TOPICAL REVIEW

Machine Learning Based Diagnostic Paradigm in Viral and Non-Viral Hepatocellular Carcinoma

ARUN ASIF[1], FAHEEM AHMED[2], ZEESHAN[3], JAVED ALI KHAN[4], EMAN ALLOGMANI[5], NORA EL RASHIDY[6], SOBIA MANZOOR[7], AND MUHAMMAD SHAHID ANWAR[8]

[1]Department of Molecular and Cellular Physiology, Albany Medical College, Albany, NY 12202, USA
[2]Department of Mechatronics Engineering, Jeju National University, Jeju-si 63243, Republic of Korea
[3]Department of Electronics Engineering, Gachon University, Seongnam-si, Gyeonggi-do 13120, South Korea
[4]Department of Computer Science, School of Physics, Engineering, and Computer Science, University of Hertfordshire, AL10 9AB Hatfield, U.K.
[5]Department of Computer Science, College of Computer and Information Sciences, Majmaah University, Al-Majmaah 11952, Saudi Arabia
[6]Department of Machine Learning and Information Retrieval, Faculty of Artificial Intelligence, Kaferelshikh University, Kafr El-Shaikh 33516, Egypt
[7]Department of Healthcare Biotechnology, Atta-Ur-Rahman School of Applied Biosciences, National University of Sciences and Technology, Islamabad 44000, Pakistan
[8]Department of AI and Software, Gachon University, Seongnam-si 13120, South Korea

Corresponding authors: Sobia Manzoor (drsobia.manzoor@asab.nust.edu.pk) and Muhammad Shahid Anwar (shahidanwar786@gachon.ac.kr)

ABSTRACT Viral and non-viral hepatocellular carcinoma (HCC) is becoming predominant in developing countries. A major issue linked to HCC-related mortality rate is the late diagnosis of cancer development. Although traditional approaches to diagnosing HCC have become gold-standard, there remain several limitations due to which the confirmation of cancer progression takes a longer period. The recent emergence of artificial intelligence tools with the capacity to analyze biomedical datasets is assisting traditional diagnostic approaches for early diagnosis with certainty. Here we present a review of traditional HCC diagnostic approaches versus the use of artificial intelligence (Machine Learning and Deep Learning) for HCC diagnosis. The overview of the cancer-related databases along with the use of AI in histopathology, radiology, biomarker, and electronic health records (EHRs) based HCC diagnosis is given.

I. INTRODUCTION

Approximately, 80% of the hepatocellular carcinoma are caused by chronic viral infections including Hepatitis C Virus (HCV) and Hepatis B Virus (HBV) [1]. World Health Organization (WHO) aims to reduce the CHV incidence by 90% and mortality by 65% before 2030 according to the strategy on viral hepatitis prepared in 2016. Globally, two million deaths are caused due to hepatocellular carcinoma (HCC) [2]. Hepatitis C virus and hepatitis B virus are the major cause of chronic liver disease leading to HCC development especially in developing countries [3], [4], [5]. As early diagnosis of liver inflammation causing fibrosis and cirrhosis is crucial for better treatment outcomes [6]. The current

The associate editor coordinating the review of this manuscript and approving it for publication was Inês Domingues.

invasive method, biopsy, for liver fibrosis analysis is complicated and present the challenges of bleeding, severe pain, observer variability, sampling errors and increased chances of infection [7]. Although, another method of ultrasound elastography is emerging but it has limitations with measuring stiffness of liver tissue owing to interference of factors such as tissue inflammation, hepatic vein congestion, obesity, meal, etc., which lead to misinterpretation of acquired data [8]. Comparatively, there is a dire need of developing a less invasive method for analysis of fibrotic liver diagnosis.

Traditional statistical models based on conventional regression have failed to capture non-linear and high order interaction of predictor variables in large datasets [9]. Multiple logistic regression models have been used to analyze fibrotic liver condition [10]. However, traditional approaches lack the capacity to interlink non-linear interactions between

multiple variables which play a significant role in defining the severity of inflammation development. Recently, machine learning-assisted methods to analyze big-data are emerging to resolve issues regarding diagnosis, staging and prognosis multiple diseases.

A new scoring system based on simple clinical parameters, including blood tests and protein profiles may help in the analysis of fibrotic liver. For better clinical analysis of data, machine learning may play a pivotal role due to its adept analytical approach. Machine learning based algorithms can capture complex relationships in large datasets [11]. Continuous evidence shows the implementation of machine learning approaches with improved prediction system for diagnosis of multiple disease including breast cancer and NAFLD, and postprandial glycemic responses [12].

For better prognostic analysis of chronic viral hepatitis, the diagnostic capacity of current approaches must be enhanced using advanced tools. Machine learning has recently emerged as to better understand the diagnostic data for diseases as shown in Figure 1.
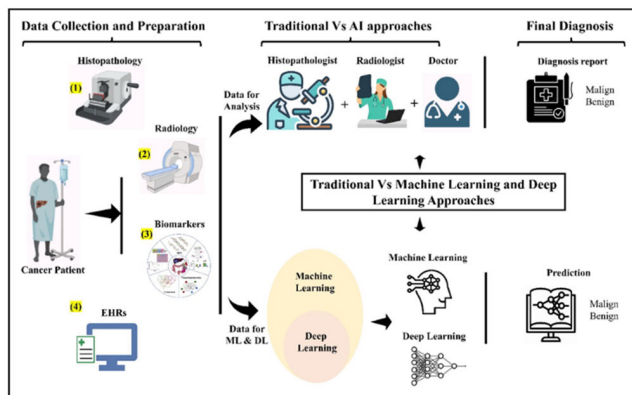
FIGURE 1. Left: Biological approaches used to obtain the data related to cancer patients (hepatocellular carcinoma), Middle: Traditional vs Machine learning and Deep learning-based approaches are shown for the diagnosis, and finally in Right: Final diagnosis results which detect and classify the data as malign, benign or can further differentiate in sub-tumor types based on model training in case of ML and DL.

II. TRADITIONAL APPROACHES IN DIAGNOSIS

Traditional approaches to diagnose hepatocellular carcinoma mainly include histopathology analysis of tissue biopsies, radiological analysis, cancer-specific biomarker analysis and electronic health records [13]. These approaches are usually time consuming and always require analytical opinion of histopathologist or physician which may include factors like error in observation or individual biasness [14], [15]. Owing to such discrepancies, the diagnosis and prognosis of a cancer is highly misjudged. Eventually, the patients' management and survival are affected by traditional approaches.

For managing liver disorders like autoimmune hepatitis and non-alcoholic steatohepatitis, Histopathology has played a fundamental role [16]. However, noninvasive techniques for HCC diagnosis are limited which makes it is obligatory to do histological scrutiny of tumor samples for masses with uncommon features on imaging or to eliminate diagnosis of cholangiocarcinoma, metastasis, or benign primary liver tumor [17]. Owing to certain limitations of noninvasive methods to diagnose cancer makes it difficult to detect cancer at an early stage. The basic method to diagnose liver disease is the histopathological analysis for defining the cancer and its stage or grade [18]. There are several limitations for histopathology as well, a significant tissue mass is usually required for the tissue sections and sample preparation for analysis [19]. Additionally, the sample analysis may be affected by the personal bias among histopathologists leading disagreements [20]. Therefore, precision in histological characterization of liver tumor samples might frequently prove to be challenging. To resolve this issue, researchers recently has applied AI techniques to support diagnosis of liver tumor samples [21]. Recent evidence suggests that by using a huge dataset of H&E-stained images from The Cancer Genome Atlas (TCGA), a convolutional neural network (CNN) was designed to differentiate between neighboring healthy liver tissue around the cancerous liver tissue to detect HCC with approximately 90% accuracy [22]. In another study, Kiani et al. developed a CNN based algorithm to distinguish stained tissue images for their classification as HCC and cholangiocarcinoma with approximately 80% accuracy [23]. These algorithms with accuracy of $\sim$ 90% on validation in combination with observations of histopathologists outperformed their individual accuracy of human analysis. So, instead of just relying on AI alone or pathologists alone we should use AI to enhance accuracy instead of replacing it for traditional histological diagnosis. Research revealed that incorrect prediction can negatively affect final diagnosis made by pathologists, which infers that a cautious approach should be taken when using AI models exclusively designed to automate the process of HCC diagnosis [24].

Another method to analyze liver cancer development is the use of radiology, by application of ultrasound analysis of liver [25]. Traditionally, a two-dimensional ultrasound image (B-mode ultrasound) is developed from the abdominal cavity which has become a clinical guideline for liver tumor detection [26], [119]. Although it has become a common practice to analyze the potential tumor with ultrasound radiation, there have been well defined limitations to this process. Mainly, B-mode ultrasonography is 50% to 60% accurate due to underlying factors such as, patient's body posture, equipment quality, operator experience, etc [27]. Owing to such limitations after initial analysis by ultrasound imaging, patients are referred to contrast-enhanced computed tomography (CT) imaging or magnetic resonance imaging (MRI) [28]. These radiological imaging techniques assist in confirming the pathognomonic features for HCC detection yet there remain some limitations. MRI and CT imaging often lack the capacity to identify liver nodules at early stages leading to subsequent follow-up imaging or liver

biopsy recommendations [29]. Consequently, radiological analysis causes unreasonable stress for patients by delaying the process of HCC diagnosis. Fig. 3 shows a schematic representation of commonly used diagnosis methods for HCC diagnosis.

The recent rise of artificial intelligence in healthcare by make use of machine learning and deep learning tools for genomics, proteomics, statistical and image-based analysis may play a pivotal role in HCC diagnosis in future [30], [31], [32], [33]. Implementation of AI in addition to clinical diagnosis may speed-up the process of diagnosis at earliest stages of HCC development with correct identification of liver lesions and cancerous tissue [34], [35].
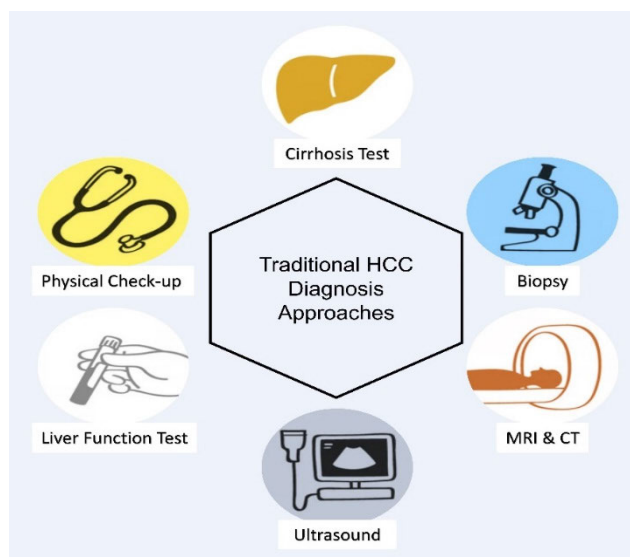
FIGURE 2. A schematic overview of the traditional approaches to diagnose HCC.

