



REVIEW

The Role of Artificial Intelligence in Improving Diagnostic Accuracy in Medical Imaging: A Review

Omar Sabri¹, Bassam Al-Shargabi^{2,*} and Abdelrahman Abuarqoub²

¹Zekelman School of Information Technology, St. Clair College, Windsor, ON N9A 6S4, Canada

²Cardiff School of Technologies, Cardiff Metropolitan University, Cardiff, CF5 2YB, UK

*Corresponding Author: Bassam Al-Shargabi. Email: bal-shargabi@cardiffmet.ac.uk

Received: 22 April 2025; Accepted: 08 August 2025; Published: 23 September 2025

ABSTRACT: This review comprehensively analyzes advancements in artificial intelligence, particularly machine learning and deep learning, in medical imaging, focusing on their transformative role in enhancing diagnostic accuracy. Our in-depth analysis of 138 selected studies reveals that artificial intelligence (AI) algorithms frequently achieve diagnostic performance comparable to, and often surpassing, that of human experts, excelling in complex pattern recognition. Key findings include earlier detection of conditions like skin cancer and diabetic retinopathy, alongside radiologist-level performance for pneumonia detection on chest X-rays. These technologies profoundly transform imaging by significantly improving processes in classification, segmentation, and sequential analysis across diverse modalities such as X-rays, Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and ultrasound. Specific advancements with Convolutional Neural Networks, Recurrent Neural Networks, and ensemble learning techniques have facilitated more precise diagnosis, prediction, and therapy planning. Notably, Generative Adversarial Networks address limited data through augmentation, while transfer learning efficiently adapts models for scarce labeled datasets, and Reinforcement Learning shows promise in optimizing treatment protocols, collectively advancing patient care. Methodologically, a systematic review (2015–2024) used Scopus and Web of Science databases, yielding 7982 initial records. Of these, 1189 underwent bibliometric analysis using the R package ‘Bibliometrix’, and 138 were comprehensively reviewed for specific findings. Research output surged over the decade, led by Institute of Electrical and Electronics Engineers (IEEE) Access (19.1%). China dominates publication volume (36.1%), while the United States of America (USA) leads total citations (5605), and Hong Kong exhibits the highest average (55.60). Challenges include rigorous validation, regulatory clarity, and fostering clinician trust. This study highlights significant emerging trends and crucial future research directions for successful AI implementation in healthcare.

KEYWORDS: Artificial intelligence; artificial intelligence applications; deep learning; medical imaging; diagnostic accuracy; bibliometric analysis

1 Introduction

Artificial intelligence (AI) is rapidly emerging as a transformative force across the entire healthcare spectrum, offering unprecedented opportunities to enhance clinical efficiency, improve patient outcomes, and redefine medical practice. It has been successfully applied in image analysis within radiology, pathology, and dermatology, achieving diagnostic speeds that surpass and accuracy that parallels those of medical experts. AI has the potential to optimise the care trajectory for chronic disease patients, suggest precision therapies, reduce medical errors, and enhance subject enrolment in clinical trials. Additionally, AI’s application of natural language processing (NLP) to scientific literature and electronic medical records will



significantly impact medical practice. Machine learning (ML) from medical data can help mitigate clinical errors caused by human cognitive biases, thereby enhancing patient care. Despite these advancements, physicians will remain essential in cognitive medical practice [1]. In recent years, the integration of AI technologies into medical imaging diagnostics has revolutionised healthcare practices, offering promising avenues for enhanced diagnostic accuracy, patient care, and treatment outcomes [2,3]. This transformative shift emphasises the potential of AI to address enduring challenges in medical imaging, ranging from improving diagnostic precision to streamlining clinical workflows [4–6]. Furthermore, as AI continues to advance, leveraging ML algorithms, deep learning (DL) techniques, and computer vision methodologies, it presents unique opportunities for innovation in medical imaging diagnostics [7]. The convergence of AI and medical imaging holds particular significance in addressing the diagnostic complexities and demands of modern healthcare systems. Medical imaging modalities such as X-ray, Computed Tomography (CT), magnetic resonance imaging (MRI), and ultrasound play pivotal roles in disease detection, treatment planning, and therapeutic monitoring [8].

Against this backdrop, the adoption of AI technologies in medical imaging diagnostics offers a paradigm shift, empowering healthcare professionals with advanced computational tools for image analysis, pattern recognition, and decision support [9,10]. AI-driven approaches hold huge potential to augment human capabilities, enabling faster, more accurate diagnoses, personalised treatment strategies, and improved patient outcomes. Moreover, AI algorithms can leverage vast datasets and learn from iterative feedback loops, continuously refining their performance and adaptability to diverse clinical scenarios [4,11]. In this context, medical imaging's role in diagnosing and predicting various medical conditions is paramount [12]. Traditionally, radiologists have borne the sole responsibility for identifying pathology and predicting outcomes. However, challenges such as fatigue, human perception limitations, and cognitive overload can lead to overlooking subtle details in the images. The integration of AI tools into medical imaging has significantly enhanced diagnostic accuracy and alleviated the workload on human radiologists [13]. AI applications in breast cancer detection and diagnostic imaging have brought about remarkable advancements, promising better patient outcomes in the future. AI aids in the early detection of diseases like COVID-19, precise measurement of tumors and organ volumes, and facilitates the use of AI in smaller medical facilities, reducing the need for specialized expertise [14]. Many prevalent and severe diseases now have the potential for early diagnosis and prognosis using AI alone. The introduction of AI holds the potential to transform healthcare quality, reduce costs, personalize drug therapies, enhance care coordination, and reduce clinical burdens.

However, despite the recognition of AI's potential, there remain several critical research gaps in the literature on AI applications in medical imaging diagnostics. This review aims to address the research gap by comprehensive evaluation of existing literature on AI applications in medical imaging diagnostics. This review aims to address these gaps by evaluating existing literature comprehensively. Currently, there is a limited synthesis of evidence as individual studies have examined the performance of AI algorithms without integrating findings across studies [10,15,16]. The review will consolidate the evidence to assess the effectiveness and clinical impact of AI-enabled diagnostic tools comprehensively. For a more detailed foundational understanding of AI's historical development and its specific applications within medical imaging, readers are referred to [Section 3: Background and Literature Review](#).

To this end, the following research questions guide our investigation:

- (1) *How have ML and DL technologies improved diagnostic accuracy in medical imaging, and what challenges remain for their real-world implementation?*
- (2) *What are the existing gaps in the literature regarding AI applications in medical imaging diagnostics, and how can future research address these gaps?*

- (3) *How has the research output on AI applications in medical imaging diagnostics evolved between 2015 and 2024?*
- (4) *What factors contribute to the decline in citation impact per article, despite the growth in research output on AI and DL applications in medical imaging?*
- (5) *What are the emerging themes and trending topics in the application of AI to medical imaging diagnostics?*

By addressing these questions, the review aims to highlight the strengths and limitations of AI, ML, and DL diagnostic tools, provide insights into the current state of research on AI applications in medical imaging diagnostics, and facilitate evidence-based decision-making in clinical practice, healthcare policy, and future research. The primary objectives of this review are to investigate the performance of AI algorithms, ML, and DL methods in medical imaging diagnostics, compare the diagnostic accuracy of AI models with traditional diagnostic methods, and assess the clinical impact and potential benefits of AI-enabled diagnostic tools on patient outcomes and healthcare delivery.

To provide a comprehensive investigation into the literature on ML and DL in medical imaging diagnostics, this review is structured as follows: [Section 2](#) outlines the methodology employed in the review, ensuring a thorough examination of relevant studies; [Section 3](#) provides a detailed background and literature review, setting the context for the research; [Section 4](#) presents the results and analysis, highlighting key findings and trends; and finally, [Section 5](#) offers a discussion of the implications of these findings, drawing meaningful conclusions and suggesting directions for future research.

2 Methods

The investigation into AI applications for improving diagnostic accuracy was carried out using a systematic and structured approach, combining bibliometric methods with recognized review protocols. Notably, we employed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework to guide the search, screening, and selection of relevant studies, ensuring transparency and reproducibility throughout the review process [17,18]. We utilized the R package “Bibliometrix” as the primary analytical tool, given its robust capabilities for evaluating scientific literature across multiple disciplines [19–24]. Additionally, the Biblioshiny web interface, an extension of Bibliometrix, facilitated a more user-friendly environment for conducting detailed bibliometric analysis, including co-citation, keyword co-occurrence, and thematic evolution. Articles were sourced from two major academic databases—Scopus and Web of Science—chosen for their wide coverage and indexing of high-impact journals. We focused on publications from reputable, peer-reviewed sources to ensure the inclusion of high-quality and relevant research. The review process involved a time-interval analysis to examine trends over time and a thematic mapping of research development across geographical regions and disciplines. Bibliometric analysis was employed to identify key research trends, influential publications, emerging themes, and collaboration networks. This method has become increasingly valuable in highlighting the frequency and impact of scholarly output, as well as the evolution of topics within a specific research domain [25–27]. This allows for valuable insights into particular research areas [20]. This dataset was analyzed using bibliometric methods with the R package Biblioshiny, facilitating efficient filtering, organization, and visualization of trends [19,28]. By leveraging this approach, we identified key themes and influential authors that might be overlooked in manual reviews. This section is partitioned into two subcategories: [Section 2.1](#), titled “Search Strategy and Study Selection,” outlines the process we used to identify and choose relevant research. [Section 2.2](#), titled “Data Extraction and Analysis,” provides an explanation of how we extracted and processed the data.

2.1 Search Strategy

To identify relevant studies focusing on the role of AI in improving diagnostic accuracy in medical imaging, we developed a comprehensive search strategy using a combination of keywords. These keywords were selected based on preliminary literature scans and expert consultation to capture the core concepts of the review: “artificial intelligence,” “deep learning,” “machine learning,” “medical imaging,” and “diagnostic accuracy.” We systematically combined these terms using Boolean operators (e.g., AND, OR) to construct search queries. An example of a search string used across databases was: ((“artificial intelligence” OR “deep learning” OR “machine learning”) AND (“medical imaging” OR “radiology” OR “image analysis”) AND (“diagnostic accuracy” OR “diagnosis” OR “detection” OR “prediction”)).

Articles were sourced from two major academic databases: Scopus and Web of Science. These databases were chosen for their extensive coverage of peer-reviewed scientific literature across various disciplines, including medicine, computer science, and engineering, thus ensuring a broad yet high-quality inclusion of relevant research. Our literature search was conducted on 15 September 2024, and was limited to English-language articles and reviews published from 1 January 2015, onwards (up to the search date of 15 September 2024).

The search strategy yielded an initial 7982 records from the Scopus and Web of Science databases. These records were then imported into a reference management tool (e.g., EndNote, Zotero) for deduplication and initial filtering based on publication type. During this initial filtering, 2538 records were excluded as they were identified as books, editors’ notes, conference proceedings, or other non-journal article formats. This process resulted in a refined set of 5444 unique records for subsequent screening.

2.2 Inclusion and Exclusion Criteria

Following the initial identification and pre-screening, the remaining 5444 unique records underwent a rigorous screening process based on predefined inclusion and exclusion criteria. Studies were included if they were original research articles or review articles published in English between 1 January 2015 and 15 September 2024. A core criterion was their focus on the application of Artificial Intelligence, encompassing ML and DL techniques, specifically in medical imaging for diagnostic purposes. Furthermore, included studies had to report on or directly discuss diagnostic accuracy, performance, or outcomes (such as sensitivity, specificity Area Under Curve(AUC), or overall accuracy) related to medical image analysis, and involve human or animal medical imaging data (including, but not limited to, X-ray, CT, MRI, Ultrasound, Pathology slides, and Fundus images). Conversely, studies were excluded if they were conference abstracts, posters, editorials, letters to the editor, commentaries, or book chapters (as these were largely filtered in an initial step). Unpublished papers, such as preprints not yet peer-reviewed by the time of the search, were also excluded. Other exclusion criteria included studies not primarily focused on diagnostic applications of AI in medical imaging (e.g., those solely on image reconstruction, image segmentation without a diagnostic outcome, or treatment planning without a diagnostic component), studies that did not report diagnostic performance metrics or discuss diagnostic accuracy, and non-English language publications.

2.3 Study Selection Process

The screening process involved two independent reviewers. In the first phase of eligibility screening, titles and abstracts of the 5444 unique records were screened against the inclusion and exclusion criteria. During this step, 163 records were excluded as non-English journal articles, leading to 5281 records proceeding to the next stage.

In the second phase, the titles and abstracts of these 5281 records underwent a more detailed assessment for relevance. 4092 records were excluded at this stage based on their titles, abstracts, or a preliminary full-text review if needed, as they did not meet the core relevance criteria for AI in diagnostic imaging. Any conflicts or disagreements between the two reviewers during these screening phases were resolved through discussion to reach a consensus. If consensus could not be reached, a third senior reviewer was consulted for final arbitration. This comprehensive screening identified 1189 potentially relevant studies for full-text review.

In the final phase of inclusion, the full texts of these 1189 studies were retrieved and meticulously assessed against the complete inclusion and exclusion criteria. Again, two reviewers independently conducted this full-text review, with disagreements resolved by discussion or third-party arbitration. This systematic approach ensured that only the most relevant and methodologically sound studies were selected.

The final dataset for our analysis was then partitioned: all 1189 identified studies formed the basis for the bibliometric analysis, providing a broad overview of publication trends and patterns within the field. From this larger set of 1189 studies, 138 studies were specifically selected for an in-depth “Comprehensive Review.” These 138 studies were chosen based on their direct relevance to diagnostic accuracy outcomes, significant methodological contributions, and representativeness of the diverse AI applications discussed. This allowed for a more focused qualitative synthesis while maintaining the breadth of the bibliometric trends. [Fig. 1](#) illustrates the detailed study screening and selection process, consistent with PRISMA guidelines.

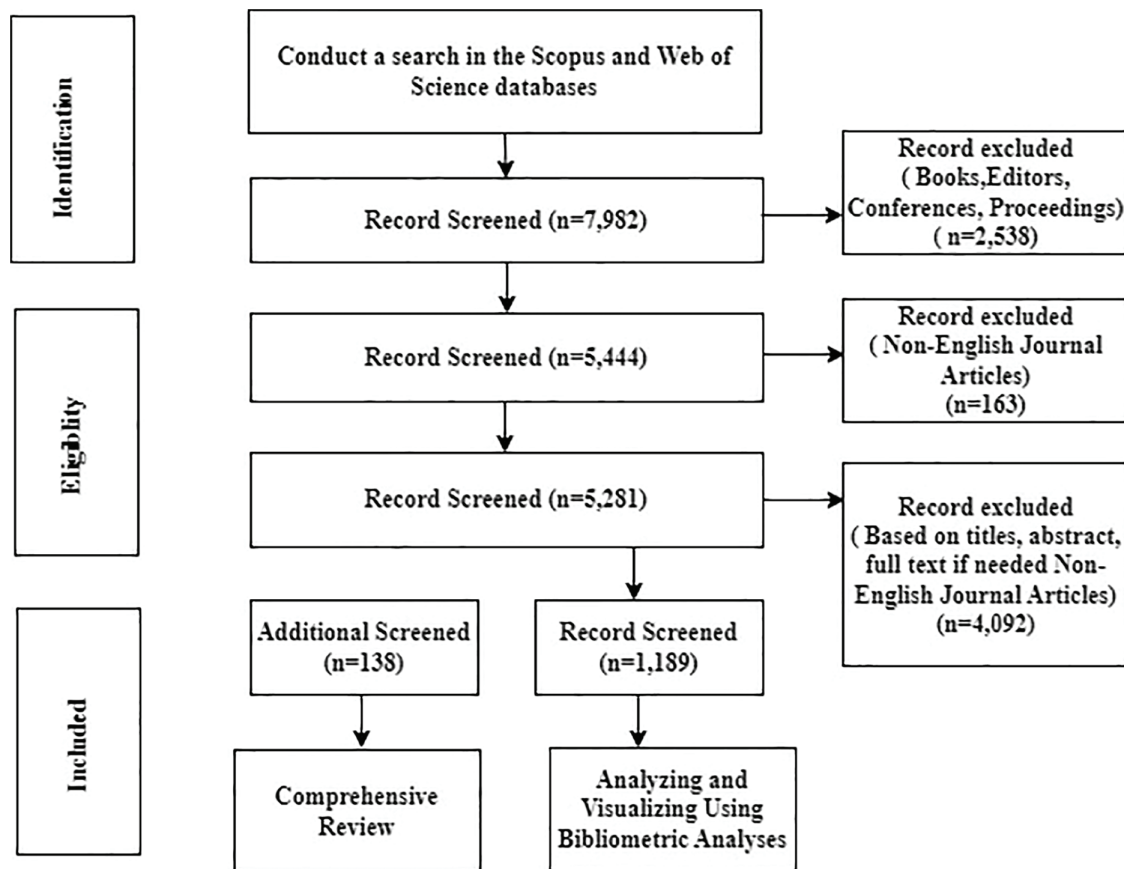


Figure 1: Review process

2.4 Data Extraction and Analysis

Data extraction was performed systematically using a standardized form to gather information on study characteristics (e.g., year of publication, country of origin, type of medical imaging modality, disease area), details of AI models used (e.g., ML/DL technique, architecture), and reported diagnostic performance metrics (e.g., sensitivity, specificity, accuracy, AUC). To maintain accuracy and minimize bias, two reviewers independently conducted this data extraction process [29,30].

