SYSTEMATIC REVIEW

A systematic literature review: exploring the challenges of ensemble model for medical imaging

Muhamad Rodhi Supriyadi^{1,2}, Azurah Bte A. Samah^{1*}, Jemie Muliadi², Raja Azman Raja Awang³, Noor Huda Ismail³, Hairudin Abdul Majid¹, Mohd Shahizan Bin Othman¹ and Siti Zaiton Binti Mohd Hashim¹

Abstract

Background Medical imaging has been essential and has provided clinicians with useful information about the human body to diagnose various health issues. Early diagnosis of diseases based on medical imaging can mitigate the risk of severe consequences and enhance long-term health outcomes. Nevertheless, the task of diagnosing diseases based on medical imaging can be challenging due to the exclusive ability of clinicians to interpret the outcomes of medical imaging, which is time-consuming and susceptible to human fallibility. The ensemble model has the potential to enhance the accuracy of diagnoses of diseases based on medical imaging by analyzing vast volumes of data and identifying trends that may not be immediately apparent to doctors. However, it takes a lot of memory and processing resources to train and maintain several ensemble models. These challenges highlight the necessity of effective and scalable ensemble models that can manage the intricacies of medical imaging assignments.

Methods This study employed an SLR technique to explore the latest advancements and approaches. By conducting a thorough and systematic search of Scopus and Web of Science databases in accordance with the principles outlined in the PRISMA, employing keywords namely ensemble model and medical imaging.

Results This study included a total of 75 papers that were published between 2019 and 2024. The categorization, methodologies, and use of medical imaging were key factors examined in the analysis of the 30 cited papers included in this study, with a focus on diagnosing diseases.

Conclusions Researchers have observed the emergence of an ensemble model for disease diagnosis using medical imaging since it has demonstrated improved accuracy and may guide future studies by highlighting the limitations of the ensemble model.

Keywords Systematic literature review, Ensemble model, Medical imaging, Artificial intelligence, Deep learning, Machine learning

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creative.commons.go/licenses/by-nc-nd/4.0/.

Azurah Bte A. Samah azurah@utm.my ¹Faculty of Computing, Universiti Teknologi Malaysia, Skudai, Johor 81310, Malaysia ²Research Center for Artificial Intelligent and Cyber Security, National Research and Innovation Agency, Bandung 40135, Indonesia

³School of Dental Sciences, Universiti Sains Malaysia, Kota Bharu,

Kelantan 16150, Malaysia

*Correspondence:







Background

The field of medicine is currently advancing and playing an increasingly significant role in the healing of human beings. Currently, there is a substantial volume of medical data available. However, it is imperative to effectively utilize this vast amount of medical data to make meaningful contributions to the field of medicine [1]. Despite the vast volume of medical data, numerous challenges persist: medical data were heterogeneous, encompassing several types of information, such as medical imaging, data privacy, regulatory concerns, and the ethical implications [2]. In recent years, there has been significant advancement in the field of automated medical image processing. Cai et al. [3] have explained that medical imaging has played a pivotal role in the diagnosis and treatment of many diseases, providing doctors with important knowledge about the human body. With the introduction of various imaging techniques such as X-rays, CT, MRI, ultrasound images, and color medical images, medical professionals now have the ability to obtain high-quality, detailed images that provide valuable information about inside body components, anatomical structures, and physiological functions [4]. X-rays are two-dimensional imaging and are usually used for detecting infections like pneumonia and COVID-19 [5]. CT imaging is a medical imaging technique that utilizes X-rays [6]. Ultrasound imaging is a non-invasive medical technique that utilizes high-frequency sound waves [7]. MRI involves the use of magnetic fields and radio waves [1]. Color medical images include dermatological images, histopathology slides, and other color-based medical images like dental color images [5].

Deshmukh [8] has explored that doctors, radiologists, and other healthcare practitioners have faced considerable hurdles due to the increased accessibility and intricacy of medical imaging data. Medical image interpretation is a demanding and time-consuming process that necessitates a high level of competence and significant training, especially for doctors and radiologists. In addition, the large amount of imaging data generated can provide difficulty in identifying minor alterations that can potentially suggest the existence of a disease. AI has emerged as a possible approach to tackle these difficulties, employing algorithms like ML, and DL, until ensemble models automatically analyze and interpret medical imaging. Akst [9] has defined AI as the capacity of robots to demonstrate a unique type of intelligence. ML is a subfield of AI that uses algorithms to generate predictions by analyzing a dataset. At that time, DL is a subdivision of machine learning that employes a deep neural network, consisting of numerous computational layers, to examine incoming data. Kumar and Harish [10] have clarified that ensemble models are formed by amalgamating numerous forecasts produced by distinct models to yield a unified ultimate projection. The advantages of using an ensemble, namely improving accuracy and robustness by combining multiple models, effective control of complex communications embedded in medical imaging, better averaging and education of overfitting, as well as better feature representation for tasks welding or joining [11].

Ensemble models are applicable in both classification and regression scenarios. These models can employ either similar methods or select techniques from distinct categories [12]. Mohammed and Kora [13] have outlined that the ensemble model exhibits three distinct characteristics that significantly impact its performance. The initial characteristic is the dependence on proficient baseline models, regardless of whether they are sequential or parallel. The second characteristic is the combination method or fusion, which entails the selection of a suitable procedure to merge the output of the baseline classifier using various weighted voting or meta-learning techniques. The third characteristic is the presence of diverse underlying classifiers, which might be either homogeneous or heterogeneous.

Tao et al. [14] have defined a homogeneous ensemble model as an ensemble model when several models are constructed using the same dataset (like medical imaging) and the same learning algorithm. A heterogeneous ensemble model is created by building a model utilizing several techniques such as neural network machine, support vector, and random forest. Sequential ensemble models learn sequentially due to data dependencies, namely medical imaging, so the second and subsequent models sequentially correct errors made by the first model [15]. Whereas in parallel ensemble models are generated simultaneously, so the errors made by one model are different from those found in other independent models [16].

The utilization of AI in medical imaging, particularly through the use of ensemble models, has the potential to revolutionize the sector by enabling more accurate and efficient identification, segmentation, and classification of diseases. Ensemble AI possesses the capacity to analyze vast amounts of medical imaging data and identify minute changes that may indicate the existence of a disease [11]. Table 1 summarizes the related study on disease diagnosis based on medical imaging using ensemble models that highlight the limitations of previous research compared to our study.

