Open Access

The added value of using artificial intelligence in adult chest X-rays for nodules and masses detection in daily radiology practice

Suzan Farouk¹⁽¹⁾, Ahmed M. Osman^{1*}⁽¹⁾, Shrouk M. Awadallah¹⁽¹⁾ and Ahmed S. Abdelrahman¹⁽¹⁾

Abstract

Background Pulmonary nodule detection in CXR is challenging. Recently, the use of artificial intelligence (AI) has been a major attraction. The current study aimed to evaluate the diagnostic performance of the AI in the detection of pulmonary nodules or masses on CXR compared to the radiologist's interpretation and to assess its impact on the reporting process. The current study included 150 patients who had CXR interpreted by radiologists and by AI software.

Results CT detected pulmonary nodules in 99 cases (66%) while the visual model of analysis, as well as AI, detected nodules among 92 cases (61.3%) compared to 93 (62%) cases detected by combined visual/AI model. A total of 216 nodules were detected by CT (64.4% solid and 31.5% GG). Only 188 nodules were detected by the AI while 170 nodules were detected by visual analysis. As per case classification or nodule analysis, the AI showed the highest area under curve (AUC) (0.890, 95% CI) and (0.875, 95% CI), respectively, followed by the combined visual/AI model. Regarding the nodules' texture, the AI model's sensitivity for solid nodules was 91.4% which was greater than the combined visual/AI and visual models alone, while in GG nodules, the combined visual/AI model's sensitivity was higher than the AI and visual models. The probability score using the combined visual/AI model was significantly higher than using the visual model alone (P value = 0.001).

Conclusions The use of the AI model in CXR interpretation regarding nodules and masses detection helps in more accurate decision-making and increases the diagnostic performance affecting the patient's morbidity and mortality.

Keywords Artificial intelligence (AI), Nodules, Masses, Chest X-ray (CXR)

Background

Chest radiography (CXR) is usually the first step in the assessment of almost any patient complaining of respiratory symptoms, in a pre-occupational setting and even in the routine follow-up of most cancer patients [1]. Also, multiple large clinical trials have been carried out for assessing its possible use as a screening tool for lung

Ahmed M. Osman

¹ Faculty of Medicine, Ain Shams University, Abassia, Ramses street, Cairo 11591, Egypt cancer detection in high-risk populations, especially with its low cost and availability [2].

Pulmonary nodule detection in CXR is still challenging as some nodules can be obscured by the anatomical structures while others not detected because of the nodule size and nature [3]. This might explain the results of the National Lung Screening Trial (NLST) which revealed the lower sensitivity of the CXRs compared to low-dose CT (LDCT) in nodule detection [4].

LDCT is considered the screening modality for lung cancer among high-risk populations in most countries [5]. However, LDCT has higher cost and a higher radiation dose compared with CXR with higher false-positive



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

^{*}Correspondence:

Dr_osman80@yahoo.com; Dr_aosman@med.asu.edu.eg

nodules detection leading to additional laboratory workup and increased patient anxiety [6].

CXR can play an important role in detecting lung cancers presenting as incidental pulmonary nodules with a better prognosis as they are likely to have earlier stages of cancer and smaller tumors [7].

Recently, the artificial intelligence (AI) algorithms for the detection of pulmonary nodules in CXRs images by optimizing data based on previous observations and deep learning techniques have been a major attraction for medical applications [8] Some studies have demonstrated the ability of the AI algorithm to improve speed and accuracy in image reading, increase the detection rate of pulmonary nodules, and reduce radiologist errors [9, 10].

Despite the wide steps in the research on the use of such artificial intelligence algorithms in thoracic imaging have been taken over the last few decades, we rarely see them in clinical practice [8].

Aim of the work

This study aimed to *primarily* evaluate the diagnostic performance and efficacy of the artificial intelligence (AI) software for automatic detection of pulmonary nodules or masses on CXR compared to the radiologist's interpretation using the computed tomography (CT) as a reference being more accurate compared to CXR.

Furthermore, we aimed to assess the impact of using the AI software in the diagnostic performance of the reporting process.

