European Journal of Radiology 187 (2025) 112077

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Accuracy of a deep neural network for automated pulmonary embolism detection on dedicated CT pulmonary angiograms

Emese Zsarnoczay [a,b], Saikiran Rapaka [c], U. Joseph Schoepf [a], Chiara Gnasso [a,d], Milan Vecsey-Nagy [a,e], Thomas M. Todoran [f], Muhammad Taha Hagar [a,g], Dmitrij Kravchenko [a,h], Giuseppe Tremamunno [a,i], Joseph Parkwood Griffith [a], Nicola Fink [a,j], Sydney Derrick [a], Meredith Bowman [a], Henry Sam [a], Mikayla Tiller [a], Kathleen Godoy [a], Florin Condrea [k], Puneet Sharma [c], Jim O'Doherty [a,l], Pal Maurovich-Horvat [b], Tilman Emrich [a,m,n,*], Akos Varga-Szemes [a]

[a] Department of Radiology and Radiological Science, Medical University of South Carolina, 25 Courtenay Drive, Charleston, SC 29425, USA
[b] Radiology Department, Medical Imaging Centre, Semmelweis University, Korányi Sándor utca 2, 1083 Budapest, Hungary
[c] Siemens Healthineers, 755 College Rd E, Princeton, NJ 08540, USA
[d] Experimental Imaging Center, IRCCS San Raffaele Scientific Institute, Via Olgettina, 60, 20132 Milan, Italy
[e] Heart and Vascular Center, Semmelweis University, Gaál József út 9, 1122 Budapest, Hungary
[f] Division of Cardiology, Department of Medicine, Medical University of South Carolina, 25 Courtenay Drive, Charleston, SC 29425, USA
[g] Department of Diagnostic and Interventional Radiology, Medical Center, Faculty of Medicine, University of Freiburg, Hugstetter Straße 55, Freiburg im Breisgau 79106, Germany
[h] Department of Diagnostic and Interventional Radiology, University Hospital Bonn, Venusberg-Campus 1, 53127 Bonn, Germany
[i] Department of Medical Surgical Sciences and Translational Medicine, Sapienza University of Rome - Radiology Unit - Sant'Andrea University Hospital, Via di Grottarossa, 1035-1039, 00189 Rome, Italy
[j] Department of Radiology, University Hospital, LMU Munich, Marchioninistraße 15, 81377 Munich, Germany
[k] Siemens Healthineers, Nine, Bulevardul Gării 13A, Braşov 500227, Romania
[l] Siemens Medical Solutions, 40 Liberty Blvd, Malvern, PA 19355, USA
[m] Department of Diagnostic and Interventional Radiology, University Medical Center of the Johannes Gutenberg-University, Langenbeckstraße 1, Mainz 55131, Germany
[n] German Centre for Cardiovascular Research, Mainz, Germany

ARTICLE INFO

Keywords:
Pulmonary Embolism
Artificial intelligence
Deep neural network
Computed tomography

ABSTRACT

Purpose: To assess the performance of a Deep Neural Network (DNN)-based prototype algorithm for automated PE detection on CTPA scans.
Methods: Patients who had previously undergone CTPA with three different systems (SOMATOM Force, go.Top, and Definition AS; Siemens Healthineers, Forchheim, Germany) because of suspected PE from September 2022 to January 2023 were retrospectively enrolled in this study (n = 1,000, 58.8 % women). For detailed evaluation, all PE were divided into three location-based subgroups: central arteries, lobar branches, and peripheral regions. Clinical reports served as ground truth. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy were determined to evaluate the performance of DNN-based PE detection.

Abbreviations: 3D, three-dimensional; AI, artificial intelligence; BMI, body mass index; CNN, convolutional neural network; CTPA, computed tomography pulmonary angiography; DNN, deep neural network; NPV, negative predictive value; PE, pulmonary embolism; PPV, positive predictive value.

* Corresponding author at: Department of Diagnostic and Interventional Radiology, University Medical Center of the Johannes Gutenberg-University, Langenbeckstraße 1, Mainz 55131, Germany.

E-mail addresses: zsarnoczay.emese@semmelweis.hu (E. Zsarnoczay), saikiran.rapaka@siemens-healthineers.com (S. Rapaka), schoepf@musc.edu (U.J. Schoepf), GNASSO.CHIARA@hsr.it (C. Gnasso), vecseyna@musc.edu (M. Vecsey-Nagy), todoran@musc.edu (T.M. Todoran), hagarmu@musc.edu (M.T. Hagar), kravchen@musc.edu (D. Kravchenko), giuseppe.tremamunno@uniroma1.it (G. Tremamunno), jpg201@musc.edu (J.P. Griffith), Nicola.Fink@med.uni-muenchen.de (N. Fink), derricsy@musc.edu (S. Derrick), bowmamer@musc.edu (M. Bowman), samh@musc.edu (H. Sam), tillerm@musc.edu (M. Tiller), mkatieg1@gmail.com (K. Godoy), florin.condrea@siemens.com (F. Condrea), sharma.puneet@siemens-healthineers.com (P. Sharma), james.odoherty@siemens-healthineers.com (J. O'Doherty), maurovich-horvat.pal@semmelweis.hu (P. Maurovich-Horvat), tilman.emrich@unimedizin-mainz.de (T. Emrich), vargaasz@musc.edu (A. Varga-Szemes).

https://doi.org/10.1016/j.ejrad.2025.112077
Received 22 December 2024; Received in revised form 11 March 2025; Accepted 26 March 2025
Available online 29 March 2025

Results: Cases were excluded due to incomplete data (n = 32), inconclusive report (n = 17), insufficient contrast detected in the pulmonary trunk (n = 40), or failure of the preprocessing algorithms (n = 8). Therefore, the final cohort included 903 cases with a PE prevalence of 12 % (n = 110). The model achieved a sensitivity, specificity, PPV, and NPV of 84.6, 95.1, 70.5, and 97.8 %, respectively, and delivered an overall accuracy of 93.8 %. Among the false positive cases (n = 39), common sources of error included lung masses, pneumonia, and contrast flow artifacts. Common sources of false negatives (n = 17) included chronic and subsegmental PEs.
Conclusion: The proposed DNN-based algorithm provides excellent performance for the detection of PE, suggesting its potential utility to support radiologists in clinical reading and exam prioritization.