A study by Akinbo and Daramola [17] has presented a literature review with a drawback: no detailed explanation of the paper's results review and no specified data about ensemble machine learning algorithms for the prediction and classification of medical images. Whereas, a study by Imran and N [11] has explained that literature review during 2021–2023 is only based on CT imaging and is a small total of papers that have been reviewed.

Table 1	Related study	on disease diagnosis base	d on medical imaging	using ensemble models
---------	---------------	---------------------------	----------------------	-----------------------

No.	Author (Year)	Year range	Focus of study	Total of review	Limitation
				paper	
1.	Akinbo and Daramola (2021)[17]	Not specified	Ensemble ML based on medical imaging	Not specified	There is no detailed explana- tion of the results of the paper review
2.	Imran and N (2024)[11]	2021-2023	Ensemble ML and DL based on computed tomography medical images	13	There is a small quantity of pa- pers that have been reviewed.
3.	Our study	2019–2024	Ensemble ML and DL based on medical images. Using PRISMA techniques and providing metadata analysis information	30	-

The studies lack a precise specification of the periods involved in the SLR. With the increasing popularity of ensemble model-based diagnosis using medical imaging, applying SLR with PRISMA is expected to address the gaps in existing studies.

The growing number of studies on disease diagnosis utilizing ensemble models in medical imaging underscores the necessity for a comprehensive evaluation of current understanding. A SLR was performed using the Scopus and WoS databases. This review analyzed a total of 75 papers and conducted a more detailed examination of 30 papers. The purpose of the metadata analysis was to identify papers by year, authors of most relevance, sources of most relevance, affiliations of most relevance, and popular topics. Furthermore, a thorough examination of 30 papers was undertaken to resolve the following queries: What are the current ensemble model-based approaches for disease diagnosis using medical imaging? The purpose of the SLR is to provide researchers with a comprehensive overview of the most recent techniques and advancements in the field. It also attempts to identify areas of knowledge that can be improved by developing a more advanced disease diagnostic system based on medical imaging, employing ensemble models.

Methods

The current SLR is conducted using the guidelines given in the PRISMA [18]. The SLR formulates research inquiries before conducting a methodical search, selection, and evaluation of studies to gather relevant information [19]. This strategy was selected because of its renowned ability to deliver an accurate and dependable combination of academic material, and it was well-acknowledged in several study disciplines. The meta-analyses included only the research items that satisfied the criteria for eligibility. In this study, a systematic review was conducted and reported using the PRISMA techniques. The study employed the PRISMA checklist to guarantee the inclusion of all pertinent information, while the flow diagram was implemented to record the process of selecting studies [20].

Data identification

This study conducted an extensive investigation of Scopus's integrated and WoS databases, which encompass all prominent publishers. The WoS and Scopus databases are considered reliable for the SLR due to the exceptional quality of their indexed content, which is comprehensive and widely used in meta-analysis studies [21]. But both have limitations, like data bias to favor certain fields [22]. To address potential data bias, we also considered incorporating additional databases, such as IEEE, for computer science. The search encompassed the most recent studies, the period from 2019 to 2024, reflecting up-todate advancements and trends in the ensemble models based on medical imaging. The study utilized the main keywords such as "ensemble model" and "medical imaging" to identify pertinent publications. It ensured the systematic review was relevant to current practices and technologies.

Screening preliminary data and establishing eligibility

This SLR utilized targeted keywords and a query containing "medical imaging" AND "ensemble model" AND "diagnosis" OR "detection" OR "classification" OR "prediction" OR "segmentation" to conduct comprehensive searches in the Scopus and WoS databases. The first Scopus search identified 271 papers, whereas the WoS search found 51 papers. By implementing the time frame of 2019-2024 and applying additional filters such as document type (articles and conference papers), subject area (computer science), language, duplicate documents, open access, and keyword constraints, the total number of publications was reduced to 75 papers. Subsequently, the 75 remaining different papers were evaluated, and the most relevant material was retrieved using a standardized extraction template. The analysis omitted studies that were unrelated to the ensemble model or primarily focused on medical imaging.

In addition, book chapters, research using non-human subjects, and reviews were excluded. Ultimately, a total of 30 comprehensive research papers satisfied the criteria for inclusion as depicted in Fig. 1, and were subsequently incorporated into the review. A flowchart was



Fig. 1 The PRISMA technique was used in this study

constructed to illustrate the process of selecting studies, which encompasses the search query and the criteria for inclusion. The study selection approach was comprehensive and meticulous, guaranteeing the inclusion of the most pertinent and up-to-date studies on utilizing ensemble models for diagnosing diseases through medical imaging.

The ensemble models were utilized to establish the inclusion and exclusion criteria, guaranteeing that the review encompassed all essential aspects of disease diagnosis using medical imaging. Articles and conference papers that meet the inclusion requirements are those published between 2019 and 2024, which focus on disease diagnosis using ensemble models and are based on original research findings or empirical data. Excluded from consideration were papers that had not yet been published in English, ensemble model studies that were not related to disease diagnosis using medical imaging, published literature prior to 2019, research that was not in the form of articles or conference papers, duplicate studies, preliminary data studies, and studies with unclear or ambiguous conclusions. Ensemble models were employed to analyze the most pertinent and current publications on the diagnosis of disease using medical imaging. Irrelevant research that did not meet the specified conditions was excluded. The aim of the literature review was to employ ensemble models in order to investigate the most recent and relevant publications about the diagnosis of diseases using medical imaging.

Observation and findings

This section will analyze the findings and observations derived from the assessment of the metadata. The findings are derived from a comprehensive analysis of 75 research papers, encompassing both their metadata and content.

Analysis of metadata

Metadata analysis facilitates comprehension of research literature by extracting information pertaining to the scholarly process, including authors, publications, journals, and other relevant factors. Metadata analysis was conducted on a total of 75 research papers. The papers were categorized according to several criteria, such as published papers by year, authors of most relevance, sources of most relevance, affiliations of most relevance, and popular topics.

Published papers by year

Figure 2 shows that a total of 75 papers were examined to determine the number of studies conducted in the past 6 years that focused on the diagnosis of diseases



Fig. 2 Published papers on disease diagnosis based on medical imaging using ensemble



Fig. 3 Authors of most relevance on disease diagnosis based on medical imaging using ensemble

using ensemble models in medical imaging. The publication sector is consistently growing and has witnessed annual increases. The largest growth in articles occurred in 2023. As of this month in 2024, there have been a total of 12 published papers, with the expectation of a further increase by the end of the year. In 2023, around 21 new papers were published, while in 2022, just 18 new papers were published.

