Enhancing Diagnostic Accuracy in Lung Cancer: Evaluating the Performance of a Cutting-edge Artificial Intelligence Algorithm for Pulmonary Nodule Detection in Chest Radiographs

Mohsen Mehrabi (PhD)[1]*, Mojtaba Askari (PhD)[1]

[1]Radiation Application Research School, Nuclear Science and Technology Research Institute (NSTRI), Tehran, Iran

ABSTRACT

Background: Lung cancer remains a major global health challenge, and early diagnosis critically improves outcomes. Artificial Intelligence (AI) is increasingly used to enhance diagnostic accuracy in medical imaging, particularly for thoracic diseases.

Objective: This study evaluated a cutting-edge AI algorithm for detecting pulmonary nodules in chest radiographs to improve early lung cancer diagnosis.

Material and Methods: In this analytical and experimental investigation, two datasets were used: Group1 (112120 frontal-view chest radiographs labeled for 14 thoracic diseases, including nodules) and Group2 (800 radiographs evenly split into positive and negative cases, with expert consensus labeling). The proprietary AI algorithm is based on CheXNet, a 121-layer convolutional neural network pretrained on ImageNet to initialize weights, enabling transfer learning for superior feature extraction. This model was chosen for its validated performance in thoracic disease detection, particularly in chest X-ray analysis. The datasets were divided into Type A (combining Group1 and Group2) and Type B (using only Group2) to assess performance under different conditions. A five-fold cross-validation was employed.

Results: The model achieved an Area Under the Curve (AUC) of 0.73 for Type A and 0.782 for Type B. Sensitivity and specificity metrics were 0.736 and 0.592 for Type A, and 0.716 and 0.734 for Type B, respectively. Notably, the algorithm outperformed radiologists in some cases, demonstrating its capability to detect subtle or challenging nodules.

Conclusion: This AI algorithm shows promise as a clinical decision support tool, especially where access to experienced radiologists is limited. Further research is needed to refine the algorithm and validate use in real-world clinical workflows.

Keywords

Artificial Intelligence; Deep Learning; Diagnosis; Lung Cancer; Pulmonary Nodule Detection; Chest Radiograph Interpretation

Introduction

The rapid evolution of medical imaging in recent years can be attributed to the advancements in deep learning algorithms and the accessibility of extensive datasets. These advancements have caused algorithms to outperform medical professionals in various medical imaging tasks, including the identification of hemorrhage and diabetic retinopathy [1,2]. However, the early detection of lung cancer

*Corresponding author:
Mohsen Mehrabi
Radiation Application Research School, Nuclear Science and Technology Research Institute (NSTRI), Tehran, Iran
E-mail:
msmehrabi@aeoi.org.ir

Received: 12 October 2024
Accepted: 21 April 2025

remains a significant challenge in the medical field, as it continues to be the leading cause of cancer-related deaths worldwide [3].

To address this urgent issue, several clinical trials have been conducted using low-dose Computed Tomography (CT) scans, such as the Nederlands-Leuvens Screening ONderzoek (NELSON) trial and the National Lung Screening Trial [4,5]. Studies comparing Low-Dose Computed Tomography (LDCT) and chest radiographs have consistently shown the superior sensitivity of LDCT in detecting early-stage lung cancers, leading to reduced mortality rates in high-risk populations [6,7]. While CT scans have shown superior ability in detecting pulmonary nodules, thoracic X-ray continues to be the leading diagnostic tool for screening and detecting lung lesions [8-10].

Pulmonary nodules often serve as the early imaging indicators of pulmonary malignancy. Nonetheless, they may evade detection in conventional chest radiographs due to their subtle radiographic opacity, subcentimeter dimensions, or anatomical localization in regions obscured by superimposed structures (e.g., retrocardiac, hilar, or subdiaphragmatic zones), where limited tissue contrast and projective summation artifacts reduce diagnostic conspicuity. In recent years, numerous Computer-Aided Detection (CAD) studies have focused on the detection of lung opacities on thoracic X-ray [11,12].