1. Introduction

Pulmonary embolism (PE) is a serious and potentially fatal condition, and with its high incidence and mortality rate, it ranks as the third most common cardiovascular syndrome worldwide [1]. With the advancement in imaging technology, including improved image quality, the sensitivity of PE detection has increased [2]. CT pulmonary angiography (CTPA) is currently the first-line modality for diagnosing PE due to its excellent accuracy and low rate of inconclusive results [3,4]. Now, clinicians are more likely to suspect and begin a diagnostic evaluation for PE than they were in the past, considering the high spatial and temporal resolutions, wider availability, non-invasion, and rapid acquisition of CTPA [5]. However, the thorough evaluation of CTPA scans is tedious, time-consuming, and requires review by an expert radiologist [6]. Furthermore, in emergency settings, multiple life-threatening conditions must be ruled out alongside PE, such as acute coronary syndrome, aortic dissection, and pneumonia, contributing to radiologists' workload [7–9]. While improved PE detection can reduce this burden, comprehensive assessment of other critical findings remains essential. Additionally, oncologic patients, though at relatively lower immediate risk, still face increased PE incidence due to malignancy-associated thrombosis, making early detection crucial for optimal management. Given the high mortality of PE, prompt diagnosis and treatment is essential to prevent life-threatening complications, as delayed detection and diagnosis significantly increases the risk of death within 30 days [10,11]. The overall performance of radiologists in detecting PE shows a sensitivity from 67 % to 87 % and a specificity from 89 % to 99 % when consensus reading is considered as reference standard [12,13], with higher accuracy for central than peripheral PE [ranging from 74 % to 86 %) [13].

Deep learning, a type of machine learning, has been increasingly used in medical image analysis due to technical advances in computer science and the growing body of medical data. A well-trained neural network has the potential to support radiology workflow by prioritizing urgent exams on the worklist such as those with the presence of PE, thereby expediting the diagnostic process [14,15]. However, despite the advancements and promising results of deep learning in medical imaging, there remain significant challenges that necessitate further investigation: the generalizability of algorithms across different scanners, populations and clinical settings, as well as their integration into routine clinical workflows, is still under active exploration. Furthermore, the potential for deep learning models to overlook subtle or atypical cases of PE, particularly in complex or high-noise environments, mandates a dedicated evaluation.

Therefore, this study aimed to assess the performance of a Deep Neural Network (DNN)-based prototype algorithm for automated PE detection on CTPA scans consecutively acquired at a single university hospital.

2. Materials and Methods

2.1. Study population

The local Institutional Review Board approved this retrospective, single-center, Health Insurance Portability and Accountability Act-compliant study and waived the requirement for informed consent. Deidentified image data of all consecutive patients who had undergone CTPA because of suspected PE from September 2022 to January 2023 were retrospectively included in this study. Cases were excluded due to incomplete data transfer, inconclusive reports, failure of preprocessing algorithms, and insufficient attenuation (< 180 Hounsfield Unit) at the level of the pulmonary trunk. Demographic and clinical patient characteristics were collected from medical records.

2.2. Image acquisition and reconstruction

The CTPA scans were acquired using three different systems (SOMATOM Force, go.Top, and Definition AS; Siemens Healthineers, Forchheim, Germany) according to the institutional standard protocols, typically with a spiral acquisition using the following parameters: collimation 0.6 mm, pitch 0.55, tube voltage 90–140 kVp, and reference mAs 20–120. Iodinated contrast (iopromide, Ultravist 370 mgI/ml, Bayer Healthcare; or iohexol, Omnipaque 350, GE Healthcare) was administered at a standard injection rate of 4–5 ml/s and volume of 60–90 ml. Images were reconstructed using the following parameters: Advanced Modeled Iterative Reconstruction [ADMIRE, Siemens] with a strength level of 1, a Br36 kernel, at a slice thickness of 1 mm.

2.3. AI model for PE analysis

The computational pipeline used for PE detection utilized multiple deep-learning models and was trained with a dataset of 5,165 CT volumes containing 1,876 PE-positive studies. The embolic regions in the PE-positive cases were annotated by human experts with radiologist confirmation before being used as ground truth for training the model.

In the first step, convolutional neural network (CNN) based models are used to perform lung lobe, heart, and pulmonary artery segmentations. These models are used to crop the input image to the relevant region of interest. A sample input image and the corresponding organ segmentations are shown in Fig. 1. The pulmonary artery segmentation model is used to identify an ROI in the pulmonary trunk to evaluate the opacification of the pulmonary artery. If the average intensity in the ROI is less than 180 HU or if any of the segmentation algorithms fail, the input volume is not processed further.

The PE detection model is a fully convolutional three-dimensional (3D) DynUNet implementation with additional skip connections in the downsampling layers [16]. The network comprises an encoder, bottleneck, and decoder, followed by a segmentation head, and utilizes 3D convolutions followed by instance normalization and LeakeReLU activation functions. Dropout is used for regularization. The network processed the input study and produces an output probability map representing the likelihood of each voxel belonging to a region with PE. The probability map is thresholded to label an entire study as being positive or negative for the presence of PE.

2.4. Reference standard

Clinical reports served as the ground-truth. The clinical reports were validated by board-certified radiologists. During the study, data was extracted from the reports to stratify cases as positive or negative.

Furthermore, all reports indicated the location of the PE which was also collected from the reports. For false positive and false negative AI results, cases were further reviewed by a board-certified radiologist and a radiology resident with 6 and 3 years of experience, respectively, to confirm ground truth.

2.5. Image analysis

The AI assessment of CTPA studies was performed. Subgroup analyses were conducted to assess the algorithm's performance in different settings that may affect its accuracy. First, body mass index (BMI)-based analysis was performed in non-obese vs. obese patients by dividing the cohort into two subgroups, BMI $<$ vs. $\geq$ 30 kg/m^2. Second, PE location-based analysis was carried out by dividing the pulmonary artery system into the following three levels: (1) central arteries (main pulmonary artery, right and left pulmonary artery), (2) lobar branches (right upper lobe, right middle lobe, right lower lobe, left upper lobe, lingula, and left lower lobe), and (3) peripheral regions (segmental and subsegmental branches) [13,17]. In cases with multiple emboli, the location of the most proximal PE was considered.

2.6. Statistical analysis

The Shapiro-Wilk test was used to assess the continuous variables for normality. Depending on the normality distribution, continuous data are reported as means ± standard deviation (SD) or as medians with interquartile ranges. Categorial variables were described as absolute frequencies and proportions. Significance tests were conducted using the Mann–Whitney tests for continuous variables and the Chi-square test for categorical variables. A two-tailed p-value $<$ 0.05 was considered as statistically significant. The comparison between the results provided by the radiology reports and those automatically computed by AI was performed according to the numbers of the true positive, true negative, false positive and false negative values. Sensitivity, specificity, accuracy, and area under the receiver operating characteristic curve (ROC AUC) were calculated for the total population. The 95 % confidence intervals (95 % CI) were computed. Statistical analysis was performed using MedCalc (version 23.0.2, Ostend, Belgium).