III. IMPLEMENTATION OF MACHINE LEARNING AND DEEP LEARNING

Machine learning (ML) and deep learning (DL) are sub-groups of artificial intelligence (AI). ML uses various statistical, probabilistic, and optimization methods that help algorithms learn and apply complex patterns from given large, noisy, and complex datasets [36]. This feature of the ML as well as DL is best suited for medical application where an acceptable generalization is obtained by searching through an n-dimensional space for a given set of biological data using variety of ML and DL techniques and algorithms [37]. Every learning process in ML can be divided in to two phases: (i) approximation of non-classified dependencies from given data (ii) and use of the approximated estimation dependencies to predict novel results from the systems [38]. Further, ML methods can be classified into two common main methods such as (i) supervised learning and (ii) unsupervised learning [39]. In supervised ML the data fed to ML algorithm for training is labeled whereas the training data for unsupervised ML is unlabeled. It is up to the ML model used to find

and apply the underlying patterns in the data and in case of supervised learning this leads to classification and regression. Similarly, in unsupervised learning it refers to clustering and association [40]. The unprecedented advances in cloud and GPU computing have revolutionized the use of ML in medical applications such as cancer diagnosis and detection [41]. More focus is paid on the use of AI in Histopathology, radiology, biomarker-based diagnostic, and use of statistical models in diagnostic of cancers. All these approaches are step by step discussed with benefits, challenges, and future recommendations.

A. AI IN HISTOPATHOLOGY

Histopathology refers to the diagnosis and study of the diseases of tissues where the examination of tissues or cells is carried out using microscope (Figure 1). Biopsy tissues obtained and processed overnight are visualized using sophisticated imaging techniques using microscope. In histopathology laboratory, a single diagnosis typically requires a histopathologist with experience of more than 10 years to correctly identify diseased tissue through microscope [42]. Thus, despite being widely used, traditional histopathology techniques suffer from limitations such as time consuming, requires high level of responsibility, and sufficient expertise. To overcome these limitations, ML and DL methods are recently used in cancer diagnosis. However, at first, technological advances and use of ML methods were limited to quality assurance certain research applications [43]. Unprecedent development in AI and ML methods with rise in computational power has led to realization of AI tools that can be readily deployed to accurately diagnose cancer disease; is remarkable advancement. Additionally, for evaluation of AI models in Hepatocellular Carcinoma (HCC) typically involves several key metrics, which are generally used to assess the performance of machine learning and artificial intelligence systems in medical imaging and diagnosis. Accuracy: This is the most straightforward metric, representing the proportion of correct predictions (both true positives and true negatives) made by the model out of all predictions. Accuracy is a good initial indicator but can be misleading in imbalanced datasets where one class is much more prevalent. Sensitivity (True Positive Rate): Sensitivity measures the proportion of actual positive cases (patients with HCC) correctly identified by the AI model. This metric is crucial in medical diagnostics to ensure that patients with the disease are not missing. Specificity (True Negative Rate): Specificity assesses the proportion of actual negative cases (healthy individuals or those without HCC) correctly identified. High specificity means the model is good at avoiding false alarms. Precision (Positive Predictive Value): This metric evaluates the proportion of positive identifications that were actually correct. In the context of HCC, it reflects how many of the patients identified by the model as having HCC actually have the disease. Recall: This is another term for sensitivity. It measures the model's ability to detect all

relevant instances (all patients with HCC). F1 Score: The F1 score is the harmonic mean of precision and recall. It is a useful measure when you need to balance precision and recall, particularly in cases of uneven class distribution. Area Under the Receiver Operating Characteristic (ROC) Curve (AUC-ROC): This is a performance measurement for classification problems at various thresholds settings. The ROC is a probability curve, and the AUC represents the degree or measure of separability. It tells how much the model is capable of distinguishing between classes. Area Under the Precision-Recall Curve (AUC-PR): This metric is used particularly in cases where there is a significant imbalance in the observations between the two classes. Confusion Matrix: While not a single metric, the confusion matrix is a table layout that allows visualization of the performance of an algorithm. It shows the counts of true positives, false positives, true negatives, and false negatives. Thus, the applications can be extended to detect and count cells and to identify the morphology of cancerous cells for early diagnosis. ML methods can also be used to classify the tissues images obtained from microscopy as healthy or cancerous. Such methods as shown in Table 1 (AI in Histopathology) can be applied to whole images, where the trained algorithms automatically extract the patterns and based on the previous training.

In this connection, a DL based architecture called NucleiSegNet [15] was developed for diagnostic of the hepatocellular carcinoma by locating the nuclei using the image segmentation of the histopathological data (Figure 3(a) [44]. Nuclei segmentation is referred to as important requirement to be fulfilled for the stained histopathological images prior to thinking about the computer-supported diagnostic systems for cancer such as training DL Model as shown in Figure 3(b). Automated segmentation of the nuclei of the histopathological images enables the qualitative and quantitative analysis at scale. This problem is handled by a study which presents a powerful DL-based network architecture trained for the nuclei segmentation of images obtained by H&E-stained liver cancer histopathology [45]. This is carried out in three steps consisting of residual block, bottle neck, and attention decoder block. The novel proposed block here is residual block which enables the extractions of high-level semantic maps followed by efficient object localization enabled by attention decoder [46]. Here the false positives are decreased, and the performance is improved. Upon deployment for nuclei segmentation, NucleiSegNet showed results better than the state of art methods for nuclei segmentation. The deployment generally follows the procedure as shown in Figure 3(c). This work also contributed by introducing the liver dataset named KMC consisting of the H&E staining images of liver cancer histopathology with labeled nuclei. Like KMC, the other databases containing the information about the cancers are summarized in Table 2 [47]. Similarly, another study done in proposes a new DL-based framework (LiverNet) for classification of cancer. They handle it as multi-class classification of

liver hepatocellular carcinoma (HCC) tumor histopathology images [48]. It enables the automatic diagnosis of four subtypes of liver hepatocellular carcinoma using the liver histopathology images as training samples. The multi-class classification places the data in one of the four categories: non-cancerous class, low sub-type liver HCC tumor, medium sub-type liver HCC tumor, and high sub-type liver HCC tumor. Moreover, in LiverNet, BreastNet architecture has been extended by addition of atrous spatial pyramid pooling (ASPP) which enables the multi-scale feature extraction in histopathological image data [48]. Additionally, study results demonstrate that the LiverNet outperforms the already existing, multi-class classification, state-of-art architectures for histopathological images by 2%. The datasets used in this study are TCGA-LIHC and KMC whereas the KMC was labeled by NucleiSegNet [49]. The study claims to be the first to provide the proof-of-concept demonstration of multi-class handling for histopathological data. The accuracy obtained is 90-93% on the KMC liver dataset with fewer parameters and floating points. The motivation behind the KMC was the shortage of properly labeled histopathological data [50]. Apart from them, a similar study presents a DL based framework called HistoCAE [31] in which a multi-resolution CAE based framework is used for image reconstruction [51]. It is further continued by patch-based classification of each histopathology image as tumor and non-tumor.

Despite much progress reported in the form DL and ML based developed frameworks for AI applications in histopathology (Figure 3), there are still certain challenges in automated analysis of histopathological images. These challenges to a greater extent are solved by the advanced CNN and RNN architectures such as shown in Figure 3(b). However, the challenges are: first, representation of vivid clinical features is still difficult to attain because the morphological features of histopathological images differ from patient to patient which makes it hard to find the generalizable data patterns [52]. Second, the labeled data even available today is scarce whereas the pixel size of the image (Standard: $100,000 \times 100,000$) is too large to handle. There are numerous cancerous regions and it's difficult to annotate all of them (Third) and finally, the noisy nature of the histopathological images caused by various reasons [53]. To handle these long-standing problems patch-based methods were utilized for large images [54]. One such study is done in where a patch-based CNN classifier was trained for the classification of tissue image (whole slide) [55]. Here special attention was paid to the batch size and parameters because of their effect on performance of networks. Many of the studies tried to solve the problem of scarcity of labeled images and one such study is presented in where global labels were used histopathology images to perform classification of liver cancer [56]. Moreover, the patch-features along with transfer learning were used to realize the features on patch-level and were later combined with multiple instances learning to scale these features to image-level for classification [57]. This

method solves the problem of data scarcity for histological images of liver cancer. Owing to the binary nature of the classification, the images are classified either as abnormal or normal and helps in early diagnosis of liver cancer. This study, considering three articles, proposes the development of non-invasive AI enabled biomedical image data processing tool to handle the otherwise impractical and time-consuming volume of the biomedical data (Figure 1). The ML methods upon fully developed thus can be used to characterize the tissue at various scales such as tumor appearance in histopathology, radiology, and the cell morphology analysis for automatic diagnosis of malignancy. Summary of these studies for application in diagnostic of viral cancers (Hepatocellular carcinoma) is given in Table 1, whereas Table 2 mentions about some important databses. Moreover, the degree of malignancy in liver cancer can be categorized into three types: poorly, moderately, and well differentiated. The differentiation of different levels is important for diagnosis and treatment of liver cancer. To fulfil this need, an attention (DL) mechanism-based classification study is presented in [31] where five different classified models were used and SENet model achieved the highest classification accuracy of 95.27%. Thus, this study, like others, demonstrate the potential of ML and DL based methods in sorting out the time consuming and laborious problems faced in previously existing manual techniques (Figure 3).
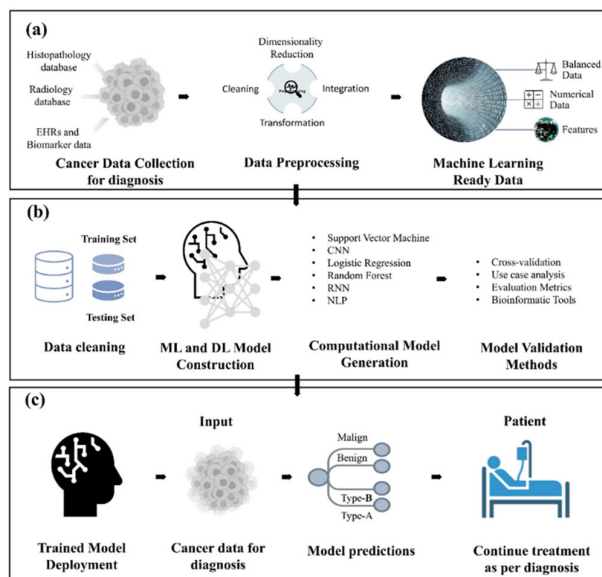
FIGURE 3. A process of cancer diagnosis using ML and DL techniques. a) shows the process of collecting cancer related data from different databases and once the data is collected then it is cleaned to make it ready for training the models, b) shows the splitting of the data into training and testing followed by the ML and DL model development, training, and testing, and finally c) shows the deployment of trained models to predict weather given new data (patient data) is classified as cancer or not-cancer.

B. AI IN RADIOLOGY

Radiology like histopathology uses medical imaging (from different parts of the body) to diagnose disease in humans and animal. In diagnostic radiology [32] some of the most widely used imaging exams include x-ray, MRI, ultrasound, CT scan, and PET scan (Figure 3 (a)).

