Emerging multifaceted application of artificial intelligence in chest radiography: a narrative review

Victor Chigbundu Nwaiwu^, Sreemoy Kanti Das

Lincoln University College, Petaling Jaya, Selangor, Malaysia

Contributions: (I) Conception and design: VC Nwaiwu; (II) Administrative support: SK Das; (III) Provision of study materials or patients: Both authors; (IV) Collection and assembly of data: VC Nwaiwu; (V) Data analysis and interpretation: Both authors; (VI) Manuscript writing: Both authors; (VII) Final approval of manuscript: Both authors.

Correspondence to: Victor Chigbundu Nwaiwu, BSc, PgCert, MSc. Lincoln University College, Wisma Lincoln, 12-18, Jalan SS 6/12, Ss 6 47301 Petaling Jaya, Selangor, Malaysia. Email: vcnwaiwu@lincoln.edu.my.

Background and Objective: Chest radiography, otherwise known as chest X-ray (CXR) is the most in demand and widely performed investigation in radiology department; owing to the multiple combining effect of rise in prevalence of respiratory diseases globally and the growing need of heath assessment for pre-employment, pre-operative and migration purposes. However, this task is already proving overwhelming, placing an immeasurable burden of workload on radiographers, radiologist, and the entire health system; this has resulted in long waiting time, fatigue-based technical error, interpretation error, reporting delays and backlogs. To ameliorate this predicament, medical imaging has witnessed the introduction of artificial intelligence (AI). Thus, with the raid evolutionary trend in technology, this article seeks to review current state of evidence on AI use in CXR and level of progress made to minimize these errors and delays. In addition, point out challenges, as well as unfold areas for future research to better detection rates and improve overall clinical outcomes.

Methods: A search for relevant literature that focuses on AI in CXR was conducted with the help of certain keywords [machine learning (ML), chest radiography, deep learning (DL), natural language processing (NLP), expert system (ES) and fuzzy logic (FL)]. Thereafter, a narrative logical approach to technically analysing and synthesizing findings across domains of AI (ML, DL, NLP, ES, FL) and robot technologies as it relates to CXR done.

Key Content and Findings: A thorough evaluation of the substance of evidence these studies bring to enhance overall workflow and health outcomes show that ML is very useful in performing administrative and imaging tasks such as exam scheduling, worklist management and image acquisition. On the other hand, DL is better suited for classification tasks on a broad spectrum of chest anomalies in CXR. However, a hybrid approach involving ML-DL, FL-DL and NLP-DL/ML technologies seems to further improve reporting accuracy and offer more insights into CXR interpretation. Further studies on training and refining models for clinical use in this perspective is demanded.

Conclusions: AI still in its early stages; this review to serve as road map to implementation and policy making, guide routine practice and improve clinical governance.

Keywords: Artificial intelligence (AI); natural language processing (NLP); fuzzy logic (FL); chest radiography; machine/deep learning (ML/DL)

Received: 10 March 2024; Accepted: 29 May 2024; Published online: 15 July 2024.
doi: 10.21037/jmai-24-67
View this article at: https://dx.doi.org/10.21037/jmai-24-67

^ ORCID: 0009-0008-9981-6881.

Introduction

Chest radiography is a painless non-invasive diagnostic test that uses conventional X-rays for evaluation of the airways, pulmonary parenchyma and vessels, mediastinum, heart, pleura and indeed the entire chest wall (1). Radiographic techniques such as posteroanterior (PA), anteroposterior (AP), obliques and lateral projections gotten in erect, supine and decubitus positions can achieve this (2,3). Chest X-ray (CXR) is the most performed investigation in radiography unit and one of the widely conducted diagnostic imaging test (both adults and children) in many countries, accounting for nearly half of the entire radiographic images obtained in routine practice (44%) (2,4-6). CXR, thanks to its high availability, low cost, execution at the patients' bed, represents the first-line examination both in the emergency and in the standard settings (7).

CXR supports diagnosis, treatment, and management of thoracic- related diseases and is required by clinicians in patients presenting with chest pains, shortness of breath, cough, metastatic diseases and malignancy, and any intending medical surgical procedure (8). It plays a huge diversifying role in the diagnosis and management of rising respiratory disorders witnessed worldwide, for recruitment and pre-operative checks, migration purposes and periodic health checkups. These situations have placed an exaggerate burden on radiographers, radiologists, and the entire health system, leading to long waiting time, fatigue-based technical and interpretation errors, reporting delays, backlogs, and longer turnaround time (9,10). In fact, according to Pouraliakbar (11), CXR remains the most common radiographic yet one of the most difficult to interpret; a high misinterpretation rate of 30% reported in the recent findings of Kaviani *et al.* (6). Also evident is the lack of diagnostic expertise in rural areas of the world where radiologists are unavailable (9,10).

In a bid to meet this growing demand, medical imaging has witnessed the advert of artificial intelligence (AI). With a steep increase in medical images acquired and a vast amount of image reviews observed in the last decade, AI (comprising different technologies that is based on advanced algorithms and learning system) is projected to be one of the major disrupting forces in radiology in future health care practice (12,13). Also is a huge expectation in delivering timely and accurate interpretation of chest radiographs obtained such as high-level confidence in differentiating normal from abnormal, and further characterizing the abnormal findings. This is anticipated to guide clinicians on proper clinical evaluation, treatment and management, and appropriate follow up (8). However, with AI still very much in the early stages, we shall in this narrative review discuss each domain of AI and its applicability to achieving this in CXR. It is envisaged that research gaps for future studies will be unravelled to support routine clinical implementation, policy making, address possible limitations and enhance clinical governance. We present this article in accordance with the Narrative Review reporting checklist (available at https://jmai.amegroups.com/article/view/10.21037/jmai-24-67/rc).

Methods

A search for all relevant literature that focuses on major AI techniques in chest radiography was done on 31st December 2023 by a team of experienced radiographers and radiologist in several databases (PubMed, Web of Science, Google Scholar, MEDLINE, ScienceDirect, Cochrane library, PLOS, and Scopus). Certain keywords aided this process (artificial intelligence, chest radiography, machine learning, deep learning, natural language processing, expert system, fuzzy logic), filtering studies only in English, conducted any year, and of any design; synthesizing evidence per AI domain in a narrative way that highlights progress made, current benefits, challenges and future opportunities to improve detection rates and patient outcomes (*Table 1*).

Discussion

This section covers discussion on key findings from existing literature across AI domains as well as robotic technologies. *Table 2* summarises the strength, limitations and gaps from all discussed literature in this review.

Machine learning (ML)

ML focuses on building a machine with the ability to learn from data and experience through algorithms, which are the engines that power ML, informing the computer how to learn to operate on its own (14). With the advent of digital radiography (DR), there has been great excitement in the last decade developing AI applications, with potential benefits of ML in chest radiography spanning across optimizing all steps in the imaging chain-exam scheduling, worklist management, image acquisition and image interpretation. A combination of algorithms for administrative, non-interpretive and quality improvement

Table 1 The search strategy summary

Items	Specification
Date of search	31st December 2023
Databases and other sources searched	PubMed, Web of Science, Google Scholar, MEDLINE, ScienceDirect, Cochrane library, PLOS, Scopus
Search terms used	Artificial intelligence, chest radiography, machine learning, deep learning, natural language processing, expert system, fuzzy logic, robotics
Timeframe	Studies conducted any year (no timeframe)
Inclusion criteria	Any study type, but studies only in English
Selection process	Selection process was conducted by a team of radiographers and radiologist, who did this together, pulling knowledge in their areas of expertise as it relates to chest radiography jointly

purposes is majorly a function of ML, a subset of AI. The desire to do this is clear, owing to the growth in imaging volumes, number of images per study and vast amount of medical information available through electronic medical records (15).

Exam scheduling

ML algorithms has the potential to support a variety of non-diagnostic tasks for quality improvement purposes such as order entry support, patient scheduling and resource allocation through incorporation of ML algorithms into electronic health record systems (16).

Pierre *et al.* (17) demonstrated integration of ML and NLP tools into the scheduling software which automatically is known as "concept of the electronic round trip". This usually begins with the referring provider (or the patient) placing an order for a radiologic examination electronically, which flows over to the electronic health record (EHR) and radiology information system (RIS) where applicable, followed by the imaging device, and then the viewer or picture archiving and communication system (PACS). Hence no need to re-enter data at any step of the cycle.