The full set of 1189 studies was then analyzed using the Bibliometrix R package, allowing us to summarize publication trends, identify key research themes, influential authors, and collaboration networks within the field. This quantitative bibliometric analysis facilitated efficient filtering, organization, and visualization of trends [19,28]. By leveraging this approach, we identified overarching patterns and influential works that might be overlooked in manual reviews, adding quantitative rigor to our insights into citation trends and key research areas.

Our analysis also included a qualitative synthesis of the diagnostic performance of various AI algorithms across different medical imaging modalities and disease types. This synthesis primarily drew insights from the extracted data of the 138 studies selected for the Comprehensive Review. We considered the potential for meta-analyses and subgroup analyses based on factors like imaging modality, disease type, and AI model architecture, though a formal meta-analysis was beyond the scope of this critical review. Additionally, we examined potential biases within the literature, such as publication and selective reporting biases, to contextualize our findings. As our review synthesized existing published literature, ethical approval was not required. This comprehensive and methodologically sound approach is designed to provide a robust overview of the field, with Fig. 1 illustrating the study screening and selection process.

3 Background and Literature Review

This discussion centres on the fundamental concepts and recent advancements in the application of AI to medical imaging, exploring how ML and DL are transforming diagnostic methods and enhancing patient outcomes. Section 3.1 provides an overview of AI's role in medical imaging, setting the stage for more specific discussions. Section 3.2 focuses on ML in medical imaging, examining various ML techniques and their contributions to the analysis of medical images. Section 3.3 then shifts to DL in medical imaging, highlighting how advanced DL algorithms are extending the capabilities of image interpretation and increasing diagnostic precision. Together, these sections offer a comprehensive overview of the evolving landscape of AI technologies and their impact on medical diagnostics.

3.1 Overview

AI has become a game-changer in healthcare, particularly in medical imaging, where it's making significant strides in diagnostic accuracy and patient care [2,31–33]. There's been a notable increase in interest in harnessing AI to boost diagnostic precision, enhance patient outcomes, and streamline healthcare processes [5]. This literature review aims to provide a comprehensive overview of the current state of research on AI applications in medical imaging diagnostics, focusing on key trends, challenges, and opportunities in this rapidly evolving field.

The integration of AI into medical imaging spans several areas, including radiology, pathology, and ophthalmology. AI-driven methods have shown impressive results in enhancing diagnostic accuracy and efficiency across various imaging modalities. For instance, deep neural networks have achieved dermatologist-level performance in skin cancer classification [12,34], demonstrated dermatologist-level classification of skin cancer using deep neural networks, achieving diagnostic performance comparable to

expert clinicians. Similarly, reference [4] developed a DL algorithm for detecting diabetic retinopathy in retinal fundus photographs, highlighting AI's efficacy in ophthalmic diagnostics. AI algorithms, particularly those based on DL techniques, have shown remarkable capabilities in processing, and interpreting complex medical images, such as X-rays, CT scans, MRI, and ultrasound scans [8,35,36]. These AI-powered systems can detect abnormalities, assist in diagnosis, and provide quantitative assessments of disease severity with high accuracy and efficiency. By leveraging large datasets and sophisticated neural networks, AI models can learn intricate patterns and features from medical images, enabling more accurate and consistent diagnostic interpretations in radiology, the advent of DL techniques has revolutionised image interpretation and analysis. Rajpurkar et al. [37] introduced CheXNet, a deep-learning model for pneumonia detection on chest X-rays, achieving radiologist-level performance. Litjens et al. [5] surveyed DL methods in medical image analysis, highlighting its widespread adoption and impact across diverse clinical domains. Pathology has also witnessed significant advancements in AI-enabled diagnostics [38]. Ting et al. [39] developed a DL system for diabetic retinopathy detection in multiethnic populations, demonstrating robust performance across different demographic groups. Hosny et al. [40] reviewed the role of AI in radiology, emphasizing its potential to enhance diagnostic accuracy, workflow efficiency, and patient care. They also applied deep learning techniques to radiomics data for lung cancer prognostication across multiple cohorts [41].

However, the clinical implementation of AI technologies in medical imaging requires rigorous validation, regulatory approval, and ongoing monitoring to ensure their safety, effectiveness, and ethical use in clinical practice. Medical imaging plays a crucial role in disease diagnosis, treatment planning, and monitoring patient response to therapy. However, traditional diagnostic approaches are often subject to limitations, including observer variability, diagnostic errors, and inefficiencies in image interpretation [42–44]. These challenges underscore the need for innovative solutions to enhance diagnostic accuracy and efficiency in medical imaging.

Although AI has shown great progress in medical imaging, a major obstacle to its practical adoption is the lack of transparency in many advanced models—especially DL algorithms—which often operate as “black boxes. While these models can achieve impressive accuracy, their internal decision-making processes are often opaque, making it difficult for medical professionals to understand why a particular diagnosis or prediction was made. This lack of transparency can hinder the adoption of AI in clinical settings, where trust, accountability, and the ability to verify decisions are paramount.

To address this, Explainable Artificial Intelligence (XAI) has emerged as a crucial field of research [45]. XAI aims to make AI models more transparent, interpretable, and understandable to humans. In medical imaging, XAI methodologies are vital for several reasons. First, they enable clinicians to gain insights into the features or patterns an AI model deems most relevant for a diagnosis, fostering trust and confidence in AI's recommendations. This is particularly important for high-stakes decisions, such as cancer detection or disease progression prediction, where erroneous or unexplainable outputs could have serious consequences for patient care. Significant contributions in this area, notably by Parola et al. [46] and Cimino et al. [47], highlight the development of human-centered XAI solutions for clinical users. Their work introduces methodologies like Informed Deep Learning (IDL) and Case-Based Reasoning (CBR) to effectively integrate medical domain expertise and provide intuitive visual output explanations. This approach not only enhances diagnostic accuracy, even with imperfect images, labeling inaccuracies, and artifacts from low-cost equipment, but also ensures explanations are congruent with clinical user demands and support post-hoc analysis for improved human-AI collaboration in decision-making. Second, XAI can facilitate error detection and model improvement. By understanding the reasoning behind an AI's misclassification, researchers and developers can identify biases in training data, refine model architectures, or adjust parameters to enhance performance and reliability. Third, XAI supports regulatory compliance and ethical considerations. As AI

deployment in healthcare becomes more widespread, there is a growing need for models that can be audited, validated, and justified in accordance with ethical guidelines and legal requirements [48]. XAI techniques, such as data explainability, model explainability, and post-hoc explainability, provide mechanisms to achieve this by offering insights into input data, internal model workings, and justifications for specific outputs, respectively [45]. The ultimate goal of XAI in medical imaging is not merely to provide explanations, but to contribute to the development of trustworthy AI systems that seamlessly integrate with clinical workflows, empower healthcare professionals, and ultimately improve patient outcomes by offering both accuracy and clarity in diagnostic decision-making.

DL and ML are both subsets of AI, but they differ significantly in their approaches to data analysis [35,49,50]. DL, characterized by its use of artificial neural networks with multiple layers, excels at extracting complex patterns from large datasets [38,50]. This makes it particularly well-suited for tasks like image and speech recognition. Conversely, ML algorithms rely on predefined features and rules to make predictions [51–53]. While often interpretable, ML models may be less accurate than DL models for highly complex problems. The authors provide a summary of the general comparison between ML and DL, as shown in Table 1, outlining their differences and strengths.

Table 1: General comparison between DL and ML

Aspect	DL	ML
Definition	A subset of ML that uses complex neural networks with multiple layers to identify intricate patterns in data.	Uses various algorithms to analyse data and make predictions or decisions based on observed patterns.
Data requirements	Needs large amounts of labeled data for effective training.	Can work with smaller datasets and may also need labeled data for supervised learning.
Feature engineering	Automatically extract features from raw data, eliminating the need for manual feature engineering.	Requires manual extraction of relevant features from the data.
Model complexity	Models are highly complex with multiple layers and parameters.	Models can vary in complexity, from simple algorithms to more advanced ones.
Performance	Often achieve high accuracy in tasks like image recognition and NLP.	Performance varies depending on the algorithm and the quality of feature engineering.
Interpretability	Models are often seen as “black boxes,” making their decision-making process harder to understand.	ML models can be more interpretable, allowing for a better understanding of the decision-making process.
Training time	Training can be very computationally intensive and time-consuming.	Typically has faster training times compared to DL models.
Application areas	Commonly used in fields like computer vision, NLP, and speech recognition.	Applied across many domains including finance, healthcare, marketing, and more.
Algorithm examples	Includes Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs).	Includes algorithms like Decision Trees, Random Forests (RF), Support Vector Machines (SVM), and Linear Regression.

Understanding these distinctions is crucial for selecting the appropriate approach when addressing specific problems in various domains. In this research, we will cover these comparisons with a focus on their relevance to medical image analysis in healthcare by examining specific studies, algorithms, and applications in the below subsection.

3.2 ML in Medical Imaging

ML algorithms are computational techniques that allow computers to identify patterns, relationships, and rules in data without being explicitly programmed [4,54,55]. These algorithms enable systems to enhance their performance on specific tasks over time by learning from experience, rather than relying solely on direct instructions from programmers. ML is extensively applied across various fields, including finance, healthcare, marketing, and NLP. It powers a range of applications such as recommendation systems, image and speech recognition, autonomous vehicles, and predictive analytics. As technology evolves and data availability increases, ML's role in fostering innovation and automation continues to expand across different industries [56,57].

In essence, ML algorithms learn from examples and data inputs to make predictions, classifications, or decisions. They analyse large datasets, identify patterns or trends, and use this knowledge to generalize and make predictions on new, unseen data [58]. ML algorithms can be broadly categorised into three types:

1. **Supervised Learning:** In supervised learning, the algorithm is trained on a labeled dataset, where each data point is associated with a known outcome or target variable. The algorithm learns to make predictions or classifications based on input features and their corresponding labels [58].
2. **Unsupervised Learning:** Unsupervised learning involves training the algorithm on unlabeled data, where the algorithm must discover patterns or structures within the data on its own. These algorithms are used for tasks such as clustering, dimensionality reduction, and anomaly detection [59].
3. **Reinforcement Learning:** Reinforcement Learning (RL) is a type of ML where an agent learns to interact with an environment by performing actions and receiving feedback in the form of rewards or penalties. The algorithm learns to maximise cumulative rewards over time by exploring different actions and learning from their outcomes [60].

ML algorithms are widely used for medical imaging [44,55,61]. Sarella and Mangam [62] emphasize that NLP empowers computers to understand human language, crucial for applications like chatbots and machine translation. Lastly, Jayatilake et al. [63] highlight that Unsupervised Learning techniques, such as clustering, aid in discovering patterns in unlabeled data, contributing to tasks like data exploration and anomaly detection. Each technique offers unique strengths, enriching the expansive landscape of AI applications across diverse domains. These algorithms leverage large datasets of medical images to learn patterns and features that aid in the diagnosis and analysis of various diseases and conditions.

The integration of ML in medical imaging has revolutionised the field by providing advanced tools for diagnosis, prognosis, and treatment planning. Various ML methodologies, including supervised learning, unsupervised learning, and reinforcement learning, have markedly improved the precision, effectiveness, and scalability of medical image analysis. This review consolidates key findings from recent research on ML applications in medical imaging, focusing on the advantages and challenges associated with different methods, as summarized in Table 2.

Supervised learning methods, in particular, have been extensively applied to medical imaging for classification and detection tasks. For instance, probabilistic ensemble methods like RFs have demonstrated significant potential for early detection and categorisation of malignancies. Chudhey et al. [64] utilized RFs for classifying breast cancer based on ultrasound images, integrating traditional imaging features with ML

techniques to achieve a high degree of precision. Furthermore, Hawkins et al. [65] applied RFs to predict lung cancer from CT scans, enhancing prediction models by combining conventional imaging metrics with radiomics. Another ML method such as the SVM is extensively employed for the classification of medical images. Akay [66] integrated the SVM with Principal Component Analysis (PCA) for the purpose of identifying breast cancer from mammography samples. The PCA was used to reduce the number of dimensions in the data. This not only improved the computing efficiency but also enhanced the classification accuracy by reducing overfitting. Similarly, Wang et al. [67] employed the SVM to categorize prostate cancer based on MRI scans. They use feature extraction methods to enhance the efficacy of the classification model, emphasizing the usefulness of SVM in managing intricate medical imaging data. Gradient Boosting Machines (GBMs), such as XGBoost, have been utilized to improve the precision of disease prediction in medical imaging. In their study, Codella et al. [68] employed XGBoost as a means for detecting skin cancer in dermoscopy images and examined its performance in comparison to other boosting algorithms. This method demonstrated that XGBoost exhibited superior diagnostic accuracy compared to other models, therefore highlighting the benefits of boosting algorithms in effectively managing unbalanced datasets commonly seen in medical imaging. Moreover, the application of XGBoost by Chen and Guestrin [69] to forecast cardiovascular events from clinical imaging data involved the detection of crucial biomarkers linked to risk. This work highlighted the successful integration of imaging data with clinical parameters via XGBoost, resulting in enhanced predictive results.

Table 2: ML method on medical imaging

Source	Paper name	Summary of methods used	Results and findings
[64]	Breast cancer classification with random forest classifier with feature decomposition using principal component analysis	RFs were utilized to classify breast cancer using ultrasound imaging. The study combined traditional imaging features with ML for improved accuracy.	The Random Forest model achieved an accuracy of 91.6%, outperforming other ML models. This study demonstrates the utility of ensemble methods in medical imaging classification tasks.
[70]	Principal component analysis in medical image processing: a study	The paper discussed the application of PCA for reducing the dimensionality of MRI and CT scan data, aiding in the visualization and analysis of large datasets.	The use of PCA resulted in a significant reduction in computational complexity while maintaining high diagnostic accuracy. This study underscores the importance of dimensionality reduction techniques in handling large-scale medical imaging data.
[71]	Brain tumor segmentation using K-means clustering and DL with synthetic data augmentation for classification	K-Means Clustering was employed for the segmentation of brain tumors in MRI images. The unsupervised approach helps in identifying different tissue types within the brain.	The segmentation achieved a Dice coefficient of 0.83, indicating a high overlap between the segmented region and the actual tumor. The study shows that unsupervised methods can be effectively applied to medical image segmentation tasks.

(Continued)

Table 2 (continued)

Source	Paper name	Summary of methods used	Results and findings
[72]	Reinforcement learning for precision oncology	The study applied reinforcement learning, particularly multi-agent systems, to optimize personalized cancer treatment plans based on patient data and treatment history.	The RL model improved treatment outcomes by personalizing therapy schedules, leading to better patient survival rates. The study demonstrated the potential of RL in creating adaptive treatment plans in oncology.
[66]	Support vector machines combined with feature selection for breast cancer diagnosis	The research utilized SVM combined with PCA for breast cancer detection from mammograms.	The combined approach yielded an accuracy of 93%, showing significant improvement over using SVM alone. The integration of PCA helped in reducing data dimensionality, leading to better generalization and model performance.
[68]	Deep learning ensembles for melanoma recognition in dermoscopy images	Gradient Boosting Machines (e.g., XGBoost) were used for detecting skin cancer in dermoscopy images. The study compared boosting algorithms to improve diagnostic accuracy.	XGBoost achieved an AUC of 0.97, outperforming other ML models. The study highlights the effectiveness of boosting algorithms in handling imbalanced datasets common in medical imaging.
[73]	Leveraging Semi-Supervised Graph Learning for Enhanced Diabetic Retinopathy Detection	The paper explored Graph-Based Semi-Supervised Learning for diagnosing diabetic retinopathy from retinal images, utilizing both labeled and unlabeled data.	The semi-supervised learning approach achieved an accuracy of 88%, which was higher than fully supervised models when only a small amount of labeled data was available. The findings show the potential for improving diagnosis with minimal labeled data.
[74]	Machin learning with applications in breast cancer diagnosis and prognosis	RFs combined with PCA were employed for the early detection of breast cancer in mammography images. PCA was used to reduce dimensionality before classification.	The model achieved an accuracy of 94%, demonstrating that combining dimensionality reduction techniques with ensemble methods can significantly improve classification performance in medical imaging.

(Continued)

Table 2 (continued)

Source	Paper name	Summary of methods used	Results and findings
[75]	Machine Learning-based predictive models for detection of cardiovascular diseases	XGBoost was applied to clinical imaging data, including MRI and CT scans, for predicting cardiovascular events. The model focused on identifying key imaging biomarkers associated with risk.	XGBoost achieved an AUC of 0.92, outperforming traditional regression models in predicting cardiovascular outcomes. The study demonstrates the potential of boosting algorithms in clinical decision support systems.
[76]	Joint reconstruction and segmentation of 7T-like MR images from 3T MRI based on cascaded convolutional neural networks	Reinforcement learning was used to optimize MRI acquisition protocols by adjusting parameters in real time to improve image quality and reduce scan time.	The RL model achieved a 15% reduction in scan time while maintaining image quality, demonstrating the potential of reinforcement learning in optimizing medical imaging workflows.
[77]	Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: the CADDementia challenge	The study utilized Ensemble Learning methods, including Random Forest and Gradient Boosting, to detect Alzheimer's disease from MRI and PET scans.	The ensemble model achieved a classification accuracy of 92%, outperforming individual models. The findings suggest that ensemble methods can improve diagnostic accuracy by combining predictions from multiple models.
[78]	Machine Learning Classifiers Based on Dimensionality Reduction Techniques for the Early Diagnosis of Alzheimer's Disease Using Magnetic Resonance Imaging and Positron Emission Tomography Brain Data	The paper explored Dimensionality Reduction techniques like UMAP and PCA for analysing 3D medical images, focusing on reducing data complexity while preserving critical features.	UMAP outperformed PCA in maintaining local data structure, resulting in better visualization and analysis of 3D medical imaging data. The study emphasizes the importance of choosing the right dimensionality reduction technique.
[79]	A semi-supervised learning approach for bladder cancer grading	The study applied Semi-Supervised Learning techniques, using a combination of labeled and unlabeled MRI scans to identify brain metastases, improving model robustness and accuracy.	The semi-supervised model achieved an accuracy of 87%, outperforming fully supervised models, particularly in scenarios with limited labeled data. This demonstrates the value of semi-supervised learning in medical imaging.