Methods

Patients

This was a retrospective study that was done in the Radiology Department of Ain Shams University. A total of 150 patients who came to do CXR followed by chest CT over the period from 05/2021 to 07/2022 were retrospectively selected from the PACS. The Ethical Committee of the Faculty of Medicine- Ain Shams University approved this retrospective study and the use of patients' data archived in the picture archiving and communication system (PACS). The informed consent was waived being a retrospective study.

The inclusion criteria were any patients above 18 years old who did a CXR for any respiratory complaint or preoccupational. Also, patients with known cancer elsewhere in the body came from the oncology department for lung metastases screening. No sex predilection.

The chest radiographs included in this study comprised patients without as well as with different lung diseases such as mild free and encysted effusion, cardiomegaly, consolidation, emphysema, sub-segmental and segmental atelectasis as well as calcified granuloma to be more matching with the encountered in real clinical practice. The excluded patients were patients with poor image quality, image artifacts caused by any instrument such as wires, and patients with massive pleural effusion, or massive fibrosis. Also, patients with more than 6 nodules were excluded (maximum one nodule per lung zone on either side was the best statistically for better reflection of the nodule detection and assigning the nodules to lung zone in statistics). Any patients with unavailable CT within 2 weeks from the CXR were not included in the study. Lastly, patients with nodules <5 mm were excluded (being insignificant).

Chest X-ray (CXR)

All the CXRs were selected from the PACS fulfilling the inclusion criteria and all were in a postero-anterior view (PA view). The CXR was reviewed by two radiologists with 6–8 years of experience in consensus. A training set of 20 chest radiographs unrelated to the study collected data was given to each radiologist to get the radiologist more familiar with the study. The radiologists were instructed to mark any significant nodule (by a circle) in the chest radiographs including pulmonary nodule (nodule = / > 5mm, non-calcified), pleural nodule/nodular thickening, or hilar lymphadenopathy. If multiple nodules are noted, they were instructed to highlight the most obvious nodules. The following points were required for the CXR assessment:

- Presence or absence of pulmonary nodules or masses.
- The number and location (right versus left and the affected upper/middle/lower lung zone) of the pulmonary nodules or masses.
- The degree of confidence regarding the diagnosis of a nodule presence: confidence score (1 to 4) for each nodule based on how confident they were about the nodule [1=uncertain (0-25% abnormality confidence score), 2=mildly confident (26-50%), 3=moderately confident (51-75%), 4=very confident (76-100%)]. The normalized confidence score for the readers was calculated by the mean confidence score for both readers.
- The data obtained from the radiologists alone were collected.

After 3 weeks the radiologists were given the same set of radiographs in different randomized arrangements and this time the results of the AI algorithm analysis were exposed to them while reading. Again the results of their reading after the aid of the AI algorithm were collected (nodules number, location, and confidence score).

Artificial intelligence (AI) software

- In the current study, Lunit INSIGHT CXR (version 2.4.11.0; Lunit Inc.) was used which is a commercially available deep learning-based algorithm that is FDA cleared as well as Korea Food and Drug Administration approved [11]. This version was installed in our department in 2022. It is a deep convolutional neural network that uses residual neural network 34 (ResNet-34) as its foundational architecture with a self-attention mechanism to generate a more distinguishable image representation [12]. This algorithm was trained using 72,704 normal CXRs and 12,408 abnormal CXRs with lung nodules or masses obtained from multiple hospitals in South Korea that included both digital and computed radiographs annotated by at least 1 of 15 board-certified radiologists [13, 14].
- This AI algorithm uses Digital Imaging and Communications in Medicine (DICOM) file as input and produces for each radiograph a heat (abnormality) colored hue map overlaid on the input chest radiographs identifying the locations of the abnormality (using ascending color scale blue, green, yellow and red according to the probability) as well as a separate abnormality score (0–100%) for each of these 10 CXR abnormalities including atelectasis, calcification, cardiomegaly, consolidation, fibrosis, mediastinal widening, nodule or mass, pleural effusion, pneumoperitoneum, and pneumothorax. This algorithm does not produce an output map when the abnormality score is below 15% (the operating point chosen using Youden criteria in the internal validation set) [15, 16].
- The entire studied CXRs underwent assessment using the AI software without any radiologist interference or modification. The results were recorded separately by another radiologist after being instructed to annotate each nodule appearing on the probability color hue map by its location and color hue considering the color hue to be representative of the probability/confidence of AI for the diagnosis of the presence of each nodule [giving a corresponding number and probability range to each color score as follows: Blue=1 (0–25%), green=2 (26–50%), yellow=3 (51–75%), red=4 (76–100%)].