In this study (18), a scheduler was created to schedule CXRs for patients; the data entry areas of the user interface permitting request for CXR, preferred time and location. This database successfully searched utilizing a "nearest neighbor algorithm" to either match the criteria or return the best alternative to the scheduler; to accept, decline or request another time. In spite of the progress, this field appears to still very much be in the embryonic stages, most ML uses theoretical in development or limited to a particular institution (15). Thus, as new AI software for scheduling imaging appointments is developed, it is essential that robust validation is undertaken in compliance with evidence-based medical imaging (19).

Worklist management

ML algorithms can also successfully predict wait times, appointment delays and no shows for patients scheduled to undergo imaging examination, based on environmental and patient-related factors (20,21). Nelson *et al.* (22) illustrated ML combination with large-scale data allows for creation of rich complex high dimensional complex models, able to predict not only attendance (that allows for targeted intervention) but matching detailed appointment and patient characteristics useful to infer systemic modifiable hospital causes of non-attendance by patients.

Chong *et al.* (23) trained a model to predict patients with the highest risk of missing their appointment, and these patients receiving a phone call reminder. It was found to decrease the no-show rate from 19.3% to 15.9% during its 6-month period of deployment, as missed appointments is usually associated with increased health care cost and high risk of poor health outcomes (24).

Furthermore, in the recent studies of Baltruschat *et al.* (25), an ML model that performed worklist prioritization for critical findings in chest radiographs was invented. A realistic simulation framework developed based on convolutional neural networks (CNNs) in which worklist were rearranged by AI, reporting based on urgency level of CXR findings instead of the usual serial arrangement [first in first out (FIFO)]; and this was found to be associated with a reduction in average report turnaround time (RTATs) for all critical CXR findings.

As seen in Annarumma *et al.*'s study (26), similar findings were observed in the studies carried out by a research team led by Giovanni Montana, extracting about 470,388 adult CXRs acquired from 2007–2017 at Guy's and St Thomas' NHS Hospitals. Having understood the findings of reporting radiologist, formulated an algorithm that can infer priority level for each radiograph (as critical, urgent, non-urgent and normal). This team applied the algorithm comprising of computer vision ML models and NLP system, which was discovered to cut down the average review time from 11 to 2.7 days for critical CXRs and from 7.6 to 4.1 days for urgent CXRs. Elsewhere in the study of Nabulsi *et al.* (27), AI was found to prioritize abnormal cases in a simulated workflow, turnaround time for abnormal cases reduced by 7–28%. These results are indicative of how such models can greatly reduce delays in acting on abnormal CXRs.

In spite of the low risk and few rules/regulations associated with this multiple software, its limited availability as well as lack of a well-documented research/evidence is stalling implementation in clinical practice; majority of conducted studies being retrospective and using historically labelled data to train and test algorithms. This gap is of great concern as AI performance is projected to be likely worse when encountering real-world data (28,29).

Image acquisition

Several ML programs developed have greatly assisted radiographers in positioning, repeat rate reduction of radiographic studies due to technique, image noise improvement and radiation dose reduction.

Gang *et al.* (30) established that ML algorithms are able to detect inadequate positioning as well as assist in automated positioning during radiographic examinations; with an accompanying 16% reduction in radiation exposure using automated positioning.

Siemens Healthineers introduced a new X-ray system known as Ysio X.Pree, the world's first intelligent X-ray system with integrated AI for optimizing daily routine of chest image acquisition in radiography; preparing for X-ray image acquisition using a live 3D camera. The AI-based algorithm automatically detects the thoracic region and sets the optimal acquisition area (collimation), focusing radiation only to the area of interest, with the goal of acquiring an image containing all necessary details at the least possible exposure. This system comprises AI functionality and other intelligent tools for image acquisition in what is known as myExam Companion (31).

Similar innovations can be seen in Carestream's AI-based software for CXR, comprising of smart positioning, smart technique and smart collimation. Smart positioning consists mainly of two RGBD (RGB & depth) cameras, pose-detection algorithm, a classifier, a hub, two controllers, a console PC (personal computer), markers and preparation areas. This system offers automatic bucky height adjustment and correct positioning check on shoulder-height, contact with bucky, pose, center alignment, tilt, orientation and hand position as seen in the console display. In addition, a video assist display providing patient with exam information along with a picture on how to position themselves next to the equipment. Smart technique uses an RGBD camera that captures patient information and applies AI algorithms to detect patient thickness and produce proper exposure technique in mobile or fixed tabletop DR machines devoid of Automatic exposure control (AEC). Smart collimation utilizes camera data on shoulder width and height during chest PA, and automatically adjusts collimator blades to the desired field for different patients (32).

According to GE HealthCare (33), ML algorithms embedded into the imaging system and operating in parallel to image acquisition have great potential to improve radiology image acquisition process by improving image quality and efficiency during radiologic examinations. Hence, Radiologic technologists can now select an anatomic part of the body, body size and projection from a list of presets and the anatomic specific algorithms allow for increase receptor sensitivity, dose reduction and optimization of image acquisition (34).

Furthermore, it is interesting to discover in the studies of Lee *et al.* (35) that noise reduction algorithms can effectively reduce radiation dose while same time maintaining image quality; Jin *et al.* (36) demonstrating two types: traditional denoising algorithm and deep learning methods based on neural network.

Elsewhere, Wuni *et al.* (37) in research asserted that AI has the capacity of impacting radiographers daily work (diagnostic radiography, 79.6% and therapeutic radiography, 88.9%) by standardizing some aspects of patient care and technical factors of radiography practice such as patient identification, image processing and dose optimization. However, there is a need for sufficient scientific evidence and proof, which at the moment is lacking.

Image interpretation

Since after the first attempt to establish computer aided detection (CAD) systems [comprising image preprocessing, extracting region of interest (ROI) regions, extracting ROI features and classifying disease according to the features] in the 1960s, it has gained popularity in radiology due to its ability to assist radiologist detect suspicious lesions often missed and improving the accuracy of their detection (38).

The earliest existed CAD system made use of the ML based methods, otherwise referred to as pixel-based method for the detection of lung cancer, with the feature extraction intuitive. In chest images for example, each pixel is assigned a corresponding anatomical structure such as lung, heart, mediastinum, diaphragm (segmentation); and "a classifier" [e.g., k-nearest neighbor (KNN), linear discriminant analysis (LDA), etc.] classifies each pixel based on various features inputted into it such as grayscale value of each pixel, spatial location information and texture statistical information (39).

During coronavirus disease 2019 (COVID-19), in his analysis built a ML model, utilizing support vector machine (SVM) classifier algorithm (one of the most widely used supervised ML approaches owing to its accuracy and less computational power demand); anticipated that such rapid computer-aided diagnostic approach boosting high performance metrics would be helpful in control of the pandemic. This SVM, trained using histogram of oriented gradients (HOG) descriptor aided the detection of COVID-19 in CXR, with a sensitivity and specificity of 97.92% and 98.91% respectively (40).

Findings from a recently conducted systematic review in 2023 on tuberculosis, one of the commonest infectious diseases revealed high potential of ML in detecting tuberculosis, with an average accuracy and sensitivity of 93.71% and 92.55% respectively. The Radiologist's report was utilized as reference standard in most included studies under review; SVM, KNN and random forest (RF) among the popular ML approaches employed (41).

Similar results were obtained in this recent study, a quadratic SVM model of ML automating the early detection process of pneumonia in CXRs with an accuracy of 97.58% and a smaller model classification time (42). Therefore, it can therefore be deduced that classical ML shows great performance on a small amount of data, iterating quickly and trying different techniques in a very short time (43). However, as radiological imaging data continued growing at a disproportionate rate compared with available trained readers, the main challenge was to determine in a robust way features for other chest pathologies aside lung cancer, tuberculosis and COVID-19; this then opened up new opportunities for advancing CAD system in medical imaging (44) leading us to deep learning.

Deep learning (DL)

DL is the most prevalent for detection, characterization and monitoring of diseases. DL does not require the process of feature extraction and disease classification as seen in the traditional CADs, but instead utilizes any neural networks' architecture with deep layers such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), Transformer, autoencoder (45-47). CNNs, designed to mimic the way human brain process information to make a final decision, comprises series of layers that successfully map image inputs to desired end points. The first layer (convolutional layer) detects and extracts features, the second layer (pooling layer) performs feature aggregation by selecting and reducing the number of features, and the third layer (fully connected layer) integrates all features extracted by the previous layer (47,48). A summarized workflow of DL-based chest Xray detection system as seen in a recent review conducted by Rehman *et al.* (49) are:

(I) Image acquisition: generating digital chest images, a primary requirement for model training.
(II) Pre-processing: techniques to improve quality of chest images e.g., image enhancement, ROI detection, bone suppression and edge detection.
(III) Feature extraction: distinctive, substantial, and concise information that aids in distinguishing one phenomenon from another.
(IV) Feature selection: selection of the most optimum and relevant features from the original features and removing irrelevant/noisy features to optimize model performance.
(V) Classification: categorizing a given set of data into classes.