Despite being widely adopted, traditional radiology suffers from some limitations such as lack of reproducibility, especially for textural features of hepatocellular carcinoma [33] and these changes negatively affect the radiomic signatures. There is no standard protocol on the count of texture features as well as on the development of prediction models which hampers the generalizability of the results. A major limitation is lack of data and restricted data sharing between institutions because of legal concerns, make it more challenging. Thanks to the rise in computational power and amazing progress in DL and ML frameworks (Algorithms and tools) many of the mentioned concerns are now solved. Lot of work combining the radiology or radiomics with ML and DL has been already reported. However, scarcity of data or more accurately properly labeled data is still a big challenge.

Moreover, radiomics with DL integration have recently been used in image-based diagnosis as well as prognosis of various liver diseases [58] and hepatocellular carcinoma is one among them. One such study presented a multi-network-based DL model for risk prediction (Similar to Symptoms) of liver transplantation in case of hepatocellular carcinoma [59]. The database was constructed by extracting magnetic resonance (MR) images from picture archiving and communications system (PACS) followed by the extraction of pathology images [60]. Additionally, a deep learning-based framework utilizing ResNet-18 and support vector machine models was developed to predict microvascular invasion (MVI) in hepatocellular carcinoma using CT images of arterial phase from 309 patients at China Medical University Hospital and 164 patients referred from 54 different hospitals. The models combined CT images (Figure 4-(1-6)) and patients' clinical factors (Figure 4-(7-8)), with the ResNet-18 model exhibiting the highest accuracy, achieving an AUC of 0.845. Its performance was consistent in external validation with an AUC of 0.777. The model's effectiveness was further confirmed through Grad-CAM visualization, demonstrating its ability to focus on relevant imaging features for MVI prediction (Figure 4) [116].

Similarly, another study used MR images for automatic diagnosis of hepatocellular carcinoma. The MRIs used in the said study were multiphasic with improved-contrast using the T1-weighted breadth hold sequences [61]. Additionally, a systematic analysis of over 1500 genome-wide DNA methylation arrays [120] was conducted across multiple studies to identify a distinctive methylation signature for HCC. This study utilized a machine learning pipeline to pinpoint differentially methylated regions in HCC, linked to the repression of genes [122] associated with cancer progression. A unique signature comprising 38 DNA methylation regions was developed, yielding a high-precision HCC detection score. This score demonstrated remarkable efficacy, correctly identifying 96% of HCC tissue samples with 98% precision in an independent dataset. It also effectively distinguished cell-free

TABLE 1. Application of AI (ML and DL) for diagnosis of Hepatocellular carcinoma using AI in histopathology, radiology, biomarker based, and EHRs based diagnosis.

	Method	Datasets	Evaluation criteria	Reference
		1) AI in Histopathology		
1	Deep Learning based multi-class classification	TCGA-LIHC dataset KMC Dataset	Accuracy and F1 score	[48]
2	Deep learning-based architecture for segmentation	H&E-stained histopathology images and KMC liver	Dice and Jaccard Index, F1 score	[75]
3	Deep learning-based classification on Global labels	TCGA,	F1 score, Recall, Precision	[56]
4	Deep learning attention mechanism	liver cancer histopathological images were obtained from research dataset	Precision, Recall, F1 Score,	[76]
5	Convolutional autoencoder	PAIP challenge 2019 - MICCAI 2019	Accuracy, F1 score, Precision, Recall	[77]
6	ASPPU-Net	TCGA Triple negative breast cancer dataset	Aggregated Jaccard index, F1 score	[78]
		2) AI in Radiology		
7	Deep learning multi-network model	MRI Images	AUC, F1 score	[79]
8	Deep convolutional neural network (DCNN) with U-net architecture	MRI Images	Accuracy, precision, dice similarity coefficient (DSC)	[80]
9	Extreme Learning Machine (ELM)	CT Images	ROC Curve, Sensitivity, False positive rate	[81]
10	Image Processing Algorithm – Fuzzy linguistic constant (FLC)	CT Images	Accuracy, structural similarity index	[66]
11	SVM and ELM	Ultrasound tissue data	Accuracy	[82]
12	Attention mechanism guided deep learning algorithm	367 two-dimensional ultrasound images from 367 individual livers	ROC-AUC	[83]
13	Deep learning-based mechanism	4386-slice computed tomography (CT) images	Sensitivity and false positives per patient	[84]
14	Deep learning-based algorithm	1350 multiphase CT scans of 1280 hepatic malignancies	Sensitivity, and false positive rate	[85]
3)		*Biomarker based disease prediction*		
15	Automated machine learning for diagnostic	TG-GATEs	Precision, Recall	[86]
16	Machine learning in TCGA	TCGA, GEO	Log2FC, P-value, FDR	[87]
17	cfDNA scoring system	Korea Biobank Network	Accuracy	[88]
18	Multiple feature selection – Akaike information criterion	TCGA	-	[89]
19	Analysis of gene co-expression network	GEO	Accuracy, sensitivity, and specificity	[90]
20	Machine learning of GWAS databases	GWAS, GEO, TCGA	Accuracy	[91]
21	SVM based machine learning method	GEO	Accuracy, sensitivity, and specificity	[92]
		4) EHRs based Diagnosis		
20	Deep learning model	A total of 52 983 samples (98% men)	Specificity and Sensitivity	[93]

TABLE 1. *(Continued.)* Application of AI (ML and DL) for diagnosis of Hepatocellular carcinoma using AI in histopathology, radiology, biomarker based, and EHRs based diagnosis.

21	Ontological features based EHRs	112 abdominal computed tomography imaging examination reports, consisting of 59 cases of hepatocellular carcinoma (HCC)	Sensitivity and Accuracy	[94]
22	Machine learning and ridge regression scores	n = 86,804; 6,821 HCC (Dataset) n = 37,202; 2,875 HCC (Validation)	AUROC	[95]
23	Deep learning model for minimal feature extraction	47,945 individuals containing 9553 patients with HCC	AUROC, Sensitivity, and Specificity	[96]
24	Algorithmic validation for an observational study	Patients = 339 HCCs	Accuracy, Sensitivity, and Specificity	[97]
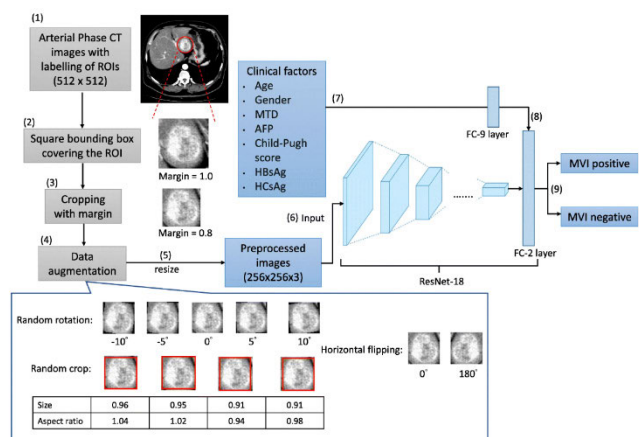
FIGURE 4. This flow chart illustrates the development of a ResNet-18 based model for preoperative MVI prediction in HCC, detailing steps from labeling and preprocessing CT images to integrating clinical factors and model training. The process includes image augmentation techniques like rotation, cropping, and flipping, with examples provided. The final model predicts MVI status by combining image data with clinical factors [116].
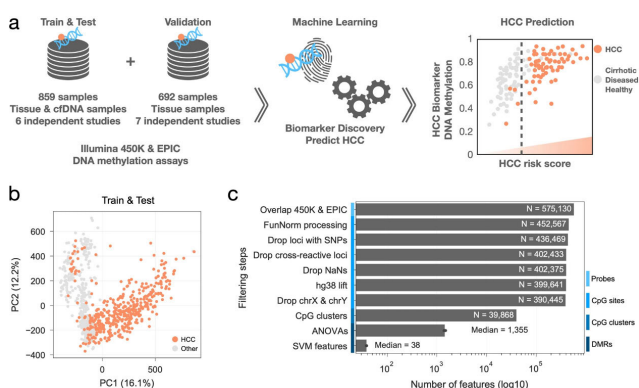
FIGURE 5. (a) A diagram showing datasets assembled for discovering HCC DNA methylation biomarkers using machine learning and constructing an HCC risk score. (b) PCA of Train & Test DNA methylation dataset highlighting HCC samples with explained variances. (c) Steps in feature reduction, including probes, CpG sites, clusters, and DMRs, in the processing and feature discovery pipeline [117].

DNA (cfDNA) of tumor samples from healthy controls and identified cfDNA from patients with other tumors, including colorectal cancer [117], [121] as shown in Figure 5.

DL model was consisted of deep convolutional neural network (DCNN) (also shown in Figure 3(b)) with U-net architecture and was trained using train-test split method where training data was 70%, followed by 15% as validation data, and remaining 15% was used as testing data [62]. To realize the effectiveness of DCNN-Unet, comparison was made between model prediction and manual work. 73~75% accuracy was obtained on test and validation sets respectively with dice similarity coefficients (DSC) adjusted to >0.2. Thus, it makes the process of diagnosis much faster, reliable, and error free. However, there is no external validation performed to validate the work against the bias and data leakage issues. One of the reasons for this inherent limitation is data scarcity (suitable). Another problem in dealing with medical imaging is, high resolution and vast feature space which makes it challenging for the image segmentation to identify and use the valuable features [63]. To sort this problem, an ensemble based extreme learning machine (ELM) study is done in which works on liver tumor diagnosis using the random feature space [64]. In this study, tumor detection is handled as binary classification problem with two-classes. Moreover, the ELM used here, is equipped with fast learning (than SVM and other algorithms) and commendable generalization and have been used already as a single-hidden-layer feed-forward neural network (SLFN) architectures.

Additionally, past years have witnessed the fruitful application of image processing techniques applied on medical imaging for diagnostics. Advances in imaging algorithms extend support to radiologists for timely diagnosis of hepatocellular carcinoma and other related disorders [65]. Apart from conventional DL and ML algorithms, fuzzy models are also used for diagnosis in hepatocellular carcinoma. One such example study is given in where a new fuzzy linguistic constant (FLC) is introduced for computer-aided automated diagnosis of the liver cancer from CT images (Low contrast) [66]. Fuzzy membership functional is developed to classify the contrast-enhanced images as cancerous and non-cancerous [123]. To assess the extracted features, structural similarity index is utilized which in turn tells whether the tumor is malign or benign. This work overcomes the limitation and performs the external validation of the

TABLE 2. Summary of cancer related databases.