3. Results

3.1. Patient population

In total, 1,000 cases were included in the study. The flowchart of the study is represented in Fig. 2. Scans were excluded due to incomplete data transfer (n = 32), inconclusive radiology report (n = 17), insufficient contrast detected in the pulmonary trunk (n = 40), or failure of preprocessing algorithms (n = 8). The final cohort included 903 cases (age 58 years; 58.9 % female). The AI model successfully analyzed all cases of the final cohort. The prevalence of PE was 12.2 % (n = 110). The proportion of individuals with limited mobility (37.3 % vs. 25.3 %, p = 0.009), active malignancy (37.3 % vs. 26.6 %, p = 0.020), surgery within one month (34.5 % vs. 20.9 %, p = 0.002), a history of clotting or clotting disorder (55.5 % vs. 28.0 %, p $<$ 0.001), and hemoptysis (7.3 % vs 3.3 %, p = 0.040) were significantly higher in patients with PE. Detailed patient characteristics are shown in Table 1.

3.2. Overall performance analysis

A sensitivity of 84.6 % (95 % CI: 76.4 %-90.7 %), specificity of 95.1 % (95 % CI: 93.3 %-96.5 %), and an accuracy of 93.8 % (95 % CI: 92.0 %-95.3 %) were achieved for the detection of PE. The performance of the algorithm for the entire cohort are presented in detail in Table 2.

ROC curve showed an AUC of 0.935 (95 % CI: 91.7 %-95.0 %) for PE detection (Fig. 3).

3.3. Obesity-based subgroup analysis

The prevalence of PE was similar in obese and non-obese patients (12.2 % vs. 11.9 %, p = 0.737, respectively). In the obese and non-obese groups, the accuracy for detecting PE was 92.5 % (95 % CI: 89.2 %–95.1 %) and 94.6 % (95 % CI: 92.4 %–96.2 %), respectively. Further details of

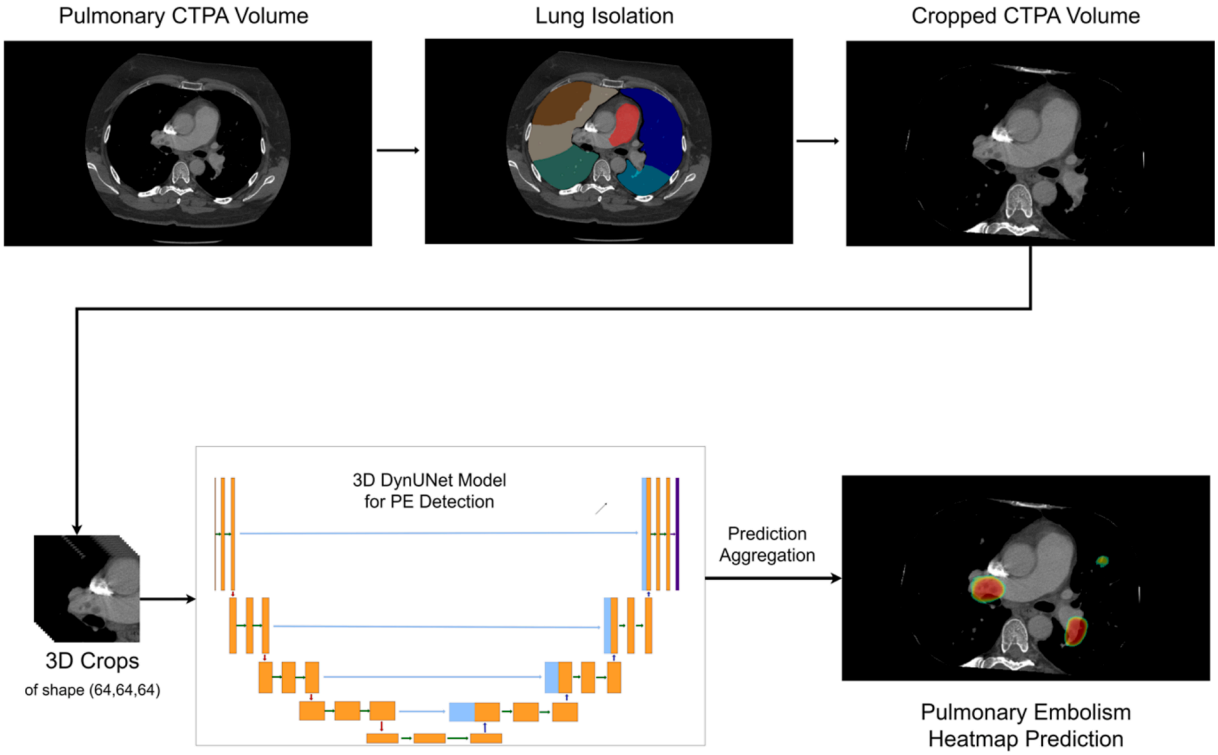

Fig. 1. The architecture diagram shows the pipeline used for pulmonary embolism detection. CTPA = CT pulmonary angiography, PE = pulmonary detection.

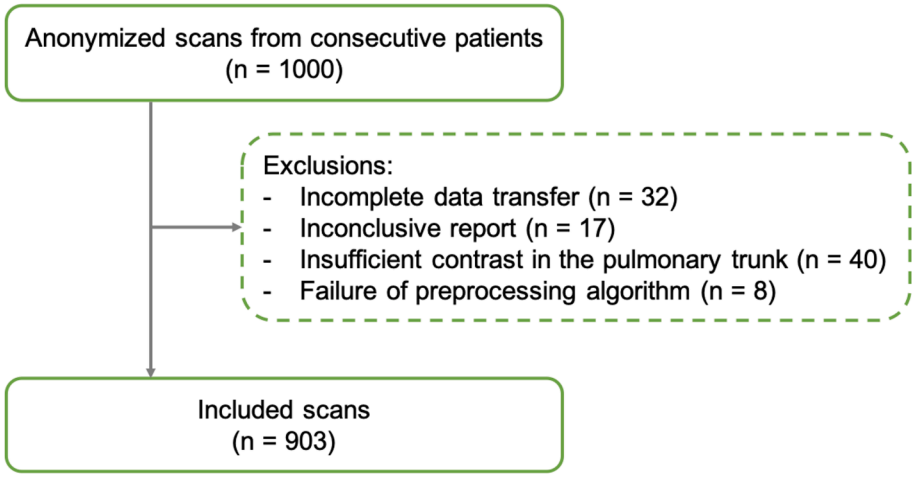

Fig. 2. Flowchart of the study.