With an increasing number of AI algorithms for triaging, detecting and classification purposes on benign and malignant CXR anomalies (50,51), DL has been integral to CXR analysis in several studies:

- ❖ For pulmonary opacities, pleural effusion, hilar prominence and enlarged cardiac silhouette classification (Qure AI) (52): no statistical difference between DL model used and test radiologist, AUC ranging from 0.837–0.929 and 0.693–0.923 respectively. The accuracy of DL in the evaluation of pulmonary and hilar abnormalities found to be limited in images with presence of chest wall implanted devices.
- ❖ For atelectasis, cardiomegaly, emphysema, hiatal hernia, pneumonia, pleural effusion, pulmonary

masses and nodules classification (CheXNeXt) (10): radiologist achieved statistically significant higher AUC than CheXNeXt on cardiomegaly, emphysema and hiatal hernia (0.888 *vs.* 0.831, 0.911 *vs.* 0.704, 0.985 *vs.* 0.851 respectively); CheXNeXt performed better than radiologist in detecting atelectasis (AUC of 0.862 *vs.* 0.808); no statistically significant difference in AUCs for other abnormalities.

- ❖ For CXR classification as either pulmonary tuberculosis or normal (AlexNet and GoogLeNet) (53): although AUCs for both pretrained models were greater than the untrained (P<0.001), best performance recorded when both models were taken as a whole (AUC of 0.99). Furthermore, a radiologist-augmented approach further improved accuracy (sensitivity and specificity of 97.3% and 100% respectively) in cases of disagreement among classifiers.
- ❖ For detection of malignant lung nodules on CXR (54): deep CNN (DCNN) software (built upon a modified version of ResNet-50) improved average sensitivity of radiologist in detecting malignant pulmonary nodules on CXR from 65.1% to 70.3% before and after DCNN use respectively, as well as reduction in number of false- positive marks per chest radiograph from an initial 0.20 to 0.18, P<0.001.
- ❖ For classification of relevant specific chest abnormalities that spans across pulmonary parenchyma diseases, pleural diseases and mediastinal diseases (55): DL algorithm had AUC of 0.95 for identification of chest radiographs with clinically relevant abnormalities; a sensitivity and specificity of 88.7% *vs.* 69.6% and 81.6% *vs.* 90.3% at high-sensitivity cut off and high-specificity cut off respectively. Also to point is the improved sensitivity of radiology residents after using DL algorithm (from 65.6–73.4%, P=0.003), although a slight reduction in specificity was recorded (from 98.1–94.3%, P<0.001).
- ❖ For examining the effect of DL algorithms on interpretation of CXR as normal or abnormal (56): a significant difference in performance between radiologist working with and without DL-based assisted technology observed (AUC 0.801, P<0.001). In addition, higher diagnostic sensitivity (68.21% *vs.* 61.13%, P<0.001), specificity (92.76% *vs.* 91.98%, P=0.577) and accuracy (76.37% *vs.* 71.39%, P<0.001) among experienced radiologist working with and without this technology.
- ❖ Trained DL models (ResNet34, ResNet50, VGG-19 and DenseNet169) for classification of CXR as either normal or pneumonia during the era of COVID-19 (57): these models exceeded 84% average accuracy on pneumonia which is promising. Best performing model was DenseNet169 (average classification accuracy of 95.72%), with reported classification accuracies of 97.97%, 96.62% and 92.57% for bacterial, viral and normal respectively.
- ❖ A systematic review to ascertain the application of DL algorithms on CXR analysis of pneumonia and COVID-19 (58): VGG (26.5%), ResNet (20.6%), MobileNet (14.7%) and DenseNet (14.7%) were among the most common CNN architectural models of DL used, with a mean accuracy of 93.62%, 92.13%, 95.16% and 84.25% respectively.
- ❖ DL techniques in CXR utilizing a multi-label classifier (DenseNet) to classify abnormalities based on 14 predefined labels that includes atelectasis, cardiomegaly, pleural effusion, consolidation, pneumonia, lung lesion, edema, lung opacities, pneumothorax (59): an approximated accuracy and AUC of 91% and 0.8 recorded for the 3 variations of DenseNet; This multi-label classifier in comparison with a benchmark improved classification of the 14 labels.
- ❖ DCNN for early detection or absence of Tuberculosis in CXR (60): the accuracy, sensitivity, specificity and AUC for the three adopted DCNN algorithms (ResNet, VGG and AlexNet) was given as follows: 96.73%, 95.50%, 98.05%, 0.9944 for ResNet; 94.96%, 94.20%, 95.78%, 0.9902 for VGG; 95.06%, 93.20% 97.08%, 0.9917 for AlexNet. While the general performance of all models was impressive (all above 94% in key metrics measurement), ResNet was found to outperform other models, showing excellent diagnostic ability amidst stratification analysis by sex, age and symptom.
- ❖ A hybrid DL approach to detecting pneumonia in CXR (61): this hybrid system (CNN architecture together with a ML classifier- SVM, KNN or RF) as oppose the traditional CNN architecture achieved a higher performance in accuracy, with short classification consumption time.
- ❖ A proposed C19D-Net model (DL techniques applying the inceptionV4 architecture and ML multiclass SVM classifier) in classifying COVID-19 and other kinds of pneumonia in CXR (62):

Table 2 Summary of the strengths, limitations, and gaps from all discussed literature across different AI domains in this review

S/N	Literature	AI domain	Strength(s)	Limitation(s)	Gap(s)
1.	Pierre *et al.* [2023]	ML & NLP	Extensive review on AI application to support and streamline daily radiology practice, engaging several studies	Issues of quality, safety and undeveloped sophisticated active algorithm monitoring mechanisms not covered	Impactful AI application in less investigated areas e.g., prognostic inference, assisted tumour grade classification, finding contraindications for imaging
2.	Langer [2002]	ML	A well-constructed model performing exam rescheduling, archiving studies and providing DICOM worklist	–	Additional work on the DICOM/HL7 interface broker for effective communication, and completion of the base DICOM service classes needed
3.	Nelson *et al.* [2019]	ML	A model built with open-source tools, estimated, and validated on conventional hardware	Model trained only on routinely collected administrative data	Developing complex models that reflects the multiplex interplay of patient, environmental and operational causal factors
4.	Chong *et al.* [2020]	ML	Empirical approach to developing state-of-the-art model, with moderate amount of data acquired from frontline information technology systems	–	A highly complex model that will report on quality improvement metrics sought
5.	Baltruschat *et al.* [2021]	ML	Development of a realistic clinical workflow simulator based on empirical data and use of state-of-the-art convolution neural network allowed for precise assessment	Open-i dataset upon which the CNN was trained included mainly out-patients in contrast to the predominantly stationary patient collective of the hospital	Smart worklist prioritization with more pathologies and different degree of manifestation
6.	Annarumma *et al.* [2019]	DL & NLP	The system offered real time prioritization	Clinical risk from delayed reporting of cases falsely classified as normal, high chance of AI performance exaggerated since findings grouped into categories, prioritization system takes account only image findings without its clinical context, absence of inpatient radiographs in simulation	More studies to reduce false negative rate to minimum as well as misclassification rate. Also, further work in developing a multiresolution architecture that allows optimal image sizes to be selected automatically
7.	Nabulsi *et al.* [2021]	DL	Extensive evaluation on models' generalizability to multiple datasets, different countries and population groups, and unseen diseases	Wide range of CXR abnormalities not represented, only CXR viewed without referencing additional clinical or patent data, absence of historical reporting timing information	Prospective studies to consider various degrees of urgency for different diseases and provide patient's clinical information so the true effect will be determined
8.	Gang *et al.* [2021]	ML	Adopted AI based automatic patient positioning technique reduced cross-infection risks between patients and medical workers	Confounding bias due to relatively small sample size. Also, scan protocols not fully optimized and iterative reconstruction algorithms not used due to confusion during the pandemic	–
9.	Siemens Healthineers [2020]	ML	Easy user interface boosting AI functionality and intelligent tools	Lack of documented empirical evidence	–
10.	Sun [2021]	ML	System comprising several smart and Intelligent tools	Lack of documented empirical evidence	–
11.	GE HealthCare [2023]	ML	Seamless integration of AI solutions in the device	Lack of documented empirical evidence	–
12.	Lee *et al.* [2020]	ML	Applied advanced algorithm that sustained noise suppression performance without degrading visual content	Relatively small amounts of subjects, study population did not have diverse abnormal findings	System compatibility, detector type and prerequisites for raw data be investigated
13.	Wuni, Botwe and Akudjedu [2021]	ML	Opportunities, challenges, and way forward of AI explored	Limited to Ghana	More research on AI strategies by national societies and regulatory bodies to harmonise the implementation efforts
14.	Erdaw and Tachbele [2021]	ML	Easy integration of algorithm into the clinical system	Retrospective images used	Study DL on prospective images, clinical and sociodemographic data from COVID patients. Also, AI application in predicting prognosis and treatment outcomes of patients
15.	Hansun *et al.* [2023]	ML	A thorough systematic approach	Only three databases were used, data volume and quality concerns	Proper data curation, transfer learning and multimodal approaches required
16.	Barakat, Awad and Abu-Nabah [2023]	ML	Practical and less computationally expensive novel approach	Limited to the binary classification of pneumonia using pediatric X-rays	Studies refining optimal feature-extraction scheme
17.	Singh *et al.* [2018]	DL	Unbiased selection of CXR in the study as none of the test radiologist were involved in the selection process	Pre-hoc power analysis (to determine the number of CXR and test radiologists required to assess the DL algorithm) performance) not performed. Also is the combined evaluation of different types of pulmonary opacities rather than as separate categories	Published guidelines on the most appropriate cut-off values for DL algorithm and how such deviations would affect the performance of DL
18.	Rajpurkar *et al.* [2018]	DL	DL algorithm internally validated after rigorous training. Concurrent detection of up to 14 CXR pathologies	Neither the DL algorithm nor the radiologists were permitted to use patient history or review prior examinations. Evaluation limited to a dataset from a single institution	Determine the feasibility of outcomes in a prospective clinical setting. Also, the need to address generalizability of these algorithms to datasets from other institutions
19.	Lakhani and Sundaram [2017]	DL	Augmentation with multiple preprocessing techniques further improved accuracy	Limited to only tuberculosis. Only PA CXR images used	–

