narra j

Review Article

Artificial intelligence-based deep learning algorithms for ground-glass opacity nodule detection: A review

Henil P. Shah[1], Agha SAH. Naqvi[2], Parth Rajput[1], Hanan Ambra[3], Harrini Venkatesh[4], Junaid Saleem[5], Sudarshan Saravanan[6], Mayur Wanjari[7] and Gaurav Mittal[8]*

[1]GMERS Medical College, Gujarat, India; [2]Pakistan Institute of Medical Science, Islamabad, Pakistan; [3]KMCT Medical College, Kerala, India; [4]Sri Ramachandra Institute of Higher Education and Research (SRIHER), Chennai, India; [5]Federal Medical College, Islamabad, Pakistan; [6]Govt. Stanley Medical College, Chennai, India; [7]Department of Research and Development, Datta Meghe Institute of Higher Education and Research, Wardha, India; [8]Mahatma Gandhi Institute of Medical Sciences, Sevagram, India

*Corresponding author: mittalgaurav742002@gmail.com

Abstract

Ground-glass opacities (GGOs) are hazy opacities on chest computed tomography (CT) scans that can indicate various lung diseases, including early COVID-19, pneumonia, and lung cancer. Artificial intelligence (AI) is a promising tool for analyzing medical images, such as chest CT scans. The aim of this study was to evaluate AI models' performance in detecting GGO nodules using metrics like accuracy, sensitivity, specificity, F1 score, area under the curve (AUC) and precision. We designed a search strategy to include reports focusing on deep learning algorithms applied to high-resolution CT scans. The search was performed on PubMed, Google Scholar, Scopus, and ScienceDirect to identify studies published between 2016 and 2024. Quality appraisal of included studies was conducted using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool, assessing the risk of bias and applicability concerns across four domains. Two reviewers independently screened studies reporting the diagnostic ability of AI-assisted CT scans in early GGO detection, where the review results were synthesized qualitatively. Out of 5,247 initially identified records, we found 18 studies matching the inclusion criteria of this study. Among evaluated models, DenseNet achieved the highest accuracy of 99.48%, though its sensitivity and specificity were not reported. WOANet showed an accuracy of 98.78%, with a sensitivity of 98.37% and high specificity of 99.19%, excelling particularly in specificity without compromising sensitivity. In conclusion, AI models can potentially detect GGO on chest CT scans. Future research should focus on developing hybrid models that integrate various AI approaches to improve clinical applicability.

Keywords: Ground glass opacity, deep neural network, high-resolution CT-scan, X-ray image, pulmonary nodule

Introduction

Ground-glass opacity (GGO) refers to a hazy, unclear opacity on computed tomography (CT) that does not cover the underlying bronchial tissues or pulmonary vascular arteries in radiology [1]. It is characterized by partial air space-filling, thickening of interstitial, partial collapse of alveoli, normal typical expiration, or increased amount of blood in capillaries [2]. GGO frequently appears in pulmonary hemorrhage, pulmonary oedema, and acute interstitial pneumonia (AIP). Hence an accurate diagnosis is critical to the prognosis and management of the illness [3-5].

Received: August 22, 2024 - Accepted: January 30, 2025 - Published online: March 5, 2025

Detecting GGOs on chest CT scans is a well-recognized challenge, even for skilled radiologists. Their faint, shadow-like appearance and small size can easily go unnoticed, making early detection vital for better patient outcomes. Deep learning-based models have the capability to address these challenges by enhancing efficiency and diagnostic throughput in a non-invasive manner [6]. For this reason, it is imperative to create and implement plans that will help medical professionals identify GGO in an accurate and timely manner.

Artificial intelligence (AI), also known as deep learning machines, is the oldest and largest field in computer science. It deals with all aspects of simulating neural networks for practical problems and creating computers that are able to think and learn like humans [7]. It has emerged as a promising tool for detecting GGO. Integrating AI and imaging methods helps in precise diagnosis by giving a few positive outcomes. A delay in intervention can happen due to the cost and requirement of medical personnel and equipment for diagnosing and concluding disease identification. AI offers precise empirical solutions for these issues, which require less work and money. Efficient and precise identification of GGOs for diagnosing and prioritizing COVID-19 patients may enhance productivity and preserve resources in pandemic-ravaged nations [8].

A type of AI model called deep neural network (DNN), is comprised of an input and an output layer. In the context of a neural network, when a sample is provided as an input, each unit (neuron) computes its activation based on the weighted inputs it receives from the preceding layer [9]. An overview of the implementation steps of the supervised deep learning algorithm is presented in **Figure 1**. They demonstrated the capability of surpassing human accuracy in numerous sectors of life. The efficacy of DNNs lies in their ability to extract complex features from raw data after extensive training on labeled datasets, resulting in a proficient representation of an input domain [10].

GGO is commonly categorized into two main groups: part-solid nodules and pure GGO, though variations and overlaps may exist [11]. DNNs have been increasingly applied to medical imaging, including the classification and follow-up of GGOs. Understanding the conservative follow-up of GGOs needs to comprehend the GGOs' natural history [12]. According to a previous study, some lesions with GGO develop gradually, and some remain unbothered for a long period [13]. For patients with pure GGO nodules and favorable characteristics, wedge resection is often preferred over lobectomy [14]. The response of different GGO classifications according to size and associated strategies is presented in **Figure 2**. Multiple GGO nodules do not necessarily indicate detrimental perseverance. Asamura *et al.* proposed patients who have numerous lesions ought to be considered candidates for surgery, although having a reserved lung parenchymal volume [15].

In clinical practice, the proficiency of individual clinicians (radiologists, pathologists, and so on) determines the accuracy of the detection and diagnosis of cancer and/or many other diseases. In response to this clinical issue, numerous computer-aided detection and diagnosis (CAD) schemes have been developed and tested to help clinicians make accurate and objective diagnostic decisions by helping them read medical images more quickly [16]. Prior research on deep learning algorithms for detecting GGOs in chest imaging has largely focused on individual models and often lacks a standardized evaluation framework, limiting its broader applicability [17]. This study addresses these limitations by systematically comparing the performance of multiple AI models. It also identifies critical gaps, such as inconsistent evaluation methodologies and the urgent need for standardized datasets, ensuring more accurate and clinically relevant advancements in GGO detection.

Methods

Database and search strategy

A search was performed in three databases, PubMed, Scopus, and Google Scholar, to find the most recent literature. ScienceDirect was used to retrieve the additional articles. The search was restricted to articles published between January 2016 and January 2024. We verified the articles using a combination of subject and free terms. The primary key terms were "artificial intelligence," "neural networks," "deep learning," "ground glass opacity," "pulmonary nodules," and so on.