Early CAD solutions had limitations regarding high false-positive rates and low sensitivity. Nevertheless, this study endeavors to overcome these constraints through the development of an innovative Artificial Intelligence (AI) algorithm and the assessment of its efficacy in detecting pulmonary nodules in chest radiographs across varying levels of complexity. This evaluation encompasses the utilization of both abnormal and normal control images. Recent advancements in AI have demonstrated that AI-assisted systems can significantly enhance pulmonary nodule detection accuracy, often surpassing the performance of radiologists in identifying challenging cases of subtle or small nodules [13,14].

The findings of the research were truly remarkable. The AI algorithm exhibited superior performance compared to the average capability of radiologists in identifying pulmonary nodules. This breakthrough indicates that automated disease detection using chest radiographs, at the level of experienced radiologists, has the potential to yield substantial advantages in a clinical environment.

This study aimed to achieve high diagnostic accuracy using a small, standardized dataset while demonstrating the model's generalizability across heterogeneous data environments. Additionally, the integration of enhanced visualization tools, such as Class Activation Maps (CAMs), provides a new dimension of interpretability, making this work a significant step toward clinical adoption.

Material and Methods

Model Formulation

This analytical and experimental investigation aimed to evaluate the diagnostic accuracy of a cutting-edge AI model for identifying lung nodules using chest radiographs. In this study, the CheXNet model, a convolutional deep network consisting of 121 layers [15], was employed to process chest X-ray images and to estimate lung nodule probability and generate localization heatmaps.

To train the CheXNet model, the Group1 dataset was utilized, which consists of 112120 anterior-view thoracic radiographs. These images are separately annotated with as many as 14 distinct chest conditions, such as edema, consolidation, emphysema, effusion, cardiomegaly, mass, pleural thickening, infiltration, hernia, atelectasis, nodule, pneumothorax, and pneumonia. For this study, a subset of the dataset containing 2500 nodules (positive samples) and 2500 normal records (negative samples) was selected.

To normalize the output of the network, a

sigmoid function was applied, resulting in values between 0 and 1. The CheXNet model was initialized with weights from the pre-trained ImageNet model [16]. Since this is a multi-label problem, the C binary cross-entropy loss function was employed to independently treat every annotation throughout the categorization, where C=14. To address the dataset's inherent class imbalance, additional weights were incorporated into the loss function, considering the frequency of each label within every batch.

The loss function used in this study can be represented as follows:

$$L(X,l_n)=-(w_p \bullet l_n \log(p)+w_n \bullet (1-l_n) \log(1-p)) \quad (1)$$

Here, $w_p=(P_n+N_n)\div P_n$ and $w_p=(P_n+N_n)\div N_n$, where P_n and N_n represent the number of samples with the presence and absence of nodules, respectively.

Training of the CheXNet model was performed end-to-end using Adam optimization with standard parameters (β1=0.9 and β2=0.999) and a minibatch size of 16. A learning rate of 0.00001 was initially set and reduced by a factor of 10 whenever the validation loss reached a plateau during each epoch. The model with the lowest validation loss was selected as the final model.

By leveraging the capabilities of the CheXNet model, significant progress has been made in pulmonary nodule detection. This model demonstrates promising results in accurately identifying nodules in chest X-ray images, which can significantly contribute to early diagnosis and treatment planning for patients.

Data Training and Model Development

In this study, two datasets were utilized to train and evaluate the AI model: Group1 and Group2. The Group1 dataset comprises 112120 frontal chest radiographs obtained from publicly available repositories, labeled for 14 thoracic diseases, including nodules, by automated systems and human experts. These labels were verified through multiple rounds of cross-checking to ensure reliability.

The Group2 dataset consists of 800 chest radiographs collected in a standardized clinical environment, where consistent imaging protocols were applied to minimize variability. This dataset is evenly balanced, with 400 images containing pulmonary nodules (positive cases) and 400 normal images (negative cases). The labeling of pulmonary nodules in Group2 was conducted through a rigorous consensus process involving three senior radiologists, ensuring the highest accuracy for lung cancer screening applications.