Table 2 (*continued*)

Table 2 (*continued*)

S/N	Literature	AI domain	Strength(s)	Limitation(s)	Gap(s)
20.	Sim *et al.* [2020]	DL	High generalizability as image selection and review not limited to an institution or geographic region	Spectrum bias due to omission of images with benign nodules or ambiguous aspects. No set time interval between stand alone and software aided sessions leading to possible recall bias	Further studies with a crossover design to reduce the effects of demographic and radiologist factors
21.	Hwang *et al.* [2019]	DL	DL algorithm application in a clinical setting	Performed at a single institution, retrospective nature of study design, algorithm performance comparison with on-call radiology residents instead of experienced radiologist, only PA CXR used	Prospective studies to confirm if algorithm use can improve clinical workflow and patient outcomes
22.	Kim *et al.* [2021]	DL	Prospective study, cost benefit of algorithm	Simulation-based trial, selection bias, a small target range	More research using an algorithm with a broader target range, evaluating the effectiveness of DLCR on patient outcomes in the real-world setting
23.	Hammoudi *et al.* [2021]	DL	Experimental nature of study	–	Future works on models to discern between COVID-19 viral and non-COVID-19 viral pneumonia
24.	Meedeniya *et al.* [2022]	DL	A well conducted systematic review highlighting available database, trends, challenges and future research directions	Limited to pneumonia and COVID-19 conditions	Reviews on current state-of-the-art solutions
25.	Monshi, Poon and Chung [2022]	DL	Multilabel classifier with antialiasing blur pooling and parallel training	–	Investigating the use of DICOM images in detecting diseases with small and complex structure
26.	Nijiati *et al.* [2022]	DL	System was effective even without external clinical information assistance	Possibility of wrong labels for chest radiographs, limited generalizability to pediatric cases, limited study population to 15 years and above	–
27.	Masad *et al.* [2021]	DL & ML	Novel system achieved efficient performance with short classification consumption time	–	–
28.	Kaur *et al.* [2021]	DL & ML	Study supported by strong empirical evidence	Hardware restrictions	Training proposed model on larger image sets and comparing its performance to wider existing methods
29.	Fati, Senan and ElHakim [2022]	DL & ML	Compared a two-part hybrid technology with an ANN	Lack of images in the tuberculosis dataset	Extracting deep features using CNN models and integrating them into feature vectors
30.	Ahn *et al.* [2022]	DL	Used data from more than one source, study met requirements for the health insurance portability and accountability act guidelines	Only PA CXR images, power analysis not conducted to determine adequacy of sample size and number of test readers	Effect of AI-aided or unaided interpretation on detection of the non-target findings, assessing the real-world clinical chest radiograph interpretation workflow
31.	Liu *et al.* [2013]	NLP	A large sample of chest radiograph reports evaluated	CXR reports drawn from a single healthcare delivery system, tools developed to analyse reports in a retrospective, rather than a real-time setting	Real-time report indexing and querying to support the use of the tool under study at point of bedside care
32.	Xue *et al.* [2018]	NLP & DL	Experimental in nature, incorporating CNN with LSTM recurrently	Training model on a small training set	A new training strategy/evaluation metric on a larger and better dataset taking both word accuracy and grammar correctness into account
33.	Towfighi *et al.* [2019]	NLP & ML	Study utilized open-source code, free for personal or community use	Limited size of sample and frequency of positive findings	–
34.	Olthof, van Ooijen and Cornelissen [2021]	NLP & DL	Study varied prevalence within the dataset suggesting a relationship between performance and prevalence	Absence of inter-rater agreement assessment, study limited to an institution	Studies that consider variation in report size within the datasets, external validation of the model under study
35.	Yi, Kim and Lin [2022]	NLP & DL	A comparative study comprising NLP derived disease labels and radiologist review of images	–	–
36.	Bressem *et al.* [2021]	NLP & DL	Approach presented enabled use of a BERT model pre-trained in the respective language to learn the domain-specific words	Memory limitation of the hardware accelerator used	Future studies investigating an approach that allows for integration of texts
37.	Zhang *et al.* [2023]	ML	Evaluation of consecutively enrolled individuals in the clinical practice setting	Data on ethnicity and patient demographic characteristics beyond age and sex not included, study limited to a country	Future work to study the generalizability of this system in different geographic settings
38.	Siemens Healthineers [2023]	Robotics	Boosts the new true2scale body scan feature	System particularly suited for orthopedic and trauma cases, lack of sound empirical evidence	–
39.	Ajani *et al.* [2023]	Robotics & DL	Model successfully implemented on social robot as an assistive platform for radiologist, guaranteeing clinical usability	–	–

Table 2 (*continued*)

Table 2 (*continued*)

S/N	Literature	AI domain	Strength(s)	Limitation(s)	Gap(s)
40.	Chapman and Haug [1999]	Expert system	Well demonstrated comparative study between computerized techniques and physicians	Challenges creating a gold standard in a field where language (e.g., opacities it describes) is vague, hazy, and ill-define	A test to determine if the uncorrected output of the parser combines with the expert system to provide an adequate assessment of support for pneumonia in CXR reports
41.	Hassen *et al.* [2013]	Fuzzy logic	Combination of segmentation and recognition approaches, using spatial relations, database contains images from different institutions allowing for generalizability	Only PA images used	–
42.	Torres *et al.* [2014]	Fuzzy logic	Robust approach, using classical morphology operations to segment lungs thereby providing low computational complexity	–	–
43.	Suttitanawat *et al.* [2018]	Fuzzy logic	Experimental in design, novel algorithm capable of performing localization tasks	–	–
44.	Zhang *et al.* [2022]	Fuzzy logic	Adoption of a class compactness graph during manifold learning to address overfitting issues	Only Euclidean distance used during construction of the class compactness graph, lack of multi-center based external validation	–
45.	Sahin, Akdogan and Aktan [2023]	Fuzzy logic	Model comprised of three fuzzy units to further improve diagnosis	–	–
46.	Ieracitano *et al.* [2022]	Fuzzy logic & DL	Portable CXR utilized here allowed for the possibility of making recordings directly at home	Smal size of dataset	Study validation using larger and linked datasets, integration of recent facemask detection method for COVID-19 prevention and control in public
47.	Yadlapalli and Dokku [2023]	Fuzzy logic & DL	Transfer learning approach very successful	–	–

S/N, serial number; AI, artificial intelligence; ML, machine learning; NLP, natural language processing; DICOM, digital information and communications in medicine; CNN, convolutional neural network; DL, deep learning; CXR, chest X-ray; COVID-19, coronavirus disease 2019; PA, posteroanterior; DLCR, deep learning-based assistive technology on chest radiograph interpretation; LSTM, long short-term memory; BERT, bidirectional encoder representations from transformers.

achieved the highest COVID-19 detection performance accuracy of 96.24% for 4 classes-normal COVID-19, bacterial pneumonia and viral pneumonia.