Furthermore, it has become obvious that the classification problem in illness diagnosis utilizing ensemble models in medical imaging has been highly regarded. Consequently, the quantity of academic publications made available to the public in 2023 is significantly higher than in any preceding year. Conversely, it is evident that there was a very small quantity of papers published, especially between 2019 and 2020, with only a few papers being produced. Therefore, there is a growing emphasis and attention on diagnosing diseases using medical imaging employing ensemble models, which involve addressing classification problems and other data-driven considerations.

Authors of most relevance

Based on Fig. 3, Kim J and Zhang H have authored the most relevant publications among the other authors namely four documents, and have had the greatest



Fig. 4 Sources of most relevance on disease diagnosis based on medical imaging using ensemble

 Table 2
 Most frequently used words in keyword sections

Words	Frequency
Medical imaging	55
Deep learning	53
Ensemble models	24
Learning systems	22
Diagnosis	21
Classification	20
Convolutional neural network	17
Convolution	16
Image enhancement	16
Diseases	15

impact. Thus, our team performed a comparative data analysis to track the author's output over a period of time. Kim J authored a total of three publications in 2024, which collectively got 3 citations. Furthermore, Zhang H authored one publication in 2023 that garnered a total of 14 citations.

Sources of most relevance

Figure 4 shows the relevant sources, with IEEE Access having the highest number of papers at 14. In second place is Lecture Notes in Computer Science, which has a total of 5 documents. Following closely are Current Medical Imaging, Journal of Digital Imaging, and Medical Image Analysis with a total of 3 documents respectively.

The most frequently used words in titles and keywords

This study utilized bibliometric analysis in R-studio. Its purpose is to discover the most commonly occurring terms. The primary objective of our study was to locate and assess literature pertaining to ensemble models, medical imaging, and disease diagnosis. The papers predominantly utilized the terms "medical imaging", "deep learning", and "ensemble models", as evidenced in Table 2. The term "diagnosis" is ranked at order 5 while "diseases" is ranked at order 10. The keyword field contains the words "medical imaging" 55 times, followed by "deep learning" 53 times, and "ensemble models" 24 times. Similarly, by examining the Treemap depicted in Fig. 5, it is evident that the most commonly utilized terms are "medical imaging", "deep learning", and "ensemble models". In the keyword field, the term "medical imaging" accounts for 10% of the total, followed by "deep learning" at 9% and "ensemble models" at 4%. This study includes a total of 75 research publications, representing the entire proportion of this SLR. The Word Cloud is a simple technique for finding the dominant themes and important phrases in the referred articles, enabling the identification of the most common terms in a complicated context. Figure 6 exhibits word clouds produced by the software, with larger and more prominent writing indicating the most often utilized terms, while smaller and less bold type emphasizes the less typically used phrases.

Popular topics

The popular topics were derived by considering exclusively articles published between 2019 and 2024. The graphical parameters utilized the author's keywords field, setting a minimum word frequency of five and a number of words per year of three. Figure 7 provides a visual representation of the main keywords utilized in each year. The lines displayed the instances of each word's usage, while the size of the bubbles corresponded to the frequency of the term's appearance. As an illustration, the

medical imaging S5 10%	ensemble models 24 4%	classification (of information) 20 4%	diseases 15 3%	computer aided diagnosis 14 3%	humans 10 2%	machine lea 10 2%	rning image s 9 2%	egmentation CC 8 1	wid-19	eature extraction β 1%
		convolutional neural networks 17 3%	ensemble learning 15 3%	convolutional neural network 14 3%	semantic segmentation 8 1%	forecasting 7 1%	neural ne 7 1%	tworks diagn 6 1%	ostic imaging ey 6 1	ye protection %
	learning systems 22 4%	5.4	income allocation at an		article 7	features extraction 6 1%	algorithm 5 1%	artificial intelligence 5 1%	biomedical imaging 5 1%	computer 5 1%
deep learning 53 9%		convolution 16 3%	15 3%	human 13 2%	1% breast cancer	image processing 6 1%	computer added instruction 5 1%	image analysis 5 1%	karning algorithms 5 1%	learning models 5 1%
	diagnosis 21 4%	image enhancement	transfer learning 15	deep neural networks	1%	images classification 6 1%	computer vision 5 1%	machine-learning 5 1%	pneumonia 5 1%	semantics 5 1%
		3%	3%	10 2%	7 196	magnatic naconanca imaging d 1%	comparentiaed springraphy 5 1%	medical diagnosis imaging 5 Th	statistical test 5 1%	s cartilage 4 1%

Fig. 5 Treemap for most frequently used words in medical imaging publications



Fig. 6 Word cloud for most frequently used words in medical imaging publications

term that was used most often in 2023 was "ensemble models" with a frequency of 24, followed by "learning system" with 22 term frequency, and "diseases" with 15 term frequency.

Interestingly, the popular topics in research have evolved over the years. In 2019, the most frequent word was "ensemble modeling" with a term frequency 7. Then the research shifted towards exploring "image segmentation" in 2021 with a term frequency of 9, followed by "neural networks" and "eye protection". The popular topics during 2019–2022 are in 2022 namely "medical imaging" with a term frequency of 55, followed by "deep learning" with 53 term frequency, and "diagnosis" with 21 term frequency. Now, in 2024, the popular topic namely "machine learning" with a term frequency of 5. The data indicates a growing popularity in the use of medical imaging and deep learning. Following closely behind is the use of ensemble models, particularly in the field of healthcare for diagnosing diseases based on medical imaging. Comprehending and diagnosing illnesses is vital in the medical domain, and these prominent areas of study emphasize the need to utilize cutting-edge technologies to enhance patient care and therapy. In summary, the data offer intriguing insights into the changing trends in popular research areas in medical imaging and themes connected to ensemble models.

Affiliations of most relevance

Figure 8 illustrates the top 10 affiliations that are most significant globally in terms of the number of published papers. The list primarily consists of one university from



Fig. 7 Popular topics extracted from the topic of disease diagnosis based on medical imaging





Saudi Arabia (King Abdulaziz University), two universities from South Korea (Chonnam National University and Korea University College of Medicine), and two universities from China (Shantou University Medical College and Guangxi University), respectively. Additional countries, like Vietnam (Ton Duc Thang University), the United Kingdom (University of Edinburgh), Italy (University of Rome Tor Vergata), Pakistan (Comsats University Islamabad), and Turkey (Gazi University), are also present. King Abdulaziz University in Saudi Arabia has the highest number of published articles, with a total of 13. Chonnam National University follows with 5 publications. These findings indicate that King Abdulaziz University is a prominent institution in the field, maybe because of its focus on research and development.