To analyze the data and train the Convolutional Neural Network (CNN), the datasets were categorized into two types:

Type A: A combined dataset of 5000 randomly selected chest radiographs from Group1 and all 800 images from Group2, representing a diverse range of imaging conditions and nodule characteristics.

Type B: A standalone dataset consisting solely of the 800 radiographs from Group2, designed to evaluate model performance in a controlled, standardized setting.

To ensure robustness and reliability, the dataset was randomly divided into five groups, each containing an equal proportion of positive and negative data. One group was assigned as the test data, while the remaining groups were used for training. By assigning each group as test data, five sets of results were obtained, enabling cross-validation and enhancing the credibility of the findings.

The ImageNet platform was employed to initialize the weights of the proprietary AI model, developed using the CheXNet architecture. A statistical analysis was then performed to assess the algorithm's performance in radiograph interpretation, with metrics, such as sensitivity, specificity, and AUC calculated for both dataset types. The separation into Type A and Type B allowed for a detailed comparison of the algorithm's robustness across varying dataset compositions.

The AUC of the AI algorithm's Receiver

Operating Characteristic (ROC) was compared with previously reported values [15,17,18], further validating its effectiveness in pulmonary nodule detection. Additionally, the precision of radiologists' assessments was presented in a multicenter study by Homayounieh et al., [19]. This evaluation test involved the participation of nine radiologists who assessed 50 radiographs with pulmonary nodules, 25 normal radiographs, and 25 radiographs with non-nodular abnormalities.

Through the utilization of these comprehensive datasets and the development of a proprietary AI algorithm, significant advancements have been made in pulmonary nodule detection. Integrating deep learning and convolutional neural networks has enabled accurate and efficient identification of pulmonary nodules in chest X-ray images. This breakthrough technology holds great promise for improving lung cancer screening and facilitating early detection, ultimately leading to enhanced patient outcomes.

Model Interpretability

To enhance the visualization of pulmonary nodules, a cutting-edge approach was employed in this study. The pulmonary nodules were depicted as heatmap displays on each chest radiograph, accompanied by a positive probability score, expressed as an index value ranging from 0.0 to 1.0. This score provided valuable insights into the likelihood of the presence of pulmonary nodules. To achieve this, a technique known as CAMs was utilized [20]. The CAMs have emerged as a critical tool in enhancing the interpretability of deep learning models, offering heatmap visualizations that guide clinicians by highlighting relevant regions associated with predictions [21,22].

The generation of CAMs involves feeding an image into a completely trained model and extracting the feature maps produced by the last convolutional layer. By computing a weighted sum of these feature maps (f_k) using their associated weights ($w_{c,k}$), the most salient features were identified. This process can be represented mathematically as follows:

$$M_c=\sum_k w_{c,k} \bullet f_k \qquad (2)$$

The resulting map M_c was up-sampled to match the original image's dimensions, and when overlaid, it highlighted the critical attributes utilized by the AI algorithm to forecast lung nodules. This innovative approach allowed a deeper understanding of the AI algorithm's decision-making process.

It should be emphasized that the machine learning process for our novel AI system involved training data obtained exclusively from the same radiographic apparatus. Through this approach, any potential influence arising from equipment variations was effectively mitigated, thereby ensuring the reliability and consistency of the AI system's performance.

By leveraging advanced visualization techniques, such as CAMs, our study has made significant strides in improving the accuracy and interpretability of pulmonary nodule detection. The ability to identify and visualize the most relevant features utilized by the AI model provides valuable insights for radiologists and researchers, ultimately leading to more precise and reliable diagnoses.

Results

In this study, the performance of the novel AI algorithm for pulmonary nodule detection was thoroughly assessed. The evaluation involved calculating the AUC, specificity, and sensitivity for each of the five categories across both the Type B and Type A data collections. To ensure robustness, cross-validation was employed, yielding reliable results.

