Healthcare J., Vol.6, No.2, pp. 94-98, 2019. https://doi.org/10.7861/futurehosp.6-2-94

[4] C. B. Collin et al., "Computational models for clinical applications in personalized medicine—guidelines and recommendations for data integration and model validation," J. of Personalized Medicine, Vol.12, No.2, Article No.166, 2022. https://doi.org/10.3390/jpm12020166

[5] V. Ehrenstein et al., "Obtaining data from electronic health records," R. E. Gliklich, M. B. Leavy, and N. A. Dreyer (Eds.), "Tools and Technologies for Registry Interoperability, Registries for Evaluating Patient Outcomes: A User's Guide, 3rd Edition, Addendum 2," Agency for Healthcare Research and Quality, 2019. https://www.ncbi.nlm.nih.gov/books/NBK551878/ [Accessed November 12, 2022]

[6] A. Paulsen, S. Overgaard, and J. M. Lauritsen, "Quality of data entry using single entry, double entry and automated forms processing—an example based on a study of patient-reported outcomes," PLOS ONE, Vol.7, No.4, Article No.e35087, 2012. https://doi.org/10.1371/journal.pone.0035087

[7] R. J. McDonald et al., "The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload," Academic Radiology, Vol.22, No.9, pp. 1191-1198, 2015. https://doi.org/10.1016/j.acra.2015.05.007

[8] B. Bergeron, "Clinical data capture: OMR and OCR and your flatbed scanner," MedGenMed, Vol.7, No.2, Article No.66, 2005.

[9] R. Rabiei et al., "Prediction of breast cancer using machine learning approaches," J. of Biomedical Physics and Engineering, Vol.12, No.3, pp. 297-308, 2022. https://doi.org/10.31661/jbpe.v0i0.2109-1403

[10] Y. Amethiya et al., "Comparative analysis of breast cancer detection using machine learning and biosensors," Intelligent Medicine, Vol.2, No.2, pp. 69-81, 2022. https://doi.org/10.1016/j.imed.2021.08.004

[11] W. Newhauser et al., "Anonymization of DICOM electronic medical records for radiation therapy," Computers in Biology and Medicine, Vol.53, pp. 134-140, 2014. https://doi.org/10.1016/j.compbiomed.2014.07.010

[12] E. Menasalvas and C. Gonzalo-Martin, "Challenges of medical text and image processing: Machine learning approaches," A. Holzinger (Ed.), "Machine Learning for Health Informatics: State-of-the-Art and Future Challenges," pp. 221-242, Springer, 2016. https://doi.org/10.1007/978-3-319-50478-0_11

[13] H. Kawano et al., "Structure extraction from decorated characters by graph spectral decomposition and component selection criterion," J. Adv. Comput. Intell. Intell. Inform., Vol.14, No.2, pp. 179-184, 2010. https://doi.org/10.20965/jaciii.2010.p0179

[14] H. Miyao et al., "Printed Japanese character recognition using multiple commercial OCRs," J. Adv. Comput. Intell. Intell. Inform., Vol.8, No.2, pp. 200-207, 2004. https://doi.org/10.20965/jaciii.2004.p0200

[15] J. Park et al., "Multi-lingual optical character recognition system using the reinforcement learning of character segmenter," IEEE Access, Vol.8, pp. 174437-174448, 2020. https://doi.org/10.1109/ACCESS.2020.3025769

[16] P. Vcelak et al., "Identification and classification of DICOM files with burned-in text content," Int. J. of Medical Informatics, Vol.126, pp. 128-137, 2019. https://doi.org/10.1016/j.ijmedinf.2019.02.011

[17] K. Mohsenzadegan, V. Tavakkoli, and K. Kyamakya, "Deep neural network concept for a blind enhancement of document-images in the presence of multiple distortions," Applied Sciences, Vol.12, No.19, Article No.9601, 2022. https://doi.org/10.3390/app12199601

[18] G. K. Tsui and T. Chan, "Automatic selective removal of embedded patient information from image content of DICOM files," American J. of Roentgenology, Vol.198, No.4, pp. 769-772, 2012. https://doi.org/10.2214/AJR.10.6352