The data also offers a concise overview of the present research output from affiliations of most relevance across the globe, especially academic affiliations. It emphasizes the continuous endeavors of academic affiliations to generate top-notch research that can enrich

Page 9 of 15

our comprehension of diverse disciplines and lead to the creation of novel information and advancements. Gaining insight into the institutions that generate the most amount of research on a specific topic can be advantageous for researchers to make well-informed choices and coordinate their research endeavors.

Results and discussion

In this section, a total of 30 comprehensive research papers satisfied the criteria and were subsequently incorporated into the review. The study examined in this SLR has demonstrated the extensive utilization of ensemble models across many medical imaging disciplines. Most of the research has employed ensemble deep learning techniques utilizing convolutional neural networks as baseline and ensemble models as fusion. Several research studies employed ensemble machine learning techniques utilizing machine learning models as baselines, such as SVM, KNN, SMO, PCA, GLCM, etc, and ensemble models as fusion. Moreover, operations are categorized based on imaging modalities, such as X-rays, CT, MRI, ultrasound images, and color medical images. Figure 9 is the stacked bar chart showing publication trends by year and split by sub-categories of imaging modality. This shows that the modality that was always used from 2019-2024 was X-ray imaging, which increased until 2023 with a total of 21 papers [23-43]. The second position, namely CT image with a total of 3 papers [14, 44, 45]. Meanwhile, for other modalities with a total of 1 paper, namely MRI [46], ultrasound images [47], color images [48], the combination of X-rays & CT [49], the combination of X-rays & color images [50], and the combination of ultrasound images & color images [51]. In the last two years, color images and ultrasound images have begun to be of interest to be studied using AI.

In this study, ensemble models based on medical imaging have been employed in 7 studies in pneumonia detection [23, 25, 32, 38, 40-42], 2 studies in lung detection [30, 49], 2 studies in COVID-19 detection [24, 27], 1 study in brain cancer prediction [46], 1 study in orthodontic diagnosis [26], 1 study in artery vein segmentation [44], 1 study predicting MVI in HCC patients [45], 1 study in tuberculosis detection [29], 8 studies in dental disease detection [31, 33, 35-37, 39, 43, 48], 2 studies in medical image classification/segmentation [50, 51], 1 study in liver masses classification [47], 2 studies in pneumonia & COVID-19 detection [28, 34], and 1 study in COVID-19 & lung detection [14]. The information may be found in Table 3 and Fig. 10 illustrates the stacked bar chart showing publication trends by year and split by sub-categories of domain. AI is commonly used to create automated software algorithms that enhance diagnostic and data management in medical imaging [6]. Essentially, they are tools used to assist clinical decisions and help experts like doctors make more educated judgments, particularly by utilizing ensemble models. These techniques have been utilized to improve the precision of diagnosing diseases such as COVID-19 [14] and dental



Fig. 9 Publication trends by year and split by sub-categories of modality

Table 3 The research specifics on disease diagnosis based on medical imaging using ensemble AI

Authors	Algorithms		Modality	Performances	Domain
(Year)	Baseline	Fusion			
Ko et al. (2019) [23]	RetinaNet & Mask R-CNN	Weighted voting ensemble	25,684 chest x-ray (CXR) images & 28,989 lung opacity bounding boxes	Mean average precision (mAP): 0.21746	Pneumonia detection
Livieris et al. (2019)[49]	SMO, C4.5, & kNN	Weighted voting ensemble	5830 chest x-ray (Pneu- monia), 566 lung mask (Tuberculosis), & 100 CT medical	Accuracy: Pneumonia 82.53–83.49%, Tuberculosis 69.79–71.73%, and CT medi- cal 69–77%	Detecting lung abnor- malities from chest X-rays
Brunese et al. (2020)[46]	First order, shape, gray level co-occurrence matrix, gray level run length matrix & gray level size zone matrix	Weighted voting ensemble	111,205 radiomic fea- tures extracted from MRI images	Accuracy of 99%	The benign grade I and the II, III, and IV malignant brain cancer detection
Deb and Jha (2020)[24]	NASNet, MobileNet & DenseNet	Ensemble	Chest x-rays	Accuracy of 91.99%	COVID-19 detection into 3 classes: Community Ac- quired Pneumonia (CAP), normal, & COVID-19.
Mao et al. (2020)[25]	RetinaNet & Mask R-CNN	Weighted aver- age ensemble	26,684 chest x-rays	Precision of 0.808, recall of 0.813, and mAP of 0.2283	Detect pneumonia
Suhail et al. (2020)[<mark>26</mark>]	Logistic regression & neural networks	Random forest ensemble	287 cephalometric x-rays	Accuracy: primary or alterna- tive outcome agreement 93–98%	Predicting orthodontic extractions
Chandra et al.(2021) [27]	ANN, KNN, NB, DT, SVM-lin- ear kernel, SVM-radial basis function, & SVM-polynomial kernel	Majority voting ensemble	2346 chest x-rays	Accuracy of 93.41%	Automatic COVID screen- ing (ACoS) system
Golla et al.(2021) [44]	2D and 3D versions of U-Net, V-Net, & DeepVesselNet	Ensemble	Abdominal CT scans	Dice similarity coefficients (DSC) of 0.758 for veins and 0.838 for arteries	Artery and vein segmentation
Jiang et al. (2021) [<mark>45</mark>]	RRC model & 3D-CNN models	XGBoost ensemble	CT images: 7302 radiomic features and 17 radiological from 405 patients	AUROC 0.887-0.906	Predicting MVI in HCC patients
Jin et al. (2021) [28]	AlexNet, ReliefF, & SVM	Hybrid ensemble	1743 chest x-rays	Overall accuracy rate: 98.642±0.398%	Diagnosis between COVID-19, common viral pneumonia, and normal
Rahman et al. (2021) [29]	ResNet101, VGG19, & DenseNet201	XGBoost ensemble	7000 chest x-rays	Accuracy of 99.92%, precision of 99.85%, & sensitivity of 100%	Detect tuberculosis (TB) and normal
Sagor et al. (2021) [<mark>30</mark>]	Inception-v3, VGG16, & ResNet-50	XGBoost ensemble	112,120 chest x-rays	Accuracy of 88.14%	Detecting lung disease
Tao et al. (2021) [14]	AlexNet, GoogleNet, & ResNet.	Relative majority voting ensemble	7,500 lung CT images	Accuracy of 99.05%	Detecting COVID-19 between lung tumor and normal lung
Bui et al. (2022) [<mark>3</mark> 1]	VGG, Resnet, Xception, & SVM	Majority voting ensemble	95 Panoramic x-rays generate 533 tooth regions	Accuracy of 93.58%, sensitiv- ity of 93.91%, & specificity of 93.33%	Automated caries screening
Gokul et al. (2022) [<mark>32</mark>]	CAPSNet & VDSNet	Snapshot ensemble	5,856 chest x-rays	Accuracy, precision, specific- ity: 0.98 & recall: 0.96	Pneumonia diagnosis between pneumonia and normal class
lmak et al. (2022) [<mark>33</mark>]	Multi-input AlexNet	Scored-based fu- sion ensemble	340 periapical x-rays	Accuracy of 99.13%	Caries detection between caries and non caries class
lqball and Wani (2022) [34]	ResNet101, InceptionV3, MobileNetV2, NasNet, & Xception	Weighted sum ensemble	1,088 chest x-rays	Accuracy, precision, recall: 100%	Pneumonia detection between COVID-19, pneumonia, normal

