scientific reports

Check for updates

OPEN

Detection and segmentation of pulmonary embolism in 3D CT pulmonary angiography using a threshold adjustment segmentation network

Jian-cong Fan[1,2], Haoyang Luan[1,2], Yaqian Qiao[3], Yang Li[1,2]✉ & Yande Ren[3]✉

Pulmonary embolism is a life-threatening condition where early diagnosis and precise localization are crucial for improving patient outcomes. While CT pulmonary angiography (CTPA) is the primary method for detecting pulmonary embolism, existing segmentation algorithms struggle to effectively distinguish thrombi from vascular structures in complex 3D CTPA images, often leading to both false positives and false negatives. To address these challenges, the Threshold Adjustment Segmentation Network (TSNet) is proposed to enhance segmentation performance in 3D CTPA images. TSNet incorporates two core modules: the Threshold Adjustment Module (TAD) and the Geometric-Topological Axial Feature Module (GT-AFM). TAD utilizes logarithmic scaling, adaptive adjustments, and nonlinear transformations to optimize the probability distributions of thrombi and vessels, reducing false positives while improving the sensitivity of thrombus detection. GT-AFM integrates geometric features and topological information to enhance the recognition of complex vascular and thrombotic structures, improving spatial feature processing. Experimental results show that TSNet achieves a sensitivity of 0.761 and a false positives per scan of 1.273 at $\varepsilon = 0$ mm. With an increased tolerance of $\varepsilon = 5$ mm, sensitivity improves to 0.878 and false positives per scan decreases to 0.515, significantly reducing false positives. These results indicate that TSNet demonstrates superior segmentation performance under various tolerance levels, showing robustness and a well-balanced trade-off between sensitivity and false positives, making it highly promising for clinical applications.

Keywords Pulmonary embolism, CT pulmonary angiography, 3D image detection and segmentation, Threshold adjustment, Medical image processing

Pulmonary embolism is a serious medical emergency caused primarily by blood clots or other substances (e.g., fat, air, or tumor cells) obstructing the pulmonary arteries or their branches (as shown in Fig. 1). Pulmonary embolism has garnered significant attention from both the medical community and the public due to its acute onset and potentially fatal consequences. Studies indicate that even with treatment, pulmonary embolism has a mortality rate is approximately 8%, while the recurrence rate can vary between 2% and 50% depending on individual patient factors[1]. Pulmonary embolism typically originates from deep vein thrombosis in the lower extremities, Clots reach the lungs via the bloodstream and block pulmonary arteries. This disrupts lung blood flow, causing dyspnea, straining the heart, and potentially damaging lung tissue or leading to hypoxia. Severe cases may cause cardiac arrest. Common symptoms include chest pain, coughing, shortness of breath, sweating, and fainting. However, the clinical manifestations of pulmonary embolism are often non-specific and resemble those of other cardiopulmonary diseases, making misdiagnosis more likely and complicating clinical differentiation. Failure to promptly diagnose and treat pulmonary embolism can lead to severe complications, potentially threatening the patient's life. Therefore, accurate and timely diagnosis and treatment are essential

[1]College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, Shandong Province, PR China. [2]Provincial Key Laboratory for Information Technology of Wisdom Mining of Shandong Province, Shandong University of Science and Technology, Qingdao, China. [3]Department of Radiology, The Affiliated Hospital of Qingdao University, Qingdao, China. ✉email: dreyang@163.com; 8198458ryd@qdu.edu.cn