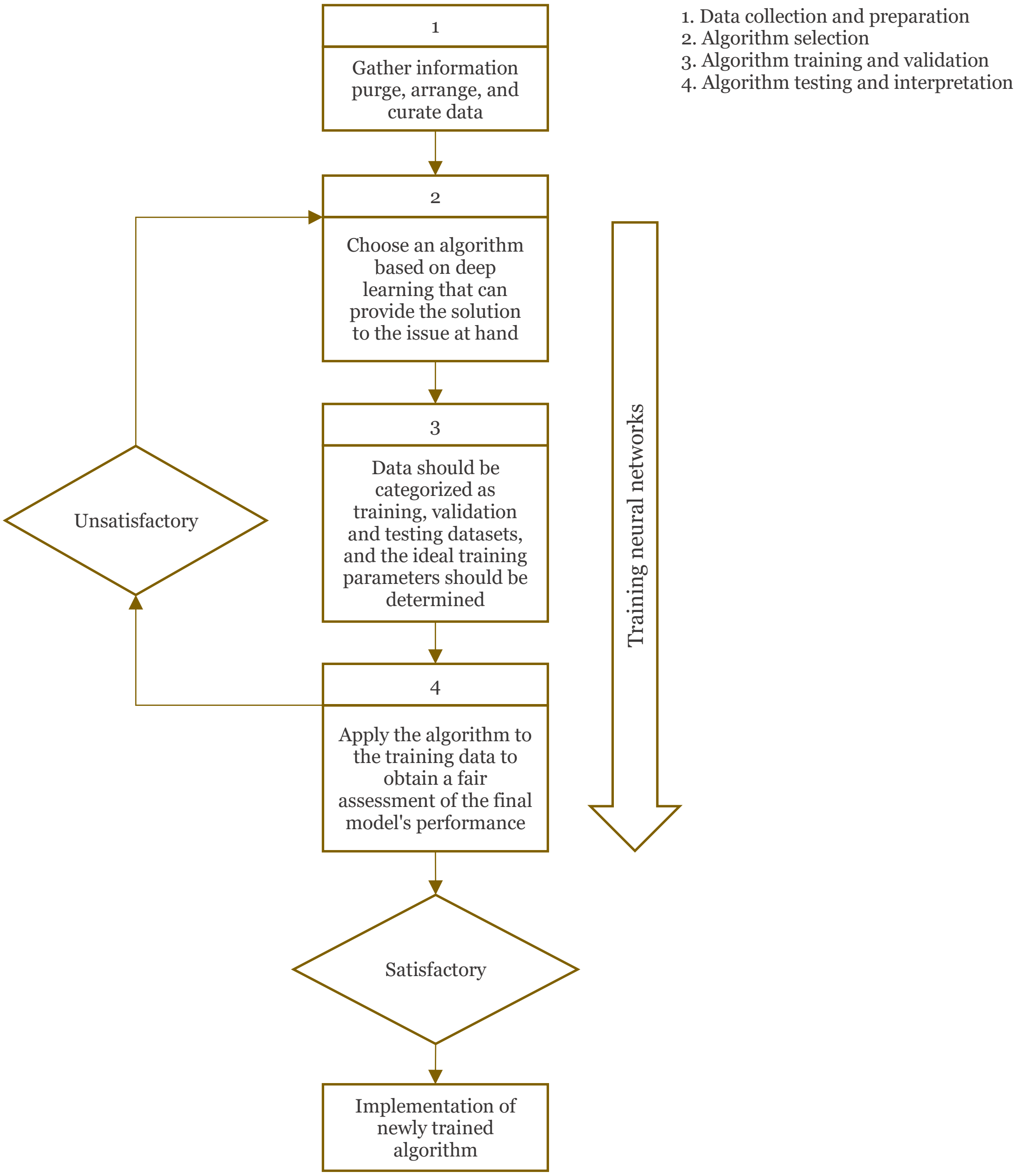

Figure 1. Overview of supervised deep learning algorithm implementation steps [18].

The research used a combination of keywords, including: ("deep learning" OR "convolutional neural network" OR "neural networks" OR "artificial intelligence" OR "machine learning") AND ("ground-glass opacity" OR "GGO") AND ("nodules" OR "lesions") AND ("high-resolution chest CT" OR "HRCT"), ("computer-aided diagnosis" OR "CAD system" OR "automated detection") AND ("GGO nodules" OR "ground-glass opacities") AND ("high-res CT" OR "chest computed tomography"), ("pulmonary nodule detection" OR "lung nodule identification") AND ("deep neural networks" OR "DL algorithms" OR "deep models") AND ("high-resolution CT" OR "HRCT").

Population, index test, reference test, and target condition (PIRT) framework

The research question, what is the effectiveness of AI-driven deep learning algorithms in identifying GGO nodules, and what are the key factors that impact their diagnostic accuracy?, was structured using the population, index test, reference test, and target condition (PIRT) framework [19]. The population included patients diagnosed with GGOs through chest imaging assessments.

The index test involved the application of deep learning neural network algorithms for detecting GGOs. The reference test relied on image readings conducted by expert physicians or experienced radiologists, with the target condition being GGO nodules.

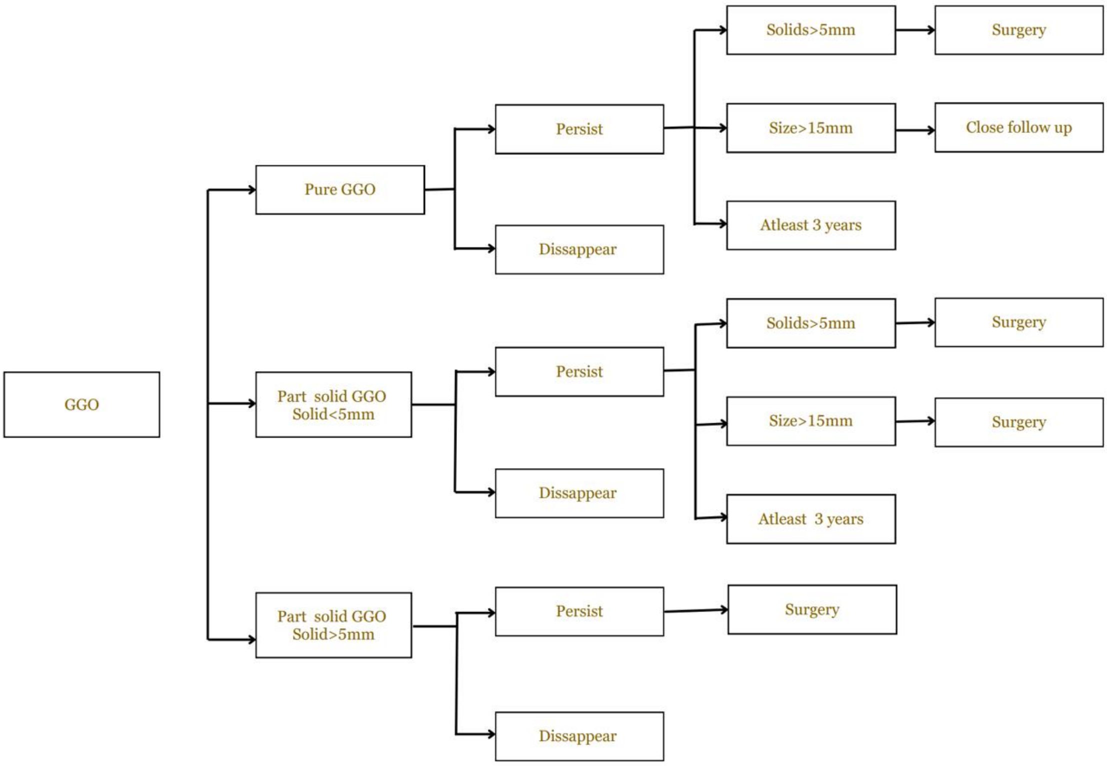

Figure 2. Follow-up algorithm for various classifications of ground-glass opacity (GGO) based on size and corresponding interventions. The scheme was adopted from a previously published report [20].