Table 3 (continued)

Authors	Algorithms		Modality	Performances	Domain
(Year)	Baseline	Fusion	-		
Jaiswal et al. (2022) [35]	ResNet50V2, ResNet101V2, MobileNetV3Small, Mo- bileNet, MobileNetV3Large, EfficientNetB0, Efficient- NetB1, & EfficientNetB2	XGBoost ensemble	500 panoramic x-rays	Accuracy: tooth wear 91.8%, periapical 92.2%, periodonti- tis 92.4%, tooth decay 93.2%, missing tooth 91.6%, and impacted tooth 90.8%	Multi oral disease classi- fication: tooth wear, peri- apical, periodontitis, tooth decay, missing tooth, and impacted tooth
Muller et al. (2022) [50]	DenseNet121, Efficient- NetB4, InceptionResNetV2, MobileNetV2, ResNeXt101, ResNet101, VGG16, Xception, & Vanilla architecture	Augmenting, bagging, & stack- ing ensemble	5,000 CHMNIST (histology), 2,905 COVID (X-ray), 25,331 ISIC (dermos- copy), & 35,126 DRD (ophthalmoscopy)	Average accuracy: CHMNIST 98%, COVID 96%, ISIC 92.75%, & DRD 81.5%	Medical image classifica- tion: CHMNIST, COVID, ISIC, and DRD
Alsubai (2023) [36]	Principle Component Analy- sis (PCA) & Chi-square (chi2)	Stacking ensemble (Ex- treme gradient boosting (XGB), random forest (RF), & extra trees classifier (ETC))	10,375 bitewing x-rays	Accuracy: 97.36%, precision: 96.14%, recall: 96,84%, & F1 score: 96.65%	Enhancing prediction of tooth caries and tooth normal
Azhari et al. (2023) [37]	ResNet50, ResNext101, & Vgg19	Ensemble U-Nets	771 bitewing x-rays: adult 554 & pediatric 217	IoU averaged of adult: zero 98%, primary 23%, moder- ate 19%, & advanced carious lesions 51% IoU averaged of pediatric: zero 97%, primary 8%, moderate 17%, & advanced carious lesions 25%	Detection of interproxi- mal carious lesions on primary and permanent dentition
Bhatt and Shah (2023) [<mark>38</mark>]	3 CNN models with kernel sizes: 3×3, 5×5, & 7×7	Weighted ensemble	5,863 chest x-rays	Accuracy, recall, precision, & F1 score: 99.23%	Pneumonia detection between pneumonia and normal
Haghanifar et al. (2023) [39]	CNN, InceptionNet, Encoder, & CheXNet,	PaXNet (Cap- sule network ensemble)	470 panoramic x-rays where the teeth total is 1229; 616 maxillary & 613 mandibular	Accuracy: 81.44% maxilla & 73.67% mandible	Tooth segmentation and dental caries detection: healthy and carious
Mabrouk et al. (2023) [40]	DenseNet169, MobileNetV2, & Vision Transformer	Ensemble learning	5,856 chest x-rays	Accuracy: 93.91%, precision: 93.96%, recall: 92.99%, & F1 score: 93.43%	Pneumonia detection between pneumonia and normal
Nakata and Siina (2023) [47]	16 CNNs (Xception, Incep- tionV3, InceptionResNetV2, ResNet50 & 101, ResNeXt50 & 101, SeResNetXt50 & 101, EfficientnetB0-B6)	Soft voting, weighted average voting, weighted hard voting, & stack- ing ensemble	Ultrasound im- ages: 6320 benign liver tumor (BLT), 2320 liver cyst (LCY), 9720 metastatic liver cancer (MLC), 7840 primary liver cancer (PLC)	Best ROC AUC: 0.944 BLT, 0.999 LCY, 0.891 MLC, & 0.903 PLC	Multiclass classification of ultrasound images of liver masses
Paul and Naskar (2023) [41]	DenseNet169, VGG16, & InceptionV3	Soft voting ensemble	5,856 children's lung chest x-ray	Accuracy of 92.79%	Children's pneumonia detection between pneu- monia and normal
Tareq et al. (2023) [48]	YOLO v5s, v5m, v5l, & v5x, ResNet50, VGG16, & DenseNet3	Ensemble	233 de-identified anterior teeth speci- mens and 1703 after augmentation	Accuracy of 86.96%	Visual diagnostics of dental caries
Dang et al. (2024) [51]	Deep neural networks (VGG16, ResNet34, & ResNet101)	Weighted ensemble	Medical image: 300 CVC-ColonDB, 912 CVC-Endo- SceneStill-2017, 808 MICCAI2015, 1000 CAMUS-ED, & 1000 CAMUS-ES	Dice score: CVC-ColonDB 0.95, CVC-EndoSceneS- till-2017 0.73, MICCAI2015 0.84, CAMUS-ED 0.94, & CAMUS-ES 0.94	Two-layer ensemble of deep learning mod- els for medical image segmentation

Table 3 (continued)

Authors	Algorithms		Modality	Performances	Domain	
(Year)	Baseline Fusion					
Gupta et al. (2024) [<mark>42</mark>]	DenseNet201, MobileNetV2, & InceptionResNetV2	Stacked ensemble	5,216 chest x-ray	Accuracy of 94%	Precise pediatric pneu- monia diagnosis between pneumonia and normal	
Marginean et al. (2024) [<mark>43</mark>]	U-Net, Feature Pyramid Network, & DeeplabV3	Ensemble learning	1,000 panoramic x-rays	Accuracy: 99.42% & mean dice coefficient 68.2%	Teeth segmentation and carious lesions segmentation	



Fig. 10 Publication trends by year and split by sub-categories of domain

caries [36]. The need for these systems is quickly growing because of their effectiveness in providing explanations and logical reasoning [1]. Clinical decision support systems that utilize AI, particularly the ensemble model, have the main objective of providing professional assistance to healthcare practitioners. While AI provides assistance to professionals, it also faces the difficulty of addressing the practical consequences of doctors' judgments [11]. In the field of medical imaging, AI mostly aids clinicians in the earliest stages of diagnosis, rather than making the final decision to proceed with a course of action. This SLR investigates the application of ensemble models in medical imaging for disease diagnosis. It evaluates the effectiveness of these models in various areas such as pneumonia/COVID-19/lung/tuberculosis detection, dental disease identification, brain cancer prediction, artery vein segmentation, predicting microvascular invasion in hepatocellular carcinoma patients, medical image classification/segmentation, and liver masses classification.