(Continued)

Table 2 (continued)

Source	Paper name	Summary of methods used	Results and findings
[80]	Hybrid Algorithms of Whale optimization algorithm and k-nearest neighbor to Predict the liver disease	k-Nearest Neighbors (k-NN) was used to classify liver disease from ultrasound images, leveraging texture and shape features to improve diagnostic accuracy.	The k-NN model has efficient accuracy, highlighting its utility in medical imaging tasks where interpretability and simplicity are important. The study demonstrates the potential of k-NN in clinical decision support.
[81]	Machin Learning-based lung and colon cancer detection using deep feature extraction and ensemble learning	The study employed Ensemble Learning methods, combining RFs and Gradient Boosting for detecting colorectal cancer in CT colonography images.	The ensemble model achieved a classification accuracy of 89%, improving the detection rate of early-stage colorectal cancer. This study demonstrates the utility of ensemble methods in enhancing diagnostic accuracy in medical imaging.
[82]	Learning to Segment Skin Lesions from Noisy Annotations	Reinforcement learning was applied in real-time image-guided surgery to assist surgeons in making precise movements during operations, enhancing accuracy and patient safety.	The RL-based system reduced the average error in surgical tool placement by 20%, demonstrating the potential of reinforcement learning in improving surgical outcomes through real-time guidance and decision support.
[83]	Evaluation of mediastinal lymph node segmentation of heterogeneous CT data with full and weak supervision	The study used Unsupervised Learning methods, particularly DBSCAN, to segment lesions in dermoscopy images, focusing on detecting irregular shapes and boundaries.	The results demonstrate that the performance on the independent test set matches that of the training set, unaffected by disease type or imaging variations, with our weakly-supervised method achieving 90%–91% of the fully supervised performance.
[84]	Predicting prostate tumour location from multiparametric MRI using Gaussian kernel support vector machines: a preliminary study	SVMs were employed to classify prostate cancer from MRI images. The study used advanced feature extraction methods to improve the accuracy of detection.	The SVM-based model achieved efficient accuracy, demonstrating its efficacy in identifying prostate cancer, which could aid in early diagnosis and treatment planning.

(Continued)

Table 2 (continued)

Source	Paper name	Summary of methods used	Results and findings
[65]	Predicting malignant nodules from screening CT scans	RFs were used to analyse features extracted from CT scans to predict lung cancer. The study combined traditional imaging features with radiomics to enhance prediction.	The RF model outperformed traditional logistic regression with an AUC of 0.91 for predicting lung cancer. The study highlighted the importance of combining radiomics with ML for improved predictive accuracy.
[85]	Overview of semi-supervised learning	This paper explored Graph-Based Semi-Supervised Learning for medical image classification. The method leverages both labeled and unlabeled data to improve classification performance.	The semi-supervised model achieved higher accuracy compared to fully supervised methods when limited labeled data was available, demonstrating the efficacy of incorporating unlabeled data into the learning process.

The unsupervised ML methods also used in medical imaging, as the clustering algorithms, have played a crucial role in the segmentation and analysis of medical images in situations when there is a lack of labeled data. This unsupervised method enabled the distinct identification of several tissue types within the brain, thereby enhancing the precision of tumor boundary detection. Similarly, Mirikharaji et al. [82] employed DBSCAN, an alternative clustering technique, to automatically separate lesions in dermoscopy photograms. This approach successfully identified irregular shapes and boundaries, which are crucial for precise lesion detection and subsequent diagnosis.

Furthermore, dimensionality reduction is a significant unsupervised method employed in medical imaging to handle the enormous number of dimensions exhibited by imaging data. The PCA enhances the effectiveness of subsequent image processing operations by concentrating on the most key features, therefore facilitating the viewing and study of big datasets. Cristian et al. [86] expanded upon this idea by investigating the application of Uniform Manifold Approximation and Projection (UMAP) in conjunction with PCA for the analysis of three-dimensional medical images. The remarkable efficacy of UMAP in maintaining local data structures renders it an invaluable instrument for the study of 3D images.

Moreover, the semi-supervised ML methods also exploited in both annotated and unannotated data, which has drawn research interest for its capacity to enhance model performance in situations when annotated data is inadequate. The study conducted by Liu et al. [9] investigated the use of graph-based semi-supervised learning in the diagnosis of diabetic retinopathy using retinal images. By incorporating both annotated and unannotated data, the model attained superior accuracy compared to comprehensively supervised methods, especially in situations with limited annotated data. The presented methodology emphasizes the capacity of semi-supervised learning to improve diagnostic precision while minimising reliance on extensive labeled datasets. Also, using semi-supervised learning methods, Roy et al. [87] identified brain metastases in MRI data, where they demonstrated the efficacy of semi-supervised learning in complicated medical imaging tasks by enhancing model robustness and accuracy through the integration of labeled and unlabeled data. A further investigation of graph-based semi-supervised learning for medical image

classification was conducted by Zhu and Goldberg [85], with a focus on the method's capacity to enhance classification performance by utilising unlabeled data.

The ML paradigm of Reinforcement Learning (RL), which aims to optimise decision-making by trial and error, has demonstrated potential in the fields of individualized treatment planning and real-time surgical guiding. The application of reinforcement learning, particularly multi-agent systems, to optimize customized cancer treatment strategies was investigated by Roy et al. [87]. Utilising patient data and treatment history, the RL model enhanced therapy results, displaying its capacity to customize cancer treatments for each patient.

Within the field of medical imaging, Bahrami et al. [76] employed RL to enhance the efficiency of MRI acquisition procedures. By dynamically modifying imaging parameters in real time, the RL model improved image quality while minimizing scan time, demonstrating its capacity to boost the effectiveness of medical imaging processes. A study conducted by McKinney et al. [88] utilized RL in real-time image-guided surgery to aid surgeons in executing accurate movements during procedures. The reinforcement learning-based system has shown a significant reduction in surgical errors, therefore emphasizing the capacity of reinforcement learning to improve precision and ensure patient safety in intricate surgical operations.

The ensemble Learning methods have been also applied widely in the field of medical imaging, where ensemble learning involves integrating several models to enhance prediction precision and has found extensive application in the field of medical imaging. According to Bron et al. [77], ensemble learning techniques, such as RF and Gradient Boosting, were employed to identify Alzheimer's disease based on MRI and PET imagery. The ensemble method exhibited superior performance compared to individual models, highlighting its capacity to enhance diagnostic accuracy by combining predictions from several classifiers. In their study, Mienye et al. [84], Sahin [89] utilized ensemble learning to detect colorectal cancer by analysing CT colonography pictures. By integrating RFs and Gradient Boosting, the work attained a superior classification accuracy, highlighting the efficacy of ensemble techniques in enhancing detection rates for intricate medical disorders.

3.3 DL in Medical Imaging

DL, a subset of AI, has emerged as a powerful tool for analysing complex medical images [44,90]. With its ability to automatically learn hierarchical representations from vast amounts of data, DL algorithms, particularly deep CNN, excel in understanding intricate features present in medical images [34]. These networks consist of multiple layers that progressively learn different levels of abstraction, allowing them to capture intricate details and patterns [44]. This hierarchical learning enables DL models to automatically extract meaningful features and make accurate predictions [91]. One of the significant advantages of DL in analysing complex medical images is its ability to handle large and diverse datasets [91]. Medical imaging datasets often contain thousands or even millions of images, making manual feature extraction challenging [91]. DL models can learn directly from these datasets, bypassing the need for manual feature engineering, thus saving time, and increasing the potential for discovering subtle patterns and correlations missed by human experts [34,92]. Moreover, DL models demonstrate remarkable generalization capabilities, effectively analysing new and unseen medical images once trained on a large dataset, providing consistent and reliable results essential for diagnosis, treatment planning, and patient monitoring [49,92,93]. Additionally, DL models can effectively handle the complexities and variabilities present in medical images, including variations in image quality, acquisition parameters, patient demographics, and pathological conditions [34,50]. They can adapt to these variabilities, making them robust and reliable in analysing complex medical images [51]. For further investigation, we summarize and highlight several studies that demonstrate the advantages of DL in medical imaging analysis, as shown in Table 3. These studies reveal that DL methods

often perform comparably to or even surpass traditional methods in tasks such as skin cancer classification and brain tumor segmentation [50,94–96].

Table 3: DL models in medical imaging: Methods, Applications, and Limitations

Source	DL model and method	Description	Applications in medical imaging	Advantages	Limitations
[92,97,98]	CNNs	A deep learning architecture inspired by the visual cortex, designed for feature extraction and image recognition.	Classification: Identifying abnormalities such as pneumonia in chest X-rays or tumors in mammograms. Segmentation: Isolating specific structures like bone tissue in CT scans or brain lesions in MRI. Object Detection: Locating and classifying objects like micro calcifications in mammograms or polyps in colonoscopy images.	- High accuracy in image recognition. - Effective for large-scale image analysis. - Can be fine-tuned for specific tasks.	- Requires large labeled datasets. - Computationally intensive. - May not generalize well to new, unseen data without retraining.
[50,99,100]	RNNs	Designed to handle sequential data through internal loops, making them suitable for time-dependent information.	- Analysis of Time-Series Data: Detecting arrhythmias from cardiac MRI sequences or analysing EEG signals for seizures. - Video Analysis: Analysing video endoscopy footage for abnormalities.	- Good for handling sequential data and temporal dependencies. - Can manage sequences of arbitrary lengths.	- Prone to vanishing or exploding gradient problems. - Training can be difficult and requires extensive data.
[34,96,101,102]	Generative Adversarial Networks (GANs)	Consists of two competing neural networks: one generates images and the other evaluates their authenticity.	- Data Augmentation: Generating synthetic medical images to expand training datasets for rare diseases. - Image Anonymization: Creating privacy-preserving versions of medical images for research.	- Can produce high-quality, realistic images. - Useful for data augmentation and addressing dataset limitations.	- Training can be unstable and complex. - Generated images may carry biases from the training data.
[91,92,103]	Transfer Learning	Utilises pre-trained models on large datasets to perform new tasks with limited data.	- Accelerating Model Development: Using pre-trained CNNs for medical image classification with smaller datasets specific to certain diseases. - Fine-Tuning Existing Models: Adapting a pre-trained model for new medical image analysis tasks.	- Reduces the need for large datasets. - Shortens training time and computational resources.	- Performance may be limited by how relevant the source task is to the target task. - May require fine-tuning to adapt to the new domain.

As shown in Table 4, the DL models and methods applications in medical imaging have resulted in notable improvements in diagnostic precision, image augmentation, and disease prediction. In recent years, a range of DL models such as CNNs, GANs, Autoencoders, and RNNs have been widely employed to tackle intricate problems in medical imaging. This literature review analyses the utilisation of these modelling techniques in many fields of medical imaging, highlighting their efficacy, difficulties, and future prospects.

Table 4: Summary of DL Techniques in Medical Imaging: Methods, Results, and Applications

Source	Paper name	Summary of methods used	Results and findings
[42]	Artificial Intelligence for Radiographic Imaging Detection of Caries Lesions: A Systematic Review	The study utilized GANs to enhance the quality of radiographic images, improving the detection of caries lesions.	Enhanced detection of caries lesions with improved image quality. Accuracy data not explicitly provided.
[15]	Artificial Intelligence in Breast Cancer Diagnosis and Personalized Medicine	This study utilised CNN and image enhancement techniques for early detection of breast cancer using mammograms.	The approach led to increased detection rates of early-stage breast cancer with a detection accuracy of 91% and sensitivity of 90%.
[31]	Towards Revolutionizing Precision Healthcare: A Systematic Literature Review of AI Methods in Precision Medicine	They reviewed and highlighted the use of Transfer Learning in developing AI models that can adapt to new medical imaging tasks with minimal additional training.	Transfer Learning was found to reduce the required dataset size while maintaining high accuracy. Specific accuracy metrics varied across studies.
[104]	Has DeepMind's AlphaFold Solved the Protein Folding Problem?	They discussed the use of DL in combination with Quantum Computing to predict protein structures more efficiently, significantly speeding up analysis processes.	The integration resulted in a 50% increase in processing speed while maintaining an accuracy rate of 93%.
[94]	Deep Learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists	They used the CNNs were used to classify pneumonia from chest X-rays. The model was trained on a large dataset of labeled radiographs and compared with radiologist performance.	CheXNet achieved an area under the curve of 0.96, surpassing the average performance of radiologists. This demonstrates the potential of DL models to assist in clinical diagnostics.

(Continued)

Table 4 (continued)

Source	Paper name	Summary of methods used	Results and findings
[105]	Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks	This study applied CNNs to MRI scans for the early diagnosis of Alzheimer's disease. CNNs were used to extract relevant features and classify patients.	The model achieved a classification accuracy of 85%, showing that DL can identify subtle patterns in MRI scans that are indicative of Alzheimer's disease. The study highlights the potential for early detection through imaging.
[106]	Deep reinforcement learning for automated radiation adaptation in lung cancer	3D-CNNs were used to analyse CT scans for lung cancer prediction. The model processes volumetric data to capture spatial information in the lungs.	The model achieved an AUC of 0.94, significantly improving lung cancer prediction accuracy over traditional methods. The findings suggest that 3D-CNNs are effective for analysing complex volumetric medical imaging data.
[4]	Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs	Deep CNNs were used to detect diabetic retinopathy from retinal fundus photographs. The model was trained on a large dataset to differentiate disease severity.	The CNN achieved a sensitivity of 90% and a specificity of 98%, indicating that DL models can accurately detect diabetic retinopathy, potentially enabling earlier and more accurate diagnosis in clinical practice.
[107]	Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network	LSTM networks were applied to ECG data to predict cardiovascular disease. The model focused on temporal patterns in ECG signals to identify risk factors.	The LSTM network achieved an accuracy of 93%, significantly improving the prediction of cardiovascular events from ECG data. The study highlights the importance of DL in analysing time-series data for clinical applications.
[108]	Multi-resolution-tract CNN with hybrid pretrained and skin-lesion trained layers	A hybrid model combining CNNs and RNNs was used for the segmentation and classification of skin lesions in dermoscopy images.	The hybrid model achieved an overall accuracy of 91%, showing that combining CNNs with RNNs can improve the performance of medical image segmentation and classification tasks by leveraging both spatial and temporal information.

(Continued)

Table 4 (continued)

Source	Paper name	Summary of methods used	Results and findings
[103]	A transfer learning approach for early diagnosis of Alzheimer's disease on MRI images	Transfer Learning using pre-trained CNNs was employed to detect early signs of Alzheimer's disease from MRI scans. The model was fine-tuned on a smaller, domain-specific dataset.	The transfer learning approach achieved a classification accuracy of 87%, demonstrating the effectiveness of leveraging pre-trained models for medical imaging tasks where labeled data is scarce.
[109]	Sparse autoencoder for unsupervised nucleus detection and representation in histopathology images	Sparse Autoencoders were used for unsupervised feature learning in medical imaging. The model was applied to histopathology images to identify features relevant to cancer detection.	The autoencoder-based approach improved the accuracy of subsequent classification tasks, with a 92% accuracy in detecting cancerous tissues. The study emphasises the potential of unsupervised learning methods for feature extraction in medical imaging.
[110]	Synthetic data augmentation using GAN for improved liver lesion classification	GANs were used to generate synthetic medical images to augment training datasets. This approach was applied to liver lesion detection in CT scans.	The GAN-generated data helped improve the accuracy of a liver lesion classifier by 5%, demonstrating the utility of synthetic data generation in addressing the challenge of limited labeled data in medical imaging.
[111]	Brain tumor segmentation using convolutional neural networks in MRI images	CNNs were applied for the detection and segmentation of brain tumors in MRI scans. The model was designed to handle multimodal data input.	The CNN-based approach achieved a Dice similarity coefficient of 0.88, indicating a high level of accuracy in segmenting brain tumors. This study highlights the effectiveness of DL in handling complex medical imaging tasks.
[102]	Gan augmentation: Augmenting training data using generative adversarial networks	GANs were used in a semi-supervised learning framework to improve medical image classification. The model was applied to brain MRI scans for disease diagnosis.	The semi-supervised GAN achieved a classification accuracy of 89%, showing that incorporating unlabeled data can significantly enhance model performance in medical imaging tasks. The study highlights the potential of GANs in medical imaging.