CT of the chest

CT chest studies were obtained within 2 weeks duration from the CXR and were used as a standard reference for validation of the radiologists' and AI algorithms results. The CT images were revised by two radiologists with 6–8 years of experience in consensus. The following parameters were recorded:

- Presence or absence of pulmonary nodules or masses.
- The number, location, size, and consistency of the detected nodules.

Statistical analysis

Data were analyzed using Statistical Package for Social Science (IBM SPSS statistics for windows, V. 25.0. Armonk, NY. USA). The qualitative data were described as numbers and percentages. The nonparametric continuous data were described as median and range. The performance of the different models was evaluated using the area under the receiver operating characteristic (ROC) curve (AUC) and the paired-sample area difference was performed to determine the difference between the two models. Cross-tabulation was used to assess the diagnostic performance of the categorical variable. Following that, the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy were calculated and the 95% confidence interval (CI) was determined. Mann-Whitney U test was used to evaluate the differences in nonparametric ordinal variables. All P values were two-tailed and P value < 0.05 was considered statistically significant.

Results

The current study included 150 patients having CXR with a patient median age of 48 years (age range: 18 to 83 years). Male patients represented 54% of the patients. CT detected pulmonary nodules in 99 cases representing 66% of the cases while the visual model of analysis, as well as AI analysis, detected nodules among 92 (61.3%) cases compared to 93 (62%) cases detected by combined visual/AI model. In all different models of interpretations, the solitary lung nodule and the right lung lobe site were the most common finding (Figs. 1, 2, 3, 4 and 5).

According to per nodule analysis, 216 nodules were detected in the CT examination with a median nodule diameter of 10 mm (5–89 mm), and the most common nodule texture was solid (64.4%) followed by GG nodule (31.5%). This was compared to 188 nodules that were detected by the AI model and only 170 nodules detected by visual CXR analysis (Table 1).

As per nodule analysis, the AI model had the highest AUC (0.875, 95% CI:0.820–0.930), followed by the visual/AI model with AUC (0.834, 95%: 0.771–0.897), while the visual model had the lowest AUC (0.772, 95%: 0.701–0.844) (Fig. 1), the paired-sample area difference revealed that the AUC of the visual model is consistently



Fig. 1 A Chest radiograph PA view showing right lower lung zone nodules (Marked by a radiologist). B Chest radiograph PA view interpreted by AI showing right lower lung zone nodules. C and D Chest CT axial cuts showing right and left variable-sized lower lung lobes nodules

lower than that of the AI model (P=0.002) and visual/AI (P=0.002), yet no significant difference in the AUC between AI and visual/AI models (P=0.185) (Table 2).

As per case (CXR) classification, the AI model showed the highest AUC (0.890, 95% CI:0.830–0.951), followed by the combined visual/AI model with AUC (0.866, 95%: 0.799–0.933), while the visual model showed the lowest AUC (0.816, 95%: 0.740–0.893) (Fig. 6), the paired-sample area difference showed that the AUC of the visual model is statistically lower than that of AI model (*P* value=0.037) and visual/AI model (*P* value=0.023), yet no significant difference in the AUC between AI and visual/AI models (*P* value=0.437) (Table 2).

As per the case (CXR) analysis, the diagnostic performance parameters of the AI model were found to be the highest followed by the combined visual/AI model with the visual model showing the least diagnostic performance. This was also the same as per nodule analysis (Table 3).

Regarding the texture and nature of the nodule, the AI model's sensitivity for solid nodule localization was 91.4% which was greater than the combined visual/AI and visual models, while in GG nodules, the combined visual/AI model's sensitivity was higher than the AI and visual models (Table 4).