Table 1
Patient characteristics.

	All cases (n = 903)	With PE (n = 110)	Without PE (n = 793)	p-value *
Clinical characteristics				
Sex (female)	532 (58.9)	60 (54.5)	472 (59.5)	0.321
Age (year)	58 (39, 70)	59 (44, 69)	57 (39, 70)	0.335
Body mass index (kg/m^2)	27.9 (23.4, 33.5)	28.5 (23.6, 32.3)	28.2 (23.5, 33.4)	0.745
Race				
White	467 (51.7)	57 (51.8)	409 (51.6)	0.962
Black	403 (44.6)	45 (40.9)	358 (45.1)	0.403
Native American	1 (0.1)	0 (0.0)	1 (0.1)	0.710
Native Hawaiian	1 (0.1)	0 (0.0)	1 (0.1)	0.710
Asian	4 (0.4)	1 (0.9)	3 (0.4)	0.432
Unknown	27 (2.9)	7 (6.4)	20 (2.5)	0.028
Comorbidities				
Limited mobility	242 (26.8)	41 (37.3)	201 (25.3)	**0.009**
Heart failure	140 (15.5)	23 (20.9)	117 (14.8)	0.096
Coronary artery disease	128 (14.2)	22 (20.0)	106 (13.4)	0.062
Prior myocardial infarction	73 (8.1)	13 (11.8)	60 (7.6)	0.129
COPD	124 (13.7)	16 (14.5)	108 (13.6)	0.792
Hypertension	526 (58.3)	63 (57.3)	463 (58.4)	0.825
Prior stroke	102 (11.3)	15 (13.6)	87 (11.0)	0.408
Active malignancy	252 (27.9)	41 (37.3)	211 (26.6)	**0.020**
Any smoking history	432 (47.8)	48 (43.6)	384 (48.4)	0.578
Surgery within 1 month	204 (22.6)	38 (34.5)	166 (20.9)	**0.002**
History of clotting or clotting disorder	283 (31.3)	61 (55.5)	222 (28.0)	**<0.001**
Symptoms				
Tachypnea	304 (33.7)	45 (40.9)	259 (32.7)	0.086
Heart rate > 100 bpm	377 (41.7)	54 (49.1)	323 (40.7)	0.096
Hemoptysis	34 (3.8)	8 (7.3)	26 (3.3)	**0.040**

Median (25th and 75th percentile) and frequencies (percentages).
bpm, beats per minute; COPD, chronic obstructive pulmonary disease.
* between patients with and without PE.

algorithm performance in the subgroups are shown in Table 3.

3.4. Location-specific subgroup analysis

The most proximal level of PE in this cohort included 21 central, 30 lobar, and 59 peripheral cases of PE. In the central arteries, all emboli were detected. In the lobar branches, 29 of 30 emboli were depicted, while the remaining one, a chronic PE was missed. In the peripheral regions, 43 of 59 emboli were correctly detected, while 16 were missed. Latter were identified as chronic subsegmental PEs and acute/chronic PEs in segmental or subsegmental arteries by further expert reading.

Table 2
Overall performance analysis of the AI model.

	All cases (n = 903)
True positive	93
False negative	17
False positive	39
True negative	754
Sensitivity	84.6 % [76.4 %-90.7 %]
Specificity	95.1 % [93.3 %-96.5 %]
Positive predictive value	70.5 % [63.5 %-76.6 %]
Negative predictive value	97.8 % [96.6 %-98.6 %]
Accuracy	93.8 % [92.0 %-95.3 %]

95% confidence intervals are in the squared brackets.

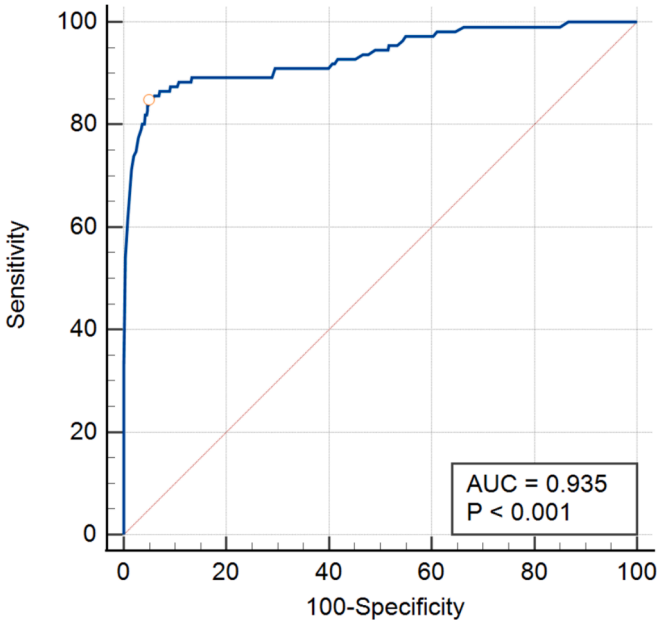

Fig. 3. Receiver-operating characteristic curve representing the overall performance of the algorithm. AUC = area under the curve.

Table 4 shows the detailed results of the location-specific analysis.

3.5. Assessment of false negative and false positive findings

There were 17 false negatives with a miss rate for PE of 15.5 % (17/

Table 3
Obese vs. non-obese subgroup analysis.

	Obese (n = 348)	Non-obese (n = 555)
True positive	37	56
False negative	7	10
False positive	19	20
True negative	285	469
Sensitivity	84.1 % [69.9 %–93.4 %]	84.8 % [73.9 %–92.5 %]
Specificity	93.8 % [90.4 %–96.2 %]	95.9 % [93.7 %–97.5 %]
Positive predictive value	66.1 % [55.3 %–75.4 %]	73.7 % [63.3 %–81.3 %]
Negative predictive value	97.6 % [95.4 %–98.8 %]	97.9 % [96.4 %–98.8 %]
Accuracy	92.5 % [89.2 %–95.1 %]	94.6 % [92.4 %–96.2 %]

95% confidence intervals are in the squared brackets.

Table 4
Location-based performance analysis.

	True positive	False negative	Sensitivity
Overall	93	17	84.6 % [76.4 %-90.7 %]
Central	21	0	100.0 % [83.9 %-100.0 %]
Lobar	29	1	96.7 % [82.8 %-99.9 %]
Peripheral	43	16	72.9 % [59.7 %-83.6 %]

95% confidence intervals are in the squared brackets.