	Database name	Type	Availability	Programmatic access
1	TCGA	molecularly characterized over 20,000 primary cancer and matched normal samples spanning 33 cancer types.	Open source	-
2	GEO	Genomics datasets, platforms etc.	Open source	Rest API
3	ClinicalTrials.gov	421,553 clinical research studies	Open source	-
4	The International Cancer Genome Consortium (ICGC)	Cancer projects 86, Cancer primary sites 22, Donor with molecular data in DCC 22,330, Total Donors 24,289, Simple somatic mutations 81,782,588	Open source	-
5	Integrative Onco Genomics	Cancer types 66, Cohorts 221, Samples 28,076, Mutations 203,003,747, Drivers 568	Open source	Rest API
6	OncoDB.HCC	Copy number, Gene Expression, QTL	Open source	Rest API
7	MutaGene	Mutation	Open source	Rest API
8	Network of Cancer Genes	411 altered in coding sequence and 69 altered in non-coding sequence	Open source	-
9	The CEO Life Sciences Consortium	clinical trial delivery	-	-
10	Oncomine	Gene Expression	Close source	-
11	Ocimum Bio Solutions	Gene Expression	Close source	-
12	CancerDR	148 anticancer drugs and their effectiveness against around 1000 cancer cell lines	-	-
13	CancerResource	A database source which contains compounds relevant to cancer and protein interactions.	Open source	-
14	COSMIC	-	Open source	-
15	Genomics of Drug Sensitivity in Cancer	1000 human cancer cell lines, 100s of compounds, and drug response data	Open source	Rest API
16	Noncode	173,112 human lncRNA transcripts, 96,411 number of lncRNA genes	Open source	Rest API
17	canSAR Black	-	Open sources	Rest API

developed models on the dataset containing the information from 179 clinical cases whereas the information for other such databases is summarized in Table 2 [67]. The class distribution is 98 and 81 for benign and malignant respectively for liver tumors. SVM applied to extracted features yield the accuracy of 98.7% whereas the presented segmentation method yields the enhanced detection value of 78% [68]. Such algorithms are valuable tools for radiologists in diagnostic of tumors. In addition to CT, some studies have focused on combination of liver ultrasound and DL to perform classification and diagnosis of hepatocellular carcinoma. Classification

of LI-RADS based ultrasound monitored reports for hepatocellular carcinoma diagnosis was performed Morgan et al. [69]. Not only in hepatocellular carcinoma, but liver ultrasound images have also been used to perform the cirrhosis diagnosis. Other numerous studies can be found demonstrating the use of computational (ML, DL etc.) methods for the diagnosis of hepatocellular carcinoma using the radiology (radiomics) data [70], [71], [72].

No doubt, a lot of progress and success have been observed in DL and ML applications to radiology, however, there are some inherent limitations impeding the true potential of the

subject. Major limitation is lack of generalization in radiomic classifiers, owing to the variance in radiomic features from person to person, dependent on protocols followed, and feature extraction approaches [73]. The limitation is partially solved by image preprocessing such as gray-level normalization and standardization of resolution. In this connection, an algorithm was introduced recently which enable the reduction of variations in radiomic features and enabled the feature analysis using multicenter image data [74]. Future research is directed to design and implement optimal, fast, and reproducible image processing algorithms to avoid the variations in future and improve the generalizations of radiomic models to enhance the diagnosis abilities in hepatocellular carcinoma.

C. BIOMARKER BASED DISEASE PREDICTION

Techniques like integration of multi-omics analysis, use of ML, and DL have become important tools for biomarker screening and disease (Figure 3) HCC diagnosis as shown in Table 1. ML and DL algorithms once trained properly can easily be deployed to diagnose the disease [118] form patient samples [98]. Variety of algorithms are trained on various biomarker related data types, and all are aligned with a common goal for early diagnosis of disease (HCC) which is necessary to treat HCC. Detection of early and potential tumors in HCC, is normally done by monitoring of vulnerable groups utilizing the abdominal ultrasonography in the presence and absence of serum analysis of α-fetoprotein (AFP) [99]. It comes with the challenge of limited sensitivity for which two current approaches are being used, first one is to use the already known empirical biomarkers (derived) and another one uses circulating nucleic acid biomarkers consisting of cell free DNA and RNA [100]. Modern molecular biology aligned with ML techniques closely allied to underlying biology of cancer [124]. The approaches are considered as promising opportunities for obtaining the timely diagnosis with additional functionalities on the plate.

To address these challenges, a study using the integration of gene expressions (database information Table 2) of cfDNA to predict the clinic-pathological response of HCC patients is done in [62]. Since HCC can be diagnosed by circulating cell-free DNA (cfDNA) and it is also referred to as predictive biomarker for HCC. The blood biomarkers are absent for early diagnosis in HCC because of which the mortality rate is very high. In this study a new ML based scoring system is proposed called cfDNA HCC. As already stated, it integrates the expression profiles of cfDNA which paved the way for the prediction of clinic-pathological response of the patients suffering from HCC. This study proves the use of known biomarkers could be useful for diagnosis of HCC. In this connection, studies have been conducted which are focusing on the discovery of robust biomarkers for HCC. One such study is done in [63] where six different methods for recursive feature elimination were used to select the gene signatures from TCGA liver cancer data (Table 2). It was hypothesized that, the genes shared among the adopted six

subsets would be regarded as robust biomarkers in HCC. Statistical interpretation for feature selection in ML was performed using the Akaike information criterion (AIC) to explain the optimization process of feature selection. The biomarkers shortlisted in this study by the process of backward logistic stepwise regression were found to contain in the already known biomarkers. Another similar study using supervised ML identifies the biomarkers levels that can be used as diagnostic tool to classify the HCC [64]. Moreover, large-scale transcriptomic data has also been used for identification of diagnostic biomarkers [65]. Previously, efforts were put to make the genetic biomarkers but due to lack of data both in quantity and diversity it was not possible. Now with the emergence of large-scale transcriptomic data this study [65] presents identification of diagnostic biomarkers in HCC. The profiled dataset contained in total 2316 positive and 1665 negative (non-tumorous) samples. These data samples were obtained from four different studies using various types of profiling techniques. Based on overlapping genes in all datasets 26 genes were found highly expressed. Different feature selection techniques were used to finally select three genes (FCN3, CLEC1B, and PRC1) as diagnostic biomarkers in HCC. In order to overcome the existing limitation this study used a systematic approach to identify genetic biomarkers for HCC diagnosis which makes it applicable to wide range of platforms.

No doubt, the rise in computational power, generation of multi-omics data at scale, advances in approaches to mine the massive datasets, and publicly available databases providing useful data has been beneficial for the discovery of diagnostic biomarkers for timely diagnosis of cancers such as liver cancer. Despite these steps, there are still some limitations that need to be addressed to uncover the full potential of biomarker-based disease prediction strategy. A major limitation is the variations found in transcriptomic signatures because of use of different platforms, protocols, and person to person variation. Future work should focus on resolving these issues and the strategies based on ML and DL which could encompass these variations and generalize well.

D. EHRS BASED DIAGNOSIS

Electronic health records (EHR) provide accurate, up to date, and comprehensive information about the patient enabling coordinated and effective patient care [101]. Sharing EHRs makes the diagnosis of disease more easily, reduces errors, and provides faster and safer care when reliable and precise patient health information is available [102]. A well-structured, end-to-end EHR system supports the critical clinical decision of disease diagnosis.

Recently, DL and ML approaches have been applied to EHRs to diagnose viral cancers such as hepatocellular carcinoma in patients. Hepatocellular carcinoma was predicted in patients with hepatitis C cirrhosis by Ioannou et al [93]. DL recurrent neural network (RNN) models were applied to raw EHRs, and the resulting model outperformed the state

of art traditional regression models in diagnosing the risk of occurring of HCC [103]. In case of data, over the years 52983 samples were collected, out of which 98% samples were collected from male patients [104]. The analysis of the EHRs revealed that the patients who developed HCC were older than the ones who did not experience it [105]. The results form RNN models showed the 80% of all HCCs that took place. RNN models are also powerful DL models for medical imaging applications as shown in Figure 3(b). Thus, the presented framework trained on EHRs outperformed conventional linear regression models. ML was used to signify EHRs and identify liver cancers which is caused by HCCs virus in more than 90% of the cases [96]. The key features demonstrated by the proposed framework was fast prediction with risk modelling [106]. In this a set of 112 abdominal CT images which contained 59 HCCs and 53 Non-HCCs were arranged from four Hong Kong based hospitals [94]. Ontological features here were extracted which formed the basis of primary predictor panel [107]. Pearson's correlation coefficient was used to quantify the relationship between every pair of HCCs and Non-HCCs samples [108]. The model trained on these features when deployed obtained 84.7% sensitivity and 88.4% overall accuracy [109]. The feature extraction using the Pearson's correlation coefficient resulted in better accuracy than without the it. Thus, the ML and DL architecture show a great potential in identifying viral cancers when properly trained on EHRs and deployed.

Moreover, one after the other methods and approaches are introduced making use of EHRs in combination with computational methods (ML and DL) (Table 2) in diagnostic of viral and non-viral hepatocellular carcinoma (Figure 3). One such state of art study is done in where a novel ML model is performing better than then already implemented risk scores in predicting the hepatocellular carcinoma in patients with chronic viral hepatitis [110]. Comprehensive clinical data is fed to ML model and based on the HCC ridge score presence or absence of viral cancer is diagnosed [111]. At the same time, this study has developed the statistical model to indirectly benchmark novel ML model. It shows that the novel ML model introduce here generate correct risk scores for HCC in patients with chronic viral hepatitis. The HCC ridge score developed by this study was more accurate than the risk scores already developed [112]. Additionally, a validation study was done in where EHRs (inpatient) were used to find out the existence of cirrhosis in patients with hepatocellular carcinoma [113]. This study did not include any AI assisted module. However, this study was able to perform the diagnosis using the inpatient EHRs. Review of these studies as mentioned in Table 1 prove that the use of EHRs coupled with computational (AI) techniques can open the door to success in diagnostic of viral cancers. However, there are certain limitations associated with the EHRs, ML, and DL models [114]. First, EHRs are raw, heterogenous, and noisy data which negatively affect the performance of the models. Second, it takes a lot of effort to make the EHRs

ML ready [115]. When it comes to ML, owing to the high dimensional and diverse nature of the data, it is hard for the linear ML models to learn and generalize the data well. In addition, the feature engineering for conventional ML is troublesome. Whereas the DL models are hard to interpret. This emphasize the need for sufficient time to be invested for the processing of EHRs to suitable format followed by the interpretability of DL models. Once these limitations are overcome and newer more diverse data adopting algorithms are developed, only then the true diagnostic potential from EHRs can be harvested.

IV. DISCUSSION AND CHALLENGES

EHRs are inherently raw, noisy, and heterogeneous, comprising unstructured text, images, lab results, and more. This diversity and inconsistency in data quality can significantly impede the performance of ML models, making it challenging to extract meaningful insights without extensive preprocessing and standardization efforts. Moreover, Data Preparation and Accessibility is also an important factor. Transforming EHRs into a format that is amenable to ML algorithms requires significant effort. The process involves data cleaning, normalization, feature extraction, and handling of missing values, which is both time-consuming and resource intensive. Furthermore, issues related to data privacy, security, and sharing can limit the accessibility of comprehensive datasets necessary for training robust ML models. While ML and DL models hold great potential in identifying complex patterns and associations within large datasets, their application in healthcare, particularly in diagnostics, is hampered by the high dimensionality and variability of medical data. Linear ML models may struggle to capture the intricate relationships present in the data, necessitating more sophisticated, yet computationally intensive, DL models. One of the significant hurdles in the adoption of DL models in clinical settings is their "black box" nature, which makes it difficult for clinicians to understand how these models arrive at a particular diagnosis or prediction. This lack of interpretability can hinder trust in AI-based diagnostic systems and poses challenges in clinical decision-making. Additionally, ensuring that ML models can generalize well to new, unseen data is a critical challenge. Models may perform well on the data they were trained on but fail to maintain accuracy when applied to data from different sources or populations. Rigorous validation studies, including external validation on diverse datasets, are essential to establish the reliability and applicability of ML models in clinical practice. Finally, the successful implementation of ML-based diagnostics in clinical settings requires seamless integration into existing workflows, with minimal disruption to clinical practices. This integration must also address clinicians' needs, providing intuitive interfaces and decision support that enhance, rather than complicate, the diagnostic process.