Table 3 shows that studies are increasingly using ensemble models in medical imaging. Deb and Jha [24] used ensemble-based combined NASNet, MobileNet, and DenseNet to assess the accuracy of COVID-19 detection from chest X-rays into 3 classes: pneumonia, normal, and COVID-19. The model has an accuracy of 91.99%, indicating a high level of accuracy, but the risk of overfitting due to the limited number of training images. In addition, this study has challenges in data quality and data heterogeneity. At the same time, Mao et al. [25], demonstrated comparable outcomes by employing a weighted average ensemble-powered CNN model (RetinaNet and mask R-CNN) for pneumonia detection. The computer-aided detection approach demonstrated a mean precision of 0.808 and a recall of 0.813. However, the study had a sample imbalance, with very few positive samples for training. Besides that, this study has challenges in distinguishing the quality of X-rays. Then, the practical application and clinical validation are insufficient.

Research for the detection of pneumonia and COVID-19 will continue to increase until 2024. Rahman et al. [29] used baseline ResNet101, VGG19, DenseNet201, and fusion XGBoost ensemble that achieved an accuracy of 99.92% for 2 classes: Tuberculosis and normal, but overfitting still occurred. Meanwhile, Iqball and Wani [34] used ResNet101, InceptionV3, and MobileNetV2 as baseline and Weighted sum ensemble as fusion, which achieved 100% accuracy, precision, and recall, respectively, for 3 classes: pneumonia, COVID-19, and normal. the study has limited or biased data can affect its performance. Bhatt and Shah [38], used 3 CNN models with kernel sizes: 3×3 , 5×5 , & 7×7 as a baseline and a weighted ensemble as fusion, which achieved accuracy, recall, precision, and F1 score of 99.23%, respectively, for 2 classes: pneumonia detection and normal detection. However, the study has data scarcity and may suffer from overfitting. And now Gupta et al. [42] used combined DenseNet201, MobileNetV2, and InceptionResNetV2, and then stacked ensemble as fusion, which achieved an accuracy of 94% for pediatric pneumonia diagnosis between pneumonia and normal. However, existing techniques may not possess the necessary level of sensitivity to pneumonia detection effectively. From 2019 to 2024 using chest x-rays for pneumonia/COVID-19 detection often use ensemble DL, where DL is used as the baseline.

Next, research for dental disease identification. Started by Suhail et al. [26] used logistic regression and neural networks as a baseline and random forest ensemble as fusion based on cephalometric x-rays for predicting orthodontic extractions, which achieved an accuracy of 93-98%, but may not capture all relevant clinical information. This study uses a dataset of 287 patients evaluated by five orthodontists, but its practical application and clinical validation in a real-world setting are still lacking. Imak et al. [33] used multi-input AlexNet and a scored-based fusion ensemble based on periapical x-rays for caries detection, which achieved an accuracy of 99.13%, but the dataset was relatively small. Alsubai [36] explained enhancing the prediction of tooth caries and tooth normal based on bitewing x-rays, which used PCA, chi2, and stacking ensemble as fusion. It achieved an accuracy of 97.36% but cannot provide a complete assessment of all mouth lesions in a single attempt. Marginean et al. [43] used U-Net, Feature Pyramid Network, and DeeplabV3 as baseline and ensemble learning, which achieved an accuracy of 99.42% based on panoramic x-rays for teeth segmentation and carious lesions segmentation but might introduce subjective bias.

A study for the detection of the benign grade I and II, III, and IV malignant brains based on MRI by Brunese et al. [46] used ML (first order, shape, gray level cooccurrence matrix, gray level run length matrix, and gray level size zone matrix) as a baseline and weighted voting ensemble as fusion, which achieved an accuracy of 99% but may not always be consistent. Research based on abdominal CT scans by Golla et al. [44] for artery and vein segmentation used 2D and 3D versions of U-Net, V-Net, and DeepVesselNet, which achieved a DSC of 0.758 for veins and 0.838 for arteries but unseen anomalies in the training data. At the same time, Jiang et al. [45] explained the RRC model, 3D-CNN models, and XGBoost for predicting MVI in HCC patients, with AUROC 0.887-0.906, so needs improvement. Nakata and Siina [47] used 16 different CNNs and ensemble (soft voting, weighted average voting, weighted hard voting, and stacking) for multiclass classification of ultrasound images of liver masses, with best ROC AUC: 0.944 BLT, 0.999 LCY, 0.891 MLC, & 0.903 PLC, but have image similarity each dataset and inadequate clinical validation. All these papers provide a forward-looking view using an ensemble based on medical imaging.

All the research in this SLR shows that the challenges and limitations that are often repeated are data scarcity, data quality, data heterogeneity, data imbalance, and difficulty to access because of confidential patient data, so the data used in the ensemble model is still relatively small and overfitting can still occur. In addition, it takes a lot of memory and processing resources to train and maintain several ensemble models. Even in the 30 papers on this SLR, there is no detailed discussion of the clinical validation and practical application. The main focus is on theoretical research and model performance analysis, with limited introduction of integrated model application cases in real clinical settings. These challenges highlight the necessity of effective and scalable ensemble models that can manage the intricacies of medical imaging assignments. To address data quality, studies emphasize preprocessing using techniques such as robust and also using augmentation methods. While for data heterogeneity using transfer learning methods or even a combination of models or carefully selecting and engineering features that are relevant to the task in medical imaging.