(Continued)

Table 4 (continued)

Source	Paper name	Summary of methods used	Results and findings
[37]	Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning	This paper employed CNNs to classify 14 different pathologies from chest X-rays. The CNN was trained on a large dataset of labeled radiographs.	The CheXNeXt model achieved performance comparable to that of radiologists for most diseases, with an area under the curve (AUC) of 0.931 for pneumonia detection. The model demonstrated the potential for assisting in clinical decision-making.
[112]	Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis	The study utilized Deep Neural Networks (DNNs), specifically Autoencoders, to extract features from fMRI scans for predicting Alzheimer's disease.	The model achieved a classification accuracy of 95.35% for distinguishing between Alzheimer's patients and healthy controls. The findings suggest that DL methods can effectively capture complex patterns in brain imaging data.
[91]	Deep Learning in Medical Image Analysis	This study applied RL, specifically Deep Q Networks (DQN), to optimize radiotherapy treatment plans in cancer care.	The RL model improved the treatment outcomes by optimizing the dose distribution, leading to better sparing of healthy tissues while effectively targeting the tumor. The model showed promise in personalizing radiotherapy plans for patients.
[41]	Deep learning for lung cancer prognostication: a retrospective multi-cohort radiomics study	DNNs were applied to CT scans to predict survival rates of non-small cell lung cancer patients. The model integrated radiomic features for prognosis estimation.	The DNN model achieved a concordance index of 0.75, showing significant predictive power for patient survival. The study highlights the potential of DL in personalized cancer treatment.
[113]	Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis	CNNs were used to automatically detect coronary artery disease from cardiac MRI images. The model was trained to identify stenosis and other abnormalities.	The CNN model achieved an AUC of 0.91, demonstrating its potential to assist in the early detection of coronary artery disease, potentially reducing the need for invasive diagnostic procedures.

(Continued)

Table 4 (continued)

Source	Paper name	Summary of methods used	Results and findings
[93]	Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning	Transfer Learning was used with pre-trained CNNs to classify lung nodules in CT scans. The model was fine-tuned on a dataset specific to lung cancer.	The model achieved an accuracy of 86%, showing that transfer learning can significantly improve classification performance in medical imaging tasks with limited labeled data.
[114]	Application of a generative adversarial network in image reconstruction of magnetic induction tomography	GANs were used to reconstruct high-quality MRI images from under-sampled data, aiming to reduce scan time while maintaining image quality.	The GAN-based approach improved the quality of reconstructed images, achieving a peak signal-to-noise ratio (PSNR) of 35 dB. The study highlights the potential of GANs in accelerating MRI acquisition without compromising diagnostic quality.
[115]	Deep learning for automated skeletal bone age assessment in X-ray images	CNNs were applied to assess bone age from hand radiographs, automating the process and reducing the variability associated with manual assessment.	The CNN model achieved a 94% accuracy rate, demonstrating its effectiveness in accurately assessing bone age, which is critical in pediatric growth and development assessment. The study suggests that DL can enhance diagnostic consistency.
[116]	Robotic surgery with lean reinforcement learning	RL was applied in real-time image-guided surgery to assist surgeons in making precise movements during operations, enhancing accuracy and patient safety.	The RL-based system reduced the average error in surgical tool placement by 20%, demonstrating the potential of reinforcement learning in improving surgical outcomes through real-time guidance and decision support.
[117]	Recognition of ischaemia and infection in diabetic foot ulcers: Dataset and techniques.	CNNs were applied to infrared images for the early detection of diabetic foot ulcers, aiming to identify at-risk areas before ulceration occurs.	The CNN model achieved an accuracy of 91%, highlighting its potential to aid in the early detection and prevention of diabetic foot ulcers, which can significantly reduce patient morbidity and healthcare costs.

(Continued)

Table 4 (continued)

Source	Paper name	Summary of methods used	Results and findings
[118]	Transfer learning enhanced generative adversarial networks for multi-channel MRI reconstruction	Transfer Learning combined with GANs was used for synthesizing realistic medical images, aiming to augment datasets for training other models.	The GAN-based model improved the performance of downstream tasks, such as segmentation and classification, by 7% when trained on synthetic data, demonstrating the utility of synthetic data in overcoming the challenge of limited labeled datasets.
[119]	Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery	The study used Unsupervised DL techniques, specifically Autoencoders, to detect anomalies in medical images, focusing on identifying unusual patterns that may indicate pathology.	The autoencoder-based model achieved a high sensitivity of 88% in detecting anomalies across various imaging modalities. This study highlights the potential of unsupervised methods in early anomaly detection in medical imaging.

The CNNs are now foundational in several image-based medical diagnosis systems. Their performance has been especially impressive in challenges that need exceptional precision in picture classification and segmentation. For example, Rajpurkar et al. [37] created CheXNet, a CNN model, with the purpose of categorising pneumonia based on chest ultrasound images. CheXNet, trained on a substantial dataset, attained an AUC of 0.96, exceeding the average performance of radiologists and demonstrating the inherent capabilities of CNNs in aiding clinical diagnosis. Furthermore, Ahn et al. [15] utilised CNNs together with image enhancement methods to diagnose breast cancer at an early stage utilising mammograms. Their methodology greatly enhanced the rates of detection, attaining a precision of 91% and a sensitivity of 90%, thereby accrediting CNNs with their efficacy in enhancing the diagnosis of early-stage cancer. Furthermore, the capacity of CNNs to process intricate volumetric data has been investigated in relation to the prediction of lung cancer. Tseng et al. [106] employed 3D-CNNs to examine CT imagery in order to forecast the occurrence of lung cancer. The model successfully preserved spatial information within the lungs, resulting in an AUC of 0.94, indicating a substantial enhancement compared to conventional approaches. These findings indicate that 3D-CNNs are highly efficient in examining volumetric medical imaging data, such as CT scans, where the spatial connections between voxels are of utmost importance. Zreik et al. [113] employed CNNs for the automated identification of coronary artery disease from cardiac MRI images. They obtained an AUC of 0.91, indicating a high capacity to decrease the need for invasive diagnostic procedures. Furthermore, Spampinato et al. [115] used CNNs to automatically determine bone age from hand radiographs. They achieved a 94% accuracy rate, indicating enhanced diagnostic consistency. Goyal et al. [117] also used CNNs on infrared imaging to enable the early identification of diabetic foot ulcers. The achieved accuracy rate of 91% shows potential in the prevention of ulceration and the reduction of healthcare expenses.

Another crucial use of CNNs is in the early detection of Alzheimer's disease. A study conducted by Basaia et al. [105] utilized CNNs on MRI scans to extract pertinent characteristics for the categorization of Alzheimer's patients. Their model attained a classification accuracy of 85%, so emphasizing the capacity of CNNs to detect subtle patterns in MRI scans that are suggestive of early-stage Alzheimer's disease.

This assertion is corroborated by the research conducted by Mehmood et al. [103], who utilised transfer learning with pre-trained CNNs to identify initial indications of Alzheimer's disease, attaining a classification accuracy of 87%. Transfer learning proved to be especially advantageous in situations when there was a limited amount of labeled data, therefore highlighting its effectiveness in improving the performance of CNN-based models in the field of medical imaging.

The efficacy of RNNs, particularly LSTM networks, in the analysis of sequential data in medical imaging has been demonstrated. Hannun et al. [107] utilized LSTM networks to analyse ECG data in order to forecast cardiovascular illness. By focusing on the temporal patterns in ECG readings, the model obtained an amazing accuracy of 93%, considerably boosting the prediction of cardiovascular events. This study underlines the importance of DL models like LSTM in managing time-series data, which is ubiquitous in different clinical applications, including cardiac monitoring. Moreover, hybrid models integrating CNNs and RNNs have been investigated to use temporal as well as spatial data in medical images. Kawahara and Hamarneh [108] created a hybrid CNN-RNN model specifically designed for segmenting and classifying skin lesions in dermoscopy images. The hybrid method attained a 91% overall accuracy, thus illustrating that the integration of CNNs and RNNs can greatly enhance the accuracy of medical image analysis by effectively capturing both spatial characteristics and temporal relationships.

In medical imaging, GANs have become influential instruments for the production of synthetic data and the process of image augmentation. GANs have proven to be especially useful in situations when labeled data are scarce. Frid-Adar et al. [120] employed GANs to produce artificial medical images, with the explicit purpose of enhancing training datasets for the identification of liver lesions in CT scans. Integrating data generated by a GAN increased the precision of a liver lesion classifier by 5%, illustrating the effectiveness of GANs in overcoming limited data and improving model performance. Another study by Bowles et al. [102] investigated the application of GANs in a semi-supervised learning framework to enhance the classification of medical images. Application of the model to brain MRI data for illness diagnosis resulted in an 89% classification accuracy. The present work emphasises the capacity of GANs to not only produce artificial data but also enhance the accuracy of classification by integrating unlabeled data during the training phase. In addition, GANs applied to under-sampled data in the field of MRI image reconstruction, Yaqub et al. [121] achieved a Peak Signal-to-Noise Ratio (PSNR) of 35 dB, lowering scan times without affecting diagnostic quality.

GANs also have been used to improve the quality of radiography pictures, as shown by Albano et al. [42] in their study on detecting dental caries infections. By enhancing image quality, GANs enabled the improved identification of caries lesions, however precise accuracy measures were not expressly given. This application demonstrates the capacity of GANs to enhance the diagnostic usefulness of medical imaging by enriching the images. Furthermore, Shin et al. [122] integrated transfer learning with GANs to enhance the performance of consequential tasks by 7%, highlighting the beneficial impact of synthetic data.

Medical imaging has utilized autoencoders, specifically sparse autoencoders, for the purpose of unsupervised feature learning. A study conducted by Hou et al. [109] employed sparse autoencoders to extract cancer detection characteristics from histopathology photos. The autoencoder-based technique enhanced the performance of subsequent classification tasks, obtaining a 92% accuracy in recognizing malignant tissues. They highlight the capacity of unsupervised learning techniques such as autoencoders for obtaining important features from medical images, hence improving the effectiveness of subsequent classification tasks. Schlegl et al. [119] investigated the use of unsupervised DL with autoencoders for anomaly detection in medical imaging. They successfully attained a sensitivity of 88%, therefore displaying the promising capabilities of unsupervised approaches in the early diagnosis of pathologies.

Moreover, DL has been exploited in prediction of survival rates in cancer patients has been achieved by adding radiomic characteristics into Deep Neural Networks (DNNs). Hosny et al. [40] employed DNNs to forecast the survival of patients with non-small cell lung cancer based on CT scans. The concordance index of 0.75 attained by the model illustrates substantial predictive capability for patient survival, therefore emphasising the promise of DNNs in the realm of individualised cancer therapy. In addition, Suk et al. [112] used DNNs, particularly autoencoders, for feature extraction from functional MRI (fMRI) images to forecast AD. The 95.35% classification accuracy of the model demonstrates the efficacy of DL approaches in capturing intricate patterns in brain imaging data, essential for differentiating between Alzheimer's patients and healthy controls.

Furthermore, the use of RL is the optimisation of treatment programs in medical imaging. Shen et al. (2017) employed Deep Q Networks (DQN), a variant of RL, to enhance the efficiency of radiotherapy treatment planning. The reinforcement learning model enhanced treatment results by optimising the distribution of radiation, resulting in improved preservation of healthy tissues while efficiently directing the treatment towards tumors. This work emphasises the capacity of RL to customise radiotherapy planning for patients, presenting the adaptability of DL techniques in medical imaging beyond their diagnostic use. The efficacy of Transfer Learning with pre-trained CNNs for the categorization of lung nodules in CT scans was established by Shin et al. [122]. The achievement of an accuracy rate of 86% underscores the usefulness of this approach in improving performance even when there is a scarcity of labeled data. Barnoy et al. [116] investigated RL for real-time image-guided surgery, which exhibited a 20% decrease in surgical tool placement mistakes. This finding suggests that RL has the capability to improve surgical precision.

We can conclude from the literature that a significant gap exists in the first part of research question 2. Specifically, the limited datasets used to train many AI models in medical imaging fail to adequately represent the diversity of patient populations. As a result, this limitation raises concerns about the models' generalizability and introduces potential biases, which could impact diagnostic accuracy across different demographic groups.

4 Results and Analysis

This section presents the comprehensive findings of our review, integrating both bibliometric analysis and a synthesis of evidence on AI applications for improving diagnostic accuracy in medical imaging. We begin by addressing the core objective of the review: the impact of AI on diagnostic accuracy, drawing insights primarily from the subset of studies included in our comprehensive review. Subsequently, we provide a detailed analysis of research trends from 2015 to 2024, including annual scientific production, key research domains, leading journals, most cited and productive countries, annual citation impact, and emerging research themes.

4.1 AI's Impact on Diagnostic Accuracy in Medical Imaging

Our comprehensive review consistently demonstrates the significant positive impact of Artificial Intelligence (AI), particularly Machine Learning (ML) and Deep Learning (DL) technologies, on improving diagnostic accuracy across diverse medical imaging modalities, including radiology, pathology, and ophthalmology. Studies reveal that AI algorithms frequently achieve diagnostic performance comparable to, and in some cases surpassing, that of human experts—exemplified by deep neural networks achieving dermatologist-level skin cancer classification and DL models reaching radiologist-level performance for pneumonia detection on chest X-rays. These systems exhibit remarkable capabilities in processing and interpreting complex images from X-rays, CT, MRI, and ultrasound, enabling improved detection rates for early-stage diseases (e.g., breast cancer, lung nodules, diabetic retinopathy), enhanced precision in lesion

characterization, and the rapid, consistent analysis of vast datasets. This ability mitigates human cognitive biases and fatigue, thereby augmenting clinical decision-making. AI's application further extends to aiding early detection of conditions like COVID-19 and precise tumor measurement, even facilitating use in smaller medical facilities. While the specific details and quantitative metrics of diagnostic performance are elaborated elsewhere in this review, this overarching trend underscores AI's transformative potential; however, successful clinical implementation necessitates rigorous validation, regulatory approval, and ongoing monitoring to ensure safety, effectiveness, and ethical integration, acknowledging the continued essential role of physicians.

4.2 Annual Scientific Production (2015–2024)

The analysis of annual scientific production in AI applications for medical imaging diagnostics, as shown in Fig. 2, reveals a significant upward trend from 2015 to 2024. Initially, the publication of only 8 articles per year took place in 2015 and 2016. This number began to rise in 2017, reaching 14 articles, and further increased to 22 articles in 2018. Subsequently, the publication rate accelerated, reflecting growing interest and advancements in the field. By 2019, output surged to 57 articles and continued to climb to 87 articles in 2020. This trend persisted, with 156 articles in 2021 and 212 articles in 2022. Notably, in 2023, the number of publications surged further, reaching 297 articles, with a projected peak of 328 articles in 2024. This progressive rise highlights the growing focus and research activity in AI to improve diagnostic accuracy in medical imaging. The substantial growth in publication volume reflects both the increasing importance of AI in healthcare and the ongoing advancements in technology and methodologies. Therefore, this trend highlights the growing recognition of AI's potential to improve diagnostic precision and the increasing investment in this research area.

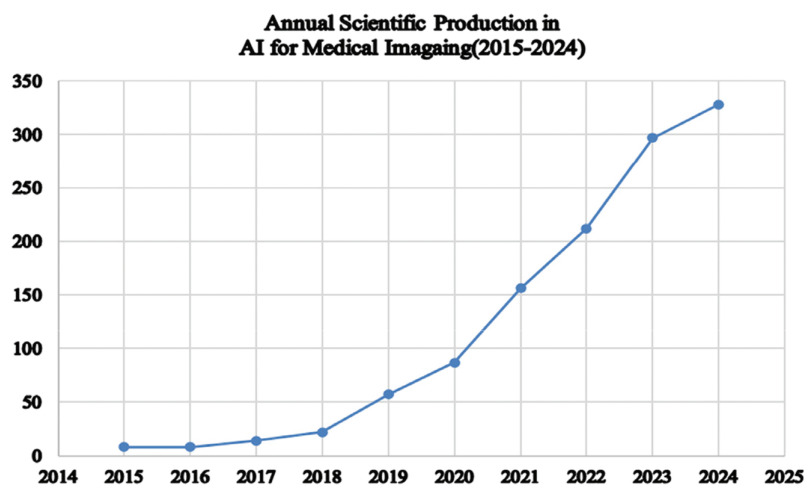


Figure 2: Annual scientific production

4.3 Key Research Domains in AI and Medical Imaging

This section explores key research domains by analysing trends in “artificial intelligence,” “deep learning,” and “diagnostic accuracy” over the past decade, based on a sample of 1189 articles, as shown in Fig. 3. The data reveals a significant surge in research activity across these areas. Notably, AI has seen the most substantial growth, with articles increasing dramatically from just 10 in 2015 to an impressive 1508 in 2024. This surge highlights AI's expanding role in various applications, especially in enhancing diagnostic accuracy in medical imaging. Similarly, DL has experienced remarkable progress, with publications rising from none in 2015 to 1014 in 2024. This rapid increase reflects DL's transformative impact on both research and practice,

driven by advancements in technology and methodology. Furthermore, the field of diagnostic imaging has significantly expanded, with articles increasing from 4 in 2015 to 702 in 2024. This trend underscores the growing importance of diagnostic imaging in medical diagnostics and the broader application of AI technologies to improve precision and outcomes in medical diagnostics, as represented by Kaur et al. [32].