The AUC, specificity, and sensitivity values for each group in both datasets are presented in Tables 1 and 2. These metrics provide valuable insights into the algorithm's performance in accurately identifying pulmonary nodules. To visualize the algorithm's discriminative ability, ROC curves were generated for both the Type B and Type A datasets, as depicted

Table 1: Accuracy after cross-validation for each outcome using the type A data collection.

Result No.	AUC	Specificity	Sensitivity
1	0.82	0.62	0.80
2	0.60	0.45	0.72
3	0.68	0.45	0.78
4	0.74	0.72	0.64
5	0.81	0.72	0.74
Average	0.73	0.592	0.736

AUC: Area Under the Curve

Table 2: Accuracy after cross-validation for each outcome using the type B data collection.

Result No.	AUC	Specificity	Sensitivity
1	0.84	0.74	0.80
2	0.63	0.69	0.51
3	0.82	0.76	0.78
4	0.78	0.72	0.71
5	0.84	0.76	0.78
Average	0.782	0.734	0.716

AUC: Area Under the Curve

in Figure 1.

For the Type A dataset, the novel AI approach demonstrated an AUC of 0.73, a specificity of 0.592, and a sensitivity of 0.736. Similarly, the Type B dataset exhibited promising results, with an AUC of 0.782, a specificity of 0.734, and a sensitivity of 0.716. These findings highlight the algorithm's effectiveness in distinguishing between positive and negative cases of pulmonary nodules.

To establish a threshold for classification, a positive probability cutoff value of 0.5 was employed by the AI algorithm. This threshold ensured a balanced approach in determining the presence of pulmonary nodules. Notably, the AI algorithm outperformed the accuracy of radiologists, as evidenced by an AUC of 0.71. Furthermore, the algorithm's performance was comparable or superior to previous reports [17,18].

Comparing the radiologist detection of pulmonary nodules, the AUC was determined to be 0.7173±0.0344, with a specificity of 0.9635±0.0198 and a sensitivity of 0.4710±0.0611 [19]. These results, presented in Table 3, emphasize the potential of the AI algorithm in improving diagnostic accuracy.

Table 4 demonstrates the competitive performance of the proposed AI algorithm, achieving an AUC of 0.782 with the Type B dataset, which contains only 800 images. This result underscores the algorithm's efficiency in extracting meaningful features from limited data, a crucial aspect when applying AI in resource-constrained clinical settings where large datasets may not be available. Compared to prior studies [15], which relied on datasets exceeding 100000 images, our study highlights the potential of achieving high diagnostic accuracy with significantly smaller yet carefully

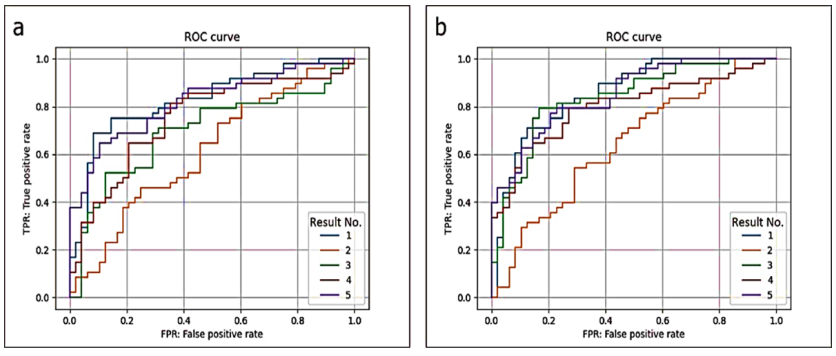

Figure 1: Receiver Operating Characteristic (ROC) curves for type A (**a**) and type B (**b**) datasets.

Table 3: Comparison of Artificial Intelligence (AI) algorithm detection accuracy with radiologist detection.