Conclusions

This study employs medical imaging to uncover the latest advancements and techniques in diagnosing disease using ensemble models and data-driven approaches. When implementing ensemble models for diagnosing diseases utilizing medical imaging in real-world scenarios, it is crucial to incorporate real patient data and utilize interpretable ensemble models to provide accurate explanations for the final predictions. A comprehensive review of 30 research papers indicates that further investigation is required to establish consistent effectiveness in healthcare environments. Despite the widespread dominance of deep learning in the field, the ensemble model continues to be extensively utilized by academics and practitioners as a method for decreasing variance (bagging) and bias (boosting) or increasing prediction (stacking). This SLR explains that researchers have observed the emergence of ensemble models for disease diagnosis using medical imaging because these models have shown increased accuracy, with the average accuracy of 30 research papers in SLR showing more than 90.14%. However, ensemble models require a lot of memory and processing resources to train and maintain them. Although ensemble models are able to overcome overfitting, in this SLR there is still overfitting due to data scarcity, data quality, data heterogeneity, data imbalance, and difficulty in accessing because patient data is confidential, so the data used in ensemble models is still relatively small. These challenges highlight the necessity of effective and scalable ensemble models that can manage the intricacies of medical imaging assignments. To mitigate this, techniques like crossvalidation, pruning, regularization, resampling, using different performance metrics, optimizing the code, and using efficient algorithms. Besides ensuring the privacy and security of patient data, AI must comply with stringent regulations to protect sensitive health information from breaches and misuse, so that maximizes its benefits while minimizing potential risks when integrating AI into clinical settings. The discussion sections may also provide guidance for future studies by highlighting the limitations of ensemble models. In the future, the utilization of the ensemble model in diagnosing diseases based on medical imaging is anticipated to reveal a multitude of unexplored prospects. In addition, it can be integrated with patient data and become a hybrid model or the practical application and clinical validation that provides recommendations for patients, but the doctor still makes the final decision.

Abbreviations

SLR	Systematic Literature Review
PRISMA	Preferred Reporting Items for Systematic Reviews and
	Meta-Analyses
CT	Computed Tomography
MRI	Magnetic Resonance Imaging
Al	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
WoS	Web of Science
MVI	Microvascular Invasion
HCC	Hepatocellular Carcinoma

Acknowledgements

The research is supported by the Faculty of Computing, Universiti Teknologi Malaysia (UTM); Research Center for Artificial Intelligence and Cyber Security, National Research and Innovation Agency (BRIN); and School of Dental Sciences, Universiti Sains Malaysia (USM).

Author contributions

M.R.S was the main author, drafted the manuscript, and performed data analysis. A.B.A.S., J.M., R.A.R.A., N.H.I., H.A.M., M.S.B.O., and S.Z.B.M.H designed and supervised the study. All authors read and approved the final manuscript.

Funding

This study was supported by the Fundamental Research Grant Scheme (FRGS/1/2023/ICT02/UTM/03/1) from the Malaysian Ministry of Higher Education.

Data availability

All data generated or analysed during this study are included in this published article and its supplementary information files.

Declarations

Ethical approval and consent to participate Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 16 October 2024 / Accepted: 7 April 2025 Published online: 18 April 2025

References

- Yeasmin MN, Amin MA, Joti TJ, Aung Z, Azim MA. Advances of Al in imagebased computer-aided diagnosis: a review. Array. 2024;23:1–23.
- Barragan-Montero A, Javaid U, Valdes G, Nguyen D, Desbordes P, Macq B, Willems S, Vandewinckele L, Holmstrom M, Lofman F, Michiels S, Souris K, Sterpin E, Lee JA. Artificial intelligence and machine learning for medical imaging: a technology review. Phys Med. 2021;83.
- Cai L, Gao J, Zhao D. A review of the application of deep learning in medical image classification and segmentation. Ann Translat Med. 2020;1–15.
- Liu F, Hernandez-Cabronero M, Sanchez V, Marcellin MW, Bilgin A.The current role of image compression standards in medical imaging. Information. 2017;8(131):1–26.
- Abhisheka B, Biswas SK, Purkayastha B, Das D, Escargueil A. Recent trend in medical imaging modalities and their applications in disease diagnosis: a review. Multimedia Tools Appl. 2024;83:43035–70.
- Azizi A, Azizi M, Nasri M. Artificial intelligence techniques in medical imaging: a systematic review. Int J Online Biomed Eng. 2023;19(17):66–97.
- 7. Zagzebski JA. Essentials of ultrasound physics. Elsevier Health Sci. 2013;1:123.
- Deshmukh PK. Improving medical image classification using ensemble learning and deep convolutional neural networks. Int J Intell Syst Appl Eng. 2023;12:106–21.
- Akst J. A primer: artificial intelligence versus neural network. Inspiring Innovation: The Scientist Exploring Life. 2019;1:65802.
- 10. Kumar HMK, Harish BS. Automatic irony detection using feature fusion and ensemble classifier. Int J Interact Multimed Artif Intell. 2019;5(7):70–9.
- Imran S, N P. A review on ensemble machine and deep learning techniques used in the classification of computed tomography medical images. Int J Health Sci Res. 2024;14(1):201–13.
- Kamal P, Ahuja S. An ensemble-based model for prediction of academic performance of students in undergrad professional course. J Eng Design Technol. 2019;17(4):769–81.
- 13. Mohammed A, Kora R. A comprehensive review on ensemble deep learning: opportunities and challenges. J King Saud Univ. 2023;35:757–74.
- Tao Z, Huiling L, Zaoli Y, Shi Q, Bingqiang H, Yali D. The ensemble deep learning model for novel COVID-19 on CT images. Appl Soft Comput J. 2021;98:1–9.
- Mangla M, Sharma N, Mohanty SN. A sequential ensemble model for software fault prediction. Innov Syst Softw Eng. 2022;18:301–08.
- Tang J, Su Q, Su B, Fong S, Cao W, Gong X. Parallel ensemble learning of convolutional neural networks and local binary patterns for face recognition. Comput. Methods Programs Biomed. 2020;197:1–7.
- 17. Akinbo RS, Daramola OA. Ensemble Machine Learning Algorithms for Prediction and Classification of Medical Images. IntechOpen. 2021;7:1–20.
- 18. McGrath TA, Alabousi M, Skidmore B, Korevaar DA, Bossuyt PMM, Moher D, Thombs B, McInnes MDF. Recommendations for reporting of systematic

reviews and meta-analyses of diagnostic test accuracy: a systematic review. System Rev. 2017;6:194.