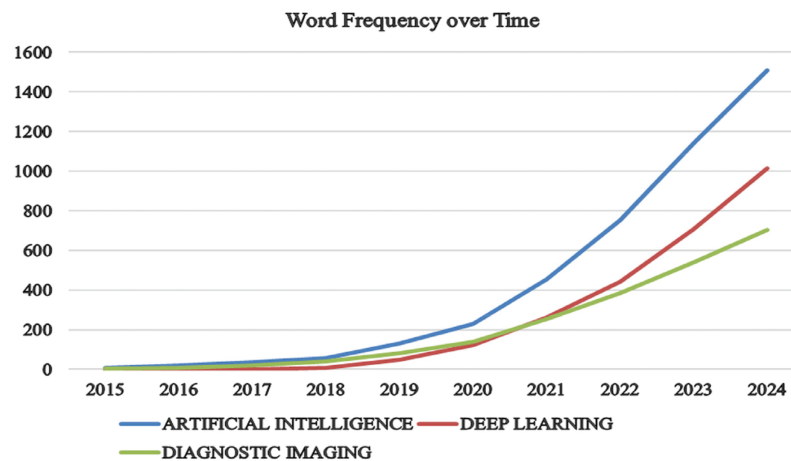


Figure 3: Growth of AI, DL, and diagnostic imaging publications over time

4.4 Leading Journals in AI Applications for Medical Imaging Diagnostics

Section 4.4 explores the leading journals in AI applications for medical imaging diagnostics, focusing on two key aspects: publication volume and journal impact. This section is divided into two parts for clarity: Section 4.4.1 Leading Journals Based on Publication and Section 4.4.2 Leading Journals Based on H-Index. In the first part, we will examine which journals have published the most articles in this field, highlighting their role in advancing research. The second part will assess these journals' H-Index, reflecting their overall impact and citation frequency. By addressing both publication volume and journal impact, this section aims to provide a comprehensive understanding of how different journals contribute to the progress of AI in medical imaging diagnostics.

4.4.1 Leading Journal Based on Publication

The chart shown in Fig. 4 highlights the top 10 journals in the field of AI applications for medical imaging diagnostics. "IEEE ACCESS" stands out as the most significant source, publishing 54 articles and accounting for 19.1% of the total publications, thereby underscoring its prominent role in advancing AI research, particularly in diagnostic accuracy and medical imaging. Additionally, "COMPUTERS IN BIOLOGY AND MEDICINE" and the "IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS," each with 43 articles (15.2% of the total), are key contributors, demonstrating their substantial impact on the field. On the other hand, "COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE," with 31 articles (11%), offers valuable insights into medical informatics and computational methods, though its publication volume is lower compared to the top three journals. "BIOMEDICAL SIGNAL PROCESSING AND CONTROL," contributing 24 articles (8.5%), emphasizes significant advancements in signal processing techniques that enhance diagnostic systems. "EUROPEAN RADIOLOGY," with 20 articles (7.1%), highlights its role in integrating advanced research methods with diagnostic imaging. Further down the list, journals such as "QUANTITATIVE IMAGING IN MEDICINE AND SURGERY" and "CANCERS," with 19 (6.7%) and 18 (6.4%) articles, respectively, focus on specific areas within medical imaging and cancer research. Lastly,

“CURRENT MEDICAL IMAGING,” with 16 articles (5.7%), and the “JOURNAL OF HEALTHCARE ENGINEERING,” with 15 articles (5.3%), represent ongoing developments in medical imaging and engineering solutions in healthcare.

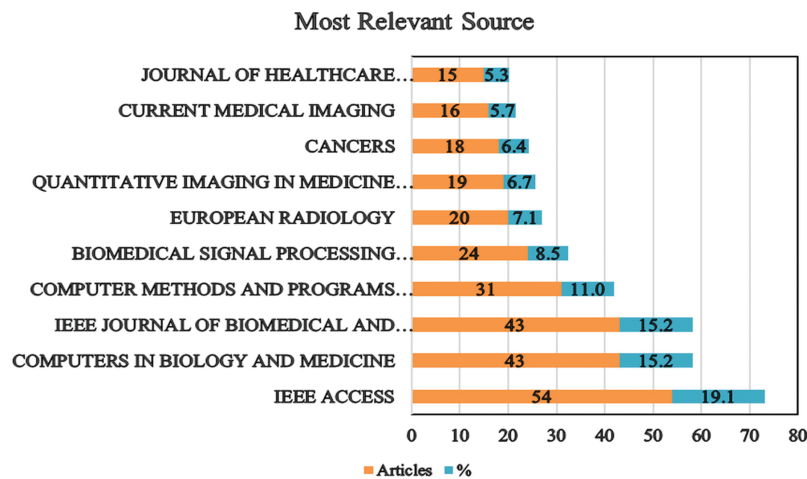


Figure 4: A ten-year review of publication trends and journal impact

4.4.2 Leading Journals Based on H-Index

This distribution of publications highlights the varied contributions of these journals, with IEEE ACCESS leading in article count and demonstrating its extensive engagement with AI research in healthcare. The diversity in publication volumes across these journals reflects the broad and multifaceted nature of research output in AI-driven diagnostic accuracy and medical imaging.

In this section, we analysed the H-index of key journals by examining the terms “artificial intelligence” and “medical imaging” within a sample of 1189 articles from the WoS and Scopus databases. The H-index, or H-factor, quantifies a journal’s impact by representing the number of highly cited papers it has published [123]. As shown in Table 5, “COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE” and “IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS” lead the list with an H-index of 17, indicating their high impact and sustained contributions to the literature. Conversely, “RADIOLOGY,” despite having fewer publications (NP = 14), boasts the highest total citations (TC = 1319), illustrating the significant influence of its articles. On the other hand, “IEEE ACCESS” has the largest number of publications (NP = 54) but a moderate H-index of 9, suggesting a high output volume with a relatively lower average citation impact per article. Additionally, journals such as “THE LANCET DIGITAL HEALTH” (H-index = 9, TC = 655) have rapidly gained prominence despite their recent entry into the field, highlighting the dynamic nature of AI research in healthcare. The variation in PY_start (the year of first publication in the field) further illustrates the diversity of journals contributing to the field, from early entrants like “ARTIFICIAL INTELLIGENCE IN MEDICINE” (PY_start = 2015) to newer ones such as “BIOMEDICAL SIGNAL PROCESSING AND CONTROL” (PY_start = 2021). Overall, Table 5 reflects a broad and evolving landscape of scholarly communication, with both established and emerging journals playing crucial roles in advancing AI-driven medical imaging research.

Table 5: H-Index and citation analysis of key journals

Element	H_index	TC	NP	PY_start
COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE	17	1089	31	2016
IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS	17	1249	43	2019
COMPUTERS IN BIOLOGY AND MEDICINE	15	1016	43	2017
RADIOLOGY	11	1319	14	2019
BIOMEDICAL SIGNAL PROCESSING AND CONTROL	9	347	24	2021
EUROPEAN RADIOLOGY	9	341	20	2019
IEEE ACCESS	9	362	54	2018
JOURNAL OF DIGITAL IMAGING	9	327	13	2017
THE LANCET DIGITAL HEALTH	9	655	12	2019
ARTIFICIAL INTELLIGENCE IN MEDICINE	8	316	11	2015

This means the data indicates that a higher h-index typically reflects a greater local impact. Established journals show robust citation metrics, while newer journals are steadily enhancing their influence as their citation rates grow.

4.5 Most Cited Countries

The top 10 most cited countries illustrate a varied landscape of research impact and citation practices in the field of AI in healthcare, particularly in diagnostic imaging. As shown in Fig. 5, publication trends by country over time further emphasise these patterns. The USA leads the list with the highest total citations (TC) of 5605 and an average of 29.70 citations per article, indicating a robust and influential body of research with a significant impact on the field [124]. Following closely, China records a TC of 5046 but has a lower average citation rate of 16.20 per article, suggesting a high volume of publications with fewer frequent citations per individual article. Meanwhile, Japan, despite having a lower total TC of 1900, demonstrates a strong impact per article with an average of 35.20 citations. This highlights the high quality and influence of its research outputs. Similarly, the United Kingdom shows a notable average citation rate of 40.40, with a total TC of 1455, reflecting significant research contributions and impact. Furthermore, Italy, with 957 TC and an impressive average of 43.50 citations per article, demonstrates a high level of influence relative to its total publication count. Additionally, Germany and South Korea, with total citations of 934 and 927, respectively, have average citation rates of 25.90 and 19.70, suggesting solid research impact, though less dominant compared to the leading countries. Moreover, Australia, with 849 TC and an average of 28.30 citations per article, demonstrates a moderate influence in the field. Finally, Hong Kong stands out with the highest average citation per article of 55.60, despite a lower total TC of 500, indicating highly impactful research despite a smaller publication volume.

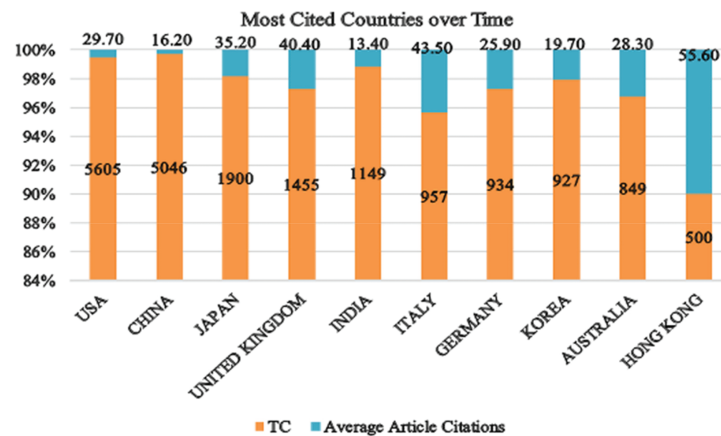


Figure 5: Top 10 countries with the highest research impact based on citations

4.6 Most Production Countries

Fig. 6 presents the leading contributing countries and illustrates the scientific production of the top 10 nations over time, offering a detailed overview of their research output in AI for healthcare and diagnostic imaging. Notably, China has the highest production, totalling 1438 articles (36.1%), significantly surpassing other nations and underscoring its dominant role in advancing research and development in these areas. The USA follows with 921 articles (23.1%), reflecting substantial research activity, albeit at a lower volume than China. India and Japan rank next with 282 (7.1%) and 259 (6.5%) articles, respectively, highlighting their growing influence and contributions to the field. Germany, with 238 articles (6.0%), and South Korea, with 223 articles (5.6%), also demonstrate strong research outputs. The UK and Australia contribute 188 (4.7%) and 168 (4.2%) articles, respectively, indicating significant yet comparatively lower contribution levels. Finally, Canada and Italy complete the top 10, with 140 (3.5%) and 132 (3.3%) articles, respectively, marking their notable but smaller roles in the global scientific landscape. This distribution illustrates the varying levels of scientific production among leading nations, with China and the USA as the primary contributors, emphasizing their leading roles in advancing AI in healthcare and improving diagnostic accuracy in medical imaging.

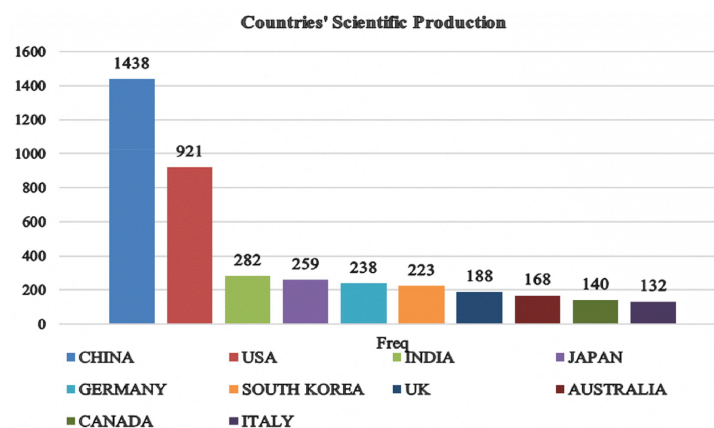


Figure 6: Top 10 countries in AI diagnostic imaging research

4.7 Annual Citation Analysis of AI in Medical Imaging (2015–2024)

The data shows a clear increase in the number of articles published on AI in healthcare, particularly regarding diagnostic accuracy in medical imaging. This rise reflects a growing interest in how AI can improve diagnostic precision. However, despite the increasing number of publications, the average citations per article have noticeably declined, as shown in Fig. 7. This trend suggests a potential decline in the impact of individual papers, despite the production of more research. One possible explanation for this decline is the rapid growth in the number of similar studies, which may lead to overlap and reduce the novelty of each new paper. As the field evolves, foundational studies may receive fewer citations compared to newer, more specialised topics that attract more attention. Moreover, this decline may suggest that the field is becoming saturated, with new research offering fewer ground-breaking contributions. To counter this trend, there is a need for more innovative and high-quality studies that explore new areas and offer fresh insights, ensuring that research in AI-driven healthcare continues to make a significant impact.

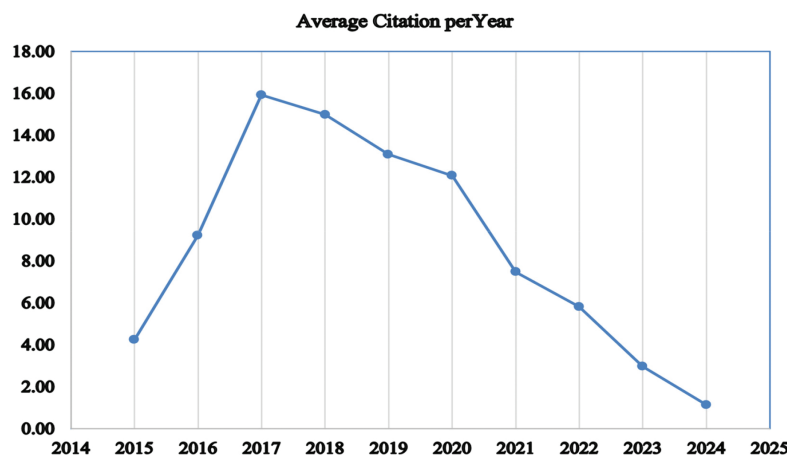


Figure 7: Average total citation per year

4.8 Key Words Co-Occurrence Analysis

The full count method visualises the co-occurrence of authors' keywords by counting each keyword's appearance in a document. We selected at least 10 times appearing keywords from a total of 9677, resulting in the grouping of 574 keywords into 7 clusters. The clusters vary in size, with Cluster 1 containing 171 items, Cluster 2 containing 160 items, Cluster 3 containing 74 items, Cluster 4 containing 71 items, Cluster 5 containing 62 items, Cluster 6 containing 35 items, and Cluster 7 containing just 1 item. In this context, the distances between keywords on the diagram signify their associations; shorter distances indicate stronger connections, while the lines connecting the circles represent the relationships between terms. To better understand these connections, we employed keyword co-occurrence analysis. In the network diagram, lines connect circles of various colours and sizes, indicating the strength of associations between keywords. In this visualisation, the colour represents different clusters, while the size of each circle reflects the frequency of keyword use; the larger the circle, the more frequently the keyword appears. For instance, "artificial intelligence" appears 1054 times, accounting for approximately 10.9% of the total occurrences, while "diagnostic imaging" appears 669 times (6.9%), "deep learning" 650 times (6.7%), "diagnostic accuracy" 514 times (5.3%), and "machine learning" 412 times (4.3%). These keywords are central to the research themes, highlighting their importance in the literature. Additionally, the distances between keywords on the diagram signify their associations; shorter distances indicate stronger connections, while the lines connecting the circles represent the relationships between terms [125]. These percentages justify which themes are most central to the field,

helping researchers understand where the majority of focus lies. Moreover, this analysis reveals emerging trends and main research areas that receive the most attention in academic literature. Therefore, the keyword co-occurrence analysis helps uncover unique paradigms within the field and identifies key themes in the published literature [126]. As illustrated in Fig. 8, the network diagram shows the interrelationships among research topics in AI and healthcare by clustering frequently co-occurring keywords. Different colours distinguish each cluster, each representing a thematic area. For example, the close grouping of terms like “artificial intelligence,” “convolutional neural networks,” and “medical imaging” demonstrates their strong interconnections and frequent use in AI studies, as supported by Jimma [125].