To overcome these challenges and fully realize the potential of ML in the diagnosis of HCC, several future directions

can be pursued such as advancements in data processing aimed to developing more sophisticated algorithms for data preprocessing and feature engineering can help mitigate the issues of data quality and heterogeneity, making EHRs more amenable to ML analysis. Similarly, adopting federated learning approaches can address data privacy and accessibility issues by enabling ML models to be trained across multiple decentralized data sources without needing to share the data directly and investing in research and development of explainable AI models can enhance the interpretability and transparency of ML-based diagnostic tools, fostering trust and acceptance among healthcare professionals. Finally, close collaboration between computer scientists, data scientists, clinicians, and other healthcare professionals is crucial to ensure that ML models are developed with a deep understanding of clinical needs and constraints, leading to more effective and user-friendly diagnostic tools.

V. CONCLUSION

Incorporating Artificial Intelligence (AI) into oncological diagnostics, particularly for hepatocellular carcinoma (HCC), enhances traditional methods by leveraging advanced Machine Learning (ML) and Deep Learning (DL) techniques alongside increased computational power to analyze extensive datasets. This synergy facilitates the nuanced examination of image-based data in conjunction with biomarker datasets, crucial for early detection of both viral and non-viral HCC, with AI algorithms designed to supplement rather than replace existing clinical methodologies. The integration of AI not only promises to refine diagnostic accuracy but also to contribute to the prognostic assessment, guiding more personalized treatment approaches. The path forward should focus on developing interpretable AI models that engender clinician and patient trust, ensuring seamless integration into clinical workflows, conducting rigorous validation studies to ensure efficacy and reliability across diverse populations, and addressing ethical and regulatory considerations to safeguard patient privacy and data security. Such efforts necessitate a multidisciplinary approach, underscoring the importance of collaboration between data scientists, clinicians, ethicists, and policymakers to fully realize the potential of AI in transforming HCC diagnostics and patient care.

REFERENCES

[1] H. B. El-Serag, "Epidemiology of viral hepatitis and hepatocellular carcinoma," *Gastroenterology*, vol. 142, no. 6, pp. 1264–1273, May 2012.

[2] J. D. Yang, P. Hainaut, G. J. Gores, A. Amadou, A. Plymoth, and L. R. Roberts, "A global view of hepatocellular carcinoma: Trends, risk, prevention and management," *Nature Rev. Gastroenterol. Hepatol.*, vol. 16, no. 10, pp. 589–604, Oct. 2019.

[3] I. Sghaier, S. Zidi, L. Mouelhi, E. Ghazoueni, E. Brochot, W. Almawi, and B. Loueslati, "TLR3 and TLR4 SNP variants in the liver disease resulting from hepatitis B virus and hepatitis C virus infection," *Brit. J. Biomed. Sci.*, vol. 76, no. 1, pp. 35–41, Jan. 2019.

[4] M. Khalid, S. Manzoor, H. Ahmad, A. Asif, T. A. Bangash, A. Latif, and S. Jaleel, "Purinoceptor expression in hepatocellular virus (HCV)-induced and non-HCV hepatocellular carcinoma: An insight into the proviral role of the P2X4 receptor," *Mol. Biol. Rep.*, vol. 45, no. 6, pp. 2625–2630, Dec. 2018.

[5] A. Asif, M. Khalid, S. Manzoor, H. Ahmad, and A. U. Rehman, "Role of purinergic receptors in hepatobiliary carcinoma in Pakistani population: An approach towards proinflammatory role of P2X4 and P2X7 receptors," *Purinergic Signalling*, vol. 15, no. 3, pp. 367–374, Sep. 2019.

[6] T. Huang, J. Behary, and A. Zekry, "Non-alcoholic fatty liver disease: A review of epidemiology, risk factors, diagnosis and management," *Internal Med. J.*, vol. 50, no. 9, pp. 1038–1047, 2020.

[7] K. Hamesch and P. Strnad, "Non-invasive assessment and management of liver involvement in adults with Alpha-1 antitrypsin deficiency," *Chronic Obstructive Pulmonary Diseases: J. COPD Found.*, vol. 7, no. 3, pp. 260–271, 2020.

[8] K. Patel and G. Sebastiani, "Limitations of non-invasive tests for assessment of liver fibrosis," *JHEP Rep.*, vol. 2, no. 2, Apr. 2020, Art. no. 100067.

[9] Z. Zhang, Y. Zhao, A. Canes, D. Steinberg, and O. Lyashevska, "Predictive analytics with gradient boosting in clinical medicine," *Ann. Translational Med.*, vol. 7, no. 7, p. 152, Apr. 2019.

[10] Y. Masugi, T. Abe, H. Tsujikawa, K. Effendi, A. Hashiguchi, M. Abe, Y. Imai, K. Hino, S. Hige, M. Kawanaka, G. Yamada, M. Kage, M. Korenaga, Y. Hiasa, M. Mizokami, and M. Sakamoto, "Quantitative assessment of liver fibrosis reveals a nonlinear association with fibrosis stage in nonalcoholic fatty liver disease," *Hepatology Commun.*, vol. 2, no. 1, pp. 58–68, 2018.

[11] K. Y. Ngiam and I. W. Khor, "Big data and machine learning algorithms for health-care delivery," *Lancet Oncol.*, vol. 20, no. 5, pp. e262–e273, May 2019.

[12] M. Subramanian, A. Wojtusciszyn, L. Favre, S. Boughorbel, J. Shan, K. B. Letaief, N. Pitteloud, and L. Chouchane, "Precision medicine in the era of artificial intelligence: Implications in chronic disease management," *J. Transl. Med.*, vol. 18, no. 1, pp. 1–12, Dec. 2020.

[13] H. B. El-Serag, J. A. Marrero, L. Rudolph, and K. R. Reddy, "Diagnosis and treatment of hepatocellular carcinoma," *Gastroenterology*, vol. 134, no. 6, pp. 1752–1763, 2008.

[14] M. Sherman, "Hepatocellular carcinoma: Epidemiology, surveillance, and diagnosis," *Seminars Liver Disease*, vol. 30, no. 1, pp. 3–16, Feb. 2010.

[15] J. Neuberger and O. Cain, "The need for alternatives to liver biopsies: Non-invasive analytics and diagnostics," *Hepatic Med., Evidence Res.*, vol. 13, pp. 59–69, Jun. 2021.

[16] S. G. Hübscher, "Histological assessment of non-alcoholic fatty liver disease," *Histopathology*, vol. 49, no. 5, pp. 450–465, Nov. 2006.

[17] V. Granata, R. Grassi, R. Fusco, A. Belli, C. Cutolo, S. Pradella, G. Grazzini, M. La Porta, M. C. Brunese, F. De Muzio, A. Ottaiano, A. Avallone, F. Izzo, and A. Petrillo, "Diagnostic evaluation and ablation treatments assessment in hepatocellular carcinoma," *Infectious Agents Cancer*, vol. 16, no. 1, pp. 1–22, Dec. 2021.

[18] P. Angulo, D. E. Kleiner, S. Dam-Larsen, L. A. Adams, E. S. Bjornsson, P. Charatcharoenwitthaya, P. R. Mills, J. C. Keach, H. D. Lafferty, A. Stahler, S. Haflidadottir, and F. Bendtsen, "Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease," *Gastroenterology*, vol. 149, no. 2, pp. 389–397, Aug. 2015.

[19] E. C. Stack, C. Wang, K. A. Roman, and C. C. Hoyt, "Multiplexed immunohistochemistry, imaging, and quantitation: A review, with an assessment of tyramide signal amplification, multispectral imaging and multiplex analysis," *Methods*, vol. 70, no. 1, pp. 46–58, Nov. 2014.

[20] B. C. Asay, B. B. Edwards, J. Andrews, M. E. Ramey, J. D. Richard, B. K. Podell, J. F. M. Gutiérrez, C. B. Frank, F. Magunda, G. T. Robertson, M. Lyons, A. Ben-Hur, and A. J. Lenaerts, "Digital image analysis of heterogeneous tuberculosis pulmonary pathology in non-clinical animal models using deep convolutional neural networks," *Sci. Rep.*, vol. 10, no. 1, pp. 1–14, Apr. 2020.

[21] D. Patel, Y. Shah, N. Thakkar, K. Shah, and M. Shah, "Implementation of artificial intelligence techniques for cancer detection," *Augmented Hum. Res.*, vol. 5, no. 1, pp. 1–10, Dec. 2020.

[22] H. Liao, Y. Long, R. Han, W. Wang, L. Xu, M. Liao, Z. Zhang, Z. Wu, X. Shang, X. Li, J. Peng, K. Yuan, and Y. Zeng, "Deep learning-based classification and mutation prediction from histopathological images of hepatocellular carcinoma," *Clin. Transl. Med.*, vol. 10, no. 2, pp. 310–318, Jun. 2020.

[23] A. Kiani et al., "Impact of a deep learning assistant on the histopathologic classification of liver cancer," *NPJ Digit. Med.*, vol. 3, no. 1, pp. 1–6, Feb. 2020.

[24] C. Balsano, A. Alisi, M. R. Brunetto, P. Invernizzi, P. Burra, F. Piscaglia, and D. Alvaro, "The application of artificial intelligence in hepatology: A systematic review," *Digestive Liver Disease*, vol. 54, no. 3, pp. 299–308, 2021.

[25] J. Ricke, M. Seidensticker, and K. Mohnike, "Noninvasive diagnosis of hepatocellular carcinoma in cirrhotic liver: Current guidelines and future prospects for radiological imaging," *Liver Cancer*, vol. 1, no. 1, pp. 51–58, 2012.

[26] T. Banzato, E. Fiore, M. Morgante, E. Manuali, and A. Zotti, "Texture analysis of B-mode ultrasound images to stage hepatic lipidosis in the dairy cow: A methodological study," *Res. Veterinary Sci.*, vol. 108, pp. 71–75, Oct. 2016.

[27] A. Jamthikar, D. Gupta, N. N. Khanna, T. Araki, L. Saba, A. Nicolaides, A. Sharma, T. Omerzu, H. S. Suri, A. Gupta, S. Mavrogeni, M. Turk, J. R. Laird, A. Protogerou, P. P. Sfikakis, G. D. Kitas, V. Viswanathan, G. Pareek, M. Miner, and J. S. Suri, "A special report on changing trends in preventive stroke/cardiovascular risk assessment via B-mode ultrasonography," in *Cognitive Informatics, Computer Modelling, and Cognitive Science*. New York, NY, USA: Springer, 2020, pp. 291–318.