- ❖ Deep and hybrid learning technique for tuberculosis detection in CXR (63): in the first approach, hybridizing two CNN models (ResNet-50 and GoogLeNet) with ML SVM classifier achieved superior results in detecting tuberculosis. In the second approach, application of artificial neural network (ANN) + ResNet-50 + Gray level co-occurrence matrix (GLCM) + discrete wavelet transform (DWT) + local binary pattern algorithms (LBP) further improved performance (accuracy, sensitivity, specificity and AUC) of over 90% in both datasets under study.
- ❖ Established in a staged reading session (with and without AI) improved readers performance of CXR using AI (64); but the superior sensitivity of AI-based application according to (65) seems to be accompanied with higher false-detection-rates, and this appears worrying. However, with these studies mostly retrospective, it is important to note that DL is still very much in the early stages, with many uses still theoretical in development or limited to a location or single institution (15). Further studies required to validate DL use in real-world setting for CXR interpretation before implementation in routine clinical practice.

Natural language processing (NLP)

NLP is concerned with ability of computers to understand texts and spoken words just as humans; combining computational linguistics, rule-based modeling of human language, ML and DL models, and working together to process human language in the form of text or voice data (66). Certain radiology reports contain imaging findings and diagnosis of radiologists in an unstructured natural text form (67), and this cannot be processed in CNN (where image classification relies on supervised training that is rooted on expert annotation), although other DL techniques can e.g., RNNs (68). NLP has the ability to recognize semantics (meaning) and context and generate medical reports. It can further assess spatial information in radiology reports, segmentation of report, detect actionable findings and produce image annotation (69-71).

In an earlier study (72), developed an NLP-based software package that supported extraction of semantic information from large data collections, and applied to CXR reports to automatically identify pneumonia among intensive care unit patients. A lexicon was formed to categorize pneumonia related terms and uncertainty profiles, assigning interpretations (positive or negative) according to each report's query profile. This NLP algorithm demonstrated an impressive 92.7% sensitivity, 91.1% specificity, 93.3% positive predictive value and 90.3% negative predictive value.

Years later, Xue *et al.* presented a multimodal recurrent NLP model incorporating CNNs with long short-term memory (LSTM) and applied it to a CXR public data set to generate the imaging description parameters and impression sentences of CXR reports (73). It was found to produce high-level conclusive impressions as well as detailed descriptive findings to support the conclusion, maintaining coherency among generated sentences.

Towfighi *et al.* (74) developed an approach comprising of ML and NLP to identify presence of opacity, absence of opacity, follow-up report, and presence of endotracheal tube in CXR. The entire model was trained from 1,000 retrieved plain film CXR reports and classified according to the above mentioned 4 labels. An impressive accuracy, precision, recall and AUC of 0.84, 0.94, 0.81 and 0.86 respectively was observed in the model responsible for identifying cases without opacity. More results included a low precision value of 0.38 seen in the follow-up label model, an accurate classification of the endotracheal tube model recorded in the only case of intubation.

Olthof *et al.* (75) went on to introduce BERT (bidirectional encoder representations from transformers), a transformer-based language model built on learning the contextual relationships between words in text; this produced a superior performance [compared to other model architectures such as fully connected neural network (Dense), a LSTM recurrent neural network and a CNN] during a single-label classification task involving radiologist-annotated chest radiographs dataset. In their more recent search in 2022, reiterated the promising role of transformer-based language models in accurate text classification. In the findings of this retrospective study, demonstrated annotated datasets of radiology request and reports used to train, test and even compared five newly developed transformer-based NLP models (BERTje, RobBERT, BERT-multilingual, BERT-clinical and BERT-base) for multilabel classification in chest imaging. The RobBERT model performed best, producing AUC values ranging from 0.808 for chest imaging request and 0.746 for report items, with an AUC value of 0.95 for classification of normal reports (76).

It is interesting to see a slight deviation in this comparative study assessing the level of agreement between NLP and radiologist-curated labels for possible Tuberculosis detection on CXR, utilizing 2 approaches (77): NLP-derived disease labels and radiologist-review of images. Findings revealed a poor level of agreement between NLP and radiologist-derived findings for possible cases of tuberculosis on CXR, with a kappa coefficient of 0.34.

A DL natural language BERT model pre-trained on 3.8 million text reports yielded highly accurate classification of CXR, AUC of 0.98, 0.97, 0.97 and 0.99 for congestion, effusion, consolidation and pneumothorax respectively (78). This result surpassed the accuracy of previous approaches, with comparatively little annotation effort.

Zhang *et al.* (67) recently generated a BERT model similar to that of (75,76), built on a transformer mechanism that learns the contextual relationships between words in text, and is able to identify language entities and span, as well as semantic type of entities and semantic relationships between language entities. In an independent prospective test, this model produced significantly shorter reporting time (283 seconds) and highest similarity to final reports from radiologist [0.69 (0.24) mean/standard deviation (SD) bilingual evaluation score] in comparison with normal template [387 seconds; 0.37 (0.09)] and rule-based model [296 seconds; 0.37 (0.09)]; maintaining high level consistency with radiologist reports on a 23-label system of abnormalities (in the lungs, mediastinum, pleura and thorax). In spite of these, variation in radiology reports and the highly challenging task of creating labels for each image seems to limit its routine use in clinical practice. Recent research advances propose NLP use in conjunction with DL, yet scarcity of studies exist.

Robotics

refers to a system where robots are built and programmed to perform specific duties without further human intervention (79). Robotics has grown in heaps and bounds improving healthcare services, robotic-assisted radiography involving the use of robots to aid image acquisition process; this includes X-ray imaging where positioning the source and detector is done by robots, ultrasound imaging where the probe is held by a robot, endoscopy where robots enable the controlled trajectory of the imaging system for increased aperture and volumetric/tomographic imaging as well as track medical instrumentation, etc. (80,81).

Multitom Rax, an innovation from Siemens represents a significant improvement beyond traditional X-ray (82). This robotic X-ray system, utilizing AI allows for an unparalleled positioning flexibility and unique automated workflows around the patient thereby guarantying expansion in precision and patient experience. AI features "true2scale body scan" and "Real 3D" enable full-body images obtained [standing, sitting or supine at ALARA (as low as reasonably achievable) dose] and offer more insights in improving diagnostic confidence respectively.

One of the earliest applications of robotics in CXR was during the COVID-19 pandemic, a robotic framework proposed and given the task of classifying positive and negative COVID-19 patients based on CXR (83). This work consisted of two key parts: Transfer learning, utilizing learning models (GoogleNet and SqueezeNet) to screen positive and negative cases in CXR and CT images; and most importantly explainability of the model's decision, demonstrated using Class Activation Mapping (CAM) and Gradient-weighted Class Activation Mapping (Grad-CAM). A test accuracy, sensitivity and specificity of 90.90%, 94.70% and 87.20% were obtained for SqueezeNet whereas 96.40%, 95.50% and 97.40% for GoogleNet respectively. Transfer learning according to Reedy (84) refers to a research problem in ML where knowledge obtained while solving a problem is stored and used to solve another related problem. This "decision making" process represents a vital aspect and is of clinical relevance as previously reviewed DL contributions to robotics setup unable to provide this. A major reason for such is likely the high computational demands of these models compared to low computational capabilities of clinical devices, and this is yet to be addressed. Extensive creative research by way of creating models, testing and analysis of numerical data is expected in this regard.

Expert system (ES)

ES is a computer program that learns and attempts to reciprocate judgment and decision-making ability of humans using AI technologies, with the intention to offer suggestions and complement rather than replace human experts (85). The main components of an ES are knowledge base, inference engine, user interface. ES works by accumulating experience and facts on a knowledge base, integrate them with an inference or rules engine and provides an answer to the problem; it relies solely on having a good knowledge base (86).

ES application in CXR was originally reported in a comparative study identifying CXR reports that supports acute bacterial pneumonia (87), the performance of two computerized techniques constructed from expert knowledge (ES) matched to two computerized techniques that learn rules and structure from data (ML). It was discovered that ML performed in same capacity as ES and physician, all three techniques performing better than a baseline keyword search.