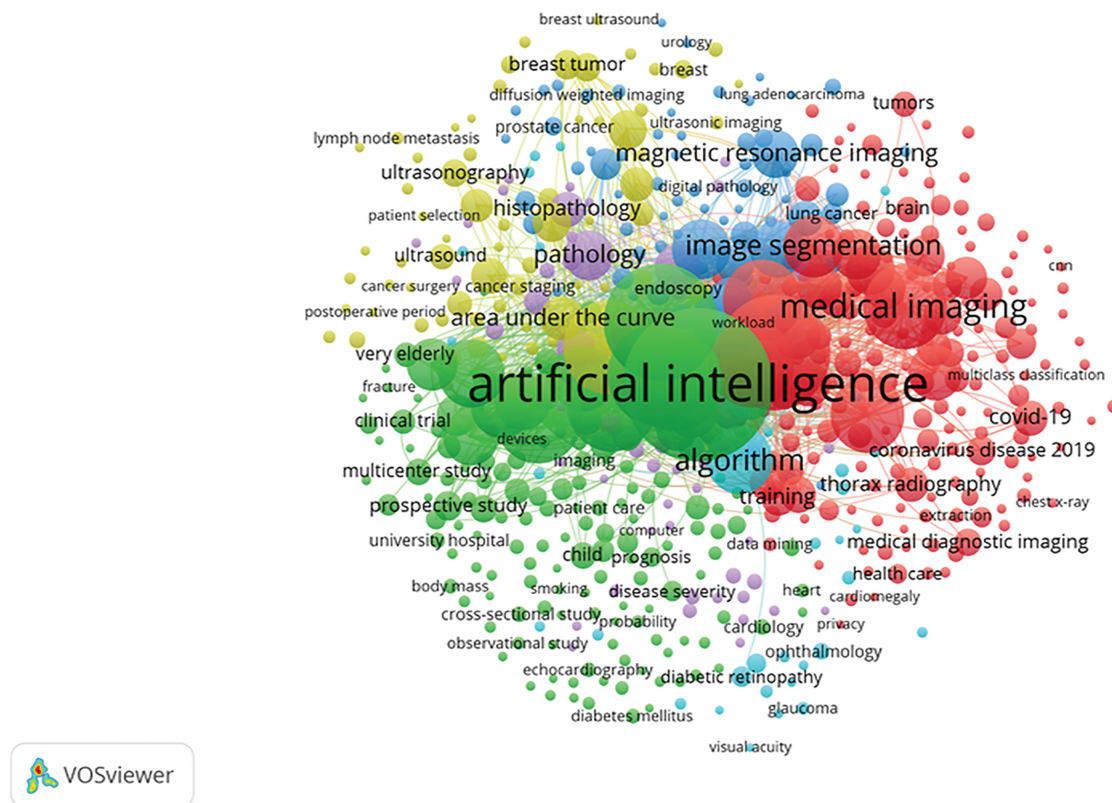


Figure 8: The co-occurrence network created with VOSviewer

By justifying the importance of keyword analysis and cluster formation, we provide a clearer picture of the research landscape. This understanding helps identify major research trends, areas of focus, and potential gaps in the literature, guiding future studies and policy decisions. Furthermore, understanding the connections between different topics can aid in developing more integrated and comprehensive approaches to applying AI in healthcare.

5 Discussion

This section is divided into two parts. Section 5.1 will provide a comprehensive review of the literature to address the first research question, focusing on how ML and DL technologies have improved diagnostic accuracy in medical imaging. Section 5.2 will present an analysis of the results and findings through a bibliometric approach, examining trends and patterns revealed by the data.

5.1 Literature Discussion

Within the evolution of AI in cancer diagnosis (2015–2024), supervised learning techniques such as RFs and SVMs have played a foundational role, consistently demonstrating their efficacy in the classification of intricate medical images. For instance, early in this period, studies like those by Hawkins et al. [65] and Chudhey et al. [64] illustrated the effective use of RFs models in the detection of breast and lung malignancies. In these particular studies, RF models notably surpassed conventional diagnostic methods in prediction accuracy, establishing their utility as a robust alternative. Furthermore, the combined use of SVMs with dimensionality reduction methods such as PCA has significantly improved the precision of prostate and breast cancer detection [66,127]. Beyond individual models, the integration of several models using ensemble techniques has proven to be an especially successful approach in enhancing diagnostic accuracy. Studies conducted by Bron et al. [77], Mienye and Sun [128], Lima et al. [8], Schlegl et al. [129], and Zhou et al. [38] demonstrated that ensemble learning approaches, which integrate models such as RF and Gradient Boosting, exhibit superior performance compared to individual models in the detection of intricate diseases like Alzheimer's disease and colorectal cancer. The widespread implementation of ML and DL techniques in medical imaging has propelled considerable progress across multiple fields, enhancing diagnostic precision, enabling earlier disease detection, and supporting personalized treatment strategies. The analyzed literature strongly indicates that both traditional ML and advanced DL models have become integral components of contemporary medical imaging, providing solutions to previously intractable challenges [35]. Our findings suggest that supervised learning remains the most frequently applied method for medical applications, a trend further supported by Karpov et al. [130].

DL models, specifically CNNs, have completely transformed the process of image classification and segmentation in medical imaging. For example, CheXNet has demonstrated that CNNs can achieve the same or even better performance than human radiologists in detecting pneumonia from chest X-rays [3,94]. Moreover, CNNs show significant potential in identifying diabetic retinopathy, as evidenced by Gulshan et al. [4], who reported a high level of sensitivity and specificity. Additionally, CNNs have shown promise in predicting Alzheimer's disease; for instance, Basaia et al. [105] demonstrated that models could detect small alterations in MRI images. Furthermore, the use of 3D-CNNs in volumetric data analysis, such as CT scans, has led to notable improvements in lung cancer prediction [51,131,132]. In addition, CNN models have improved the early diagnosis of breast cancer, achieving a high level of accuracy, thereby underscoring their continued significance in the field of medical diagnostics. Meanwhile, RL has demonstrated its advantages in optimising treatment plans and the real-time decision-making process in diagnostic imaging. One study by Shen et al. [91] showed that DQN can make radiotherapy treatment plans more effective. This means that healthy tissues are exposed to less radiation, which leads to better outcomes. Moreover, researchers have utilised RL to enhance MRI acquisition procedures, leading to improved picture quality and shorter scan durations [76,133].

GANs have also effectively addressed the limited availability of labeled data by producing artificial images that augment training datasets. For instance, Frid-Adar et al. [110] showed that data generated by GANs can significantly enhance the performance of a classifier in detecting liver lesions. Furthermore, Sparse Autoencoders, as employed by Hou et al. [109], have contributed significantly to the extraction of unsupervised features, thereby improving the precision of future classification tasks in cancer diagnosis. Studies have demonstrated that GANs enhance the quality of radiographic images, thereby improving the accuracy in identifying dental caries lesions. The work Abbaoui et al. [31] illustrated how image enhancement using GANs can improve diagnostic outcomes, despite the lack of precise accuracy measurements.

Finally, transfer learning has gained popularity, particularly in adapting pre-trained models for novel medical imaging applications. According to Abbaoui et al. [31], the use of transfer learning in precision

medicine highlights its ability to reduce dataset size while maintaining accuracy. This approach is especially effective when labeled data is scarce, further demonstrating the adaptability of AI techniques to diverse challenges in medical imaging.

For AI models to truly mature and maintain their efficacy in diverse and evolving clinical environments, a robust and continuous feedback loop from real-world application is indispensable. While initial model development and validation often rely on retrospective datasets, the dynamic nature of patient populations, disease presentations, and imaging protocols necessitates ongoing learning. Integrating AI solutions seamlessly into existing clinical workflows provides an opportunity for real-time performance monitoring and iterative refinement. This feedback allows models to adapt to new data distributions, identify novel patterns, and correct errors that may not have been present in their original training sets. Such a system would enable AI to learn from the discrepancies between its predictions and actual clinical outcomes, continuously improving its generalizability and diagnostic precision over time. However, establishing such a feedback mechanism is complex, involving challenges like ensuring data privacy, standardizing annotation processes, and developing efficient interfaces for clinicians to provide actionable input without disrupting their demanding schedules. Overcoming these hurdles is paramount for AI in medical imaging to transition from a static diagnostic tool to a truly adaptive and ever-improving clinical partner.

Despite the promising advancements of AI in medical imaging, a significant challenge for real-world implementation lies in the limited availability of continuous and structured feedback from clinical environments. AI models often require iterative refinement based on real-world performance data and expert input to maintain and improve their diagnostic accuracy over time. However, several factors impede this crucial feedback loop. Firstly, data privacy regulations and institutional complexities often restrict the easy flow of patient data, which is essential for model validation and re-training. Secondly, many current AI applications are developed and validated using retrospective datasets, which lack the dynamism and real-time variability of clinical practice. Obtaining prospective, real-time annotations and corrections from radiologists and other specialists is labor-intensive, time-consuming, and not yet seamlessly integrated into existing clinical workflows. The absence of standardized mechanisms for collecting diverse and continuous feedback, including instances of model errors or discrepancies, limits the ability of AI systems to adapt and learn from new patient populations or evolving disease patterns. This challenge underscores the need for robust infrastructure and collaborative frameworks that can facilitate the secure and efficient collection of clinical feedback, enabling AI models to truly evolve into reliable and continuously improving diagnostic tools within the healthcare ecosystem.

According to the examined literature, the answer to the first research question, “How have ML and DL technologies improved diagnostic accuracy in medical imaging, and what challenges remain for their real-world implementation?” demonstrates that AI, particularly through ML and DL technologies, has significantly enhanced diagnostic accuracy. For instance, traditional ML methods such as RF and SVMs have proven effective for reliable disease classification. While it is well-established that ensemble ML methods generally show superior performance over single models across various domains, their application in medical imaging is particularly impactful, demonstrably improving prediction precision and robustness in complex diagnostic tasks. Additionally, DL, especially through CNNs, has set new benchmarks in high-accuracy image analysis. The advent of 3D-CNNs has further expanded CNN capabilities to handle complex volumetric data, which is crucial for intricate diagnostic tasks. Beyond these, specialized DL architectures continue to push performance boundaries. For example, innovations like ADMM-Net [1] have significantly improved reconstruction accuracy and computational speed in challenging tasks such as Compressive Sensing MRI. Similarly, Supervised Deep Sparse Coding Networks [2] have demonstrated enhanced discriminative representation learning and efficiency for image classification tasks, particularly in

data-constrained scenarios. Furthermore, advanced generative models like GANs and autoencoders have significantly enhanced model performance by addressing challenges related to dataset size and robust feature extraction. RL also holds promise for optimizing treatment plans and enabling real-time clinical decision-making. Selecting the appropriate technique for a given medical imaging challenge requires carefully balancing these diverse methodological strengths to enhance diagnostic accuracy and ultimately improve patient care.

Addressing the challenges for real-world implementation, as part of the first research question, several critical aspects emerge from the literature. These include the “black box” nature of many advanced AI models, which necessitates advancements in XAI to foster trust and enable clinical adoption [45,48]. Furthermore, the limited availability of continuous and structured feedback from dynamic clinical environments, as discussed previously, poses a significant hurdle for ongoing model refinement and adaptation. Other challenges highlighted involve ensuring data diversity and quality, addressing regulatory and ethical considerations, and achieving seamless integration into existing healthcare workflows. These technical and operational barriers are compounded by crucial socio-technical factors. For instance, fostering clinician trust in AI remains a significant hurdle, often stemming from a lack of transparency in AI’s decision-making process and concerns about accountability. Interoperability between various AI systems and existing hospital IT infrastructure also presents a major challenge, hindering seamless data flow and integration into the diagnostic workflow. Regulatory challenges, including the absence of standardized frameworks for AI model approval and post-market surveillance, further complicate widespread adoption. Overcoming these fundamental barriers is crucial for maximizing the clinical impact and widespread adoption of AI in medical imaging.

Addressing the second part of the research question, which focuses on “*how future research can address these gaps*,” involves enhancing the effectiveness and acceptance of AI applications in medical imaging. This requires focusing on several key research areas. Firstly, improving AI models to process diverse medical imaging data from various sources and patient demographics is crucial for ensuring broad applicability and accuracy. Secondly, conducting extensive trials of these AI systems in real-world clinical settings is essential for verifying their reliability and safety. Additionally, seamlessly integrating AI into existing medical workflows is vital for facilitating its adoption by healthcare professionals. Lastly, addressing ethical concerns and ensuring transparency in AI decision-making processes will be important for maintaining trust and equitable access to these technologies. Targeting these areas will help us better leverage AI to advance medical imaging and improve patient outcomes. It is important to note that many of these insights into AI’s clinical integration and challenges are not unique to diagnostic imaging. The issues of data privacy, regulatory oversight, interoperability, clinician acceptance, and the need for explainability resonate across the broader healthcare landscape, from electronic health records analysis to precision medicine and drug discovery. Therefore, advancements in addressing these barriers within medical imaging could provide transferable solutions and best practices for AI adoption in other healthcare contexts, broadening the overall impact of this research.

5.2 Discussion of Trends through Bibliometric Analysis

Bibliometric, a quantitative technique, analyses scientific literature to uncover trends and patterns within a specific research domain.

Firstly, the analysis of annual scientific production from 2015 to 2024, illustrated in Fig. 2, shows a significant upward trend in the use of AI in medical imaging. This noticeable increase in production is supported by Dhiman et al. [134]. Initially, the number of publications was modest, with only 8 articles in 2015 and 2016. However, from 2017 onwards, there was a steady increase, peaking at 297 articles in 2023 and projected to reach 328 in 2024. This increase underscores the growing importance of AI in medical imaging,

reflecting both heightened research activity and the expanding role of AI in improving diagnostic accuracy. Thus, this trend addresses the first part of Research Question 3 by showing how research output has evolved over the past decade.

In addition, key research areas, illustrated in Fig. 3, highlight trends in topics such as “artificial intelligence,” “deep learning,” and “diagnostic accuracy.” The increase in articles—from 10 on AI in 2015 to 1508 in 2024—demonstrates rapid development and growing recognition. Similarly, the rise in publications on DL (from none in 2015 to 1014 in 2024) and diagnostic imaging (from 4 in 2015 to 702 in 2024) underscores the transformative impact of these technologies. Over time, there have been more studies in the field, as presented by Pereira [111].

Furthermore, Section 4.4 reviews leading journals for AI applications in medical imaging. Section 4.4.1 identifies journals with the highest publication volume, as shown in Fig. 4. IEEE Access leads with 54 articles (19.1%), highlighting its significant role in advancing research in this field [32,134]. Other notable journals include COMPUTERS IN BIOLOGY AND MEDICINE and IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, each with 43 articles accounting (for 15.2%). Consequently, this analysis addresses Research Question 3 by identifying the most prolific journals. Section 4.4.2 evaluates the H-Index of these journals, detailed in Table 5. COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE and IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, both with an H-Index of 17, demonstrate high impact. RADIOLOGY, despite fewer publications, shows considerable influence due to its high citation count. Thus, this assessment provides insight into journal impact and addresses Research Question 3.

Section 4.5 highlights key insights into research impact, showing that the USA leads with 5605 total citations and an average of 29.70 citations per article, indicating its dominant role in AI research. China follows with 5046 total citations but has a lower average of 16.20 citations per article, reflecting its high publication volume. Hong Kong has the highest average citation rate of 55.60 per article, despite a total of 500 articles, while Japan, the UK, and Italy have strong average citation rates of 35.20, 40.40, and 43.50, respectively. This analysis underscores the varying levels of research impact among countries and addresses Research Question 3. Additionally, Fig. 6 reveals a decline in average citations per article despite increased publication output, suggesting potential field saturation [134]. This trend highlights the need for innovative and high-quality research to maintain a significant impact and effectively respond to Research Question 4.

In terms of scientific production, Fig. 7 shows the top 10 countries contributing to AI research in healthcare. China leads with 1438 articles (36.1%), reflecting its substantial investment and focus in this field. The USA follows with 921 articles (23.1%), indicating significant research activity. As noted by Kocak et al. [135] and Dhiman et al. [134], the USA and China are the leading countries in terms of productivity and collaboration. India and Japan contribute 282 (7.1%) and 259 (6.5%) articles, respectively, highlighting their growing influence and emerging roles in AI research. Germany, with 238 articles (6.0%), and South Korea, with 223 articles (5.6%), show strong research output. The UK and Australia follow with 188 (4.7%) and 168 (4.2%) articles, respectively, reflecting their important but somewhat smaller contributions. Canada and Italy round out the top 10 with 140 (3.5%) and 132 (3.3%) articles, respectively, highlighting their significant but more modest roles in the global research landscape. This distribution underscores the varying levels of scientific production among leading countries, with China and the USA standing out as the primary contributors, highlighting their key roles in advancing AI in healthcare and enhancing diagnostic accuracy in medical imaging. Most of these findings are supported by Dhiman et al. [134].

Finally, in response to Research Question 5 regarding emerging themes and trending topics in AI for medical imaging diagnostics, we employed keyword co-occurrence analysis. The network diagram, as shown in Fig. 8, displays circles of various colours and sizes, representing different clusters and keyword

frequencies. For example, “artificial intelligence” appears 1054 times (10.9% of total occurrences), while “diagnostic imaging” appears 669 times (6.9%), “deep learning” 650 times (6.7%), “diagnostic accuracy” 514 times (5.3%), and “machine learning” 412 times (4.3%). These keywords are central to the research themes, underscoring their importance in the literature.

We can conclude that these findings reveal a dynamic and expanding landscape of AI research in healthcare, marked by increasing publication volumes, evolving citation patterns, and significant global contributions, justified by the field’s rapid development and growing technological integration.

5.3 Limitations

This review, while comprehensive, is subject to several limitations that warrant acknowledgement. Firstly, our literature search was restricted to articles published in English and indexed in Scopus and Web of Science. While these databases offer extensive coverage, studies published in other languages or in regional databases might have been overlooked, potentially introducing a selection bias and limiting the global representation of research. Secondly, due to the heterogeneity in study designs, reported metrics, and the diverse range of AI models and imaging modalities encountered, we were unable to perform a formal meta-analysis. This limits our ability to provide a quantitative synthesis of diagnostic accuracy across studies and instead necessitates a qualitative synthesis of findings. Thirdly, the rapid pace of AI development means that some very recent advancements might not yet be fully reflected in the published literature covered by our search period, which concluded on 15 September 2024. Finally, while we have provided a broad overview of AI’s impact and challenges, the scope of this review did not allow for an in-depth clinical validation of specific AI applications or a detailed exploration of socio-economic impacts, which are complex and warrant dedicated investigations.