According to nodule localization; the right middle lung zone was the highest zone with lung nodule prevalence in CT, AI, visual/AI, and visual models (20.8%, 21.3%, 22.1%, and 22.4%, respectively), while the left lower lung zone was the lowest zone (14.6%, 11.2%, 11.6%, and 12.9%, respectively). There was no statistically significant difference in the distribution of lung nodules among visual and AI models or between visual and combined visual/AI models, but there was a statistically significant difference in the distribution of lung nodules between the visual model and CT (P<0.001) (Table 5).

According to the study's findings, the probability score using the combined visual/AI model was significantly higher than using the visual model alone (P value = 0.001) (Table 6).

Discussion

The current applications of AI algorithms in radiology have demonstrated their efficacy in better, rapid, and more accurate reporting with this study aiming primarily at evaluating the diagnostic performance and efficacy of AI software in automatic detection of pulmonary nodules or masses on CXR compared to the radiologist visual reading using the CT as a reference. Second, we assessed the impact of using the AI software as a helping tool



Fig. 2 A Chest radiograph PA view showing no definite nodules. B Chest radiograph PA view interpreted by AI showing low probability for nodules. C Chest CT axial cuts showing no lung nodules

with the radiologist during the reporting process for the detection of chest nodules and masses in the CXR.

The results of the current study showed that AI software had better efficacy than the radiologists' visual analysis alone in the detection of the pulmonary nodules and masses in CXR and it also improved the radiologists' performance in detecting these lesions when used in combination with the visual analysis.

The nodule incidence in the current study was 66% using the CT, 61.3% using the AI, and 62% using combined visual/AI model. This was higher as compared to Hendrix et al. [17] who reported 50% incidence of pulmonary nodule in 2019. The higher incidence on the current study can be explained by the sample size characteristics with the presence of patients with cancer and came for pulmonary metastases and this raised the incidence of nodules.

The median nodule diameter in this study was 10 mm (5–89 mm) and this was different from the results of Liang et al. [1] who detected pulmonary nodules/ masses with an average diameter of 4.872 cm with

limited ability to detect smaller lesions and this can be explained by the use of different AI algorithms in this study.

Regarding the AUC using per case (CXR) classification or even per nodule classification, the AI model showed the highest AUC with 0.890 (0.830–0.951) and 0.875 (0.820–0.930), respectively, and *P* value < 0.001, followed by the combined visual/AI model.

These results were almost similar to the study done by Choi et al. [18] which showed the improvement of performance in image classification using the AI models with AUC, 0.8679-0.9285; *P* value=0.006 as well as lesion localization with AUC, 0.8426-0.9112; *P* value=0.0287. This was also close to Jin et al. [19] that showed that the AI solution itself showed acceptable performance with AUC, 0.863-0.873; also the study of Hwang et al. [14] showed improvements in diagnostic performance in both image-wise classifications (AUC, 0.814-0.932 to 0.904-0.958; all *P*<0.005) and lesion-wise localization (AUC, 0.781-0.907 to 0.873-0.938; all *P*<0.001) with the assistance of the algorithm.



Fig. 3 A CXR PA view showing a right hilar mass (Marked by a radiologist with a confidence score of (1). B CXR PA interpreted by AI showing a right hilar mass (abnormality score 32%, level (2) and right pleural effusion. C Chest CT axial cuts showing a right hilar mass and pleural effusion

As per the case (CXR) classification, we found that the sensitivity of the radiologists to detect pulmonary nodules improved from 82.8 to 86.9%, also the specificity improved from 80.4 to 86.3%. Also, per nodule detection the sensitivity, specificity PPV, NPV, and accuracy were better in the combined visual/AI model (80.6%, 86.3%, 96.1%, 51.2%, and 81.7%, respectively), compared to visual models alone (74.1%, 80.4%, 94.1%, 42.3%, and 75.3%, respectively).

This was in agreement with Choi et al. [18] results for lesion detection which showed improvement in radiologists' sensitivity from 78.13% to 88.75% for nodule/mass and the study of Sim et al. [10] who revealed an increase in the sensitivity of radiologists from 65.1 to 70.3% with a decline in the number of false-positive findings (from 0.2 to 0.18, P<0.001) when the radiologists re-reviewed radiographs with the deep learning algorithm for pulmonary nodules.