[28] T. Ichikawa, K. Saito, N. Yoshioka, A. Tanimoto, T. Gokan, Y. Takehara, T. Kamura, T. Gabata, T. Murakami, K. Ito, S. Hirohashi, A. Nishie, Y. Saito, H. Onaya, R. Kuwatsuru, A. Morimoto, K. Ueda, M. Kurauchi, and J. Breuer, "Detection and characterization of focal liver lesions: A Japanese phase III, multicenter comparison between gadoxetic acid disodium-enhanced magnetic resonance imaging and contrast-enhanced computed tomography predominantly in patients with hepatocellular carcinoma and chronic liver disease," *Investigative Radiol.*, vol. 45, no. 3, pp. 133–141, Mar. 2010.

[29] T. Sersté, V. Barrau, V. Ozenne, M. P. Vullierme, Pierre Bedossa, Olivier Farges, D. C. Valla, V. Vilgrain, V. Paradis, and F. Degos, "Accuracy and disagreement of computed tomography and magnetic resonance imaging for the diagnosis of small hepatocellular carcinoma and dysplastic nodules: Role of biopsy," *Hepatology*, vol. 55, no. 3, pp. 800–806, 2012.

[30] F. Ahmed, A. M. Soomro, A. R. Chethikkattuveli Salih, A. Samantasinghar, A. Asif, I. S. Kang, and K. H. Choi, "A comprehensive review of artificial intelligence and network based approaches to drug repurposing in COVID-19," *Biomed. Pharmacotherapy*, vol. 153, Sep. 2022, Art. no. 113350.

[31] F. Ahmed, J. W. Lee, A. Samantasinghar, Y. S. Kim, K. H. Kim, I. S. Kang, F. H. Memon, J. H. Lim, and K. H. Choi, "SperoPredictor: An integrated machine learning and molecular docking-based drug repurposing framework with use case of COVID-19," *Frontiers Public Health*, vol. 10, pp. 90–107, Jun. 2022.

[32] Q. Lang, C. Zhong, Z. Liang, Y. Zhang, B. Wu, F. Xu, L. Cong, S. Wu, and Y. Tian, "Six application scenarios of artificial intelligence in the precise diagnosis and treatment of liver cancer," *Artif. Intell. Rev.*, vol. 54, no. 7, pp. 5307–5346, Oct. 2021.

[33] J. Calderaro and J. N. Kather, "Artificial intelligence-based pathology for gastrointestinal and hepatobiliary cancers," *Gut*, vol. 70, no. 6, pp. 1183–1193, Jun. 2021.

[34] M. Wang, F. Fu, B. Zheng, Y. Bai, Q. Wu, J. Wu, L. Sun, Q. Liu, M. Liu, Y. Yang, H. Shen, D. Kong, X. Ma, P. You, X. Li, and F. Tian, "Development of an AI system for accurately diagnose hepatocellular carcinoma from computed tomography imaging data," *Brit. J. Cancer*, vol. 125, no. 8, pp. 1111–1121, Oct. 2021.

[35] S. Agarwal, A. S. Yadav, V. Dinesh, K. S. S. Vatsav, K. S. S. Prakash, and S. Jaiswal, "By artificial intelligence algorithms and machine learning models to diagnosis cancer," *Mater. Today, Proc.*, vol. 80, pp. 2969–2975, 2023.

[36] T. Kotsiopoulos, P. Sarigiannidis, D. Ioannidis, and D. Tzovaras, "Machine learning and deep learning in smart manufacturing: The smart grid paradigm," *Comput. Sci. Rev.*, vol. 40, May 2021, Art. no. 100341.

[37] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie, and L. Farhan, "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, pp. 1–74, Mar. 2021.

[38] M. Nikoo, F. Ramezani, M. Hadzima-Nyarko, E. K. Nyarko, and M. Nikoo, "Flood-routing modeling with neural network optimized by social-based algorithm," *Natural Hazards*, vol. 82, no. 1, pp. 1–24, May 2016.

[39] S. B. Kotsiantis, I. Zaharakis, and P. Pintelas, "Supervised machine learning: A review of classification techniques," *Emerg. Artif. Intell. Appl. Comput. Eng.*, vol. 160, no. 1, pp. 3–24, 2007.

[40] D. Cintia Ganesha Putri, J.-S. Leu, and P. Seda, "Design of an unsupervised machine learning-based movie recommender system," *Symmetry*, vol. 12, no. 2, p. 185, Jan. 2020.

[41] M. Pandey, M. Fernandez, F. Gentile, O. Isayev, A. Tropsha, A. C. Stern, and A. Cherkasov, "The transformational role of GPU computing and deep learning in drug discovery," *Nature Mach. Intell.*, vol. 4, no. 3, pp. 211–221, Mar. 2022.

[42] D. C. Fernandez, R. Bhargava, S. M. Hewitt, and I. W. Levin, "Infrared spectroscopic imaging for histopathologic recognition," *Nature Biotechnol.*, vol. 23, no. 4, pp. 469–474, Apr. 2005.

[43] W. Fan, Y. Chen, J. Li, Y. Sun, J. Feng, H. Hassanin, and P. Sareh, "Machine learning applied to the design and inspection of reinforced concrete bridges: Resilient methods and emerging applications," *Structures*, vol. 33, pp. 3954–3963, Oct. 2021.

[44] A. B. Tufail, Y.-K. Ma, M. K. A. Kaabar, F. Martínez, A. R. Junejo, I. Ullah, and R. Khan, "Deep learning in cancer diagnosis and prognosis prediction: A minireview on challenges, recent trends, and future directions," *Comput. Math. Methods Med.*, vol. 2021, pp. 1–28, Oct. 2021.

[45] K. Bera, K. A. Schalper, D. L. Rimm, V. Velcheti, and A. Madabhushi, "Artificial intelligence in digital pathology—New tools for diagnosis and precision oncology," *Nature Rev. Clin. Oncol.*, vol. 16, no. 11, pp. 703–715, Nov. 2019.

[46] S. Lal, D. Das, K. Alabhya, A. Kanfade, A. Kumar, and J. Kini, "NucleiSegNet: Robust deep learning architecture for the nuclei segmentation of liver cancer histopathology images," *Comput. Biol. Med.*, vol. 128, Jan. 2021, Art. no. 104075.

[47] P. Sabitha and G. Meeragandhi, "A dual stage AlexNet-HHO-DrpXLM archetype for an effective feature extraction, classification and prediction of liver cancer based on histopathology images," *Biomed. Signal Process. Control*, vol. 77, Aug. 2022, Art. no. 103833.

[48] A. A. Aatresh, K. Alabhya, S. Lal, J. Kini, and P. P. Saxena, "LiverNet: Efficient and robust deep learning model for automatic diagnosis of subtypes of liver hepatocellular carcinoma cancer from H&E stained liver histopathology images," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 16, no. 9, pp. 1549–1563, Sep. 2021.

[49] C. A. Ferreira, T. Melo, P. Sousa, M. I. Meyer, E. Shakibapour, P. Costa, and A. Campilho, "Classification of breast cancer histology images through transfer learning using a pre-trained inception ResNet V2," in *Proc. Int. Conf. Image Anal. Recognit.* Cham, Switzerland: Springer, 2018, pp. 763–770.

[50] I. Hirra, M. Ahmad, A. Hussain, M. U. Ashraf, I. A. Saeed, S. F. Qadri, A. M. Alghamdi, and A. S. Alfakeeh, "Breast cancer classification from histopathological images using patch-based deep learning modeling," *IEEE Access*, vol. 9, pp. 24273–24287, 2021.

[51] M. Khened, A. Kori, H. Rajkumar, G. Krishnamurthi, and B. Srinivasan, "A generalized deep learning framework for whole-slide image segmentation and analysis," *Sci. Rep.*, vol. 11, no. 1, pp. 1–14, Jun. 2021.

[52] Z. Li, X. Zhang, H. Müller, and S. Zhang, "Large-scale retrieval for medical image analytics: A comprehensive review," *Med. Image Anal.*, vol. 43, pp. 66–84, Jan. 2018.

[53] N. Marini, S. Otálora, H. Müller, and M. Atzori, "Semi-supervised training of deep convolutional neural networks with heterogeneous data and few local annotations: An experiment on prostate histopathology image classification," *Med. Image Anal.*, vol. 73, Oct. 2021, Art. no. 102165.

[54] H. R. Tizhoosh and L. Pantanowitz, "Artificial intelligence and digital pathology: Challenges and opportunities," *J. Pathol. Informat.*, vol. 9, no. 1, p. 38, Jan. 2018.

[55] L. Hou, D. Samaras, T. M. Kurc, Y. Gao, J. E. Davis, and J. H. Saltz, "Patch-based convolutional neural network for whole slide tissue image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2424–2433.

[56] C. Sun, A. Xu, D. Liu, Z. Xiong, F. Zhao, and W. Ding, "Deep learning-based classification of liver cancer histopathology images using only global labels," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 6, pp. 1643–1651, Jun. 2020.

[57] K. George, S. Faziludeen, P. Sankaran, and P. Joseph K, "Breast cancer detection from biopsy images using nucleus guided transfer learning and belief based fusion," *Comput. Biol. Med.*, vol. 124, Sep. 2020, Art. no. 103954.

[58] D. Nam, J. Chapiro, V. Paradis, T. P. Seraphin, and J. N. Kather, "Artificial intelligence in liver diseases: Improving diagnostics, prognostics and response prediction," *JHEP Rep.*, vol. 4, no. 4, Apr. 2022, Art. no. 100443.

[59] J. Wang, D. Wu, M. Sun, Z. Peng, Y. Lin, H. Lin, J. Chen, T. Long, Z.-P. Li, C. Xie, B. Huang, and S.-T. Feng, "Deep segmentation feature-based radiomics improves recurrence prediction of hepatocellular carcinoma," *BME Frontiers*, vol. 2022, Jan. 2022.

[60] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3462–3471.

[61] E. E. van der Wall, H. W. Vliegen, A. de Roos, and A. V. G. Bruschke, "Magnetic resonance imaging in coronary artery disease," *Circulation*, vol. 92, no. 9, pp. 2723–2739, 1995.

[62] T. Falk et al., "U-Net: Deep learning for cell counting, detection, and morphometry," *Nature Methods*, vol. 16, no. 1, pp. 67–70, Jan. 2019.

[63] A. I. Khan et al., "CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest X-ray images," *Comput. Methods Programs Biomed.*, vol. 196, Nov. 2020, Art. no. 105581, doi: 10.1016/j.cmpb.2020.105581.

[64] W. Huang, Y. Yang, Z. Lin, G.-B. Huang, J. Zhou, Y. Duan, and W. Xiong, "Random feature subspace ensemble based extreme learning machine for liver tumor detection and segmentation," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2014, pp. 4675–4678.

[65] M. K. Santos, J. R. Ferreira, D. T. Wada, A. P. M. Tenório, M. H. Nogueira-Barbosa, and P. M. D. A. Marques, "Artificial intelligence, machine learning, computer-aided diagnosis, and radiomics: Advances in imaging towards to precision medicine," *Radiologia Brasileira*, vol. 52, no. 6, pp. 387–396, Dec. 2019.

[66] A. A. Khan and G. B. Narejo, "Analysis of abdominal computed tomography images for automatic liver cancer diagnosis using image processing algorithm," *Current Med. Imag. Formerly Current Med. Imag. Rev.*, vol. 15, no. 10, pp. 972–982, Nov. 2019.