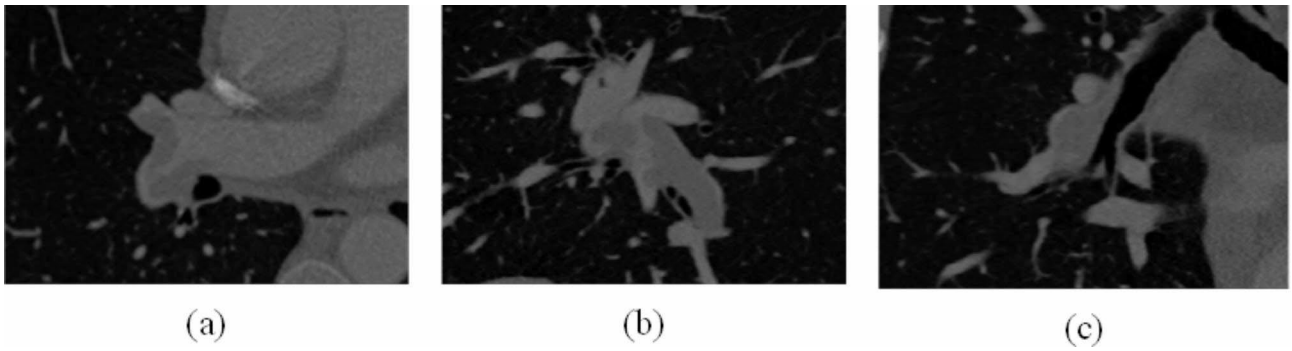

Fig. 1. Illustrates the typical presentation of pulmonary embolism as one or more filling defects within the pulmonary arteries or their branches. These defects appear as regions with reduced enhancement compared to the surrounding contrast, leading to distinct imaging features. Figures (**a**), (**b**), and (**c**) display partial filling defects of the arteries in axial, sagittal, and coronal views, respectively.

for optimal patient outcomes. In clinical practice, physicians must comprehensively assess the patient's medical history, symptoms, and relevant test results to accurately diagnose pulmonary embolism and mitigate associated risks.

In recent years, significant advances have been made in pulmonary embolism diagnostic methods. D-dimer is a widely recognized biomarker for thrombosis detection[2], as a fibrin degradation product fragment, is typically associated with thrombus formation when elevated. D-dimer levels are frequently measured during early diagnosis of pulmonary embolism to assess the likelihood of thrombosis. However, elevated D-dimer levels are not specific to pulmonary embolism, as various other diseases or conditions can also induce such elevations. Thus, D-dimer testing is typically combined with other imaging tests for the.

diagnosis of pulmonary embolism. Despite radiologic imaging advancements, chest X-rays are occasionally used for the initial screening of pulmonary embolism screening. They generally fail to directly visualize pulmonary artery clots, especially during early stages or when image quality is suboptimal, limiting diagnostic accuracy. In contrast, pulmonary arteriography offers more detailed vascular imaging, effectively revealing obstructions in the pulmonary arteries or their branches, making it especially suitable for critically ill patients or cases requiring precise diagnosis. Therefore, pulmonary arteriography continues to be a reliable method for evaluating pulmonary embolism in clinical practice.

Pulmonary arteriography was once the gold standard for the diagnosis of pulmonary embolism, but it was difficult to provide sufficient detail in the assessment of subsegmental and subsegmental arterial pulmonary embolism due to resolution limitations and contrast risk. The introduction of CTPA has significantly improved the detection of fine thrombi, and it is particularly effective in the assessment of small arterial pulmonary embolisms, which has led to improved diagnostic and therapeutic options[3]. Additionally, Magnetic Resonance Angiography (MRA), as proposed by Meaney et al., has been used as a non-invasive test to detect pulmonary embolism[4]. However, it is less favorable than CTPA in terms of spatial resolution and detection speed. CTPA is more widely adopted in clinical practice owing to its superior image clarity, rapid detection, and ease of use[5]. Although CTPA provides high-resolution images of the pulmonary vasculature, the detection of fine thrombi may be compromised by high image noise or suboptimal patient positioning, and its limitations of high-dose radiation and reliance on contrast media make computer-aided diagnostic techniques, which use automated image analysis and deep-learning algorithms to assist physicians in making a more accurate diagnosis, particularly important[6].

The paper is organized as follows: Chap. 2 introduces the related work and reviews the fields and existing methods related to this study; Chap. 3 describes our proposed method and model in detail; Chap. 4 demonstrates the experimental design and result analysis to validate the methodology; and finally, Chap. 5 summarizes the contributions of this paper and discusses future research directions.

Related work

In recent years, the UNet model has been widely adopted as a foundational framework in research[7]. The computer - aided system proposed by Jiantao Pu et al. segments arteries, applies adaptive thresholds, etc., enabling the identification of pulmonary embolism without manual outlining[8]. AlexNet, developed by Alex Krizhevsky et al.[9], significantly improves classification accuracy and optimizes the processing of large-scale image data through a deep convolutional structure and 'dropout' regularization. This model represents an important breakthrough in the field of image recognition. Yoshitaka M et al.[10] proposed a pulmonary embolism detection method based on image analysis, combining geometric features to analyze three-dimensional data and improve detection accuracy.