However, as knowledge base increased following increasing data, resulted in a proportionate increase in the processing complexity of ES, subjecting these systems to many computational problems. The expectation of the inference engine to process huge number of rules to reach a decision thrown into serious jeopardy given the challenge in verifying that decision rules are consistent with each other amidst many rules (88,89). Furthermore, overfitting and overgeneralization effects when using known facts to generalize other cases not well described in the knowledge base (90), as well as lack of a quick and efficient update of the knowledge base continue to persist (91,92). Hence it has become clear that new approaches to AI are lacking; and this is required as oppose rule-based technologies to improve health outcomes.

Fuzzy logic (FL)

FL is defined as generalization of classical logic; offering mechanism of approximation (approximate reasoning based on abundance of data) and inference (decision making) in situations of partial truth, i.e., times where decision as to whether true or false cannot be made (93). FL algorithms work on the principle of solving a problem after considering all available data.

In a paper presented by Hassen *et al.* (94), proposed spatial relation integration (represented by fuzzy subsets of the image space) in the process of segmentation of CXR; spatial relations of great help when finding contours of poorly contrasted objects or ill-defined boundaries. This automatic approach significantly performed with high accuracy for all lung structures, a strong level of agreement between automatic and manual chest radiography segmentations noted; although the recognition rate for left pericardiac was low compared with other structures.

In the research conducted by Torres *et al.* (95), produced a robust fuzzy classifier to detect cardiomegaly in chest radiographs. This method uses classical morphology operations to segment the lungs and provides a fast computation of the CTR; a 93.85% and 100% value for sensitivity and specificity respectively recorded for cardiomegaly.

In detecting one of the deadly cancers (lung nodules) in CXR images, an interval type-2 FL system was proposed by Suttitanawat *et al.* (96) based on a novel lung nodule detection algorithm. This system impressively detected per image probable locations of lung nodules, with a 0.82 true positive and 13.11 false positive rate, utilizing 4 features: D-descriptors, mean intensity of inside boundary, circularity ratio and HH (high frequency) diagonal component from wavelet transform.

It is surprising to see in the studies of Zhang *et al.* (97), an established interpretable TSK (Takagi-Sugeno-Kang) fuzzy system for COVID-19 detection using radiomics features extracted from CXR images. This technique, involving binary label matrix of training samples (with the assumption that the samples are in same class and kept in close proximity) and an after what transformation into the label space achieved a classification accuracy of over 83%, better than modern models and maintaining high interpretability.

Similar experience in COVID-19 detection was witnessed in the research conducted by Şahin *et al.* (98), a type-2 FL-based model comprising of three fuzzy units developed to aid diagnosis. The first fuzzy unit produced COVID-19 positivity as a percentage of respiratory rate, loss of smell and body temperature values; the second according to C-reactive protein, lymphocyte and D-dimer values; the third according to clinical examination and blood analysis (Third fuzzy unit simply represents outputs of the first and second fuzzy units). Under extensive evaluation, this system detected COVID-19 with 86.6% accuracy.

In a bid to address the challenges in the research of Hassen *et al.* (94), Ieracitano *et al.* (99) employed a hybrid approach, developing a FL-based DL model known as CovNNet (a neural network) to characterize COVID-19 related pneumonia and no COVID-19 pneumonia, based on extracting relevant features from portable CXR images combined and fuzzy images generated by fuzzy edge detection algorithm. The reason for the introduction of the fuzzy edge detection procedure (alongside DL, a hybrid approach) was to tackle issues such as vagueness, ambiguities and uncertainties in lung edges usually present in portable CXR images. Results showed that a combination of CXR and fuzzy features embedded within a DL approach produced a higher classification performance (81% accuracy) compared to conventional DL approaches.

An even more successful approach to classifying COVID, viral and bacterial pneumonia in CXR utilizing FL and DL was seen in the recent studies of Yadlapalli *et al.* (100). Here, ResNet 18 model attained a 97% classification accuracy, 96% precision, and 98% recall (in the case of COVID-19 detection using CXRs images). Another significant finding is the maximum sensitivity ratio of 97.1% and a 97.47% F1-score rate, the highest compared to previous techniques. However, it can be deduced that the accuracy of FL systems is usually compromised as the system works mostly on inaccurate inputs/imprecise data; thus, results are not always widely accepted. Also is the fact that there is no single systematic approach to solving problems, making it confusing as there are many solutions arising for a particular problem. A hybrid approach that involves a combination of FL systems and artificial neural networks enhances efficiency and tremendously improves overall performance (93,101).

Conclusions

A chronological narrative of AI, its domains, and applications to CXR have been well demonstrated by way of comprehensive assessment of key findings in a range of studies. Synthesis of evidence was done, highlighting the strengths, limitations, and gaps in knowledge. In a nutshell, results suggest that ML can be very useful in certain administrative tasks in CXRs such as exam scheduling and worklist management, as well as the image acquisition process. DL showed very promising signs as regards image interpretation, although a hybrid approach may offer more insight. It is highly recommended that more in-depth research be conducted to address limitations, justify implementation, and tackle policy issues.

Acknowledgments

Funding: None.

Footnote

Reporting Checklist: The authors have completed the Narrative Review reporting checklist. Available at https://jmai.amegroups.com/article/view/10.21037/jmai-24-67/rc

Peer Review File: Available at https://jmai.amegroups.com/article/view/10.21037/jmai-24-67/prf

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at https://jmai.amegroups.com/article/view/10.21037/jmai-24-67/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Speets AM, van der Graaf Y, Hoes AW, et al. Chest radiography in general practice: indications, diagnostic yield and consequences for patient management. Br J Gen Pract 2006;56:574-8.
2. Guena MN, Alapha FZ. Pertinence of Children's Chest X-ray Request Form and Practice at the Regional Hospital of Ngaoundere Cameroon. Open Journal of Radiology 2018;8:223-35.
3. Broder J. Diagnostic Imaging for the Emergency Physician. Saunders, 2011 pp. 185-296.
4. Robin Smithuis and Otto van Delden. Chest X-Ray- Basic Interpretation. Available online: https://radiologyassistant.nl/chest/chest-x-ray/basic-interpretation. (Accessed: 17 December 2023)
5. Dasegowda G, Kalra MK, Abi-Ghanem AS, et al. Suboptimal Chest Radiography and Artificial Intelligence: The Problem and the Solution. Diagnostics (Basel) 2023;13:412.
6. Kaviani P, Digumarthy SR, Bizzo BC, et al. Performance of a Chest Radiography AI Algorithm for Detection of Missed or Mislabeled Findings: A Multicenter Study. Diagnostics (Basel) 2022;12:2086.
7. Irmici G, Cè M, Caloro E, et al. Chest X-ray in Emergency