[67] R. D. Riley, J. Ensor, K. I. E. Snell, T. P. A. Debray, D. G. Altman, K. G. M. Moons, and G. S. Collins, "External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: Opportunities and challenges," *BMJ*, vol. 353, p. i4379, Jun. 2016.

[68] A. Wang, W. Zhang, and X. Wei, "A review on weed detection using ground-based machine vision and image processing techniques," *Comput. Electron. Agricult.*, vol. 158, pp. 226–240, Mar. 2019.

[69] T. A. Morgan, K. E. Maturen, N. Dahiya, M. R. M. Sun, and A. Kamaya, "U.S. LI-RADS: Ultrasound liver imaging reporting and data system for screening and surveillance of hepatocellular carcinoma," *Abdominal Radiol.*, vol. 43, no. 1, pp. 41–55, Jan. 2018.

[70] H. J. Park, B. Park, and S. S. Lee, "Radiomics and deep learning: Hepatic applications," *Korean J. Radiol.*, vol. 21, no. 4, p. 387, 2020.

[71] S. Yao, Z. Ye, Y. Wei, H.-Y. Jiang, and B. Song, "Radiomics in hepatocellular carcinoma: A state-of-the-art review," *World J. Gastrointestinal Oncol.*, vol. 13, no. 11, pp. 1599–1615, Nov. 2021.

[72] L. Wei, D. Owen, B. Rosen, X. Guo, K. Cuneo, T. S. Lawrence, R. T. Haken, and I. E. Naqa, "A deep survival interpretable radiomics model of hepatocellular carcinoma patients," *Phys. Medica*, vol. 82, pp. 295–305, Feb. 2021.

[73] V. Andrearczyk, A. Depeursinge, and H. Müller, "Neural network training for cross-protocol radiomic feature standardization in computed tomography," *J. Med. Imag.*, vol. 6, no. 2, p. 1, Jun. 2019.

[74] R. Da-ano, F. Lucia, I. Masson, R. Abgral, J. Alfieri, C. Rousseau, A. Mervoyer, C. Reinhold, O. Pradier, U. Schick, D. Visvikis, and M. Hatt, "A transfer learning approach to facilitate ComBat-based harmonization of multicentre radiomic features in new datasets," *PLoS ONE*, vol. 16, no. 7, Jul. 2021, Art. no. e0253653.

[75] M. M. Abdelsamea, U. Zidan, Z. Senousy, M. M. Gaber, E. Rakha, and M. Ilyas, "A survey on artificial intelligence in histopathology image analysis," *WIREs Data Mining Knowl. Discovery*, vol. 12, no. 6, p. e1474, Nov. 2022.

[76] C. Chen, C. Chen, M. Ma, X. Ma, X. Lv, X. Dong, Z. Yan, M. Zhu, and J. Chen, "Classification of multi-differentiated liver cancer pathological images based on deep learning attention mechanism," *BMC Med. Informat. Decis. Making*, vol. 22, no. 1, pp. 1–13, Dec. 2022.

[77] M. Roy, J. Kong, S. Kashyap, V. P. Pastore, F. Wang, K. C. L. Wong, and V. Mukherjee, "Convolutional autoencoder based model HistoCAE for segmentation of viable tumor regions in liver whole-slide images," *Sci. Rep.*, vol. 11, no. 1, pp. 1–10, Jan. 2021.

[78] A. K. Chanchal, S. Lal, and J. Kini, "High-resolution deep transferred ASPPU-Net for nuclei segmentation of histopathology images," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 16, no. 12, pp. 2159–2175, Dec. 2021.

[79] T. He, J. N. Fong, L. W. Moore, C. F. Ezeana, D. Victor, M. Divatia, M. Vasquez, R. M. Ghobrial, and S. T. C. Wong, "An imageomics and multi-network based deep learning model for risk assessment of liver transplantation for hepatocellular cancer," *Computerized Med. Imag. Graph.*, vol. 89, Apr. 2021, Art. no. 101894.

[80] S. M. Kamrul Hasan and C. A. Linte, "A modified U-Net convolutional network featuring a nearest-neighbor Re-sampling-based elastic-transformation for brain tissue characterization and segmentation," in *Proc. IEEE Western New York Image Signal Process. Workshop*, Oct. 2018, pp. 1–5.

[81] W. Huang, N. Li, Z. Lin, G.-B. Huang, W. Zong, J. Zhou, and Y. Duan, "Liver tumor detection and segmentation using kernel-based extreme learning machine," in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2013, pp. 3662–3665.

[82] M. Biswas, V. Kuppili, D. R. Edla, H. S. Suri, L. Saba, R. T. Marinhoe, J. M. Sanches, and J. S. Suri, "Symtosis: A liver ultrasound tissue characterization and risk stratification in optimized deep learning paradigm," *Comput. Methods Programs Biomed.*, vol. 155, pp. 165–177, Mar. 2018.

[83] B. Schmauch, P. Herent, P. Jehanno, O. Dehaene, C. Saillard, C. Aubé, A. Luciani, N. Lassau, and S. Jégou, "Diagnosis of focal liver lesions from ultrasound using deep learning," *Diagnostic Interventional Imag.*, vol. 100, no. 4, pp. 227–233, Apr. 2019.

[84] K. Kim, S. Kim, K. Han, H. Bae, J. Shin, and J. S. Lim, "Diagnostic performance of deep learning-based lesion detection algorithm in CT for detecting hepatic metastasis from colorectal cancer," *Korean J. Radiol.*, vol. 22, no. 6, p. 912, 2021.

[85] D. W. Kim, G. Lee, S. Y. Kim, G. Ahn, J.-G. Lee, S. S. Lee, K. W. Kim, S. H. Park, Y. J. Lee, and N. Kim, "Deep learning–based algorithm to detect primary hepatic malignancy in multiphase CT of patients at high risk for HCC," *Eur. Radiol.*, vol. 31, no. 9, pp. 7047–7057, Sep. 2021.

[86] M. Puri, "Automated machine learning diagnostic support system as a computational biomarker for detecting drug-induced liver injury patterns in whole slide liver pathology images," *ASSAY Drug Develop. Technol.*, vol. 18, no. 1, pp. 1–10, Jan. 2020.

[87] J. Lv, J. Wang, X. Shang, F. Liu, and S. Guo, "Survival prediction in patients with colon adenocarcinoma via multiomics data integration using a deep learning algorithm," *Biosci. Rep.*, vol. 40, no. 12, pp. 28–40, Dec. 2020.

[88] T. Lee, P. A. Rawding, J. Bu, S. Hyun, W. Rou, H. Jeon, S. Kim, B. Lee, L. J. Kubiatowicz, D. Kim, S. Hong, and H. Eun, "Machine-Learning-Based clinical biomarker using cell-free DNA for hepatocellular carcinoma (HCC)," *Cancers*, vol. 14, no. 9, p. 2061, Apr. 2022.

[89] Z. Zhang and Z.-P. Liu, "Robust biomarker discovery for hepatocellular carcinoma from high-throughput data by multiple feature selection methods," *BMC Med. Genomics*, vol. 14, no. S1, pp. 1–12, Aug. 2021.

[90] H. Tian, D. Guan, and J. Li, "Identifying osteosarcoma metastasis associated genes by weighted gene co-expression network analysis (WGCNA)," *Medicine*, vol. 97, no. 24, 2018, Art. no. e10781.

[91] T. Huang, L. Lan, X. Fang, P. An, J. Min, and F. Wang, "Promises and challenges of big data computing in health sciences," *Big Data Res.*, vol. 2, no. 1, pp. 2–11, Mar. 2015.

[92] S. Shetty and Y. S. Rao, "SVM based machine learning approach to identify Parkinson's disease using gait analysis," in *Proc. Int. Conf. Inventive Comput. Technol. (ICICT)*, vol. 2, Aug. 2016, pp. 1–5.

[93] G. N. Ioannou, W. Tang, L. A. Beste, M. A. Tincopa, G. L. Su, T. Van, E. B. Tapper, A. G. Singal, J. Zhu, and A. K. Waljee, "Assessment of a deep learning model to predict hepatocellular carcinoma in patients with hepatitis C cirrhosis," *JAMA Netw. Open*, vol. 3, no. 9, Sep. 2020, Art. no. e2015626.

[94] L. W. C. Chan, S. C. C. Wong, C. C. Chiau, T.-M. Chan, L. Tao, J. Feng, and K. W. H. Chiu, "Association patterns of ontological features signify electronic health records in liver cancer," *J. Healthcare Eng.*, vol. 2017, pp. 1–9, May 2017.

[95] G. L.-H. Wong, V. W.-K. Hui, Q. Tan, J. Xu, H. W. Lee, T. C.-F. Yip, B. Yang, Y.-K. Tse, C. Yin, F. Lyu, J. C.-T. Lai, G. C.-Y. Lui, H. L.-Y. Chan, P.-C. Yuen, and V. W.-S. Wong, "Novel machine learning models outperform risk scores in predicting hepatocellular carcinoma in patients with chronic viral hepatitis," *JHEP Rep.*, vol. 4, no. 3, Mar. 2022, Art. no. 100441.

[96] C.-W. Liang, H.-C. Yang, M. M. Islam, P. A. A. Nguyen, Y.-T. Feng, Z. Y. Hou, C.-W. Huang, T. N. Poly, and Y.-C.-J. Li, "Predicting hepatocellular carcinoma with minimal features from electronic health records: Development of a deep learning model," *JMIR Cancer*, vol. 7, no. 4, Oct. 2021, Art. no. e19812.

[97] R. J. Driver, V. Balachandrakumar, A. Burton, J. Shearer, A. Downing, T. Cross, E. Morris, and I. A. Rowe, "Validation of an algorithm using inpatient electronic health records to determine the presence and severity of cirrhosis in patients with hepatocellular carcinoma in England: An observational study," *BMJ Open*, vol. 9, no. 7, Jul. 2019, Art. no. e028571.

[98] J. C. Ahn, A. Connell, D. A. Simonetto, C. Hughes, and V. H. Shah, "Application of artificial intelligence for the diagnosis and treatment of liver diseases," *Hepatology*, vol. 73, no. 6, pp. 2546–2563, 2021.

[99] M. Ishii, "Simultaneous measurements of serum α-fetoprotein and protein induced by vitamin K absence for detecting hepatocellular carcinoma," *Amer. J. Gastroenterol.*, vol. 95, no. 4, pp. 1036–1040, Apr. 2000.

[100] P. Johnson, Q. Zhou, D. Y. Dao, and Y. M. D. Lo, "Circulating biomarkers in the diagnosis and management of hepatocellular carcinoma," *Nature Rev. Gastroenterol. Hepatol.*, vol. 19, no. 10, pp. 670–681, Oct. 2022.

[101] S. Silow-Carroll, J. N. Edwards, and D. Rodin, "Using electronic health records to improve quality and efficiency: The experiences of leading hospitals," *Issue Brief, Commonw Fund*, vol. 17, no. 1, p. 40, 2012.

[102] A. A. Siyal, A. Z. Junejo, M. Zawish, K. Ahmed, A. Khalil, and G. Soursou, "Applications of blockchain technology in medicine and healthcare: Challenges and future perspectives," *Cryptography*, vol. 3, no. 1, p. 3, Jan. 2019.