Investigation	AUC	Specificity	Sensitivity
Homayounieh et al. [19]	0.717±0.034	0.9635±0.0198	0.471±0.0611
Type A	0.730±0.042	0.592±0.028	0.736±0.041
Type B	0.782±0.045	0.734±0.032	0.716±0.038

AUC: Area Under the Curve

Table 4: Comparison of previous reports' Area Under the Curve (AUC) values with Artificial Intelligence (AI) algorithms

Investigation	AUC	Sensitivity	Specificity
Rajporkar et al. [15]	0.78	0.616	0.634
Yao et al. [18]	0.717	0.665	0.667
Wang et al. [17]	0.671	0.592	0.636
Type B	0.782	0.716	0.734
Type A	0.73	0.736	0.592

AUC: Area Under the Curve

curated datasets.

Furthermore, the consistent performance across Type A (AUC=0.73) and Type B (AUC=0.782) datasets demonstrates the robustness and adaptability of the algorithm. Type A represents a diverse dataset combining large-scale data (Group1) with standardized clinical data (Group2), while Type B exclusively uses the standardized Group2 dataset. This dual evaluation provides compelling evidence of the algorithm's generalizability across both diverse and controlled environments, addressing a key challenge in deploying AI systems in heterogeneous clinical workflows.

The integration of CAMs adds a novel dimension to the study by enhancing the interpretability of the algorithm's predictions. Unlike earlier studies that primarily focused on classification accuracy, our approach provides visual heatmaps highlighting the most relevant regions in the chest radiographs associated with pulmonary nodules. This interpretability is vital for clinical adoption, as it allows radiologists to understand and validate the AI model's decision-making process. The CAMs not only improve confidence in AI-assisted diagnosis but also facilitate the identification of subtle nodules that may otherwise be overlooked by human observers.

Overall, the combination of robust performance with limited data, strong generalization across different dataset compositions, and enhanced interpretability through CAMs establishes this study as a significant step forward in the application of AI for pulmonary nodule detection. These contributions address critical barriers to integrating AI into real-world clinical practice, offering a practical and scalable solution to improve diagnostic accuracy in lung cancer screening.

In this study, lung nodules were seen using heatmaps shown on a monitor screen. These heatmaps indicated the location of any lung nodules present in each roentgenogram (Figure 2). It is vital to highlight that heatmaps were generated based on the AI algorithm's analysis of radiographic images.

The heatmaps exhibited a direct correspondence to the location of pulmonary nodules, effectively highlighting their presence.

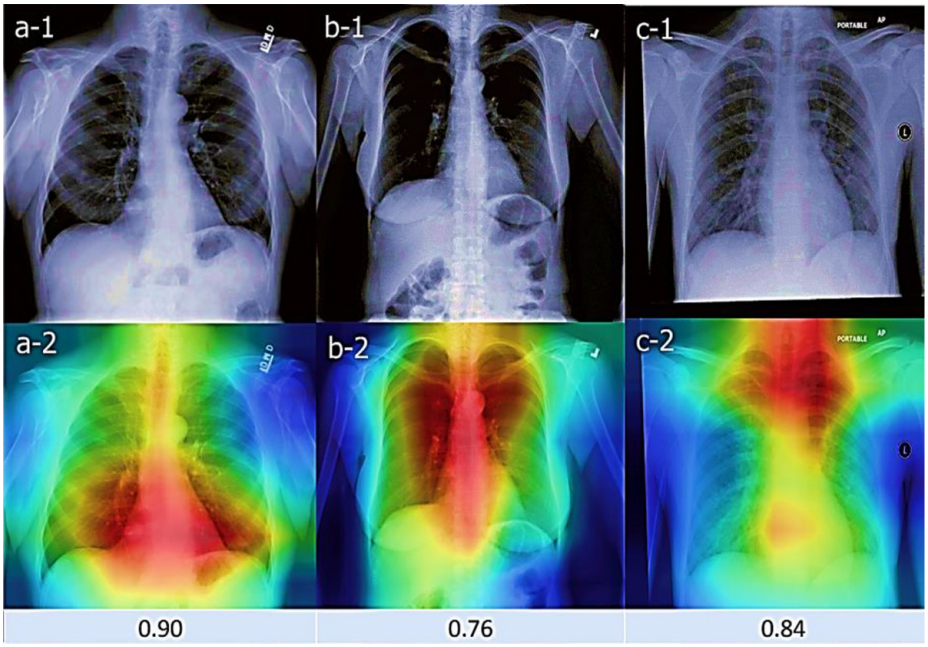

Figure 2: Examples from a health examination center's datasets used in this study. The proposed artificial intelligence algorithm accurately detected pulmonary nodules and highlighted the most indicative regions in the images (**a-1, a-2**), even identifying nodules missed by physicians (**b-1, b-2**); however, it also produced false-positive results that need improvement (**c-1, c-2**), with corresponding positive probability scores shown in **a-2, b-2, c-2**, respectively.