Eligibility criteria

For inclusion, the following specific criteria were used: (1) specifically focused studies on high-resolution CT scans of the chest that aim to identify and categorize GGO nodules; (2) studies that employ deep learning algorithms as the primary method for detection and classification; (3) studies involving human subjects or clinical data; (4) studies published in the English language (5) studies with adequate and complete data for analysis; (6) studies with a sample size of at least 10 participants; (7) studies that provide explicit details on the deep learning algorithms used and the methodology employed; (8) studies that compare the effectiveness of deep learning algorithms with other relevant methods, if available; (9) studies conducted on high-resolution chest CT scans or similar imaging modalities; (10) studies that provide measurements for accuracy, such as area under the curve (AUC) of receiver operating characteristic (ROC), positive predictive value, specificity, negative predictive value, and sensitivity; (11) studies employing commercial AI solutions.

The exclusion criteria were as follows: (1) studies that were not based on deep learning algorithms; (2) studies that did not include human subjects or utilized clinical data; (3) studies published in languages other than English; (4) studies with incomplete or insufficient data for analysis; (5) research involving fewer than 10 individuals in the sample; (6) research that did not provide explicit details of the deep learning algorithms used or the methodology employed; (7) studies that primarily focused on the performance of non-deep learning algorithms or traditional image processing techniques; (8) studies that primarily investigated the performance of deep learning algorithms on low-resolution chest CT scans.

Screening and selection

Study selection was conducted based on predefined inclusion criteria. Two reviewers independently screened the search results by reviewing titles and abstracts. Studies deemed potentially eligible were then subjected to full-text review for final inclusion. Duplicates were identified and removed, and all bibliographic records and retrieved data were stored and documented using the Rayyan app (Rayyan Systems Inc., Doha, Qatar) [21]. Any discrepancies during the selection process were resolved through discussion between the two reviewers, who re-evaluated the full-text articles. If disagreements persisted, a third reviewer was consulted to reach a consensus.

Quality appraisal

To determine the likelihood of bias in the included studies, the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool [22] was utilized. The studies showed minimal risk of bias in participant flow, patient selection, index test application, and overall risk. The domains assessed were: (1) patient selection: evaluating how participants were selected, focusing on sampling methods, disease severity spectrum, and exclusion criteria; (2) index test: assessing its description, reproducibility, and whether it was interpreted blinded; (3) reference standard: evaluating its appropriateness, blinded interpretation, and reliability in confirming diagnoses; (4) flow and timing: assessing patient flow through the study and the timing between index test and reference standard. Quality appraisal using QUADAS-2 tool was performed by two independent reviewers, and discrepancies were resolved through discussion or consultation with the third reviewer.

Data extraction

The following information for extraction: initial author, year, name, study type, primary goal, kind of DNN used, the criteria for inclusion, and result (indicating sensitivity, specificity, accuracy, AUC, F1 score, precision, recall). Two reviewers independently extracted the data into a prepared table. Discrepancies in the extracted data were resolved through comprehensive discussions and the assistance of a third reviewer.

Data synthesis

In the data synthesis, the extracted information was categorized according to the diagnostic performance metrics, namely accuracy, sensitivity, and specificity. We made sure that the values were estimated from the validated model, with uniform estimation method. The performance of each AI model was compared based on the three metrics which were presented in a bar graph. All review authors were involved in discussions on the diagnostic performance of the AI model, along with its strengths and weaknesses.

Results

Searching results

In the initial stage, it involved identifying 5,247 studies across all four databases: PubMed (n=820), Google Scholar (n=87), Scopus (n=4,327) and ScienceDirect (n=13). A total of 740 duplicated studies were eliminated. Thereafter, 4,507 abstracts and study headings were screened; of which, 4,403 were eliminated because they did not meet the eligibility criteria for inclusion. As a result, 35 full-text publications were examined, and eligibility was evaluated. Seventeen articles were eliminated because they did not meet the inclusion requirements or the entire text could not be found. The screening and selection process were summarized and presented in **Figure 3**.

Characteristics of the included studies

Characteristics of the included studies each focusing on the application of AI models for the early detection of GGO are presented in **Table 1**. COVIDiag model employed databases from different nations to assess the effectiveness of their model for routine procedures [23]. CovAI-Net [24] and Context Learning CNN [25] achieved notably high specificity levels, suggesting the potential for reducing false positives in clinical settings. Meanwhile, COVID-Net CT-2 [26] achieved exceptionally high sensitivity in cases of GGO detection for early diagnosis of COVID-19. On the other hand, a study [27] showcases the effectiveness of AI models, specifically CNN like DenseNet-121, in improving the overall accuracy of COVID-19 prediction. Rather than using a traditional dataset for transfer learning, CovXNet employed a wider dataset encompassing X-rays from both normal and other non-COVID pneumonia patients [28]. The uAI-ChestCare automated the process of delineating the whole 3D region of interest (ROI) to diagnose lesions by drawing the tumor border on a series of axial lung window pictures [29].