[19] E. Monteiro, C. Costa, and J. L. Oliveira, "A machine learning methodology for medical imaging anonymization," 2015 37th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 1381-1384, 2015. https://doi.org/10.1109/EMBC.2015.7318626

[20] Y. Ma and Y. Wang, "Text detection in medical images using local feature extraction and supervised learning," 2015 12th Int. Conf. on Fuzzy Systems and Knowledge Discovery (FSKD), pp. 953-958, 2015. https://doi.org/10.1109/FSKD.2015.7382072

[21] C. Reul et al., "Expectation-driven text extraction from medical ultrasound images," A. Hoerbst et al. (Eds.), "Exploring Complexity in Health: An Interdisciplinary Systems Approach," pp. 712-716, IOS Press, 2016. https://doi.org/10.3233/978-1-61499-678-1-712

[22] E. Monteiro, C. Costa, and J. L. Oliveira, "A de-identification pipeline for ultrasound medical images in DICOM format," J. of Medical Systems, Vol.41, No.5, Article No.89, 2017. https://doi.org/10.1007/s10916-017-0736-1

[23] J. M. Silva et al., "Controlled searching in reversibly de-identified medical imaging archives," J. of Biomedical Informatics, Vol.77, pp. 81-90, 2018. https://doi.org/10.1016/j.jbi.2017.12.002

[24] X. Xu, W. Wang, and Q. Liu, "Medical image character recognition based on multi-scale neural convolutional network," 2021 Int. Conf. on Security, Pattern Analysis, and Cybernetics (SPAC), pp. 408-412, 2021. https://doi.org/10.1109/SPAC53836.2021.9539999

[25] M. Antunes, R. Machado, and A. Silva, "Anonymization of burned-in annotations in ultrasound imaging," Eletrónica e Telecomunicações, Vol.5, No.3, pp. 360-364, 2011.

[26] J. P. Segal and R. Hansen, "Medical images, social media and consent," Nature Reviews Gastroenterology & Hepatology, Vol.18, No.8, pp. 517-518, 2021. https://doi.org/10.1038/s41575-021-00453-1

[27] A. Badano et al., "Consistency and standardization of color in medical imaging: A consensus report," J. of Digital Imaging, Vol.28, No.1, pp. 41-52, 2015. https://doi.org/10.1007/s10278-014-9721-0

[28] M. Kociołek, M. Strzelecki, and R. Obuchowicz, "Does image normalization and intensity resolution impact texture classification?," Computerized Medical Imaging and Graphics, Vol.81, Article No.101716, 2020. https://doi.org/10.1016/j.compmedimag.2020.101716

[29] L. Maier-Hein et al., "Why rankings of biomedical image analysis competitions should be interpreted with care," Nature Communications, Vol.9, No.1, Article No.5217, 2018. https://doi.org/10.1038/s41467-018-07619-7

[30] M. Aljabri et al., "Towards a better understanding of annotation tools for medical imaging: A survey," Multimedia Tools and Applications, Vol.81, No.18, pp. 25877-25911, 2022. https://doi.org/10.1007/s11042-022-12100-1

[31] Y. Li, B. Sixou, and F. Peyrin, "A review of the deep learning methods for medical images super resolution problems," IRBM, Vol.42, No.2, pp. 120-133, 2021. https://doi.org/10.1016/j.irbm.2020.08.004

[32] H. Michalak and K. Okarma, "Improvement of image binarization methods using image preprocessing with local entropy filtering for alphanumerical character recognition purposes," Entropy, Vol.21, No.6, Article No.562, 2019. https://doi.org/10.3390/e21060562

[33] W. Bieniecki, S. Grabowski, and W. Rozenberg, "Image preprocessing for improving OCR accuracy," 2007 Int. Conf. on Perspective Technologies and Methods in MEMS Design, pp. 75-80, 2007. https://doi.org/10.1109/MEMSTECH.2007.4283429

[34] S. Nomura et al., "Morphological preprocessing method to thresholding degraded word images," Pattern Recognition Letters, Vol.30, No.8, pp. 729-744, 2009. https://doi.org/10.1016/j.patrec.2009.03.008

[35] J. M. D. Delgado and L. Oyedele, "Deep learning with small datasets: Using autoencoders to address limited datasets in construction management," Applied Soft Computing, Vol.112, Article No.107836, 2021. https://doi.org/10.1016/j.asoc.2021.107836