Nam et al. [8] also reported that the AI algorithm system showed significantly higher specificity than that of the pooled radiologists (94% vs 78%; P=0.01). The sensitivity and specificity demonstrated by the radiologists were significantly improved when they were aided by the algorithm (P<0.001 and P=0.01, respectively), and the algorithm showed good sensitivity that was higher than that of the thoracic radiologists (69.6% vs 47.0%; P<0.001) as well as higher specificity (94.0% vs 78.0%; P=0.01).

The current study also showed that the diagnostic performance parameters of the AI were the highest compared to the visual analysis alone or even the combined visual/AI model with acceptable good performance compared to the CT results and this agreed with the study of Kim et al. [20] that demonstrated an overall fair to good AI stand-alone performance to determine the presence of referable thoracic abnormalities.



Fig. 4 A Chest radiograph PA view showing a right middle lung zone mass (Marked by a radiologist). B CXR PA view interpreted by AI showing a right middle lung zone mass. C Chest CT axial cuts showing a right lower lung lobe mass



Fig. 5 A Chest radiograph PA view showing right lower lung zone nodules (Marked by a radiologist). B Chest radiograph PA view interpreted by AI showing right lower lung zone nodules (probability level 4) and left lower lung zone nodules (probability level 2). C Chest CT axial cuts showing right and left variable-sized lower lung lobes nodules

Regarding the nature of the nodules detected, the AI model showed the highest sensitivity for solid nodule localization, while the visual/AI model showed a higher

sensitivity for GG nodule localization. Lastly, for the cavitating nodules, the AI and AI/visual models had the same sensitivity which was higher than the visual model alone.

	СТ	AI	Visual/Al	Visual
Age (median and range)	48 (18–83)			
Sex				
Male	81 (54%)			
Female	69 (46%)			
Nodule size (median and range)	10 (5–89)			
Per Case (CXR) classification				
Positive	99/150 (66%)	92/150 (61.3%)	93/150 (62%)	92/150 (61.3%)
Number of nodules				
1 nodule	62/99 (62.6%)	57/92 (62%)	61/93 (65.6%)	62/92 (67.4%)
2 nodules	12/99 (12.1%)	12/92 (13%)	9/93 (9.7%)	9/92 (9.8%)
3 nodules	4/99 (4%)	5/92 (5.4%)	8/93 (8.6%)	9/92 (9.8%)
4 nodules	3/99 (3%)	6/92 (6.5%)	4/93 (4.3%)	4/92 (4.3%)
5 nodules	2/99 (2%)	4/92 (4.3%)	4/93 (4.3%)	1/92 (1.1%)
6 nodules	16/99 (16.3%)	8/92 (8.7%)	7/93 (7.5%)	7/92 (7.6%)
CXR nodule laterality				
Right lung lobe	41/99 (41.4%)	42/92 (45.7%)	40/93 (43%)	39/92 (42.4%)
Left lung lobe	36/99 (36.4%)	30/92 (32.6%)	33/93 (35.5%)	34/92 (37%)
Bilateral lung lobes	22/99 (22.2%)	20/92 (21.7%)	20/93 (21.5%)	19/92 (20.6%)
Number of the nodule	216	188	181	170
Nodule texture				
Solid	139/216 (64.4%)			
GG	68/216 (31.5%)			
Cystic lesion	9/216 (4.1%)			

Table 1 Demographic data of the patients, number, and characteristics of the nodule using different models of interpretation

Table 2 Comparison between the AUC of different models of interpretation using per case (CXR) and per nodule analysis

	AUC (95% CI)	<i>P</i> value	Al vs visual	Al versus visual/Al	Visual versus visual/Al
Per CXR			0.037*	0.437	0.023*
AI	0.890 (0.830-0.951)	< 0.001*			
Visual/Al	0.866 (0.799–0.933)	< 0.001*			
Visual	0.816 (0.740-0.893)	< 0.001*			
Per nodule			0.002*	0.185	0.002*
AI	0.875 (0.820–0.930)	< 0.001*			
Visual/Al	0.834 (0.771–0.897)	< 0.001*			
Visual	0.772 (0.701–0.844)	< 0.001*			

*Significant

And this agreed with the study of Liang et al. [1] who reported that of the 47 pulmonary nodules in their study, 39 (82.97%) were solid nodules and eight (17.02%) were part-solid nodules, and Nam et al. [8] whom study showed 80% solid nodules (149 out of 187) versus 14% subsolid nodules (26 of 187) and 6.4% cavitations [12].