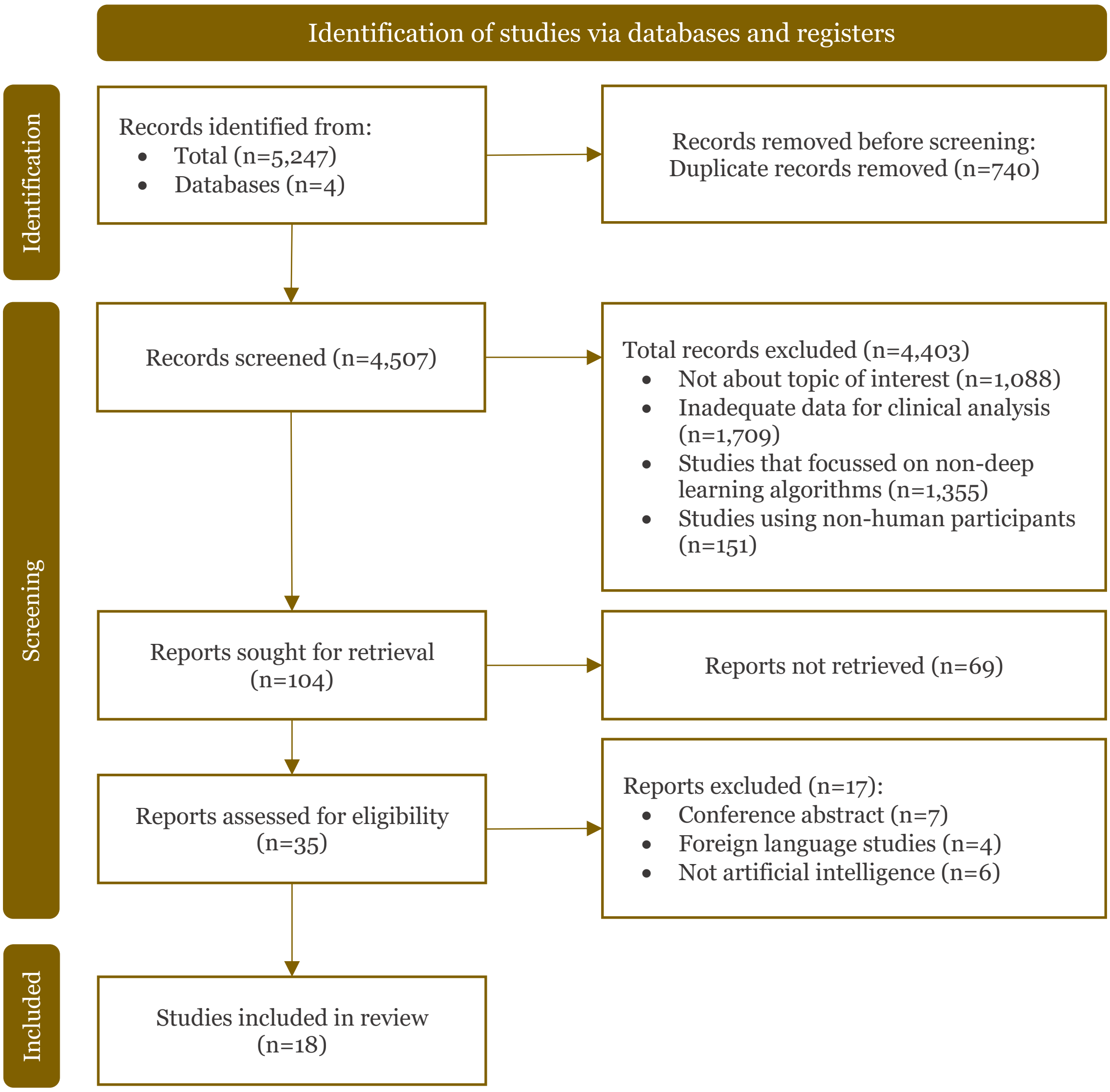

Figure 3. Flow diagram illustrating the study selection based on eligibility criteria.

Quality of the included studies

There were only three studies that are completely free from risk of bias in all domains [27,30,31]. Other studies received high risk or some concerns marks. Reasons for biases in selected studies are as follows: (1) the patient sample was not representative of the target population, chosen non-consecutively or randomly; (2) deep learning models for COVID-19 identification from CT scans were poorly described or standardized, leading to bias; (3) inconsistent application of the reference standard across studies, affecting diagnosis accuracy; (4) timing inconsistencies between the reference standard and index test led to misclassification. The summary and its traffic light plot of the quality assessment conducted for the included studies are presented in **Figure 4**.

Accuracy of the AI-based GGO identification

The sensitivity and specificity of AI for the detection of GGO were reported to be 71–99.1% and 77–100%, respectively [23-25,28-30,32-35]. Similarly, the accuracy of AI for GGO detection varies from 78.97% to 99% [23-30,32-34,36-38]. These algorithms were trained on a variety of image features extracted from chest CT scans, potentially including nodule size, shape, texture, and location. For comparison, traditional HRCT scans have demonstrated sensitivity and specificity ranging from 41% to 52% and 56% to 63%, respectively [39].

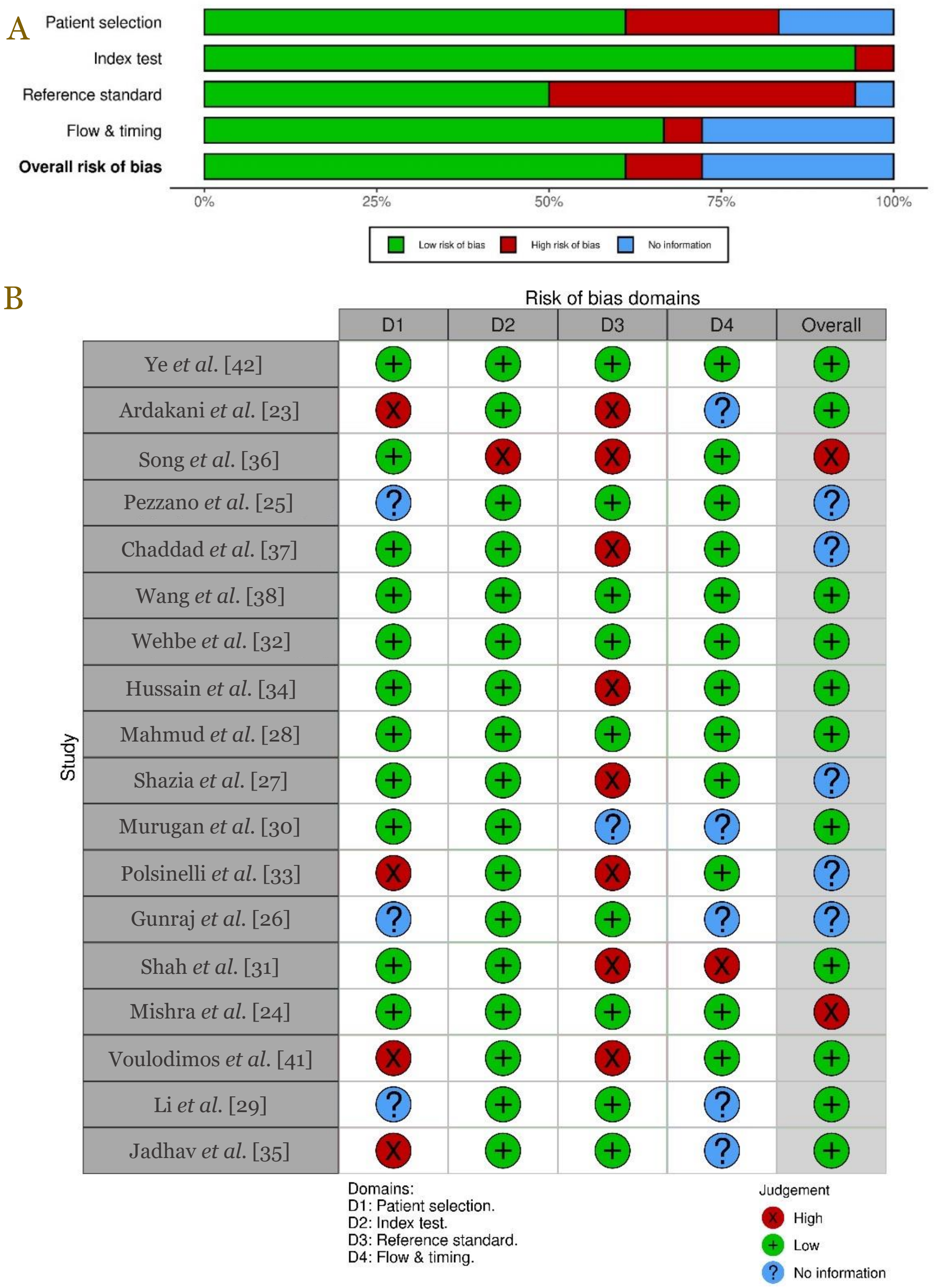

Figure 4. Summary plot (A) and traffic light plot (B) for the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) result.