- 19. Chitu O. A Guide to Conducting a Standalone Systematic Literature Review. Commun Assoc Inf Syst. 2015;37:879–910.
- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Bmj. 2021;372:1–9.
- 21. Fahimnia B, Sarkis J, Davarzani H. Green supply chain management: a review and bibliometric analysis. Int J Prod Econ. 2015;162:101–14.
- 22. Zhu J, Liu W. A tale of two databases: the use of web of science and scopus in academic papers. Scientometrics. 2020;123:321–35.
- Ko H, Ha H, Cho H, Seo K, Lee J. Pneumonia detection with weighted voting ensemble of CNN models. In 2nd International Conference on Artificial Intelligence and Big Data; 2019.
- 24. Deb SD, Jha RK. COVID-19 detection from chest X-Ray images using ensemble of CNN models. In International Conference on Power, Instrumentation, Control and Computing (PICC); 2020.
- Mao L, Yumeng T, Lina C, Pneumonia detection in chest X-rays: a deep learning approach based on ensemble RetinaNet and Mask R-CNN. In Eighth International Conference on Advanced Cloud and Big Data (CBD); 2020.
- Suhail Y, Upadhyay M, Chhibber A, Kshitiz. Machine learning for the diagnosis of orthodontic extractions: a computational analysis using ensemble learning. Bioengineering. 2020;7(55):1–13.
- 27. Chandra TB, Verma K, Singh BK, Jain D, Netam SS. Coronavirus disease (COVID-19) detection in Chest X-Ray images using majority voting based classifier ensemble. Expert Syst Appl. 2021;165:1–13.
- Jin W, Dong S, Dong C, Ye X. Hybrid ensemble model for differential diagnosis between COVID-19 and common viral pneumonia by chest X-ray radiograph. Comput Biol Med. 2021;131:1–12.
- Rahman M, Cao Y, Sun X, Li B, Hao Y. Deep pre-trained networks as a feature extractor with XGBoost to detect tuberculosis from chest X-ray. Comput. Electr. Eng. 2021;93:1–16.
- Sagor MK, Dipto SM, Jahan I, Chowdhury S, Reza MT, Alam MA. An efficient deep learning approach for detecting lung disease from chest x-ray images using transfer learning and ensemble modeling. In Asia-Pacific Conference on Computer Science and Data Engineering (CSDE); 2021.
- 31. Bui TH, Hamamoto K, Paing MP. Automated caries screening using ensemble deep learning on panoramic radiographs. Entropy. 2022;24:1–12.
- 32. Gokul G, Kumaratharan LR, Devi. Ensembling framework for pneumonia detection in chest X-ray images. In International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN); 2022.
- Imak A, Celebi A, Siddique K, Turkoglu M, Sengur A, Salam I. Dental caries detection using score-based multi-input deep convolutional neural network. IEEE Access. 2022;10:18320–29.
- Iqball T, Wani MA. COVID-19 and pneumonia detection using deep weighted ensemble model. In 9th International Conference on Computing for Sustainable Global Development (INDIACom); 2022.
- Jaiswal P, Katkar V, Bhirud SG. Multi oral disease classification from panoramic radiograph using transfer learning and XGBoost. Int J Adv Comp Sci Appl. 2022;13(12):239–49.
- 36. Alsubai S. Enhancing prediction of tooth caries using significant features and multi-model classifier. Peer J Comput Sci. 2023;9:1–24.
- Azhari AA, Helal N, Sabri LM, Abduljawad A. Artificial intelligence (AI) in restorative dentistry: performance of AI models designed for detection of interproximal carious lesions on primary and permanent dentition. Digital Health. 2023;9:1–13.

- Bhatt H, Shah M. A convolutional neural network ensemble model for pneumonia detection using chest X-ray images. Healthcare Analytics. 2023;3:1–6.
- Haghanifar A, Majdabadi MM, Haghanifar S, Choi Y, Ko S-B. PaXNet: tooth segmentation and dental caries detection in panoramic X-ray using ensemble transfer learning and capsule classifier. Multimedia Tools Appl. 2023;82:27659–79.
- Mabrouk A, Redondo RP, Dahou A, Elaziz MA, Kayed M. Pneumonia detection on chest X-ray images using ensemble of deep convolutional neural networks. arXiv. 2023;12(13):1–10.
- Paul M, Naskar R. Deep learning enabled pneumonia detection from chest X-rays: a transfer learning based ensemble classification approach. In 3rd International Conference on Smart Technologies for Power, Energy and Control (STPEC); 2023.
- 42. Gupta SJ, Kumar R, Singh BP, Bansal S, Kaur H, Pandey MC. Unleashing the power of deep convolutional neural networks in a stacked ensemble model for precise paediatric pneumonia diagnosis from chest radiographs. In International Students' Conference on Electrical, Electronics and Computer Science; 2024.
- Marginean AC, Muresanu S, Hedesiu M, Diosan L. Teeth segmentation and carious lesions segmentation in panoramic x-ray images using cariseg, a networks ensemble. Heliyon. 2024;10:1–14.
- 44. Golla A-K, Bauer DF, Schmidt R, Russ T, Norenberg D, Chung K, Tonnes, C, Schad LR, Zollner FG. Convolutional neural network ensemble segmentation with ratio-based sampling for the arteries and veins in abdominal CT scans. IEEE Trans Biomed Eng. 2021;68(5):1518–26.
- Jiang Y, Cao S, Cao S, Chen J, Wang G, Shi W, Deng Y, Cheng N, Ma K, Zeng K, Yan X, Yang H, Huan W, Tang W, Zheng Y, Shao C, Wang J, Yang Y, Chen G. Preoperative identification of microvascular invasion in hepatocellular carcinoma by XGBoost and deep learning. J Cancer Res Clin Oncol. 2021;147:821–33.
- Brunese L, Mercaldo F, Reginelli A, Santone A. An ensemble learning approach for brain cancer detection exploiting radiomic features. Comput Methods Programs Biomed. 2020;185:1–16.
- Nakata N, Siina T.Ensemble learning of multiple models using deep learning for multiclass classification of ultrasound images of hepatic masses. Bioengineering. 2023;10(69):1–20.
- Tareq A, Faisal MI, Islam MS, Rafa NS, Chowdhury T, Ahmed S, Farook TH, Mohammed N, Dudley J. Visual diagnostics of dental caries through deep learning of non-standardised photographs using a hybrid YOLO ensemble and transfer learning model. Int J Environ Res Public Health. 2023;20:1–13.
- Livieris IE, Kanavos A, Tampakas V, Pintelas P. A weighted voting ensemble self-labeled algorithm for the detection of lung abnormalities from X-Rays. Algorithms. 2019;12(64):1–15.
- Muller D, Soto-Rey I, Kramer F. An analysis on ensemble learning optimized medical image classification with deep convolutional neural networks. IEEEAccess. 2022;10:66467–80.
- Dang T, Nguyen TT, McCall J, Elyan E, Moreno-García CF. Two-layer ensemble of deep learning models for medical image segmentation. Cogni Comput. 2024;16:1141–60.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.