Radiology: What Artificial Intelligence Applications Are Available?. Diagnostics (Basel) 2023;13:216.
8. Klein JS, Rosado-de-Christenson ML. A Systematic Approach to Chest Radiographic Analysis. In: Hodler J, Kubik-Huch RA, von Schulthess GK, eds. Diseases of the Chest, Breast, Heart and Vessels 2019-2022: Diagnostic and Interventional Imaging. Cham (CH): Springer; February 20, 2019:1-16.
9. Calandriello L, Walsh SLF. Artificial intelligence for thoracic radiology: from research tool to clinical practice. Eur Respir J 2021;57:2100625.
10. Rajpurkar P, Irvin J, Ball RL, et al. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLoS Med 2018;15:e1002686.
11. Pouraliakbar H. Chest Radiography in Cardiovascular Disease. In: Majid M, Azin A, Majid H, (eds). Practical Cardiology. Elsevier, pp. 2018: 113-130.
12. Fasterholdt I, Naghavi-Behzad M, Rasmussen BSB, et al. Value assessment of artificial intelligence in medical imaging: a scoping review. BMC Med Imaging 2022;22:187.
13. Julie, S. (2021) Using Artificial Intelligence to Improve Radiology Workflow. Available online: https://www.thedoctors.com/articles/using-artificial-intelligence-to-improve-radiology-workflow/ (Assessed: 25 December 2023)
14. Oracle (2023). What is Machine learning? Available online: https://www.oracle.com/ng/artificial-intelligence/machine-learning/what-is-machine-learning/ (Assessed: 29 December 2023)
15. Kapoor N, Lacson R, Khorasani R. Workflow Applications of Artificial Intelligence in Radiology and an Overview of Available Tools. J Am Coll Radiol 2020;17:1363-70.
16. Ranschaert E, Topff L, Pianykh O. Optimization of Radiology Workflow with Artificial Intelligence. Radiol Clin North Am 2021;59:955-66.
17. Pierre K, Haneberg AG, Kwak S, et al. Applications of Artificial Intelligence in the Radiology Roundtrip: Process Streamlining, Workflow Optimization, and Beyond. Semin Roentgenol 2023;58:158-69.
18. Langer S. OpenRIMS: an open architecture radiology informatics management system. J Digit Imaging 2002;15:91-7.
19. Dlama J. The Application Artificial Intelligence (AI) in Radiation Protection and Medical Imaging: How prepared are Radiographers. Journal of Medical Imaging and Radiation Sciences 2022;53:61-5.
20. Curtis C, Liu C, Bollerman TJ, et al. Machine Learning for Predicting Patient Wait Times and Appointment Delays. J Am Coll Radiol 2018;15:1310-6.
21. Mieloszyk RJ, Rosenbaum JI, Hall CS, et al. Environmental Factors Predictive of No-Show Visits in Radiology: Observations of Three Million Outpatient Imaging Visits Over 16 Years. J Am Coll Radiol 2019;16:554-9.
22. Nelson A, Herron D, Rees G, et al. Predicting scheduled hospital attendance with artificial intelligence. NPJ Digit Med 2019;2:26.
23. Chong LR, Tsai KT, Lee LL, et al. Artificial Intelligence Predictive Analytics in the Management of Outpatient MRI Appointment No-Shows. AJR Am J Roentgenol 2020;215:1155-62.
24. Holdship J, Dhanoa H, Hopper A, et al. Prediction of Hospital Outpatient Attendance in UK Hospitals: A Retrospective Study Applying Machine Learning to Routinely Collected Data for Patients of All Ages. MedRxiv preprint, 2022pp. 2-10.
25. Baltruschat I, Steinmeister L, Nickisch H, et al. Smart chest X-ray worklist prioritization using artificial intelligence: a clinical workflow simulation. Eur Radiol 2021;31:3837-45.
26. Annarumma M, Withey SJ, Bakewell RJ, et al. Automated Triaging of Adult Chest Radiographs with Deep Artificial Neural Networks. Radiology 2019;291:272.
27. Nabulsi Z, Sellergren A, Jamshy S, et al. Deep learning for distinguishing normal versus abnormal chest radiographs and generalization to two unseen diseases tuberculosis and COVID-19. Sci Rep 2021;11:15523.
28. Kelly CJ, Karthikesalingam A, Suleyman M, et al. Key challenges for delivering clinical impact with artificial intelligence. BMC Med 2019;17:195.
29. Nguyen NH, Nguyen HQ, Nguyen NT, et al. Deployment and validation of an AI system for detecting abnormal chest radiographs in clinical settings. Front Digit Health 2022;4:890759.
30. Gang Y, Chen X, Li H, et al. A comparison between manual and artificial intelligence-based automatic positioning in CT imaging for COVID-19 patients. Eur Radiol 2021;31:6049-58.
31. Siemens Healthineers (2020) Siemens Healthineers uses artificial intelligence to take X-ray diagnostics to a new level. Available online: https://www.siemens-healthineers.com/press/releases/ysioxpree-ai-chest.html (Assessed: 28 December 2023)
32. Sun L. (2021) Applying AI in Radiology to optimize

workflow for Chest X-rays. Available online: https://www.carestream.com/blog/2021/09/13/applying-ai-in-radiology-to-optimize-workflow/ (Assessed: 27 December 2023)
33. GE HealthCare (2023) AI in X-ray: Workflow solutions to improve care delivery. Available online: https://www.gehealthcare.com/insights/article/ai-in-xray-workflow-solutions-to-improve-care-delivery (Assessed: 8 December 2023)
34. Wilson DU, Bailey MQ, Craig J. The role of artificial intelligence in clinical imaging and workflows. Vet Radiol Ultrasound 2022;63 Suppl 1:897-902.
35. Lee W, Lee S, Chong S, et al. Radiation dose reduction and improvement of image quality in digital chest radiography by new spatial noise reduction algorithm. PLoS One 2020;15:e0228609.
36. Jin Y, Jiang XB, Wei Z, et al. Chest X-ray image denoising method based on deep convolution neural network. IET Image Process 2019;13:1970-8.
37. Wuni AR, Botwe BO, Akudjedu TN. Impact of artificial intelligence on clinical radiography practice: Futuristic prospects in a low resource setting. Radiography (Lond) 2021;27 Suppl 1:S69-73.
38. Qin C, Yao D, Shi Y, Song Z. Computer-aided detection in chest radiography based on artificial intelligence: a survey. Biomed Eng Online 2018;17:113.
39. McNitt-Gray MF, Saure JW, Huang HK, et al. Pattern classification approach to segmentation of chest radiographs. Proc. SPIE 1898, Medical Imaging 1993: Image Processing.
40. Erdaw Y, Tachbele E. Machine Learning Model Applied on Chest X-Ray Images Enables Automatic Detection of COVID-19 Cases with High Accuracy. Int J Gen Med 2021;14:4923-31.
41. Hansun S, Argha A, Liaw ST, et al. Machine and Deep Learning for Tuberculosis Detection on Chest X-Rays: Systematic Literature Review. J Med Internet Res 2023;25:e43154.
42. Barakat N, Awad M, Abu-Nabah BA. A machine learning approach on chest X-rays for pediatric pneumonia detection. Digit Health 2023;9:20552076231180008.
43. Sethy PK, Behera SK, Ratha PK, et al. Detection of Coronavirus disease (COVID-19) based on deep features and support vector machine. Int J Math Eng Manag Sci 2020;5:643-51.
44. Hosny A, Parmar C, Quackenbush J, et al. Artificial intelligence in radiology. Nat Rev Cancer 2018;18:500-10.
45. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436-44.
46. Miotto R, Wang F, Wang S, et al. Deep learning for healthcare: review, opportunities and challenges. Brief Bioinform 2018;19:1236-46.
47. Santos MK, Ferreira Júnior JR, Wada DT, et al. Artificial intelligence, machine learning, computer-aided diagnosis, and radiomics: advances in imaging towards to precision medicine. Radiol Bras 2019;52:387-96.
48. Serena Low WC, Chuah JH, Tee CATH, et al. An Overview of Deep Learning Techniques on Chest X-Ray and CT Scan Identification of COVID-19. Comput Math Methods Med 2021;2021:5528144.
49. Rehman A, Khan A, Fatima G, et al. Review on chest pathogies detection systems using deep learning techniques. Artif Intell Rev 2023;56:12607-53.
50. Li B, Kang G, Cheng K, et al. Attention-Guided Convolutional Neural Network for Detecting Pneumonia on Chest X-Rays. Annu Int Conf IEEE Eng Med Biol Soc 2019;2019:4851-4.
51. Wu JT, Wong KCL, Gur Y, et al. Comparison of Chest Radiograph Interpretations by Artificial Intelligence Algorithm vs Radiology Residents. JAMA Netw Open 2020;3:e2022779.
52. Singh R, Kalra MK, Nitiwarangkul C, et al. Deep learning in chest radiography: Detection of findings and presence of change. PLoS One 2018;13:e0204155.
53. Lakhani P, Sundaram B. Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. Radiology 2017;284:574-82.
54. Sim Y, Chung MJ, Kotter E, et al. Deep Convolutional Neural Network-based Software Improves Radiologist Detection of Malignant Lung Nodules on Chest Radiographs. Radiology 2020;294:199-209.
55. Hwang EJ, Nam JG, Lim WH, et al. Deep Learning for Chest Radiograph Diagnosis in the Emergency Department. Radiology 2019;293:573-80.
56. Kim JH, Han SG, Cho A, et al. Effect of deep learning-based assistive technology use on chest radiograph interpretation by emergency department physicians: a prospective interventional simulation-based study. BMC Med Inform Decis Mak 2021;21:311.
57. Hammoudi K, Benhabiles H, Melkemi M, et al. Deep Learning on Chest X-ray Images to Detect and Evaluate Pneumonia Cases at the Era of COVID-19. J Med Syst 2021;45:75.
58. Meedeniya D, Kumarasinghe H, Kolonne S, et al. Chest X-ray analysis empowered with deep learning: A systematic