Comparative illustrations of various neural network architectures based on accuracy, specificity, and sensitivity are presented in **Figure 5** and **Figure 6**. DenseNet-121 [27] demonstrates consistent excellence across all three metrics, making it reliable for GGO detection. Additionally, specialized models such as WOANet [30], uAI-ChestCare [29], CoroDet [34], and DeepCOVID-XR [32] also show strong performance, particularly in sensitivity and accuracy, enabling effective identification of subtle GGO abnormalities. Models like AlexNet and NASNet-Mobile [37] exhibit comparatively lower performance, rendering them less suitable for accurate GGO detection. Likewise, SqueezeNet [33] displays reduced sensitivity, which diminishes its diagnostic effectiveness.

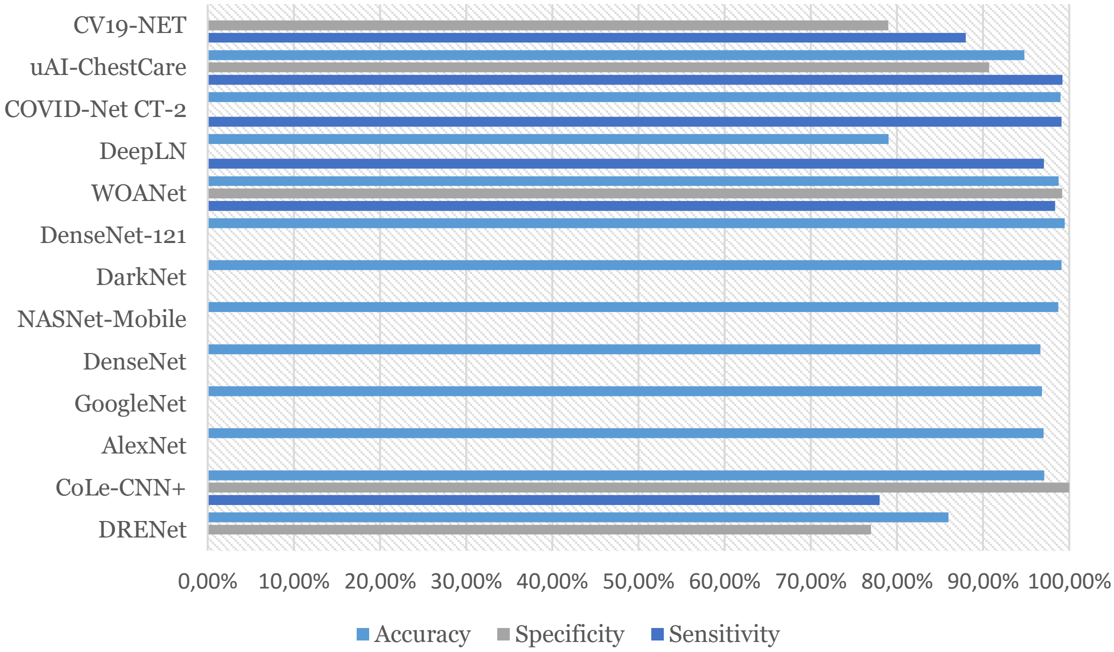

Figure 5. Evaluation of accuracy, specificity, and sensitivity across different neural network models used in medical image analysis. The figure provides a detailed comparison of performance metrics, showcasing the relative efficiency of each model [25-27,29,30,35-38].

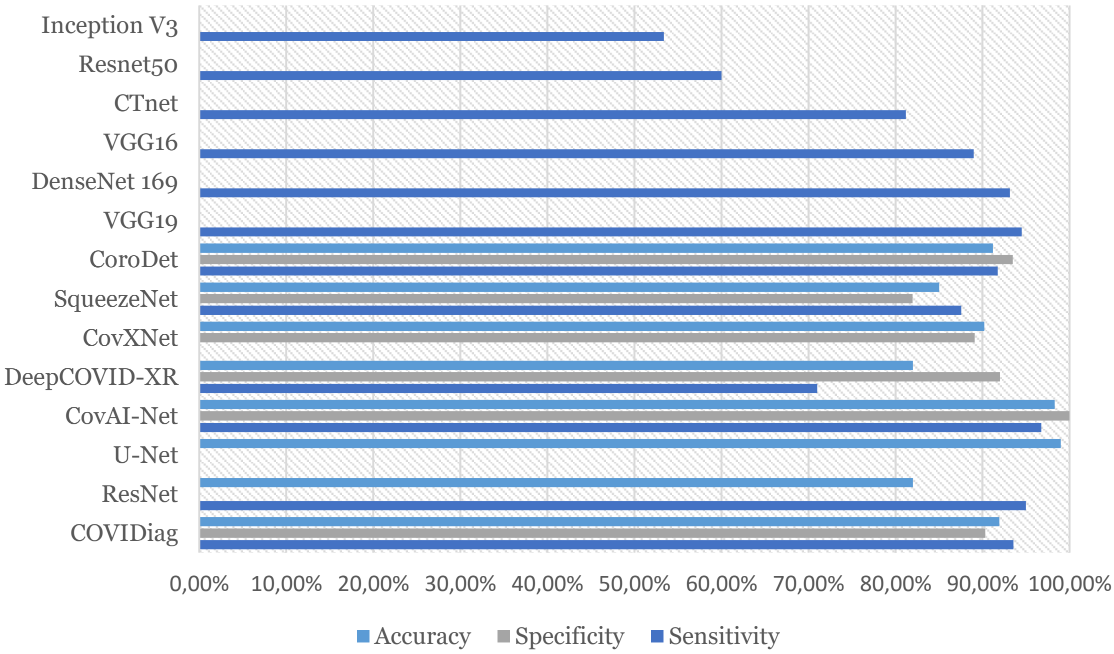

Figure 6. Evaluation of accuracy, specificity, and sensitivity across different neural network models used in medical image analysis. The figure provides a detailed comparison of performance metrics, showcasing the relative efficiency of each model [23,28,31-34,40-42].

Shah et al. Narra J 2025; 5 (1): e1361 - http://doi.org/10.52225/narra.v5i1.1361

Table 1. Characteristics and outcomes of the included research/review

Study characteristics			Imaging data, n		Diagnostic parameters						
First author	Year	Study design	Training phase	Testing phase	Type	Algorithm type	Resolution (in pixels)	Accuracy	Sensitivity	Specificity	Others
Ye *et al.* [42]	2019	Cross-sectional	2,421	593	Pre-trained	ResNet and ResNet-trained models	Not specified	ResNet: 82.00%; pre-trained ResNet: 87.00%	95.00%	NR	F score ResNet: 85.52%; pre-trained ResNet: 87.77%
Ardakani *et al.* [23]	2020	Cross-sectional	488	124	Pre-trained	COVIDiag	Not specified	91.94%	93.54%	90.32%	AUC: 96.50%
Voulodimos *et al.* [41]	2020	Case-control	798	141	Pre-trained	U-Net	630×630	99.00%	NR	NR	F1-score: 89.00%; precision: 91.00%; recall: 89.00%
Mishra *et al.* [24]	2020	Cohort	15,024	86	Pre-trained	CovAI-Net (Inception, DenseNet, Xception)	224×224	98.31%	96.74%	100%	F1-score: 98.34%; precision: 100%
Wehbe *et al.* [32]	2020	Cross-sectional	14,788	2,214	Pre-trained	DeepCOVID-XR	224×224, 331×331	82.00%	71.00%	92.00%	AUC: 88.00%
Mahmud *et al.* [28]	2020	Cross-sectional	5,856	305	Pre-trained	CovXNet	Not specified	90.20%	NR	89.10%	AUC: 91.10%; precision: 90.80%; F1-score: 90.40%
Polsinelli *et al.* [33]	2020	Cross-sectional	Dataset arrangement 1: 1,646; dataset arrangement 2: 1,906	Dataset arrangement 1: 388; dataset arrangement 2: 203	Pre-trained	SqueezeNet	Not specified	85.03%	87.55%	91.95%	F1-score: 86.20%; precision: 85.01%
Hussain *et al.* [34]	2020	Cross-sectional	For 4-class classification: 2,100; for 3-class classification: 2,100; for 2-class classification: 1,300	Not specified; typically, a portion of the dataset was used for testing.	Pre-trained	CoroDet	Not specified	91.20%	91.76%	93.48%	Precision: 92.04%; recall: 91.90%; F1-score: 90.04%
Shah *et al.* [31]	2021	Cross-sectional	664	74	Custom, pre-trained	VGG-19 and CT-Net10	128×128 to 224×224	VGG19: 94.52%; DenseNet 169: 93.15%; VGG16: 89.00%; CTnet: 82.10%; Resnet: 60.00%; Inception V3: 53.4%	NR	NR	NR
Song *et al.* [36]	2021	Cross-sectional	897	385	Pre-trained	DRENet	14×14, 7×7	86.00%	NR	NR	AUC: 95.00%; precision: 79.00%