For the location of lung nodules, the right middle lung zone was the highest in all different models while the left lower lung zone had the lowest prevalence. This was quite different from the study of Nam et al. [8] who showed that the RLL had the highest number of nodules detected and the RML had the lowest number. But we agreed with them regarding the AI higher detection rate than readers for obscured nodules and pooled readers successfully detected more than 10% more nodules when assisted with the algorithm for the hilar and apical nodules.

The degree of confidence (the probability score) in the current study was significantly higher with the visual/AI



Fig. 6 The ROC curve for the different models of interpretation using the per case (CXR) analysis (A) and per nodule analysis (B)

models	lable 3	The diagnostic	performance	of the differer	nt parameters o	f CXR analysis	including the A	l, combined visi	Jal/Al, and the visual
	models								

	Sensitivity	Specificity	PPV	NPV	Accuracy
Per CXR					
AI	87.9% (87/99) [80–94]	90.2% (46/51) [79–97]	94.6% (87/92) [88–88]	79.3% (46/58) [69–87]	88.7% (133/150) [82–93]
Visual /AI	86.9% (86/99) [79–93]	86.3% (44/51) [74–94]	92.5% (86/93) [86–96]	77.2% (44/57) [67–85]	86.7% (130/150) [80–92]
Visual	82.8% (82/99) [74–90]	80.4% (41/51) [67–90]	89.1% (82/92) [82–94]	70.7% (41/58) [61–79]	82% (123/150) [75–88]
Per nodule*					
AI	84.7% (183/216) [79–89]	90.2% (46/51) [79–97]	97.3% (183/188) [94–99]	58.2% (46/79) [50–66]	85.8% (229/267) [81–90]
Visual /Al	80.6% (174/216) [75–86]	86.3% (44/51) [74–94]	96.1% (174/181) [93–98]	51.2% (44/86) [44–58]	81.7% (218/267) [76–86]
Visual	74.1% (160/216) [68–80]	80.4% (41/51) [67–90]	94.1% (160/170) [90–97]	42.3% (41/97) [36–49]	75.3% (201/267) [70–80]

Numbers in parentheses are proportions; numbers in brackets are 95% confidence intervals

*For per nodule localization, the number of negative CXR was used as the denominator for specificity

Table 4 The sensitivity and false negative rate of the AI, combined visual/AI, and Visual models according to CT nodule texture

	Solid nodule		GG nodule		Cystic nodule		
	Sensitivity	False negative rate	Sensitivity	False negative rate	Sensitivity	False negative rate	
AI	91.4% (127/139)	8.6% (12/139)	70.6 (48/68)	29.4% (20/68)	88.9% (8/9)	11.1% (1/9)	
Visual/Al	83.5% (116/139)	16.5% (23/139)	73.5% (50/68)	26.5% (18/68)	88.9% (8/9)	11.1% (1/9)	
Visual	79.1% (110/139)	20.9% (29/139)	63.2% (43/68)	36.8% (25/68)	77.8% (7/9)	22.2% (2/9)	

model than that with the visual model alone. This was also demonstrated by Liang et al. [1] who reported that for plain chest films with the lowest probability score, radiologists can exclude lesions with greater confidence.

The difference of the results between the visual analysis and AI can be explained by the presence of blind or obscured areas on chest radiographs for both the radiologist and AI algorithms and this is a limitation of the X-ray itself. These obscured areas showed more limitation to the visual analysis. Also, AI due to the deep learning showed higher ability to detect less dense nodules and smaller nodules compared to the visual assessment.