review. Appl Soft Comput 2022;126:109319.
59. Monshi MMA, Poon J, Chung V. Distributed Deep Learning for Multi-Label Chest Radiography Classification. In: Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications 2022;4:949-56.
60. Nijiati M, Ma J, Hu C, et al. Artificial Intelligence Assisting the Early Detection of Active Pulmonary Tuberculosis From Chest X-Rays: A Population-Based Study. Front Mol Biosci 2022;9:874475.
61. Masad IS, Alqudah A, Alqudah AM, et al. A hybrid deep learning approach towards building an intelligent system for pneumonia detection in chest X-ray images. International Journal of Electrical and Computer Engineering 2021;11:5530-40.
62. Kaur P, Harnal S, Tiwari R, et al. A Hybrid Convolutional Neural Network Model for Diagnosis of COVID-19 Using Chest X-ray Images. Int J Environ Res Public Health 2021;18:12191.
63. Fati SM, Senan EM, ElHakim N. Deep and hybrid learning technique for early detection of tuberculosis based on X-ray images using feature fusion. Appl Sci 2022;12:7092.
64. Ahn JS, Ebrahimian S, McDermott S, et al. Association of Artificial Intelligence-Aided Chest Radiograph Interpretation With Reader Performance and Efficiency. JAMA Netw Open 2022;5:e2229289.
65. Niehoff JH, Kalaitzidis J, Kroeger JR, et al. Evaluation of the clinical performance of an AI-based application for the automated analysis of chest X-rays. Sci Rep 2023;13:3680.
66. IBM (2023). What is natural language processing (NLP)? Available online: https://www.ibm.com/topics/natural-language-processing#:~:text=Natural%20language%20processing%20(NLP)%20refers,same%20way%20human%20beings%20can. (Assessed: 2 December 2023)
67. Zhang Y, Liu M, Zhang L, et al. Comparison of Chest Radiograph Captions Based on Natural Language Processing vs Completed by Radiologists. JAMA Netw Open 2023;6:e2255113.
68. Jiang B, Zhang Y, Zhang L, et al. Human-recognizable CT image features of subsolid lung nodules associated with diagnosis and classification by convolutional neural networks. Eur Radiol 2021;31:7303-15.
69. Kuling G, Curpen B, Martel AL. BI-RADS BERT and Using Section Segmentation to Understand Radiology Reports. J Imaging 2022;8:131.
70. Datta S, Roberts K. A Hybrid Deep Learning Approach for Spatial Trigger Extraction from Radiology Reports. Proc Conf Empir Methods Nat Lang Process 2020;2020:50-5.
71. Nakamura Y, Hanaoka S, Nomura Y, et al. Automatic detection of actionable radiology reports using bidirectional encoder representations from transformers. BMC Med Inform Decis Mak 2021;21:262.
72. Liu V, Clark MP, Mendoza M, et al. Automated identification of pneumonia in chest radiograph reports in critically ill patients. BMC Med Inform Decis Mak 2013;13:90.
73. Xue Y, Xu T, Long R, et al. Multimodal recurrent model with attention for automated radiology report generation. In: Frangi A, Schnabel J, Davatzikos C,et al. (eds) Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. MICCAI 2018. Lecture Notes in Computer Science, 2018 pp. 457-466.
74. Towfighi S, Agarwal A, Mak DYF, et al. Labelling chest x-ray reports using an open-source NLP and ML tool for text data binary classification. medRxiv 19012518. doi: 10.1101/19012518.
75. Olthof AW, van Ooijen PMA, Cornelissen LJ. Deep Learning-Based Natural Language Processing in Radiology: The Impact of Report Complexity, Disease Prevalence, Dataset Size, and Algorithm Type on Model Performance. J Med Syst 2021;45:91.
76. Olthof AW, van Ooijen PM, Cornelissen LJ. The natural language processing of radiology requests and reports of chest imaging: Comparing five transformer models' multilabel classification and a proof-of-concept study. Health Informatics J 2022;28:14604582221131198.
77. Yi PH, Kim TK, Lin CT. Comparison of radiologist versus natural language processing-based image annotations for deep learning system for tuberculosis screening on chest radiographs. Clin Imaging 2022;87:34-7.
78. Bressem KK, Adams LC, Gaudin RA, et al. Highly accurate classification of chest radiographic reports using a deep learning natural language model pre-trained on 3.8 million text reports. Bioinformatics 2021;36:5255-61.
79. Martin A. (2021) Robotics and artificial intelligence: The role of AI in robots. Available online: https://aibusiness.com/verticals/robotics-and-artificial-intelligence-the-role-of-ai-in-robots (Assessed: 7 December 2023)
80. Salcudean SE, Moradi H, Black DG, et al. Robot-Assisted Medical Imaging: A Review. Proceedings of the IEEE 2022;110:951-67.
81. Habuza T, Navaz AN, Hashim F, et al. AI applications in robotics, diagnostic image analysis and precision medicine:

Current limitations, future trends, guidelines on CAD systems for medicine. Informatics in Medicine Unlocked 2021;24:100596.
82. Siemens Healthineers (2023) Robotic X-ray. Available online: https://www.siemens-healthineers.com/robotic-x-ray (Assessed: 4 December 2023)
83. Ajani OS, Obasekore H, Kang BY, et al. Robotic Assistance in Radiology: A Covid-19 Scenario. IEEE Access 2023;11:49785-93.
84. Reedy P. Strategic Leadership in Digital Evidence. Academic Press, 2020 pp. 145-150.
85. Yadav AK, Shukla R, Singh TR. Machine learning in expert systems for disease diagnostics in human healthcare. In: Kumae P, Kumar Y, Tawhid MA. (eds) Machine Learning, Big Data, and IoT for Medical Informatics. Academic Press, 2021 pp. 179-200.
86. Lutkevich B. (2022) Expert system. Available online: https://www.techtarget.com/searchenterpriseai/definition/expert-system. (Assessed: 7 December 2023)
87. Chapman WW, Haug PJ. Comparing expert systems for identifying chest x-ray reports that support pneumonia. Proc AMIA Symp 1999;216-20.
88. Lenat DA, Feigenbaum EA. On the thresholds of knowledge. Artificial Intelligence 1991;47:185-250.
89. Bezem M. Consistency of rule-based expert systems. In: Lusk E, Overbeek R. (eds) 9th International Conference on Automated Deduction. CADE 1988. Lecture Notes in Computer Science, Berlin: Springer, pp. 151-161.
90. Pham HNA, Triantaphyllou E. The Impact of Overfitting and Overgeneralization on the classification Accuracy in Data Mining. In: Maimon O, Rokach L. (eds) Soft Computing for Knowledge Discovery and Data Mining. Boston: Springer, 2008 pp. 391-431.
91. Shan N, Ziarko W. Data-based acquisition and incremental modification of classification rules. Computational Intelligence 199;11:357-70.
92. Hendriks PHJ, Vriens DJ. Knowledge-based systems and knowledge management: friends or foe? Information and Management 1999;35:113-25.
93. Vlamou E, Papadopoulos B. Fuzzy logic systems and medical applications. AIMS Neurosci 2019;6:266-72.
94. Hassen DB, Taleb H, Yaakoub IB, et al. A Fuzzy Approach to Chest Radiography Segmentation Involving Spatial Relations. International Journal of Computer Applications, Special issue on Novel aspects of Digital Imaging Applications (DIA) 2013 pp. 40-45.
95. Torres F, Rosales AJ, Gallegos-Funes FJ, et al. A robust neuro-fuzzy classifier for the detection of cardiomegaly in digital chest radiographies. Dyna (Medellin, Colombia) 2014;81:35-41.
96. Suttitanawat K, Apinum U, Sansanee A, et al. Lung Nodule Detection from Chest X-Ray Images Using Interval Type-2 Fuzzy Logic System. 8th IEEE International Conference on Control System, Computing and Engineering (ICCSCE) Penang, Malaysia, 2018 pp. 223-6.
97. Zhang Y, Yang D, Lam S, et al. Radiomics-Based Detection of COVID-19 from Chest X-ray Using Interpretable Soft Label-Driven TSK Fuzzy Classifier. Diagnostics (Basel) 2022;12:2613.
98. Şahin I, Akdogan E, Aktan ME. A type-2 fuzzy rule-based model for diagnosis of COVID-19. Turkish Journal of Electrical Engineering and Computer Sciences 2023;31:39-52.
99. Ieracitano C, Mammone N, Versaci M, et al. A fuzzy-enhanced deep learning approach for early detection of Covid-19 pneumonia from portable chest X-ray images. Neurocomputing (Amst) 2022;481:202-15.
100. Yadlapalli P, Dokku B. Application of Fuzzy Deep Neural Networks for Covid 19 diagnosis through chest radiographs. F1000Research 2023;12:60.
101. Kechit, G. (2022) Fuzzy Logic in Artificial Intelligence: Applications, Advantages & Disadvantages. Available online: https://www.upgrad.com/blog/fuzzy-login-in-artificial-intelligence/

doi: 10.21037/jmai-24-67
Cite this article as: Nwaiwu VC, Das SK. Emerging multifaceted application of artificial intelligence in chest radiography: a narrative review. J Med Artif Intell 2024;7:39.