Study characteristics			Imaging data, n		Diagnostic parameters						
First author	Year	Study design	Training phase	Testing phase	Type	Algorithm type	Resolution (in pixels)	Accuracy	Sensitivity	Specificity	Others
Pezzano *et al.* [25]	2021	Case-control	2,947	368	Pre-trained	CoLe-CNN+	256×256	97.10%	78.00%	100%	Precision: 100%
Chaddad *et al.* [37]	2021	Cohort	1,016	254	Pre-trained	CNNs including: AlexNet, DenseNet, GoogleNet, NASNet Mobile, ResNet18, and DarkNet	512×512	AlexNet: 97.04%; GoogleNet: 96.84%; DenseNet: 96.66%; NASNet: 98.72%; DarkNet: 99.09%	NR	NR	AlexNet AUC: 99.28%; GoogleNet AUC: 98.25%; DenseNet AUC: 98.12%; NASNet-Mobile AUC: 99.25%; DarkNet AUC: 99.89%
Shazia *et al.* [27]	2021	Experimental comparative	2,757	4,405	Pre-trained	DenseNet-121	224×224	99.48%	NR	NR	F1-score: 99.49%; precision: 99.54%
Murugan *et al.* [30]	2021	Cross-sectional	2,214	246	Pre-trained	WOANet: Whale optimized deep neural network based on ResNet-50	224×224×3	98.78%	98.37%	99.19%	Precision: 99.18; F1-score: 98.37
Wang *et al.* [38]	2022	Retrospective (single centre)	7,160	1,790	Pre-trained	DeepLN	512×512	Test set: 79.02%	Test set: 80.80%	NR	AUC: 88.58%
Gunraj *et al.* [26]	2022	Cohort	COVIDx CT-2A: 169,264; COVIDx CT-2B: 175,445	COVIDx CT-2A: 25,658; COVIDx CT-2B: 25,658	Pre-trained	COVID-Net CT-2	Not specified	99.00%	99.10%	99.40%	NR
Li *et al.* [29]	2023	Retrospective observational	78	33	Pre-trained	uAI-ChestCare	512×512	94.80%	90.70%	100%	AUC in the training set: 99.20%; AUC of validation set: 97.50%
Jadhav *et al.* [35]	2023	Cohort	5,236	5,869	Pre-trained	CV19-NET	Not specified	NR	88.00%	79.00%	AUC: 92.00%

AUC: area under the curve; CNNs: convolutional neural networks; CoLe-CNN: context-learning convolutional neural network; DeepLN: deep learning network; DenseNet: dense convolutional network; DRENet: details relation extraction neural network; NASNet: neural architecture search network; NR: not reported; ResNet: residual network; VGG: visual geometry group; WOANet: whale optimization algorithm network

Discussion

The present review suggested that the sensitivity and specificity of AI for the detection of GGO ranged from moderate (75–90%) to high (>90%). As for the accuracy, it varies considerably across different algorithms, ranging from 79% to 100%. Variability in reported metrics (accuracy, sensitivity, specificity) can be attributed to differences in the datasets used, the specific algorithms applied, and the criteria for evaluation. This variability underscores the importance of standardizing datasets and benchmarking processes to ensure consistent and comparable evaluations of AI models. Two previous studies performed cross-validation to assess the effectiveness and reliability of AI models [36,37].

Advancements in segmentation, early detection methods, and deep learning techniques for classification have accelerated work in this domain. Previous study has demonstrated potential in developing AI-based methods for GGO segmentation [43]. Different algorithms, such as those in radiomics and deep learning, are making significant contributions; these algorithms hold promise not only in differentiating benign from malignant nodules but also in predicting the prognosis of small-cell lung cancer and pneumonia cases [44-48]. However, contradictory results were obtained regarding the performance of AI in GGO screening and diagnosis, with some studies reporting poor performance and others reporting better performance compared to traditional methods. This might be attributed to the limitations in earlier AI methodologies. Conversely, newer AI approaches appear to demonstrate improved detection matrices [49].

Detecting GGOs on chest CT scans is notoriously difficult, even for experienced radiologists. Their faint and tiny shadow appearance can easily be missed, making early detection crucial for patient outcomes. The accuracy of traditional diagnosis may be influenced by various factors, including the presence of benign lesions (necrosis, inflammation, tuberculosis), diverse lung image textures, and radiologist experience [50]. The susceptibility of AI algorithms to variations in underlying data can lead to inconsistent outcomes. Computational limitations associated with processing speed and memory requirements might pose practical challenges for real-world implementation, especially when dealing with large datasets [51]. Additionally, small sample sizes in some studies reduce their statistical power, potentially affecting the generalizability of findings to larger populations[52]. The study also highlights the potential for false-negative results, particularly for early-stage COVID-19 patients with negative CT findings, necessitating further consideration.

Due to the inherent difficulty in diagnosing GGOs, early detection and management are crucial, as delays can significantly impact a patient's quality of life. Currently, diagnosis of GGO nodules relies on high-resolution CT (HRCT), bronchoscopy with biopsy, and MRI. However, studies have shown that even skilled pulmonologists can struggle with accurate diagnosis [43,50,53-56]. AI has the potential to enhance diagnostic accuracy, reduce clinicians' workload, and improve treatment outcomes [57,58]. Evidence suggests that AI may outperform human expertise in recognizing specific patterns relevant to GGO detection [59]. This highlights the potential value of standardized AI-based diagnostic tools [60]. The claim that AI improves the effectiveness of diagnosis, reduces clinicians' workload, and enhances treatment and prognosis is supported by these findings.

The absence of data on crucial demographic factors (age, sex), administered treatments, and overall survival rates could restrict the effectiveness and generalizability of AI-based models in real-world healthcare settings [61,62]. While the application of AI models in clinical settings holds significant promise, ensuring their validity is crucial for broader adoption. Furthermore, radiomics, a field with immense potential, has yet to achieve broad clinical integration [63]. Large-scale data collection and sharing initiatives are needed to create comprehensive, standardized healthcare datasets [64].