110 positive cases). False negative cases mostly corresponded to small chronic PEs in subsegmental arteries (n = 8) and probably acute PEs in segmental or subsegmental branches (n = 7). Other causes were low image quality due to high BMI (n = 1) and pneumonia with lymphadenopathy (n = 1).

There were 39 false positive cases. False positive cases were caused by turbulent flow in the arteries and vascular branching (n = 9), beam hardening artifacts (n = 4), breathing artifacts (n = 4), consolidation (n = 6), pulmonary mass (n = 6), lymphadenopathy (n = 2), low image quality (n = 2), and other pulmonary pathologies (pneumothorax, atelectasis, and missing right pulmonary artery) (n = 6). Fig. 4 presents examples of a true-positive, a false-positive, and a false-negative finding. (Fig. 5, Fig. 6, Fig. 7).

4. Discussion

PE is a life-threatening condition with high incidence and mortality, making prompt and accurate diagnosis essential. While CTPA is the gold standard for PE detection, its evaluation is time-consuming and adds to radiologists' workload. This study aimed to assess the performance of a DNN-based algorithm for automated PE detection on a large number of CTPA scans.

The algorithm showed high sensitivity (84.6 %) and specificity (95.1 %) for the detection of PE, with excellent overall accuracy (93.8 %). Location-specific analysis revealed excellent sensitivity for detecting central (100 %) and lobar PE (96.7 %), while accurately capturing acute peripheral PE as well. False negative classification occurred in 17 out of 754 cases and was limited to lobar and peripheral branches. The software did not miss any PE in the central pulmonary arteries.

PE is a potentially life-threatening condition; thus, prompt diagnosis and initiation of therapy are essential to improve outcomes. The number of CTPA scans continues to increase around 3–4 % annually due to fear of missing PE, increased accessibility to CT systems, and financial considerations [18]. Therefore, an AI-based algorithm that can triage and prioritize studies in the radiologist's worklist or even serve as a potential second reader may improve the time to early diagnosis [19]. A successful AI model, however, must balance sensitivity and specificity by setting optimal operating thresholds. Minimizing false positives when employing AI for worklist triage is crucial to prevent unnecessary alerts and disruptions to the radiologist's workflow, ensuring that only legitimate cases of PE are flagged and expedited [20]. But even more important is reducing false negative cases to guarantee no PE-positive scans are overlooked.

The tested algorithm demonstrated an overall sensitivity of 84.6 % and a specificity of 95.1 %, which aligns with previous literature and a recent *meta*-analysis including five studies evaluating different AI algorithms, with a pooled sensitivity of 88.0 % and a specificity of 86.0 % [12,21,22]. Furthermore, the radiologist's sensitivity for detecting PE ranges from 67 % to 87 % with a specificity of 89 % to 99 %, which is also consistent with the performance of the algorithm tested in our study [13].

The present study showed that all central emboli were accurately identified, and only one lobar embolus was missed, leading to a 100 % and 96.7 % sensitivity, respectively. At peripheral locations, the algorithm achieved a sensitivity of 72.9 %; however, it is important to note that eight of these were small chronic PEs and seven were probably acute lesions in segmental or subsegmental branches. The visual assessment was limited in some of these cases due to the small caliber, which prevented the distinction between acute and chronic PE. In recent years, a significant rise in the incidence of subsegmental PEs has been identified, owing to the ability of current high-resolution CT scanners to detect filling defects even in subsegmental arteries with a diameter of 2–3 mm [23]. Since this, however, is coupled with a minimal decline in mortality, it implies that a considerable portion of the additional, mostly peripheral emboli being identified may not be clinically significant [24]. There is only a handful of studies researching the performance of automated PE detection on a location-based level. In a recent study, the performance of a convolutional neural network for PE detection was investigated and showed a sensitivity of 94 % for central, 94 % for lobar, and 80 % for peripheral emboli [25]. However, for their training and internal validation set, CTPA scans with extensive chest abnormalities were excluded, the training set only included PE positive scans and the external validation set only included 250 CTPA scans, whereas the algorithm tested in our study included 5,165 CTPA scans with around a prevalence of 36 % of PE positive cases.

The subgroup analysis revealed minimal to negligible impact on performance concerning BMI. This is important because CT image quality typically reduces in patients with high BMI due to the elevated X-ray absorption rates in adipose tissue, thereby increasing image noise

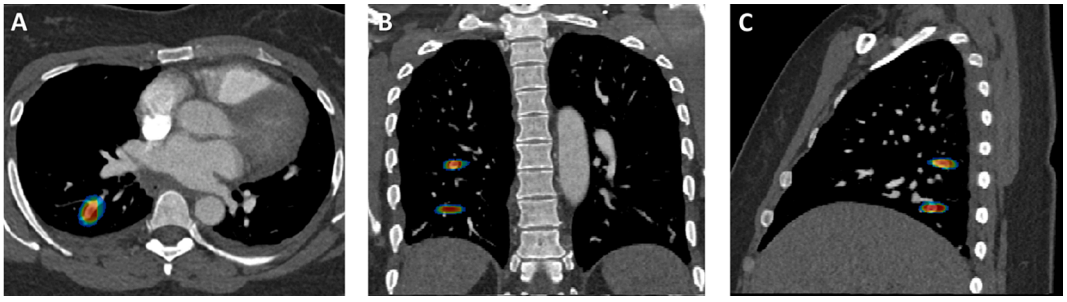

Fig. 4. Representative true positive case of a patient with right multiple right lower lobe subsegmental emboli using the AI algorithm for PE detection in axial (A), coronal (B) and sagittal (C) planes.

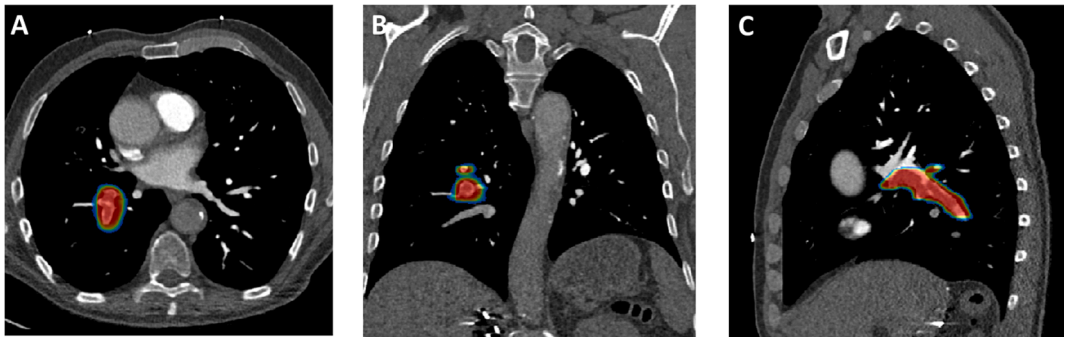

Fig. 5. Representative true positive case of a patient large right pulmonary embolism, branching into the right middle and lower lobe lobar branches using the AI algorithm for PE detection in axial (A), coronal (B) and sagittal (C) planes.