However, it is worth mentioning that Figure 2-c-2 displays a false-positive result, where the heatmap indicated the possibility of a chest nodule that was not actually present. This serves as a reminder of the importance of considering false-positive findings in interpreting heatmaps.

A positive probability score was assigned to each case to evaluate further the likelihood of chest nodules (Figure 2). A cutoff value 0.5 was established, with a score of ≥0.5, indicating a positive finding. Notably, Figure 2-c-2 demonstrates a false-positive result, as the positive probability score exceeded the cutoff value, suggesting the presence of a chest nodule that was not confirmed.

These findings underscore the importance of cautious interpretation when utilizing heatmaps and positive probability scores for pulmonary nodule detection. While these visualization techniques offer valuable insights, false-positive results can occur, emphasizing the need for comprehensive analysis and clinical correlation.

Discussion

Chest radiography is widely recognized as the primary imaging technique for diagnosing thoracic conditions in the medical community. It offers accessibility, cost-effectiveness, and lower radiation exposure than chest CT scans. However, studies have shown that a significant percentage of lung cancers are missed during the initial reading of chest radiographs. To address this issue, low-dose CT scans are recommended for improved lung cancer detection. The development of AI algorithms capable of analyzing radiography and CT images has emerged as a potential solution. In this study, two innovative AI algorithms, Type B and Type A, were developed and evaluated to validate the accuracy of AI using a smaller number of radiographs. The results showed that the novel AI algorithms, notably the Type

B algorithm, demonstrated better performance in terms of the Area Under the Curve (AUC) and sensitivity for the detection of pulmonary nodules on chest radiographs in comparison to radiologists' interpretations. However, further research is needed to assess the clinical utility of AI algorithms in real-world settings. It is worth noting that false-positive results on chest radiography are not considered critical, as it is primarily used for initial screening.

The Type B dataset, consisting of standardized health screening radiographs from the Group2 dataset, outperformed the Type A dataset, which had undergone pretraining with the larger Group1 dataset. This discrepancy may be attributed to variations in imaging environments and examinee postures in the Group1 dataset. The study also acknowledged several limitations, such as the use of case-control methods and the need for further research to assess the clinical utility of AI algorithms in detecting all abnormalities, not just pulmonary nodules.

The novelty of this study is that it shows high diagnostic accuracy can be achieved with a small and well-curated dataset, highlighting the importance of data quality. Furthermore, the dual evaluation on Type A and Type B datasets showcases the algorithm’s robustness in handling both diverse and standardized imaging conditions. The integration of CAMs for heatmap visualization and positive probability scoring enhances the interpretability of predictions, addressing a critical gap in AI-assisted diagnostic tools by facilitating clinician validation.

Conclusion

In conclusion, the AI algorithm achieves clinically relevant accuracy in detecting pulmonary nodules on frontal chest radiographs, demonstrating its potential as a scalable tool for early lung cancer detection in settings with limited radiological expertise. The accuracy of this algorithm surpasses that of radiologists, representing a significant advancement in medical imaging technology. This breakthrough can potentially revolutionize healthcare delivery, particularly in regions where access to skilled radiologists is limited.

Implementing this AI system can bridge the gap in medical imaging expertise, providing accurate and timely diagnoses to patients in underserved areas. By leveraging the capabilities of AI, healthcare professionals can enhance their diagnostic capabilities and improve patient outcomes.