6 Future Directions and Opportunities

Future research should focus on exploring innovative AI methodologies and their applications in diagnostic imaging to tackle the saturation of existing studies. Rather than simply increasing training set sizes or enhancing image quality, it’s vital to develop entirely new algorithms. By examining the impact of AI innovations on clinical outcomes and integrating diverse datasets, researchers can improve the generalizability of their findings and identify disparities in research output and technology adoption. Gathering data from lesser-known databases and regional journals will offer a more comprehensive view of the research landscape, highlighting emerging challenges and innovations that might be overlooked. To address the decline in citation impact, new research must be both high-quality and relevant to current medical imaging challenges. Collaborative efforts among researchers, clinicians, and industry stakeholders are essential for bridging the gap between theoretical research and practical application. Moreover, enhancing the interpretability of ML and DL models is crucial for gaining wider acceptance in clinical settings. By incorporating multi-modal data and expanding the use of semi-supervised and unsupervised learning techniques, along with developing reinforcement learning models for real-time decision support, we can unlock new opportunities in medical research and clinical practice.

7 Conclusion

This research takes a systematic approach to explore the advancements and challenges of ML and DL technologies in medical imaging diagnostics, framed around key questions regarding their impact on diagnostic accuracy, remaining implementation challenges, existing literature gaps, and the evolution of research output from 2015 to 2024. Our analysis began with 7982 initial records from the Scopus and Web of Science databases, which were refined to a dataset of 1189 relevant articles analyzed using

bibliometric analysis with the R package Biblioshiny. Additionally, 100 articles from the literature were thoroughly investigated to enhance the quality of findings and results. The analysis reveals that ML and DL techniques, particularly CNNs, have significantly improved diagnostic accuracy, although challenges related to interpretability and practical deployment remain. We observed a notable increase in research output, with publications rising from 8 in 2015 to a projected 328 in 2024, indicating growing interest in AI in this field. Key journals identified include *IEEE Access*, *Computers in Biology and Medicine*, and the *IEEE Journal of Biomedical and Health Informatics*, with the United States, China, and Japan as major contributors. Despite this growth, a decline in citation impact per article suggests potential saturation in the field, underscoring the need for innovative studies. Our keyword co-occurrence analysis highlights central themes such as “artificial intelligence,” “deep learning,” and “diagnostic accuracy,” reflecting a shift toward advanced diagnostic tools and emphasizing the necessity for ongoing research and development to address existing challenges and sustain progress in medical imaging diagnostics.

Acknowledgement: None.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: Conceptualization: Omar Sabri and Bassam Al-Shargabi; methodology: Omar Sabri, Bassam Al-Shargabi and Abdelrahman Abuarqoub; software: Omar Sabri and Bassam Al-Shargabi; formal analysis: Omar Sabri, Bassam Al-Shargabi and Abdelrahman Abuarqoub; writing: Omar Sabri, Bassam AL-Shargabi and Abdelrahman Abuarqoub; review and editing: Bassam Al-Shargabi and Abdelrahman Abuarqoub; visualization: Omar Sabri. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The article reviews existing studies only, which are available online on different platforms, e.g., Scopus, Google Scholar, etc.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

1. Sun J, Li H, Xu Z. Deep ADMM-Net for compressive sensing MRI. In: *Advances in neural information processing systems*. Vol. 29. Cambridge, MA, USA: mIT Press; 2016.
2. Sun X, Nasrabadi NM, Tran TD. Supervised deep sparse coding networks for image classification. *IEEE Trans Image Process*. 2019;29:405–18. doi:10.1109/TIP.2019.2928121.
3. Santosh KC, Ghosh S. CheXNet for the evidence of covid-19 using 2.3K positive chest X-rays. In: *Recent trends in image processing and pattern recognition*. Cham, Switzerland: Springer International Publishing; 2022. p. 33–41. doi:10.1007/978-3-031-07005-1_4.
4. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402–10. doi:10.1001/jama.2016.17216.
5. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42(13):60–88. doi:10.1016/j.media.2017.07.005.
6. Oduoye MO, Fatima E, Ali Muzammil M, Dave T, Irfan H, Fariha FNU, et al. Impacts of the advancement in artificial intelligence on laboratory medicine in low- and middle-income countries: challenges and recommendations—a literature review. *Health Sci Rep*. 2024;7(1):e1794. doi:10.1002/hsr2.1794.
7. Wang B, Chang Z, Li S, Hämmäläinen T. An efficient and privacy-preserving blockchain-based authentication scheme for low earth orbit satellite-assisted Internet of Things. *IEEE Trans Aerosp Electron Syst*. 2022;58(6):5153–64. doi:10.1109/TAES.2022.3187389.

8. Lima AA, Mridha MF, Das SC, Kabir MM, Islam MR, Watanobe Y. A comprehensive survey on the detection, classification, and challenges of neurological disorders. *Biology*. 2022;11(3):469. doi:10.3390/biology11030469.
9. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *Lancet Digit Health*. 2019;1(6):e271–97. doi:10.1016/S2589-7500(19)30123-2.
10. Shehab M, Abualigah L, Shambour Q, Abu-Hashem MA, Shambour MKY, Alsalibi AI, et al. Machine learning in medical applications: a review of state-of-the-art methods. *Comput Biol Med*. 2022;145(1):105458. doi:10.1016/j.compbimed.2022.105458.
11. Hinton G. Deep learning—a technology with the potential to transform health care. *JAMA*. 2018;320(11):1101–2. doi:10.1001/jama.2018.11100.
12. Mendes J, Domingues J, Aidos H, Garcia N, Matela N. AI in breast cancer imaging: a survey of different applications. *J Imaging*. 2022;8(9):228. doi:10.3390/jimaging8090228.
13. Fan KL, Thompson YLE, Chen W, Abbey CK, Samuelson FW. Use of equivalent relative utility (ERU) to evaluate artificial intelligence-enabled rule-out devices. *arXiv:2311.07736*. 2023.
14. Born J, Beymer D, Rajan D, Coy A, Mukherjee VV, Manica M, et al. On the role of artificial intelligence in medical imaging of COVID-19. *Patterns*. 2021;2(6):100269. doi:10.1016/j.patter.2021.100269.
15. Ahn JS, Shin S, Yang SA, Park EK, Kim KH, Cho SI, et al. Artificial intelligence in breast cancer diagnosis and personalized medicine. *J Breast Cancer*. 2023;26(5):405. doi:10.4048/jbc.2023.26.e45.
16. Younis HA, Eisa TAE, Nasser M, Sahib TM, Noor AA, Alyasiri OM, et al. A systematic review and meta-analysis of artificial intelligence tools in medicine and healthcare: applications, considerations, limitations, motivation and challenges. *Diagnostics*. 2024;14(1):109. doi:10.3390/diagnostics14010109.
17. Sarkis-Onofre R, Catalá-López F, Aromataris E, Lockwood C. How to properly use the PRISMA statement. *Syst Rev*. 2021;10(1):117. doi:10.1186/s13643-021-01671-z.
18. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*. 2009;339:b2700. doi:10.1136/bmj.b2700.
19. Derviş H. Bibliometric analysis using bibliometrix an R package. *J Scientometr Res*. 2020;8(3):156–60. doi:10.5530/jscires.8.3.32.
20. Volpe S, Mastroleo F, Krengli M, Jereczek-Fossa BA. Quo vadis radiomics? Bibliometric analysis of 10-year radiomics journey. *Eur Radiol*. 2023;33(10):6736–45. doi:10.1007/s00330-023-09645-6.
21. Saheb T, Saheb T, Carpenter DO. Mapping research strands of ethics of artificial intelligence in healthcare: a bibliometric and content analysis. *Comput Biol Med*. 2021;135(80):104660. doi:10.1016/j.compbimed.2021.104660.
22. Ouyang J, Zhao L, Song Y, Qu H, Du T, Shi L, et al. Trends in gut-heart axis and heart failure research (1993–2023): a bibliometric and visual analysis. *Heliyon*. 2024;10(4):e25995. doi:10.1016/j.heliyon.2024.e25995.
23. Yan S, Zhang H, Wang J. Trends and hot topics in radiology, nuclear medicine and medical imaging from 2011–2021: a bibliometric analysis of highly cited papers. *Jpn J Radiol*. 2022;40(8):847–56. doi:10.1007/s11604-022-01268-z.
24. Syed AH, Khan T. Evolution of research trends in artificial intelligence for breast cancer diagnosis and prognosis over the past two decades: a bibliometric analysis. *Front Oncol*. 2022;12:854927. doi:10.3389/fonc.2022.854927.
25. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric analysis: an overview and guidelines. *J Bus Res*. 2021;133(5):285–96. doi:10.1016/j.jbusres.2021.04.070.
26. Boyack KW, Klavans R. Co-citation analysis, bibliographic coupling, and direct citation: which citation approach represents the research front most accurately? *J Am Soc Inf Sci*. 2010;61(12):2389–404. doi:10.1002/asi.21419.
27. Alhashmi SM, Hashem IAT, Al-Qudah I. Artificial intelligence applications in healthcare: a bibliometric and topic model-based analysis. *Intell Syst Appl*. 2024;21(4):200299. doi:10.1016/j.iswa.2023.200299.
28. Aria M, Cuccurullo C. Bibliometrix: an R-tool for comprehensive science mapping analysis. *J Informetr*. 2017;11(4):959–75. doi:10.1016/j.joi.2017.08.007.
29. Nightingale A. A guide to systematic literature reviews. *Surg Oxf*. 2009;27(9):381–4. doi:10.1016/j.mpsur.2009.07.005.

30. Bai A, Si M, Xue P, Qu Y, Jiang Y. Artificial intelligence performance in detecting lymphoma from medical imaging: a systematic review and meta-analysis. *BMC Med Inform Decis Mak.* 2024;24(1):13. doi:10.1186/s12911-023-02397-9.
31. Abbaoui W, Retal S, El Bhiri B, Kharmoum N, Ziti S. Towards revolutionizing precision healthcare: a systematic literature review of artificial intelligence methods in precision medicine. *Inform Med Unlocked.* 2024;46(6):101475. doi:10.1016/j.imu.2024.101475.
32. Kaur S, Singla J, Nkenyereye L, Jha S, Prashar D, Joshi GP, et al. Medical diagnostic systems using artificial intelligence (AI) algorithms: principles and perspectives. *IEEE Access.* 2020;8:228049–69. doi:10.1109/access.2020.3042273.
33. Patel UK, Anwar A, Saleem S, Malik P, Rasul B, Patel K, et al. Artificial intelligence as an emerging technology in the current care of neurological disorders. *J Neurol.* 2021;268(5):1623–42. doi:10.1007/s00415-019-09518-3.
34. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature.* 2017;542(7639):115–8. doi:10.1038/nature21056.
35. Castiglioni I, Rundo L, Codari M, Di Leo G, Salvatore C, Interlenghi M, et al. AI applications to medical images: from machine learning to deep learning. *Phys Med.* 2021;83(76):9–24. doi:10.1016/j.ejmp.2021.02.006.
36. Philip MM, Watts J, Moeini SNM, Musheb M, McKiddie F, Welch A, et al. Comparison of semi-automatic and manual segmentation methods for tumor delineation on head and neck squamous cell carcinoma (HNSCC) positron emission tomography (PET) images. *Phys Med Biol.* 2024;69(9):095005. doi:10.1088/1361-6560/ad37ea.
37. Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, et al. Chexnet: radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv:1711.05225.* 2017.
38. Zhou SK, Greenspan H, Davatzikos C, Duncan JS, van Ginneken B, Madabhushi A, et al. A review of deep learning in medical imaging: imaging traits, technology trends, case studies with progress highlights, and future promises. *Proc IEEE Inst Electr Electron Eng.* 2021;109(5):820–38. doi:10.1109/jproc.2021.3054390.
39. Ting DSW, Cheung CY, Lim G, Tan GSW, Quang ND, Gan A, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA.* 2017;318(22):2211–23. doi:10.1001/jama.2017.18152.
40. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer.* 2018;18(8):500–10. doi:10.1038/s41568-018-0016-5.
41. Hosny A, Parmar C, Coroller TP, Grossmann P, Zeleznik R, Kumar A, et al. Deep learning for lung cancer prognostication: a retrospective multi-cohort radiomics study. *PLoS Med.* 2018;15(11):e1002711. doi:10.1371/journal.pmed.1002711.
42. Albano D, Galiano V, Basile M, Di Luca F, Gitto S, Messina C, et al. Artificial intelligence for radiographic imaging detection of caries lesions: a systematic review. *BMC Oral Health.* 2024;24(1):274. doi:10.1186/s12903-024-04046-7.
43. Alberg AJ, Park JW, Hager BW, Brock MV, Diener-West M. The use of “overall accuracy” to evaluate the validity of screening or diagnostic tests. *J Gen Intern Med.* 2004;19(5 Pt 1):460–5. doi:10.1111/j.1525-1497.2004.30091.x.
44. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521(7553):436–44. doi:10.1038/nature14539.
45. Longo L, Brcic M, Cabitza F, Choi J, Confalonieri R, Del Ser J, et al. Explainable artificial intelligence (XAI) 2.0: a manifesto of open challenges and interdisciplinary research directions. *Inf Fusion.* 2024;106(3):102301. doi:10.1016/j.inffus.2024.102301.
46. Parola M, Galatolo FA, Mantia GL, Cimino MGCA, Campisi G, Di Fede O. Towards explainable oral cancer recognition: screening on imperfect images via informed deep learning and case-based reasoning. *Comput Med Imaging Graph.* 2024;117(21):102433. doi:10.1016/j.compmedimag.2024.102433.
47. Cimino MGCA, Campisi G, Galatolo FA, Neri P, Tozzo P, Parola M, et al. Explainable screening of oral cancer via deep learning and case-based reasoning. *Smart Health.* 2025;35(4):100538. doi:10.1016/j.smhl.2024.100538.
48. Ali S, Abuhmed T, El-Sappagh S, Muhammad K, Alonso-Moral JM, Confalonieri R, et al. Explainable artificial intelligence (XAI): what we know and what is left to attain trustworthy artificial intelligence. *Inf Fusion.* 2023;99(3):101805. doi:10.1016/j.inffus.2023.101805.

49. Rana M, Bhushan M. Machine learning and deep learning approach for medical image analysis: diagnosis to detection. *Multimed Tools Appl.* 2023;82(17):26731–69. doi:10.1007/s11042-022-14305-w.
50. Suganyadevi S, Seethalakshmi V, Balasamy K. A review on deep learning in medical image analysis. *Int J Multimed Inf Retr.* 2022;11(1):19–38. doi:10.1007/s13735-021-00218-1.
51. Iqbal S, Qureshi AN, Li J, Mahmood T. On the analyses of medical images using traditional machine learning techniques and convolutional neural networks. *Arch Comput Methods Eng.* 2023;30(5):3173–233. doi:10.1007/s11831-023-09899-9.
52. Pacurari AC, Bhattarai S, Muhammad A, Avram C, Mederle AO, Rosca O, et al. Diagnostic accuracy of machine learning AI architectures in detection and classification of lung cancer: a systematic review. *Diagnostics.* 2023;13(13):2145. doi:10.3390/diagnostics13132145.
53. Pokkalla H, Pethia K, Taylor A, Glass B, Elliott H, Han L, et al. FRI003 Machine learning models identify novel histologic features predictive of clinical disease progression in patients with advanced fibrosis due to non-alcoholic steatohepatitis. *J Hepatol.* 2020;73:S402. doi:10.1016/S0168-8278(20)31289-7.
54. Wu N, Huang Z, Shen Y, Park J, Phang J, Makino T, et al. Reducing false-positive biopsies using deep neural networks that utilize both local and global image context of screening mammograms. *J Digit Imaging.* 2021;34(6):1414–23. doi:10.1007/s10278-021-00530-6.
55. Ghaffar Nia N, Kaplanoglu E, Nasab A. Evaluation of artificial intelligence techniques in disease diagnosis and prediction. *Discov Artif Intell.* 2023;3(1):5. doi:10.1007/s44163-023-00049-5.
56. Bertolini M, Mezzogori D, Neroni M, Zammori F. Machine learning for industrial applications: a comprehensive literature review. *Expert Syst Appl.* 2021;175(6):114820. doi:10.1016/j.eswa.2021.114820.
57. Rehan H. Revolutionizing America's cloud computing the pivotal role of AI in driving innovation and security. *J Artif Intell Gen Sci.* 2024;2(1):239–40. doi:10.60087/jaigs.v2i1.110.
58. Bishop CM, Nasrabadi NM. *Pattern recognition and machine learning.* New York, NY, USA: Springer; 2006.
59. Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning: data mining, inference, and prediction.* 2nd ed. Berlin/Heidelberg, Germany: Springer; 2009.
60. Sutton RS, Barto AG. *Reinforcement learning: an introduction.* Cambridge, MA, USA: mIT Press; 1998.
61. Cortes C, Vapnik V. Support-vector networks. *Mach Learn.* 1995;20(3):273–97. doi:10.1023/A:1022627411411.
62. Nathaniel Kumar Sarella P, Therissa Mangam V. AI-driven natural language processing in healthcare: transforming patient-provider communication. *Ind J Pharm Pract.* 2024;17(1):21–6. doi:10.5530/ijopp.17.1.4.
63. Jayatilake SMDAC, Ganegoda GU, Martorelli M. Involvement of machine learning tools in healthcare decision making. *J Healthc Eng.* 2021;2021(12):6679512. doi:10.1155/2021/6679512.
64. Chudhey AS, Goel M, Singh M. Breast cancer classification with random forest classifier with feature decomposition using principal component analysis. In: *Advances in data and information sciences.* Singapore: Springer Singapore; 2022. p. 111–20. doi:10.1007/978-981-16-5689-7_10.
65. Hawkins S, Wang H, Liu Y, Garcia A, Stringfield O, Krewer H, et al. Predicting malignant nodules from screening CT scans. *J Thorac Oncol.* 2016;11(12):2120–8. doi:10.1016/j.jtho.2016.07.002.
66. Akay MF. Support vector machines combined with feature selection for breast cancer diagnosis. *Expert Syst Appl.* 2009;36(2):3240–7. doi:10.1016/j.eswa.2008.01.009.
67. Wang S, Burt K, Turkbey B, Choyke P, Summers RM. Computer aided-diagnosis of prostate cancer on multi-parametric MRI: a technical review of current research. *Biomed Res Int.* 2014;2014(13):789561. doi:10.1155/2014/789561.
68. Codella NCF, Nguyen QB, Pankanti S, Gutman DA, Helba B, Halpern AC, et al. Deep learning ensembles for melanoma recognition in dermoscopy images. *IBM J Res Dev.* 2017;4(4/5):5:1–15. doi:10.1147/jrd.2017.2708299.
69. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016 Aug 13–17; San Francisco, CA, USA.* p. 785–94. doi:10.1145/2939672.2939785.
70. Nandi D, Ashour AS, Samanta S, Chakraborty S, Salem MAM, Dey N. Principal component analysis in medical image processing: a study. *Int J Image Min.* 2015;1(1):65. doi:10.1504/ijim.2015.070024.