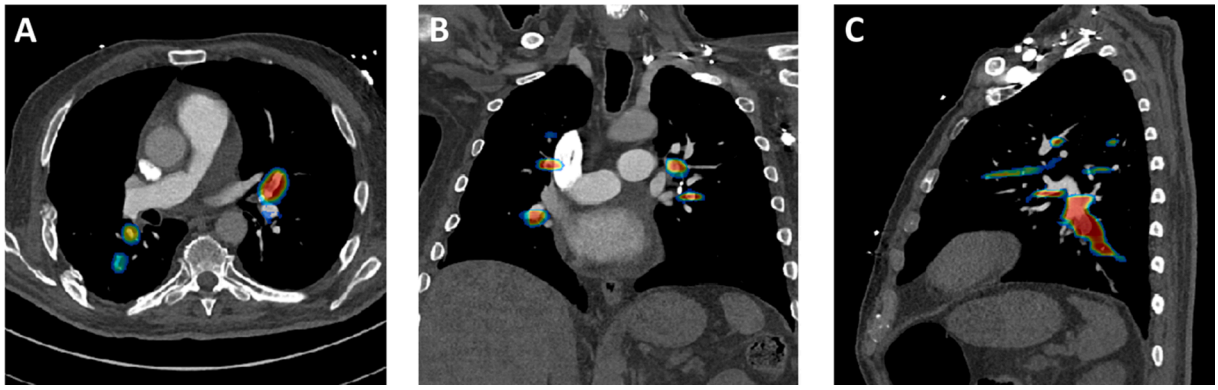

Fig. 6. Representative true positive case of a patient with multiple bilateral emboli in segmental and subsegmental branches using the AI algorithm for PE detection in axial (A), coronal (B) and sagittal (C) planes.

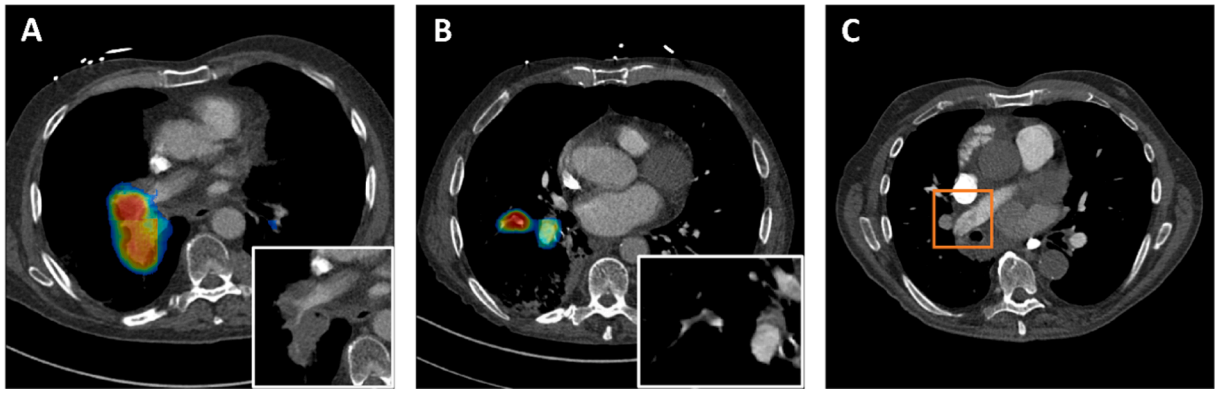

Fig. 7. Representative false positive and false negative cases. (A): flow related artifact at a bifurcation presenting as a false positive finding, (B): false positive finding in a patient with right hilar mass with lymphadenopathy, (C): patient with thromboembolic material in the right interlobar artery, which was the only lobar level false negative case from the algorithm.

[26]. Reduced spatial resolution due to a wider field of view and beam hardening artifacts due to increased body size contribute to decreased image quality, potentially compromising emboli detection [27].

Several previous studies tested various algorithms for PE detection on cohorts typically including a high number of PE positive cases [28–30]. However, as prevalence strongly influences diagnostic accuracy, this could potentially introduce a bias in evaluating diagnostic accuracy and results might become incomparable to clinical practice. To overcome this in our research, testing was performed on all consecutive pertinent CTPAs at our institution over 5 months, ensuring an accurate representation of clinical realities, including the proportion of positive to negative exams and the distribution of emboli.

Analysis of patient demographics is important when determining the generalizability of an AI algorithm. In our patient population, it was well represented that patients with PE had a higher proportion of known risk factors, e.g., limited mobility, active malignancy, surgery within 1 month, and clotting disorder. Due to the extended size of the study cohort, testing the algorithm on the entire range of PE locations and diverse disease severity was possible. Furthermore, our data, collected at an academic hospital, included consecutive patients undergoing CTPA, reflecting a real-life clinical situation. Additionally, subgroup analyses were conducted to evaluate the algorithm's performance in potentially more challenging conditions. Demonstrating the algorithm's ability to reliably detect PE, especially in an acute setting, further supports the

potential implementation of such an algorithm in radiology workflows, to aid the fast and efficient detection of PE, and enable precise patient prioritization.

This study had some limitations which should be acknowledged. First, this study was a retrospective analysis that included patients from a single center. Scans were acquired from scanner platforms from a single vendor, which might limit the generalizability of our results to patients imaged on other CT systems. Second, the classification of exams for PE presence, and thus the reference standard, was based on clinically approved reports. While any initial misclassification could impact performance metrics, PE detection on CTPA is generally straightforward. Additionally, all reports were validated by at least two radiologists, including at least one board-certified physician, minimizing the likelihood of significant influence on the results [31–33]. Third, in patients with multiple emboli, only the most proximal embolus was included for the location-based evaluation. Fourth, the tested algorithm was unable to process datasets with an average intensity < 180 HU in the pulmonary artery. Having such capability would be clinically useful to potentially label incidental PEs from e.g., oncological scans where usually venous phase scans are performed [34]. Future developments should encourage the inclusion of these cases when developing an algorithm to further improve their clinical relevance.