It is important to note that further validation and refinement of the algorithm are underway to enhance its accuracy and achieve autonomous diagnosis. This ongoing research aims to address any potential limitations and ensure the reliability and effectiveness of the AI algorithm. By continually improving the system's performance, we can maximize its potential to revolutionize pulmonary nodule detection and contribute to advancements in medical imaging technology.

The integration of AI algorithms into radiology holds great promise for the future of healthcare. As technology continues to evolve, we can expect further advancements in AI-driven diagnostic tools that will enhance the precision and efficiency of medical imaging. With continued research and development, we can realize the full potential of AI in improving healthcare outcomes and expanding access to quality medical services worldwide.

Authors’ Contribution

M. Mehrabi developed the study concept, conducted the literature review, implemented the methodology, analyzed the results, and drafted the manuscript. M. Askari provided guidance on results analyses, reviewed, and supervised the research process. All authors read, revised, and approved the final manuscript.

Ethical Approval

This study exclusively utilized publicly available datasets, specifically the NIH Chest

X-ray Dataset, which is freely accessible for research purposes. The dataset was obtained from an open-access repository and does not include any personally identifiable information, as all patient data are anonymized in compliance with ethical and privacy standards. Since the study did not involve any human subjects directly or require the collection of new data, no ethical approval or code of ethics was necessary. We confirm that all data usage adhered to the terms and conditions specified by the dataset providers.

Funding

Not Applicable

Conflict of Interest

None

Data Availability Statement

Data supporting this study are available from the corresponding author upon request.

References

1. Voets M, Møllersen K, Bongo LA. Reproduction study using public data of: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *PLoS One.* 2019;**14**(6):e0217541. doi: 10.1371/journal.pone.0217541. PubMed PMID: 31170223. PubMed PMCID: PMC6553744.
2. Wang X, Shen T, Yang S, Lan J, Xu Y, Wang M, et al. A deep learning algorithm for automatic detection and classification of acute intracranial hemorrhages in head CT scans. *Neuroimage Clin.* 2021;**32**:102785. doi: 10.1016/j.nicl.2021.102785. PubMed PMID: 34411910. PubMed PMCID: PMC8377493.
3. Mayer JE, Maurer MA, Nguyen HT. Diffuse cutaneous breast cancer metastases resembling subcutaneous nodules with no surface changes. *Cutis.* 2018;**101**(3):219-23. PubMed PMID: 29718016.
4. Tammemagi MC, Lam S. Screening for lung cancer using low dose computed tomography. *BMJ.* 2014;**348**:g2253. doi: 10.1136/bmj.g2253. PubMed PMID: 24865600.
5. Horeweg N, Van Rosmalen J, Heuvelmans MA, Van Der Aalst CM, Vliegenthart R, Scholten ET, et al. Lung cancer probability in patients with CT-detected pulmonary nodules: a prespecified analysis of data from the NELSON trial of low-dose CT screening. *Lancet Oncol.* 2014;**15**(12):1332-41. doi: 10.1016/S1470-2045(14)70389-4. PubMed PMID: 25282285.
6. Lancaster HL, Heuvelmans MA, Oudkerk M. Low-dose computed tomography lung cancer screening: Clinical evidence and implementation research. *J Intern Med.* 2022;**292**(1):68-80. doi: 10.1111/joim.13480. PubMed PMID: 35253286. PubMed PMCID: PMC9311401.
7. Jonas DE, Reuland DS, Reddy SM, Nagle M, Clark SD, Weber RP, et al. Screening for Lung Cancer With Low-Dose Computed Tomography: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. *JAMA.* 2021;**325**(10):971-87. doi: 10.1001/jama.2021.0377. PubMed PMID: 33687468.
8. Ridder K, Preuhs A, Mertins A, Joerger C. Routine Usage of AI-based Chest X-ray Reading Support in a Multi-site Medical Supply Center [Internet]. arXiv [Preprint]. 2021 [cited 2022 Oct 17]. Available from: https://arxiv.org/abs/2210.10779.
9. Nam JG, Hwang EJ, Kim J, Park N, Lee EH, Kim HJ, et al. AI Improves Nodule Detection on Chest Radiographs in a Health Screening Population: A Randomized Controlled Trial. *Radiology.* 2023;**307**(2):e221894. doi: 10.1148/radiol.221894. PubMed PMID: 36749213.
10. Altorki N, Kent M, Pasmantier M. Detection of early-stage lung cancer: computed tomographic scan or chest radiograph? *J Thorac Cardiovasc Surg.* 2001;**121**(6):1053-7. doi: 10.1067/mtc.2001.112827. PubMed PMID: 11385370.
11. Mustafa Z, Nsour H. Using Computer Vision Techniques to Automatically Detect Abnormalities in Chest X-rays. *Diagnostics (Basel).* 2023;**13**(18):2979. doi: 10.3390/diagnostics13182979. PubMed PMID: 37761345. PubMed PMCID: PMC10530162.
12. Yoo H, Kim KH, Singh R, Digumarthy SR, Kalra MK. Validation of a Deep Learning Algorithm for the Detection of Malignant Pulmonary Nodules in Chest Radiographs. *JAMA Netw Open.* 2020;**3**(9):e2017135. doi: 10.1001/jamanetworkopen.2020.17135. PubMed PMID: 32970157. PubMed PMCID: PMC7516603.
13. Peters AA, Wiescholek N, Müller M, Klaus J, Strodka F, Macek A, et al. Impact of artificial intelligence assistance on pulmonary nodule detection and localization in chest CT: a comparative study among radiologists of varying experience levels. *Sci Rep.* 2024;**14**(1):22447. doi: 10.1038/s41598-