Nodule location	CT (216)	AI (188)*	Visual/AI (181)*	Visual (170)*	Visual vs CT	Visual vs Al	Visual vs visual/Al
					< 0.001	0.180	0.130
Right upper zone	39/216 (18%)	38/188 (20.2%)	37/181 (20.4%)	32/170 (18.8%)			
Right middle zone	45/216 (20.8%)	40/188 (21.3%)	40/181 (22.1%)	38/170 (22.4%)			
Right lower zone	28/216 (12.9%)	27/188 (14.4%)	22/181 (12.2%)	23/170 (13.5%)			
Left upper zone	37/216 (17.1%)	34/188 (18%)	32/181 (17.7%)	29/170 (17.1%)			
Left middle zone	36/216 (16.6%)	28/188 (14.9%)	29/181 (16%)	26/170 (15.3%)			
Left lower zone	31/216 (14.6%)	21/188 (11.2%)	21/181 (11.6%)	22/170 (12.9%)			

Table 5 The prevalence of lung nodules among different lung zones in CT and the other different models of analysis

*The total number of lung nodules detected by models represents TP + FP

To our knowledge, there was no sufficient number of studies that focused on the radiologist's diagnostic performance with and without the aid of AI algorithms in the area of Africa. However, several limitations were encountered in the current study being a single-center retrospective study. Also, we studied only pulmonary nodules ignoring the other pulmonary abnormalities. Also, nodule characterization either benign versus malignant was not included in the aim. Lastly, we studied chest radiographs using the PA projection, without lateral or any other added projections and this may be one of the reasons for the lower diagnostic performance of the radiologists alone compared to the AI models.

Conclusions

The performance of the AI model in CXR interpretation regarding nodules and masses detection is acceptable if used alone or even as an assistant to the radiologist as it enhances the radiologist's confidence and helps in more accurate decision-making in real clinical situations. It increases the diagnostic performance of the CXR interpretation which of course will make a positive impact on the patient's morbidity and mortality. Using AI will not replace radiologists but it will improve the diagnostic

Table 6 The probability score among visual versus combined visual/AI models of nodule analysis

Probability score	Visual	Visual/AI	Z score	P value
			3.675	< 0.001
0%	97 /267 (36.3%)	86/267 (32.2%)		
1–25%	53/267 (19.9%)	22/267 (8.2%)		
26–50%	53/267 (19.9%)	52/267 (19.5%)		
51–75%	47/267 (17.6%)	54/267 (20.2%)		
76–100%	17/267 (6.3%)	53/267 (19.9%)		

performance and enhance the diagnostic accuracy especially with the uprising in the workload.

Abbreviations

Al	Artificial intelligence
AUC	Area under the curve
CT	Computed tomography
CXR	Chest X-ray
GG	Ground glass
LDCT	Low Dose Computed tomography
PACS	Picture archiving and communication system

Acknowledgements

Not applicable.

Author contributions

SF contributed to the idea, organization, data collection, and writing. AMO contributed to the idea, writing, and corresponding author. SMA helped in data collection and writing. ASA involved in statistical analysis and helped in writing and revision.

Funding

All authors declare that no funding or any source of financial interest. All authors declare that: they all read and approved the final version of the submitted manuscript. The authors declare that this is an original work, hasn't received prior publication, and isn't under consideration for publication elsewhere. The authors confirm sole submission to Egyptian Journal of Radiology and Nuclear Medicine. They all confirm that the article is not under consideration for publication elsewhere. All authors declare no subject overlap with previously published work.

Availability of data and materials

Available on request with the corresponding author.

Declarations

Ethics approval and consent to participate

Ethics approval were taken from our institute ethical committee (Faculty of Medicine-Ain shams university) and being a retrospective study, the informed written consents were waived.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 20 May 2023 Accepted: 19 August 2023 Published online: 25 August 2023