[103] T. Ching et al., "Opportunities and obstacles for deep learning in biology and medicine," *J. Roy. Soc. Interface*, vol. 15, no. 141, 2018, Art. no. 20170387.

[104] N. Salam, W. M. Al-Shaqha, and A. Azzi, "Leishmaniasis in the middle east: Incidence and epidemiology," *PLoS Neglected Tropical Diseases*, vol. 8, no. 10, p. e3208, Oct. 2014.

[105] X. Xiang, X.-M. You, J.-H. Zhong, and L.-Q. Li, "Hepatocellular carcinoma in the absence of cirrhosis in patients with chronic hepatitis B virus infection," *J. Hepatol.*, vol. 67, no. 4, pp. 885–886, Oct. 2017.

[106] U. Kc, S. Garg, J. Hilton, and J. Aryal, "A cloud-based framework for sensitivity analysis of natural hazard models," *Environ. Model. Softw.*, vol. 134, Dec. 2020, Art. no. 104800.

[107] T. Hawkins, M. Chitale, S. Luban, and D. Kihara, "PFP: Automated prediction of gene ontology functional annotations with confidence scores using protein sequence data," *Proteins, Struct., Function, Bioinf.*, vol. 74, no. 3, pp. 566–582, 2009.

[108] T.-C. Shih, Y.-J. Tien, C.-J. Wen, T.-S. Yeh, M.-C. Yu, C.-H. Huang, Y.-S. Lee, T.-C. Yen, and S.-Y. Hsieh, "MicroRNA-214 downregulation contributes to tumor angiogenesis by inducing secretion of the hepatoma-derived growth factor in human hepatoma," *J. Hepatol.*, vol. 57, no. 3, pp. 584–591, Sep. 2012.

[109] N. Verma, A. Choudhury, V. Singh, and A. Duseja, "APASL-ACLF research consortium–artificial intelligence (AARC-AI) model precisely predicts outcomes in acute-on-chronic liver failure patients," *Liver Int.*, vol. 43, no. 2, pp. 442–451, 2022.

[110] S. Chen et al., "Using quasispecies patterns of hepatitis B virus to predict hepatocellular carcinoma with deep sequencing and machine learning," *J. Infectious Diseases*, vol. 223, no. 11, pp. 1887–1896, Jun. 2021.

[111] A. Spann, A. Yasodhara, J. Kang, K. Watt, B. Wang, A. Goldenberg, and M. Bhat, "Applying machine learning in liver disease and transplantation: A comprehensive review," *Hepatology*, vol. 71, no. 3, pp. 1093–1105, 2020.

[112] T. Su, C. Wu, and J. Kao, "Artificial intelligence in precision medicine in hepatology," *J. Gastroenterol. Hepatol.*, vol. 36, no. 3, pp. 569–580, Mar. 2021.

[113] L. Lapointe-Shaw, F. Georgie, D. Carlone, O. Cerocchi, H. Chung, Y. Dewit, J. J. Feld, L. Holder, J. C. Kwong, B. Sander, and J. A. Flemming, "Identifying cirrhosis, decompensated cirrhosis and hepatocellular carcinoma in health administrative data: A validation study," *PLoS ONE*, vol. 13, no. 8, Aug. 2018, Art. no. e0201120.

[114] A. Qayyum, J. Qadir, M. Bilal, and A. Al-Fuqaha, "Secure and robust machine learning for healthcare: A survey," *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 156–180, 2021.

[115] H. Wang, E. Pujos-Guillot, B. Comte, J. L. de Miranda, V. Spiwok, I. Chorbev, and F. Castiglione, "Deep learning in systems medicine," *Briefings Bioinf.*, vol. 22, no. 2, pp. 1543–1559, 2021.

[116] S.-C. Liu, J. Lai, J.-Y. Huang, C.-F. Cho, P. H. Lee, M.-H. Lu, C.-C. Yeh, J. Yu, and W.-C. Lin, "Predicting microvascular invasion in hepatocellular carcinoma: A deep learning model validated across hospitals," *Cancer Imag.*, vol. 21, no. 1, pp. 1–16, Dec. 2021, doi: 10.1186/s40644-021-00425-3.

[117] E. Gonçalves, M. Gonçalves-Reis, J. B. Pereira-Leal, and J. Cardoso, "DNA methylation fingerprint of hepatocellular carcinoma from tissue and liquid biopsies," *Sci. Rep.*, vol. 12, no. 1, pp. 1–12, Jul. 2022, doi: 10.1038/s41598-022-15058-0.

[118] L. Pan, F. Feng, J. Wu, S. Fan, J. Han, S. Wang, L. Yang, W. Liu, C. Wang, and K. Xu, "Demethylzeylasteral targets lactate by inhibiting histone Lactylation to suppress the tumorigenicity of liver cancer stem cells," *Pharmacolog. Res.*, vol. 181, Jul. 2022, Art. no. 106270, doi: 10.1016/j.phrs.2022.106270.

[119] S. Lu, J. Yang, B. Yang, Z. Yin, M. Liu, L. Yin, and W. Zheng, "Analysis and design of surgical instrument localization algorithm," *Comput. Model. Eng. Sci.*, vol. 137, no. 1, pp. 669–685, 2023, doi: 10.32604/cmes.2023.027417.

[120] L. Chen, Y. He, J. Zhu, S. Zhao, S. Qi, X. Chen, H. Zhang, Z. Ni, Y. Zhou, G. Chen, S. Liu, and T. Xie, "The roles and mechanism of m6A RNA methylation regulators in cancer immunity," *Biomed. Pharmacotherapy*, vol. 163, Jul. 2023, Art. no. 114839, doi: 10.1016/j.biopha.2023.114839.

[121] Q. Sun, Q. Tao, T. Ming, S. Tang, H. Zhao, M. Liu, H. Yang, S. Ren, J. Lei, Y. Liang, Y. Peng, M. Wang, and H. Xu, "Berberine is a suppressor of hedgehog signaling cascade in colorectal cancer," *Phytomedicine*, vol. 114, Jun. 2023, Art. no. 154792, doi: 10.1016/j.phymed.2023.154792.

[122] H. Huang, N. Wu, Y. Liang, X. Peng, and J. Shu, "SLNL: A novel method for gene selection and phenotype classification," *Int. J. Intell. Syst.*, vol. 37, no. 9, pp. 6283–6304, Sep. 2022, doi: 10.1002/int.22844.

[123] L. Li, S. Wang, and W. Zhou, "Balance cell apoptosis and pyroptosis of caspase-3-activating chemotherapy for better antitumor therapy," *Cancers*, vol. 15, no. 1, p. 26, Dec. 2022, doi: 10.3390/cancers15010026.

[124] X. Yi, X. Guan, C. Chen, Y. Zhang, Z. Zhang, M. Li, P. Liu, A. Yu, X. Long, L. Liu, B. T. Chen, and C. Zee, "Adrenal incidentaloma: Machine learning-based quantitative texture analysis of unenhanced CT can effectively differentiate sPHEO from lipid-poor adrenal adenoma," *J. Cancer*, vol. 9, no. 19, pp. 3577–3582, 2018, doi: 10.7150/jca.26356.

ARUN ASIF is currently a dedicated Postdoctoral Researcher with the Albany Medical Center. With a background in AI and organ on a chip, his research focuses on brief description of research interests. His contributions aim to advance the understanding of specific area of research for the betterment of potential impact or applications.

FAHEEM AHMED is currently pursuing the Ph.D. degree with Jeju National University, South Korea. With a master's degree in AI in imaging and drug discovery followed by organ on a chip, his doctoral research is centered on drug discovery. He is also a dynamic Researcher. His academic journey reflects his commitment to contributing to the advancements in cancer.

ZEESHAN is currently pursuing the Ph.D. degree in electronics and AI imaging. His studies are driven by a passion for understanding and contributing to the knowledge base in cancer.

JAVED ALI KHAN received the Ph.D. degree in software engineering from Tsinghua University (QS ranked 12th), Beijing, China. He is currently a Senior Lecturer with the Foundation of Software Engineering (FSE) Group, Department of Computer Science, University of Hertfordshire, U.K. Previously, he was an Assistant Professor cum Chairperson with the Department of Software Engineering, University of Science and Technology Bannu, Pakistan. He regularly publishes papers in reputable software engineering journals and conferences. His research interests include software engineering, requirements engineering, CrowdRE, argumentation and argument mining, mining software repositories, human values in software, quantum software engineering, feedback analysis, empirical software engineering, sentiment and opinion mining, requirements prioritization, mining fake reviews, sarcasm detection, and health analytics.

EMAN ALLOGMANI received the bachelor's degree in computer science from Taibah University, Saudi Arabia, the master's degree in computer science from California Lutheran University, USA, and the Doctor of Science (D.Sc.) degree in computer science from Bowie State University, USA. Currently, she is an Assistant Professor with Majmaah University, Saudi Arabia. Her research interests include deep learning, machine learning, natural language processing, image classification, and data augmentation strategies.
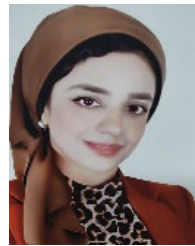
NORA EL RASHIDY received the M.S. and Ph.D. degrees from the Faculty of Computer Science and Information, Mansoura University, Egypt, in 2016 and 2020, respectively. She is currently an Associate Professor with the Machine Learning and Information Retrieval Department, Faculty of Artificial Intelligence, Kafrelsheikh University. She is also an Assistant Professor with Galala University, an Assistant Professor with the Machine Learning and Information Retrieval Department, Kafrelsheikh University, and a Manager of the Measurement and Assessment Central Unit, Kafrelsheikh University. Her research interests include machine earning, medical informatics, distributed and hybrid clinical decision support systems, big data, and cloud computing. She is a reviewer in many journals, and she is very interested in disease diagnosis and treatment research.

SOBIA MANZOOR received the Ph.D. degree from the National University of Sciences and Technology, Islamabad.

She is an accomplished Researcher. Her contributions to the field have been instrumental in viral cancers and viral infections. Additionally, she has worked on solid cancers as well.
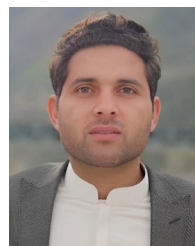
MUHAMMAD SHAHID ANWAR received the M.Sc. degree in telecommunications technology from Aston University, Birmingham, U.K., in 2012, and the Ph.D. degree in information and communication engineering from the School of Information and Electronics, Beijing Institute of Technology, Beijing, China, in 2021. He is currently an Assistant Professor with the Department of AI and Software, Gachon University, Seongnam-si, South Korea. He has authored and coauthored more than 80 publications, including IEEE, Springer, Elsevier, IET, Frontiers journals, flagship conference papers, and seven book chapters. He has been an Editorial Board Member of *CMC* and *CSSE Tech Science*, a TPC Member of QoMEX, and a Reviewer of several journals, including ACM and IEEE TRANSACTIONS. His research interests include 360-degree videos, immersive media (VR/AR), metaverse, networks QoS, the quality of experience (QoE), the evaluations of VR telemedicine and healthcare systems, AI healthcare, and machine learning-based QoE prediction models.

• • •