Our review has few limitations. Since Asian patients are the majority of the study participants, it is possible that the results cannot be applied to all global ethnic groups. There was a linguistic bias because only English-language research was chosen, but studies conducted in other languages may have found a stronger association. The inaccessibility of data from the different databases that have been used worldwide also contributes to information bias. Despite our best attempts to include studies with a big sample size of GGO images, deep learning requires huge sample amounts of training data to power the AI model. The lung GGO samples from the study that we have included in our review are still insufficient. In the future, bigger sample sizes will be demanded to assess the performance of AI models, and we anticipate that doing so will help deep learning's accuracy improve even further.

Conclusion

This study highlights the effectiveness of deep learning algorithms in identifying GGOs on high-resolution chest CT scans, showing consistently high accuracy, sensitivity, and specificity across studies. AI-based models demonstrate significant potential in assisting early and accurate detection of GGOs and related lung conditions, including COVID-19. However, limitations include biases from retrospective study designs, reliance on single-center data, and the need for diverse datasets to enhance model robustness. Standardizing datasets and benchmarking processes are crucial for ensuring consistent evaluations. With continued advancements and larger datasets, AI is poised to play an increasingly pivotal role in medical imaging, offering enhanced diagnostic and prognostic capabilities while reducing time and costs.

Ethics approval

Not required.

Acknowledgements

The authors declare no specific acknowledgements for this study.

Competing interests

All the authors declare that there are no conflicts of interest.

Funding

This study received no external funding.

Underlying data

Derived data supporting the findings of this study are available from the corresponding author on request.

Declaration of artificial intelligence use

We hereby confirm that no artificial intelligence (AI) tools or methodologies were utilized at any stage of this study, including during data collection, analysis, visualization, or manuscript preparation. All work presented in this study was conducted manually by the authors without the assistance of AI-based tools or systems.

How to cite

Shah HP, Naqvi ASAH, Rajput P, *et al*. Artificial intelligence-based deep learning algorithms for ground-glass opacity nodule detection: A review. Narra J 2025; 5 (1): e1361 - http://doi.org/10.52225/narra.v5i1.1361.

References

1. Hansell DM, Bankier AA, MacMahon H, *et al*. Fleischner Society: Glossary of terms for thoracic imaging. Radiology 2008;246(3):697-722.
2. Qi K, Wang K, Wang X, *et al*. Lung-PNet: An automated deep learning model for the diagnosis of invasive adenocarcinoma in pure ground-glass nodules on chest CT. AJR Am J Roentgenol 2024;222(1):e2329674.
3. Nowers K, Rasband JD, Berges G, *et al*. Approach to ground-glass opacification of the lung. Semin Ultrasound CT MR 2002;23(4):302-323.
4. Hewitt MG, Miller WT Jr, Reilly TJ, Simpson S. The relative frequencies of causes of widespread ground-glass opacity: a retrospective cohort. Eur J Radiol 2014;83(10):1970-1976.

5. Kishaba T, Tamaki H, Shimaoka Y, *et al.* Staging of acute exacerbation in patients with idiopathic pulmonary fibrosis. Lung 2014;192(1):141-149.
6. Paez R, Kammer MN, Tanner NT, *et al.* Update on biomarkers for the stratification of indeterminate pulmonary nodules. Chest 2023;164(4):1028-1041.
7. Poole D, Mackworth A, Goebel R. Computational intelligence: A logical approach. Oxford: Oxford University Press; 1998.
8. Saha M, Amin SB, Sharma A, *et al.* AI-driven quantification of ground glass opacities in lungs of COVID-19 patients using 3D computed tomography imaging. PLoS One 2022;17(3):e0263916.
9. Larochelle H, Bengio Y, Louradour J, Lamblin P. Exploring strategies for training deep neural networks. J Mach Learn Res 2009;1:1-40.
10. Sze V, Chen YH, Yang TJ, Emer JS. Efficient processing of deep neural networks: A tutorial and survey. Proceedings of the IEEE 2017;105(12):2295-2329.
11. Fu F, Zhang Y, Wen Z, *et al.* Distinct prognostic factors in patients with stage I non-small cell lung cancer with radiologic part-solid or solid lesions. J Thorac Oncol 2019;14(12):2133-2142.
12. Kodama K, Higashiyama M, Yokouchi H, *et al.* Natural history of pure ground-glass opacity after long-term follow-up of more than 2 years. Ann Thorac Surg 2002;73(2):386-393.
13. Silva M, Sverzellati N, Manna C, *et al.* Long-term surveillance of ground-glass nodules: Evidence from the MILD trial. J Thorac Oncol 2012;7(10):1541-1546.
14. Nakata M, Sawada S, Saeki H, *et al.* Prospective study of thoracoscopic limited resection for ground-glass opacity selected by computed tomography. Ann Thorac Surg 2003;75(5):1601-1606.
15. Asamura H, Suzuki K, Watanabe S, *et al.* A clinicopathological study of resected subcentimeter lung cancers: A favorable prognosis for ground glass opacity lesions. Ann Thorac Surg 2003;76(4):1016-1022.
16. Holland P, Spence H, Clubley A, *et al.* Reporting radiographers and their role in thoracic CT service improvement: Managing the pulmonary nodule. BJR Open 2020;2(1):20190018.
17. Delli Pizzi A, Chiarelli AM, Chiacchiaretta P, *et al.* Radiomics-based machine learning differentiates "ground-glass" opacities due to COVID-19 from acute non-COVID-19 lung disease. Sci Rep 2021;11(1):17237.
18. Torres-Velázquez M, Chen WJ, Li X, *et al.* Application and construction of deep learning networks in medical imaging. IEEE Trans Radiat Plasma Med Sci 2021;5(2):137-159.
19. Thompson M, Bruel AV. Evaluating new diagnostic tests. In: Heneghan C, Perera R, Badenoch D, editors. Diagnostic tests toolkit. London: BMJ Books; 2011.
20. Kobayashi Y, Mitsudomi T. Management of ground-glass opacities: Should all pulmonary lesions with ground-glass opacity be surgically resected?. Transl Lung Cancer Res 2013;2(5):354-363.
21. Ouzzani M, Hammady H, Fedorowicz Z, *et al.* Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016;5(1):210.
22. Whiting PF, Rutjes AW, Westwood ME, *et al.* QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155(8):529-536.
23. Ardakani AA, Acharya UR, Habibollahi S, Mohammadi A. COVIDiag: A clinical CAD system to diagnose COVID-19 pneumonia based on CT findings. Eur Radiol 2021;31(1):121-130.
24. Mishra M, Parashar V, Shimpi R. Development and evaluation of an AI System for early detection of Covid-19 pneumonia using X-ray (student consortium). 2020 IEEE Sixth International Conference on Multimedia Big Data. New Delhi: IEEE; 2020.
25. Pezzano G, Díaz O, Ripoll VR, *et al.* CoLe-CNN+: Context learning-convolutional neural network for COVID-19-ground-glass-opacities detection and segmentation. Comput Biol Med 2021;136:104689.
26. Gunraj H, Sabri A, Koff D, *et al.* COVID-Net CT-2: Enhanced Deep neural networks for detection of COVID-19 from chest CT images through bigger, more diverse learning. Front Med 2021;8:729287.
27. Shazia A, Xuan TZ, Chuah JH, *et al.* A comparative study of multiple neural network for detection of COVID-19 on chest X-ray. EURASIP J Adv Signal Process 2021;2021(1):50.
28. Mahmud T, Rahman MA, Fattah SA. CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization. Comput Biol Med 2020;122:103869.
29. Li C, Jin Y, Deng Q, *et al.* Development and validation of a nomogram based on CT texture analysis for discriminating minimally invasive adenocarcinoma from glandular precursor lesions in sub-centimeter pulmonary ground glass nodules. Oncol Lett 2024;27(1):26.