References

- Liang CH, Liu YC, Wu MT et al (2020) Identifying pulmonary nodules or masses on chest radiography using deep learning: external validation and strategies to improve clinical practice. Clin Radiol 75(1):38–45. https://doi.org/10.1016/j.crad.2019.08.005
- Yoo H, Kim KH, Singh R et al (2020) Validation of a deep learning algorithm for the detection of malignant pulmonary nodules in chest radiographs. JAMA Netw Open 3(9):e2017135–e2017135
- Gavelli G, Giampalma E (2000) Sensitivity and specificity of chest X-ray screening for lung cancer. Cancer 89:2453–2456. https://doi.org/10.1038/ sj.bjc.6604351
- Aberle DR, DeMello S, Berg CD et al (2013) Results of the two incidence screenings in the National Lung Screening Trial. N Engl J Med 369:920– 931. https://doi.org/10.1056/NEJMoa1208962
- De Koning H, Van Der Aalst C, Ten Haaf K et al (2018) Effects of volume CT lung cancer screening: mortality results of the NELSON randomizedcontrolled population-based trial. J Thorac Oncol 13(10):S185. https://doi. org/10.1016/j.jtho.2018.08.012
- Bach PB, Mirkin JN, Oliver TK et al (2012) Benefits and harms of CT screening for lung cancer: a systematic review. JAMA 307(22):2418–2429. https://doi.org/10.1001/jama.2012.5521
- Quadrelli S, Lyons G, Colt H et al (2015) Clinical characteristics and prognosis of incidentally detected lung cancers. Int J Surg Oncol. https://doi. org/10.1155/2015/287604
- Nam JG, Park S, Hwang EJ et al (2019) Development and validation of deep-learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs. Radiology 290(1):218–228. https://doi.org/10.1148/radiol.2018180237
- Ather S, Kadir T, Gleeson F (2020) Artificial intelligence and radiomics in pulmonary nodule management: current status and future applications. Clin Radiol 75(1):13–19. https://doi.org/10.1016/j.crad.2019.04.017
- Sim Y, ChungMJ KE et al (2020) Deep convolutional neural network-based software improves radiologist detection of malignant lung nodules on chest radiographs. Radiology 294(1):199–209. https://doi.org/10.1148/ radiol.2019182465
- Kim JH, Han SG, Cho A et al (2021) Effect of deep learning-based assistive technology use on chest radiograph interpretation by emergency department physicians: a prospective interventional simulationbased study. BMC Med Inf Decis Mak 21:311. https://doi.org/10.1186/ s12911-021-01679-4
- He K, Zhang X, Ren S et al (2016). Deep residual learning for image recognition. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR). Institute of Electrical and Electronics Engineers, pp 770–778. https://doi.org/10.1109/CVPR.2016.90
- Yoo H, Kim KH, Singh R et al (2020) Validation of a deep learning algorithm for the detection of malignant pulmonary nodules in chest radiographs. JAMA Netw Open 3:e2017135–e2017135. https://doi.org/10. 1001/jamanetworkopen.2020.17135
- Hwang EJ, Park S, Jin K-N et al, DLAD Development, and Evaluation Group (2019) Development and validation of deep-learning-based automated detection algorithm for major thoracic diseases on chest radiographs. JAMA Network Open. 2(3):e191095. https://doi.org/10.1001/jamanetwor kopen.2019.1095
- Yoo H, Lee SH, Arru CD et al (2021) Al-based improvement in lung cancer detection on chest radiographs: results of a multi-reader study in NLST dataset. Eur Radiol 31:9664–9674. https://doi.org/10.1007/ s00330-021-08074-7
- Nam JG, Hwang EJ, Kim DS et al (2020) Undetected lung cancer at posteroanterior chest radiography: potential role of a deep learning–based detection algorithm. Radiol Cardiothorac Imaging 2(6):e190222. https:// doi.org/10.1148/ryct.2020190222
- Hendrix W, Rutten M, Hendrix N et al (2023) Trends in the incidence of pulmonary nodules in chest computed tomography: 10-year results from two Dutch hospitals. Eur Radiol. https://doi.org/10.1007/ s00330-023-09826-3

- Choi SY, Park S, Kim M et al (2021) Evaluation of a deep learning-based computer-aided detection algorithm on chest radiographs: case-control study. Medicine 100(16):e25663. https://doi.org/10.1097/MD.000000000 025663
- Jin KN, Kim EY, Kim YJ et al (2022) Diagnostic effect of artificial intelligence solution for referable thoracic abnormalities on chest radiography: a multicenter respiratory outpatient diagnostic cohort study. Eur Radiol 32(5):3469–3479. https://doi.org/10.1007/s00330-021-08397-5
- Kim EY, Kim YJ, Choi WJ et al (2021) Performance of a deep-learning algorithm for referable thoracic abnormalities on chest radiographs: a multicenter study of a health screening cohort. PLoS ONE. https://doi. org/10.1371/journal.pone.0246472

Publisher's Note

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

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- ► Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at > springeropen.com