[36] M. Li, R. Poovendran, and S. Narayanan, "Protecting patient privacy against unauthorized release of medical images in a group communication environment," Computerized Medical Imaging and Graphics, Vol.29, No.5, pp. 367-383, 2005. https://doi.org/10.1016/j.compmedimag.2005.02.003

[37] X. Qin, F. M. Bui, and H. H. Nguyen, "Learning from an imbalanced and limited dataset and an application to medical imaging," 2019 IEEE Pacific Rim Conf. on Communications, Computers and Signal Processing (PACRIM), 2019. https://doi.org/10.1109/PACRIM47961.2019.8985057

[38] D. Pal et al., "MSHSCNN: Multi-scale hybrid-Siamese network to differentiate visually similar character classes," 2021 9th European Workshop on Visual Information Processing (EUVIP), 2021. https://doi.org/10.1109/EUVIP50544.2021.9483980

[39] P. Inkeaw et al., "Recognition of similar characters using gradient features of discriminative regions," Expert Systems with Applications, Vol.134, pp. 120-137, 2019. https://doi.org/10.1016/j.eswa.2019.05.050

[40] A. Safaei, "Text-based multi-dimensional medical images retrieval according to the features-usage correlation," Medical & Biological Engineering & Computing, Vol.59, No.10, pp. 1993-2017, 2021. https://doi.org/10.1007/s11517-021-02392-0

[41] D. Cireşan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," arXiv: 1202.2745, 2012. https://doi.org/10.48550/ARXIV.1202.2745

[42] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Communications of the ACM, Vol.60, No.6, pp. 84-90, 2017. https://doi.org/10.1145/3065386

[43] L. Lam and S. Y. Suen, "Application of majority voting to pattern recognition: An analysis of its behavior and performance," IEEE Trans. on Systems, Man, and Cybernetics – Part A: Systems and Humans, Vol.27, No.5, pp. 553-568, 1997. https://doi.org/10.1109/3468.618255

[44] J. Shlens, "A tutorial on principal component analysis," arXiv: 1404.1100, 2014. https://doi.org/10.48550/ARXIV.1404.1100

[45] Z. M. Kovács-V, "A novel architecture for high quality hand-printed character recognition," Pattern Recognition, Vol.28, No.11, pp. 1685-1692, 1995. https://doi.org/10.1016/0031-3203(95)00044-Z

Name:
Efosa Osagie

Affiliation:
Visiting Lecturer and Ph.D. Researcher, University of Hertfordshire

Address:
College Lane Campus, Hatfield, Hertfordshire AL10 9AB, United Kingdom

Brief Biographical History:
2014 Received Bachelor's degree in Materials and Production Engineering from Ambrose Alli University
2019 Received Master's degree in Information Technology from University of Bolton
2020- Ph.D. Researcher in Computer Science (Data Science and Machine Learning), University of Hertfordshire
2021- Visiting Lecturer, School of Physics, Engineering and Computer Science, University of Hertfordshire

Main Works:
• Medical image character recognition systems, medical image analysis based on deep learning, semantic similarity learning, and text-based retrieval systems.

Membership in Academic Societies:
• British Computer Society (BCS)

Name:
Wei Ji

Affiliation:
Ph.D. Supervisor and Principal Lecturer, Department of Computer Science, University of Hertfordshire

Address:
College Lane Campus, Hatfield, Hertfordshire AL10 9AB, United Kingdom

Brief Biographical History:
2003- Department of Computer Science, University of Hertfordshire

Main Works:
• Neural networks and fuzzy logic in cancer research.

Name:
Na Helian

ORCID:
0000-0001-6687-0306

Affiliation:
Supervisor and Principal Lecturer, University of Hertfordshire

Address:
College Lane Campus, Hatfield, Hertfordshire AL10 9AB, United Kingdom

Brief Biographical History:
1997-2002 University of Plymouth
2002-2008 London Metropolitan University
2008- University of Hertfordshire

Main Works:
• "Feature weighting in DBSCAN using reverse nearest neighbours," Pattern Recognition, Vol.137, Article No.109314, 2023.
• "Boosting ant colony optimization with reptile search algorithm for churn prediction," Mathematics, Vol.10, No.7, Article No.1031, 2022.