In conclusion, the proposed DNN-based algorithm provides excellent performance for the detection of PE, suggesting its potential utility to support radiologists in clinical reading and exam prioritization.

Ethics approval and consent to participate.

Approval received without consent needed.

Consent for publication

Not applicable.

Availability of data and materials.

Upon reasonable request.

Competing interests.

U. Joseph Schoepf receives institutional research support and / or personal fees from Bayer, Bracco, Elucid Bioimaging, Guerbet, HeartFlow, Inc., Keya Medical, and Siemens. Akos Varga-Szemes receives institutional research support and/or personal fees from Bayer, Elucid Bioimaging and Siemens. Tilman Emrich received a speaker fee and travel support from Siemens Medical Solutions USA Inc. Jim O'Doherty, Puneet Sharma, and Saikiran Rapaka are employees of Siemens Medical Solutions USA Inc. Florin Condrea is an employee of Siemens Healthineers.

Data sharing statement: Data generated or analyzed during the study are available from the corresponding author by request.

Funding.

Siemens Healthcare USA.

Authors' information

Not applicable.

Endnotes.

Not applicable.

CRediT authorship contribution statement

Emese Zsarnoczay: Writing – original draft, Formal analysis, Data curation, Conceptualization. **Saikiran Rapaka:** Writing – original draft, Software, Formal analysis. **U.Joseph Schoepf:** Writing – review & editing, Funding acquisition. **Chiara Gnasso:** Writing – review & editing. **Milan Vecsey-Nagy:** Writing – review & editing. **Thomas M. Todoran:** Writing – review & editing, Investigation, Data curation. **Muhammad Taha Hagar:** Writing – review & editing. **Dmitrij Kravchenko:** Writing – review & editing. **Giuseppe Tremamunno:** Writing – review & editing. **Joseph Parkwood Griffith:** Formal analysis, Data curation. **Nicola Fink:** Writing – review & editing. **Sydney Derrick:** Data curation. **Meredith Bowman:** Data curation. **Henry Sam:** Data curation. **Mikayla Tiller:** Data curation. **Kathleen Godoy:** Data curation. **Florin Condrea:** Writing – review & editing, Software, Formal analysis. **Puneet Sharma:** Writing – review & editing, Software. **Jim O'Doherty:** Writing – review & editing, Resources. **Pal Maurovich-Horvat:** Writing – review & editing. **Tilman Emrich:** Writing – review & editing, Writing – original draft, Data curation, Conceptualization. **Akos Varga-Szemes:** Writing – review & editing, Writing – original draft, Project administration, Data curation, Conceptualization.

Declaration of competing interest

U. Joseph Schoepf receives institutional research support and / or personal fees from Bayer, Bracco, Elucid Bioimaging, Guerbet, HeartFlow, Inc., Keya Medical, and Siemens. Akos Varga-Szemes receives institutional research support and/or personal fees from Bayer, Elucid Bioimaging and Siemens. Tilman Emrich received a speaker fee and travel support from Siemens Medical Solutions USA Inc. Jim O'Doherty, Puneet Sharma, and Saikiran Rapaka are employees of Siemens Medical Solutions USA Inc. Florin Condrea is an employee of Siemens Healthineers.

Acknowledgements

This project was partially funded by Siemens Healthcare USA.

References

[1] G.E. Raskob, P. Angchaisuksiri, A.N. Blanco, et al., Thrombosis: a major contributor to global disease burden, Arterioscler. Thromb. Vasc. Biol. 34 (11) (2014) 2363–2371, https://doi.org/10.1161/ATVBAHA.114.304488.

[2] S.B. Smith, J.B. Geske, P. Kathuria, et al., Analysis of national trends in admissions for pulmonary embolism, Chest 150 (1) (2016) 35–45, https://doi.org/10.1016/j.chest.2016.02.638.

[3] P.D. Stein, S.E. Fowler, L.R. Goodman, et al., Multidetector computed tomography for acute pulmonary embolism, N Engl. J. Med. 354 (22) (2006) 2317–2327, https://doi.org/10.1056/NEJMoa052367.

[4] L.J. Zhang, G.M. Lu, F.G. Meinel, A.D. McQuiston, J.G. Ravenel, U.J. Schoepf, Computed tomography of acute pulmonary embolism: state-of-the-art, Eur. Radiol. 25 (9) (2015) 2547–2557, https://doi.org/10.1007/s00330-015-3679-2.

[5] S.V. Konstantinides, G. Meyer, C. Becattini, et al., 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS): the task force for the diagnosis and management of acute pulmonary embolism of the European Society of Cardiology (ESC), Eur. Heart J. 41 (4) (2019) 543–603, https://doi.org/10.1093/eurheartj/ehz405.

[6] C. Wittram, M.M. Maher, A.J. Yoo, M.K. Kalra, J.A. Shepard, T.C. McLoud, CT angiography of pulmonary embolism: diagnostic criteria and causes of misdiagnosis, Radiographics 24 (5) (2004) 1219–1238, https://doi.org/10.1148/rg.245045008.

[7] R.J. McDonald, K.M. Schwartz, L.J. Eckel, et al., The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload, Acad Radiol. 22 (9) (2015) 1191–1198, https://doi.org/10.1016/j.acra.2015.05.007.

[8] S. Rohatgi, T.N. Hanna, C.W. Sliker, R.M. Abbott, R. Nicola, After-Hours radiology: challenges and strategies for the radiologist, AJR Am J. Roentgenol. 205 (5) (2015) 956–961, https://doi.org/10.2214/ajr.15.14605.

[9] T.N. Hanna, M.E. Zygmont, R. Peterson, et al., The effects of fatigue from overnight shifts on radiology search patterns and diagnostic performance, J. Am Coll. Radiol. 15 (12) (2018) 1709–1716, https://doi.org/10.1016/j.jacr.2017.12.019.

[10] A.A. Donato, J.J. Scheirer, M.S. Atwell, J. Gramp, R. Duszak Jr., Clinical outcomes in patients with suspected acute pulmonary embolism and negative helical computed tomographic results in whom anticoagulation was withheld, Arch Intern. Med. 163 (17) (2003) 2033–2038, https://doi.org/10.1001/archinte.163.17.2033.

[11] K.K. Kumamaru, A.R. Hunsaker, H. Kumamaru, E. George, A. Bedayat, F.J. Rybicki, Correlation between early direct communication of positive CT pulmonary angiography findings and improved clinical outcomes, Chest 144 (5) (2013) 1546–1554, https://doi.org/10.1378/chest.13-0308.