024-73435-3. PubMed PMID: 39341945. PubMed PMCID: PMC11439040.

14. Kim JY, Ryu WS, Kim D, Kim EY. Better performance of deep learning pulmonary nodule detection using chest radiography with pixel level labels in reference to computed tomography: data quality matters. *Sci Rep.* 2024;**14**(1):15967. doi: 10.1038/s41598-024-66530-y. PubMed PMID: 38987309. PubMed PMCID: PMC11237128.
15. Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, et al. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning [Internet]. arXiv [Preprint]. 2021 [cited 2017 Nov 14]. Available from: https://arxiv.org/abs/1711.05225.
16. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. Imagenet: A large-scale hierarchical image database. In IEEE conference on computer vision and pattern recognition; Miami, FL, USA: IEEE; 2009. p. 248-55.
17. Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In Proceedings of the IEEE conference on computer vision and pattern recognition; Honolulu, HI, USA: IEEE; 2017. p. 2097-106.
18. Yao L, Poblenz E, Dagunts D, Covington B, Bernard D, Lyman K. Learning to diagnose from scratch by exploiting dependencies among labels [Internet]. arXiv [Preprint]. 2017 [cited 2017 Oct 28]. Available from: https://arxiv.org/abs/1710.10501.
19. Homayounieh F, Digumarthy S, Ebrahimian S, Rueckel J, Hoppe BF, Sabel BO, et al. An Artificial Intelligence-Based Chest X-ray Model on Human Nodule Detection Accuracy From a Multicenter Study. *JAMA Netw Open.* 2021;**4**(12):e2141096. doi: 10.1001/jamanetworkopen.2021.41096. PubMed PMID: 34964851. PubMed PMCID: PMC8717119.
20. Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning deep features for discriminative localization. In Proceedings of the IEEE conference on computer vision and pattern recognition; Las Vegas, NV, USA: IEEE; 2016. p. 2921-9.
21. Wang J, Bhalerao A, Yin T, See S, He Y. CAMANet: Class Activation Map Guided Attention Network for Radiology Report Generation. *IEEE J Biomed Health Inform.* 2024;**28**(4):2199-210. doi: 10.1109/JBHI.2024.3354712. PubMed PMID: 38227409.
22. Wang K, Yin S, Wang Y, Li S. Explainable deep learning for medical image segmentation with learnable class activation mapping. In Proceedings of the 2023 2nd Asia Conference on Algorithms, Computing and Machine Learning; New York, NY, United States: Association for Computing Machinery; 2023. p. 210-5.