71. Khan AR, Khan S, Harouni M, Abbasi R, Iqbal S, Mehmood Z. Brain tumor segmentation using K-means clustering and deep learning with synthetic data augmentation for classification. *Microsc Res Tech.* 2021;84(7):1389–99. doi:10.1002/jemt.23694.
72. Eckardt JN, Wendt K, Bornhäuser M, Middeke JM. Reinforcement learning for precision oncology. *Cancers.* 2021;13(18):4624. doi:10.3390/cancers13184624.
73. Dhinakaran D, Srinivasan L, Selvaraj D, Udhaya Sankar SM. Leveraging semi-supervised graph learning for enhanced diabetic retinopathy detection. *Int J Electron Commun Eng.* 2023;10(8):9–21. doi:10.14445/23488549/ijece-v10i8p102.
74. Yue W, Wang Z, Chen H, Payne A, Liu X. Machine learning with applications in breast cancer diagnosis and prognosis. *Designs.* 2018;2(2):13. doi:10.3390/designs2020013.
75. Ogunpola A, Saeed F, Basurra S, Albarrak AM, Qasem SN. Machine learning-based predictive models for detection of cardiovascular diseases. *Diagnostics.* 2024;14(2):144. doi:10.3390/diagnostics14020144.
76. Bahrami K, Rekik I, Shi F, Shen D. Joint reconstruction and segmentation of 7T-like MR images from 3T MRI based on cascaded convolutional neural networks. In: *Medical Image Computing and Computer Assisted Intervention—MICCAI 2017.* Cham, Switzerland: Springer International Publishing; 2017. p. 764–72. doi:10.1007/978-3-319-66182-7_87.
77. Bron EE, Smits M, van der Flier WM, Vrenken H, Barkhof F, Scheltens P, et al. Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: the CADDementia challenge. *Neuroimage.* 2015;111(12):562–79. doi:10.1016/j.neuroimage.2015.01.048.
78. Lazli L. Machine learning classifiers based on dimensionality reduction techniques for the early diagnosis of Alzheimer's disease using magnetic resonance imaging and positron emission tomography brain data. In: *Computational intelligence methods for bioinformatics and biostatistics.* Cham, Switzerland: Springer International Publishing; 2022. p. 117–31. doi:10.1007/978-3-031-20837-9_10.
79. Wenger K, Tirdad K, Dela Cruz A, Mari A, Basheer M, Kuk C, et al. A semi-supervised learning approach for bladder cancer grading. *Mach Learn Appl.* 2022;9(1):100347. doi:10.1016/j.mlwa.2022.100347.
80. Hajhashemi V, Hassani Z, Dehmajnoonie I, Borna K. Hybrid Algorithms of Whale optimization algorithm and k-nearest neighbor to Predict the liver disease. *EAI Endorsed Trans Context Aware Syst Appl.* 2019;6(16):156838. doi:10.4108/eai.13-7-2018.156838.
81. Talukder MA, Islam MM, Uddin MA, Akhter A, Hasan KF, Ali Moni M. Machine learning-based lung and colon cancer detection using deep feature extraction and ensemble learning. *Expert Syst Appl.* 2022;205:117695. doi:10.1016/j.eswa.2022.117695.
82. Mirikharaji Z, Yan Y, Hamarneh G. Learning to segment skin lesions from noisy annotations. In: *Domain adaptation and representation transfer and medical image learning with less labels and imperfect data.* Cham, Switzerland: Springer International Publishing; 2019. p. 207–15.
83. Mehrtash A, Ziegler E, Idris T, Somarouthu B, Urban T, LaCasce AS, et al. Evaluation of mediastinal lymph node segmentation of heterogeneous CT data with full and weak supervision. *Comput Med Imaging Graph.* 2024;111(6):102312. doi:10.1016/j.compmedimag.2023.102312.
84. Sun Y, Reynolds H, Wraith D, Williams S, Finnegan ME, Mitchell C, et al. Predicting prostate tumour location from multiparametric MRI using Gaussian kernel support vector machines: a preliminary study. *Australas Phys Eng Sci Med.* 2017;40(1):39–49. doi:10.1007/s13246-016-0515-1.
85. Zhu X, Goldberg AB. Overview of semi-supervised learning. In: *Introduction to semi-supervised learning.* Cham, Switzerland: Springer International Publishing; 2009. p. 9–19. doi:10.1007/978-3-031-01548-9_2.
86. Cristian PM, Aarón VJ, Armando ED, Estrella MY, Daniel NR, David GV, et al. Diffusion on PCA-UMAP manifold: the impact of data structure preservation to denoise high-dimensional single-cell RNA sequencing data. *Biology.* 2024;13(7):512. doi:10.3390/biology13070512.
87. Roy S, Meena T, Lim SJ. Demystifying supervised learning in healthcare 4.0: a new reality of transforming diagnostic medicine. *Diagnostics.* 2022;12(10):2549. doi:10.3390/diagnostics12102549.
88. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. *Nature.* 2020;577(7788):89–94. doi:10.1038/s41586-019-1799-6.

89. Sahin EK. Assessing the predictive capability of ensemble tree methods for landslide susceptibility mapping using XGBoost, gradient boosting machine, and random forest. *SN Appl Sci.* 2020;2(7):1308. doi:10.1007/s42452-020-3060-1.
90. Wang X, He J, Jin Z, Yang M, Wang Y, Qu H. M2Lens: visualizing and explaining multimodal models for sentiment analysis. *IEEE Trans Vis Comput Graph.* 2022;28(1):802–12. doi:10.1109/TVCG.2021.3114794.
91. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. *Annu Rev Biomed Eng.* 2017;19(1):221–48. doi:10.1146/annurev-bioeng-071516-044442.
92. Li M, Jiang Y, Zhang Y, Zhu H. Medical image analysis using deep learning algorithms. *Front Public Health.* 2023;11:1273253. doi:10.3389/fpubh.2023.1273253.
93. Shin H-C, Tenenholz NA, Rogers JK, Schwarz CG, Senjem ML, Gunter JL, et al. Medical image synthesis for data augmentation and anonymization using generative adversarial networks. In: *Simulation and synthesis in medical imaging.* Berlin/Heidelberg, Germany: Springer; 2018. p. 1–11 doi:10.1007/978-3-030-00536-8_1.
94. Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, et al. Deep learning for chest radiograph diagnosis: a retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med.* 2018;15(11):e1002686. doi:10.1371/journal.pmed.1002686.
95. Voter AF, Larson ME, Garrett JW, Yu JJ. Diagnostic accuracy and failure mode analysis of a deep learning algorithm for the detection of cervical spine fractures. *Am J Neuroradiol.* 2021;42(8):1550–6. doi:10.3174/ajnr.a7179.
96. Aggarwal R, Sounderajah V, Martin G, Ting DSW, Karthikesalingam A, King D, et al. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. *npj Digit Med.* 2021;4(1):65. doi:10.1038/s41746-021-00438-z.
97. Jiang F, Fu Y, Gupta BB, Liang Y, Rho S, Lou F, et al. Deep learning based multi-channel intelligent attack detection for data security. *IEEE Trans Sustain Comput.* 2020;5(2):204–12. doi:10.1109/tsusc.2018.2793284.
98. Page J, Moser F, Maya M, Prasad R, Pressman B. SUN-370 validation of a deep learning based algorithm to diagnose vertebral compression fractures. *J Endocr Soc.* 2020;4(Suppl 1):SUN–370. doi:10.1210/jendso/bvaa046.333.
99. Dhiravidachelvi E, Senthil Pandi S, Prabavathi R, Bala Subramanian C. Artificial humming bird optimization-based hybrid CNN-RNN for accurate exudate classification from fundus images. *J Digit Imaging.* 2023;36(1):59–72. doi:10.1007/s10278-022-00707-7.
100. Xu J, Pan Y, Pan X, Hoi S, Yi Z, Xu Z. RegNet: self-regulated network for image classification. *IEEE Trans Neural Netw Learn Syst.* 2023;34(11):9562–7. doi:10.1109/tnnls.2022.3158966.
101. Alqahtani H, Kavakli-Thorne M, Kumar G. Applications of generative adversarial networks (GANs): an updated review. *Arch Comput Meth Eng.* 2021;28(2):525–52. doi:10.1007/s11831-019-09388-y.
102. Bowles C, Chen L, Guerrero R, Bentley P, Gunn R, Hammers A, et al. Gan augmentation: augmenting training data using generative adversarial networks. *arXiv:1810.10863.* 2018.
103. Mehmood A, Yang S, Feng Z, Wang M, Ahmad AS, Khan R, et al. A transfer learning approach for early diagnosis of Alzheimer's disease on MRI images. *Neuroscience.* 2021;460(5):43–52. doi:10.1016/j.neuroscience.2021.01.002.
104. Al-Janabi A. Has DeepMind's AlphaFold solved the protein folding problem? *Biotechniques.* 2022;72(3):73–6. doi:10.2144/btn-2022-0007.
105. Basaia S, Agosta F, Wagner L, Canu E, Magnani G, Santangelo R, et al. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *Neuroimage Clin.* 2019;21:101645. doi:10.1016/j.nicl.2018.101645.
106. Tseng HH, Luo Y, Cui S, Chien JT, Ten Haken RK, Naqa IE. Deep reinforcement learning for automated radiation adaptation in lung cancer. *Med Phys.* 2017;44(12):6690–705. doi:10.1002/mp.12625.
107. Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat Med.* 2019;25(1):65–9. doi:10.1038/s41591-018-0268-3.
108. Kawahara J, Hamarneh G. Multi-resolution-tract CNN with hybrid pretrained and skin-lesion trained layers. In: *Machine learning in medical imaging.* Cham, Switzerland: Springer International Publishing; 2016. p. 164–71. doi:10.1007/978-3-319-47157-0_20.

109. Hou L, Nguyen V, Kanevsky AB, Samaras D, Kurc TM, Zhao T, et al. Sparse autoencoder for unsupervised nucleus detection and representation in histopathology images. *Pattern Recognit.* 2019;86(8):188–200. doi:10.1016/j.patcog.2018.09.007.
110. Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing.* 2018;321(5):321–31. doi:10.1016/j.neucom.2018.09.013.
111. Pereira S, Pinto A, Alves V, Silva CA. Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Trans Med Imag.* 2016;35(5):1240–51. doi:10.1109/TMI.2016.2538465.
112. Suk HI, Lee SW, Shen D. Alzheimer's Disease Neuroimaging Initiative. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *Neuroimage.* 2014;101:569–82. doi:10.1016/j.neuroimage.2014.06.077.
113. Zreik M, Lessmann N, van Hamersvelt RW, Wolterink JM, Voskuil M, Viergever MA, et al. Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis. *Med Image Anal.* 2018;44(1):72–85. doi:10.1016/j.media.2017.11.008.
114. Yang D, Liu J, Wang Y, Xu B, Wang X. Application of a generative adversarial network in image reconstruction of magnetic induction tomography. *Sensors.* 2021;21(11):3869. doi:10.3390/s21113869.
115. Spampinato C, Palazzo S, Giordano D, Aldinucci M, Leonardi R. Deep learning for automated skeletal bone age assessment in X-ray images. *Med Image Anal.* 2017;36:41–51. doi:10.1016/j.media.2016.10.010.
116. Barnoy Y, O'Brien M, Wang W, Hager G. Robotic surgery with lean reinforcement learning. arXiv:2105.01006. 2021.
117. Goyal M, Reeves ND, Rajbhandari S, Ahmad N, Wang C, Yap MH. Recognition of ischaemia and infection in diabetic foot ulcers: dataset and techniques. *Comput Biol Med.* 2020;117(7):103616. doi:10.1016/j.compbiomed.2020.103616.
118. Lv J, Li G, Tong X, Chen W, Huang J, Wang C, et al. Transfer learning enhanced generative adversarial networks for multi-channel MRI reconstruction. *Comput Biol Med.* 2021;134(2):104504. doi:10.1016/j.compbiomed.2021.104504.
119. Schlegl T, Seebock P, Waldstein SM, Schmidt-Erfurth U, Langs G. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In: *Information processing in medical imaging*. Cham, Switzerland: Springer International Publishing; 2017. p. 146–57. doi:10.1007/978-3-319-59050-9_12.
120. Frid-Adar M, Klang E, Amitai M, Goldberger J, Greenspan H. Synthetic data augmentation using GAN for improved liver lesion classification. In: *IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*; 2018 Apr 4–7; Washington, DC, USA; 2018. p. 289–93. doi:10.1109/ISBI.2018.8363576.
121. Yaqub M, Feng J, Ahmed S, Arshid K, Bilal MA, Akhter MP, et al. GAN-TL: generative adversarial networks with transfer learning for MRI reconstruction. *Appl Sci.* 2022;12(17):8841. doi:10.3390/app12178841.
122. Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, et al. Deep convolutional neural networks for computer-aided detection: cnn architectures, dataset characteristics and transfer learning. *IEEE Trans Med Imag.* 2016;35(5):1285–98. doi:10.1109/tmi.2016.2528162.
123. Bornmann L, Daniel HD. What do we know about the h index? *J Am Soc Inf Sci.* 2007;58(9):1381–5. doi:10.1002/asi.20609.
124. Guo Y, Hao Z, Zhao S, Gong J, Yang F. Artificial intelligence in health care: bibliometric analysis. *J Med Internet Res.* 2020;22(7):e18228. doi:10.2196/18228.
125. Jimma BL. Artificial intelligence in healthcare: a bibliometric analysis. *Telematics Inform Rep.* 2023;9(Suppl. 1):100041. doi:10.1016/j.teler.2023.100041.
126. Feng Y, Zhu Q, Lai KH. Corporate social responsibility for supply chain management: a literature review and bibliometric analysis. *J Clean Prod.* 2017;158(43):296–307. doi:10.1016/j.jclepro.2017.05.018.
127. Fei B. Computer-aided diagnosis of prostate cancer with MRI. *Curr Opin Biomed Eng.* 2017;3(Pt 2):20–7. doi:10.1016/j.cobme.2017.09.009.
128. Mienye ID, Sun Y. A survey of ensemble learning: concepts, algorithms, applications, and prospects. *IEEE Access.* 2022;10(1):99129–49. doi:10.1109/access.2022.3207287.

129. Schlegl T, Seeböck P, Waldstein SM, Langs G, Schmidt-Erfurth U. F-AnoGAN: fast unsupervised anomaly detection with generative adversarial networks. *Med Image Anal.* 2019;54(3):30–44. doi:10.1016/j.media.2019.01.010.
130. Karpov OE, Pitsik EN, Kurkin SA, Maksimenko VA, Gusev AV, Shusharina NN, et al. Analysis of publication activity and research trends in the field of AI medical applications: network approach. *Int J Environ Res Public Health.* 2023;20(7):5335. doi:10.3390/ijerph20075335.
131. Alakwaa W, Nassef M, Badr A. Lung cancer detection and classification with 3D convolutional neural network (3D-CNN). *IJACSA.* 2017;8(8):409–17. doi:10.14569/ijacsa.2017.080853.
132. Majidpourkhoei R, Alilou M, Majidzadeh K, Babazadehsangar A. A novel deep learning framework for lung nodule detection in 3d CT images. *Multimed Tools Appl.* 2021;80(20):30539–55. doi:10.1007/s11042-021-11066-w.
133. Arwa EO, Folly KA. Reinforcement learning techniques for optimal power control in grid-connected microgrids: a comprehensive review. *IEEE Access.* 2020;8:208992–9007. doi:10.1109/access.2020.3038735.
134. Dhiman P, Bonkra A, Kaur A, Gulzar Y, Hamid Y, Mir MS, et al. Healthcare trust evolution with explainable artificial intelligence: bibliometric analysis. *Information.* 2023;14(10):541. doi:10.3390/info14100541.
135. Kocak B, Baessler B, Cuocolo R, Mercaldo N, Pinto Dos Santos D. Trends and statistics of artificial intelligence and radiomics research in Radiology, Nuclear Medicine, and Medical Imaging: bibliometric analysis. *Eur Radiol.* 2023;33(11):7542–55. doi:10.1007/s00330-023-09772-0.