[12] S. Soffer, E. Klang, O. Shimon, et al., Deep learning for pulmonary embolism detection on computed tomography pulmonary angiogram: a systematic review and meta-analysis, Sci. Rep. 11 (1) (2021) 15814, https://doi.org/10.1038/s41598-021-95249-3.

[13] S.J. Kligerman, J.W. Mitchell, J.W. Sechrist, A.K. Meeks, J.R. Galvin, C.S. White, Radiologist performance in the detection of pulmonary embolism: features that favor correct interpretation and risk factors for errors, J. Thorac Imaging. 33 (6) (2018) 350–357, https://doi.org/10.1097/rti.0000000000000361.

[14] P. Hamet, J. Tremblay, Artificial intelligence in medicine, Metabolism 69s (2017) S36–S40, https://doi.org/10.1016/j.metabol.2017.01.011.

[15] G. Litjens, T. Kooi, B.E. Bejnordi, et al., A survey on deep learning in medical image analysis, Med Image Anal. 42 (2017) 60–88, https://doi.org/10.1016/j.media.2017.07.005.

[16] Fabian Isensee PFJ, Simon A. A. Kohl, Jens Petersen, Klaus H. Maier-Hein. Automated Design of Deep Learning Methods for Biomedical Image Segmentation 2020.
[17] X. Li, X. Wang, X. Yang, Y. Lin, Z. Huang, Preliminary study on artificial intelligence diagnosis of pulmonary embolism based on computer in-depth study, Ann. Transl. Med. 9 (10) (2021) 838, https://doi.org/10.21037/atm-21-975.
[18] R.C. Wang, D.L. Miglioretti, E.C. Marlow, et al., Trends in imaging for suspected pulmonary embolism across US health care systems, 2004 to 2016, JAMA Netw. Open 3 (11) (2020) e2026930, https://doi.org/10.1001/jamanetworkopen.2020.26930.
[19] K. Batra, Y. Xi, S. Bhagwat, A. Espino, R.M. Peshock, Radiologist worklist reprioritization using artificial intelligence: impact on report turnaround times for CTPA examinations positive for acute pulmonary embolism, AJR Am J. Roentgenol. 221 (3) (2023) 324–333, https://doi.org/10.2214/ajr.22.28949.
[20] M. Mitka, Joint commission warns of alarm fatigue: multitude of alarms from monitoring devices problematic, JAMA 309 (22) (2013) 2315–2316, https://doi.org/10.1001/jama.2013.6032.
[21] H. Huhtanen, M. Nyman, T. Mohsen, A. Virkki, A. Karlsson, J. Hirvonen, Automated detection of pulmonary embolism from CT-angiograms using deep learning, BMC Med. Imaging 22 (1) (2022) 43, https://doi.org/10.1186/s12880-022-00763-z.
[22] P. Ajmera, A. Kharat, J. Seth, et al., A deep learning approach for automated diagnosis of pulmonary embolism on computed tomographic pulmonary angiography, BMC Med. Imaging 22 (1) (2022) 195, https://doi.org/10.1186/s12880-022-00916-0.
[23] B. Ghaye, Peripheral pulmonary embolism on multidetector CT pulmonary angiography, Jbr-Btr. 90 (2) (2007) 100–108.
[24] R.S. Wiener, L.M. Schwartz, S. Woloshin, When a test is too good: how CT pulmonary angiograms find pulmonary emboli that do not need to be found, BMJ 347 (2013) f3368, https://doi.org/10.1136/bmj.f3368.
[25] H. Zhu, G. Tao, Y. Jiang, et al., Automatic detection of pulmonary embolism on computed tomography pulmonary angiogram scan using a three-dimensional convolutional neural network, Eur. J. Radiol. 177 (2024) 111586, https://doi.org/10.1016/j.ejrad.2024.111586.
[26] D. Roggenland, S.A. Peters, S.P. Lemburg, T. Holland-Letz, V. Nicolas, C.M. Heyer, CT angiography in suspected pulmonary embolism: impact of patient characteristics and different venous lines on vessel enhancement and image quality, AJR Am J. Roentgenol. 190 (6) (2008) W351–W359, https://doi.org/10.2214/ajr.07.3402.
[27] B. Megyeri, A. Christe, S.T. Schindera, et al., Accuracy of computed tomography angiography in the detection of pulmonary embolism in patients with high body weight, Eur. J. Intern Med. 25 (8) (2014) 724–730, https://doi.org/10.1016/j.ejim.2014.07.003.
[28] K.N. Blackmon, C. Florin, L. Bogoni, et al., Computer-aided detection of pulmonary embolism at CT pulmonary angiography: can it improve performance of inexperienced readers? Eur. Radiol. 21 (6) (2011) 1214–1223, https://doi.org/10.1007/s00330-010-2050-x.
[29] K. Lahiji, S. Kligerman, J. Jeudy, C. White, Improved accuracy of pulmonary embolism computer-aided detection using iterative reconstruction compared with filtered back projection, AJR Am J. Roentgenol. 203 (4) (2014) 763–771, https://doi.org/10.2214/ajr.13.11838.
[30] G. Lee, H.Y. Lee, H. Park, et al., Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical management: State of the art, Eur. J. Radiol. 86 (2017) 297–307, https://doi.org/10.1016/j.ejrad.2016.09.005.
[31] L.M. Prevedello, B.S. Erdal, J.L. Ryu, et al., Automated critical test findings identification and online notification system using artificial intelligence in imaging, Radiology 285 (3) (2017) 923–931, https://doi.org/10.1148/radiol.2017162664.
[32] J.C. Rayan, N. Reddy, J.H. Kan, W. Zhang, A. Annapragada, Binomial classification of pediatric elbow fractures using a deep learning multiview approach emulating radiologist decision making, Radiology: Artif. Intell. 1 (1) (2019) e180015, https://doi.org/10.1148/ryai.2019180015.
[33] T. Weikert, D.J. Winkel, J. Bremerich, et al., Automated detection of pulmonary embolism in CT pulmonary angiograms using an AI-powered algorithm, Eur. Radiol. 30 (12) (2020) 6545–6553, https://doi.org/10.1007/s00330-020-06998-0.
[34] L. Topff, E.R. Ranschaert, A. Bartels-Rutten, et al., Artificial intelligence tool for detection and worklist prioritization reduces time to diagnosis of incidental pulmonary embolism at CT, Radiol Cardiothorac Imaging. 5 (2) (2023) e220163, https://doi.org/10.1148/ryct.220163.