30. Murugan R, Goel T, Mirjalili S, Chakrabartty DK. WOANet: Whale optimized deep neural network for the classification of COVID-19 from radiography images. Biocybern Biomed Eng 2021;41(4):1702-1718.
31. Shah V, Keniya R, Shridharani A, *et al*. Diagnosis of COVID-19 using CT scan images and deep learning techniques. Emerg Radiol 2021;28(3):497-505.
32. Wehbe RM, Sheng J, Dutta S, *et al*. DeepCOVID-XR: An artificial intelligence algorithm to detect COVID-19 on chest radiographs trained and tested on a large U.S. clinical data set. Radiology 2021;299(1):E167-E176.
33. Polsinelli M, Cinque L, Placidi G. A light CNN for detecting COVID-19 from CT scans of the chest. Pattern Recognit Lett 2020;140:95-100.
34. Hussain E, Hasan M, Rahman MA, *et al*. CoroDet: A deep learning based classification for COVID-19 detection using chest X-ray images. Chaos Solitons Fractals 2021;142:110495.
35. Jadhav A, Pujari S. Artificial intelligence defogging algorithm for chest CT scan images for post-COVID-19 patients infected with H3N2 virus. Int J Intell Syst Appl Eng 2023;12(2):119-128.
36. Song Y, Zheng S, Li L, *et al*. Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. IEEE/ACM Trans Comput Biol Bioinform 2021;18(6):2775-2780.
37. Chaddad A, Hassan L, Desrosiers C. Deep CNN models for predicting COVID-19 in CT and X-ray images. J Med Imaging 2021;8 Suppl 1:014502.
38. Wang C, Shao J, Xu X, *et al*. DeepLN: A multi-task AI tool to predict the imaging characteristics, malignancy and pathological subtypes in CT-detected pulmonary nodules. Front Oncol 2022;12:683792.
39. Dai J, Yu G, Yu J. Can CT imaging features of ground-glass opacity predict invasiveness? A meta-analysis. Thorac Cancer 2018;9(4):452-458.
40. Mishra M, Parashar V, Shimpi R. Development and evaluation of an AI System for early detection of Covid-19 pneumonia using X-ray (student consortium). 2020 IEEE Sixth International Conference on Multimedia Big Data. New Delhi: IEEE; 2020.
41. Voulodimos A, Protopapadakis E, Katsamenis I, *et al*. A few-shot U-Net deep learning model for COVID-19 infected area segmentation in CT images. Sensors 2021;21(6):2215.
42. Ye W, Gu W, Guo X, *et al*. Detection of pulmonary ground-glass opacity based on deep learning computer artificial intelligence. Biomed Eng Online 2019;18(1):6.
43. Zhang S, Chen X, Zhu Z, *et al*. Segmentation of small ground glass opacity pulmonary nodules based on Markov random field energy and Bayesian probability difference. Biomed Eng Online 2020;19(1):51.
44. Zhang X, Zhang Y, Zhang G, *et al*. Deep learning with radiomics for disease diagnosis and treatment: Challenges and potential. Front Oncol 2022;12:773840.
45. Braghetto A, Marturano F, Paiusco M, *et al*. Radiomics and deep learning methods for the prediction of 2-year overall survival in LUNG1 dataset. Sci Rep 2022;12(1):14132.
46. Rajaraman S, Candemir S, Kim I, *et al*. Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs. Appl Sci 2018;8(10):1715.
47. Kermany DS, Goldbaum M, Cai W, *et al*. Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell 2018;172(5):1122-1131.e9.
48. Chouhan V, Singh SK, Khamparia A, *et al*. A novel transfer learning based approach for pneumonia detection in chest X-ray images. Appl Sci 2020;10(2):559.
49. Huang T, Xu H, Wang H, *et al*. Artificial intelligence for medicine: Progress, challenges, and perspectives. Innov Med 2023;1(2):100030.
50. Liu M, Wu J, Wang N, *et al*. The value of artificial intelligence in the diagnosis of lung cancer: A systematic review and meta-analysis. PLoS One 2023;18(3):e0273445.
51. Thompson NC, Greenewald K, Lee K, Manso GF. The computational limits of deep learning. arXiv 2020;10:2007.05558v2.
52. Aksu G, Güzeller CO, Eser MT. The effect of the normalization method used in different sample sizes on the success of artificial neural network model. Int J Asst Tools Educ 2019;6(2):170-192.
53. Rani KV. Lung cancer classification using radial basis function based probabilistic neural networks. J Commun Eng Innov 2018;4(1):1-10.
54. Kido S, Hirano Y, Mabu S. Deep learning for pulmonary image analysis: Classification, detection, and segmentation. Adv Exp Med Biol 2020;1213:47-58.
55. Xiao Z, Liu B, Geng L, *et al*. Segmentation of lung nodules using improved 3D-UNet neural network. Symmetry 2020;12(11):1787.

56. El-Regaily SA, Salem MA, Aziz MHA, Roushdy MI. Survey of computer aided detection systems for lung cancer in computed tomography. Curr Med Imaging Rev 2008;14:3-18.
57. Tsakok MT, Mashar M, Pickup L, *et al*. The utility of a convolutional neural network (CNN) model score for cancer risk in indeterminate small solid pulmonary nodules, compared to clinical practice according to British Thoracic Society guidelines. Eur J Radiol 2021;137:109553.
58. Paez R, Kammer MN, Balar A, *et al*. Longitudinal lung cancer prediction convolutional neural network model improves the classification of indeterminate pulmonary nodules. Sci Rep 2023;13(1):6157.
59. Pinto-Coelho L. How artificial intelligence is shaping medical imaging technology: A survey of innovations and applications. Bioengineering 2023;10(12):1435.
60. Esmaeilzadeh P. Use of AI-based tools for healthcare purposes: A survey study from consumers' perspectives. BMC Med Inform Decis Mak 2020;20(1):170.
61. Jeyaraman M, Balaji S, Jeyaraman N, *et al*. Unraveling the ethical enigma: Artificial intelligence in healthcare. Cureus 2023;15(8):e43262.
62. Khan B, Fatima H, Qureshi A, *et al*. Drawbacks of artificial intelligence and their potential solutions in the healthcare sector. Biomed Mater Devices 2023:1-8.
63. Ibrahim A, Primakov S, Beuque M, *et al*. Radiomics for precision medicine: Current challenges, future prospects, and the proposal of a new framework. Methods 2021;188:20-29.
64. Arora A, Alderman JE, Palmer J, *et al*. The value of standards for health datasets in artificial intelligence-based applications. Nat Med 2023;29(11):2929-2938.