Accurate medical image segmentation is crucial for detecting pulmonary embolism, particularly in distinguishing it from neighboring structures such as the lung parenchyma, interstitium, pleura, lymph nodes, and muscles. Precise segmentation enables physicians to make informed clinical decisions based on detailed anatomical information. The limitations of current image resolution and segmentation algorithms continue to challenge the accurate detection of pulmonary embolism. Yuan C et al. highlighted that fine segmentation

methods enhance the visualization of vascular anatomy and effectively detect intravascular abnormalities such as thrombi[11].

With the advancement of computer vision and machine learning techniques, pulmonary artery detection and technology have made significant progress in accurately distinguishing the boundaries of pulmonary arteries. Meticulous pulmonary artery segmentation is considered to be the key to improving the detection rate of pulmonary embolism when dealing with ambiguous thrombi or thrombi with similar density to the surrounding tissues. Meticulous pulmonary artery segmentation is considered key to improving the detection rate of pulmonary embolism. An adaptive 3D voxel clustering method based on expectation-maximization analysis, proposed by Zhou C et al.[12], can be dynamically adjusted to fit different vascular features, thus improving the accuracy of detection of subtle pulmonary embolism. However, the method may lead to instability in the detection results when dealing with image noise and complex backgrounds, which in turn affects the overall diagnostic performance. Soffer S et al. proposed a classifier-guided detection method combining of deep learning techniques and attention-guided convolutional neural networks[13]. The approach achieved significant improvements in the automated diagnosis of pulmonary embolism in CTPA, particularly in detecting small thrombi. The attention mechanism improved the detection of small thrombi and enhanced the accuracy of automated diagnosis. The 3D CNN model incorporating vascular information proposed by Tajbakhsh N et al. has demonstrated significant advancements in enhancing diagnostic accuracy. However, the model has yet to achieve full end-to-end integration and still requires additional pre-processing and post-processing steps[14]. The deep neural network model developed by Huhtanen H et al. combines long short-term memory (LSTM) networks with convolutional neural networks (CNNs)[15]. The model was trained on chest X-ray and natural images and applied to binary label annotation of CTPA images, demonstrating excellent performance in pulmonary embolism diagnosis.

Similarly, Long Kun et al. proposed a probabilistic flexible Mask R-CNN model that enhances the detection of small and sub-segment thrombi by up-sampling feature maps and using anchor extraction techniques, significantly improving detection efficiency[16]. In addition, Huang S C et al. developed and evaluated an end-to-end deep learning model[17]. This model effectively improved the efficiency and accuracy of pulmonary embolism detection by fully utilizing volumetric CTPA data.

The AANet model proposed by Guo J et al.[18] focuses on pulmonary embolism detection and enhances performance via an arterial context fusion block and multi-scale feature fusion. However, effectively distinguishing pulmonary embolism from veins and soft tissues remains difficult due to their subtle differences. Pulmonary embolism detection in 3D CTPA images has long relied heavily on conventional convolutional neural network structures. However, these methods face many challenges when processing 3D medical images. Conventional CNN architectures, relying on single - scale feature extraction and ignoring critical contextual information, may experience feature loss when dealing with diverse lesions. To address some of these challenges, Islam N. U. et al.[19]proposed a framework incorporating vessel-oriented image representations and self-supervised pretraining, achieving improved pulmonary embolism detection performance and false positive reduction, particularly at the clot level. Recent studies have shown that attention mechanisms can enhance model performance by focusing on important regions in the image. Earlier attentional models mainly enhanced features in the channel or spatial dimension, but failed to fully utilize the complementarity between dimensions, and failed to handle the diversity of input features, thus limiting the comprehensive understanding of complex medical images[20].

Against the background of facing the above challenges, this paper proposes a modular architecture under the condition of reducing the false positive rate to ensure its sensitivity to a great extent.

- Threshold Adjustment Module introduces a learnable threshold adjustment parameter to logarithmically adjust the difference in predicted probability between thrombus and vessel.
- Geometric-Topological Axial Feature Module combines geometric and topological features with axial features to focus on the possible location of pulmonary embolism and to extract information about the critical structure of the vessel.
- Channel Split Swap Attention Module segments subsets of feature maps, which are merged after processing by 3D convolution and adaptive attention units, and explores feature combinations in different orders to enhance the expressive power of 3D feature maps.

Methods

The study was approved by the Medical Ethics Committee of the Affiliated Hospital of Qingdao University (No. QYFYWZLL29335), and informed consent was waived by the Medical Ethics Committee of the Affiliated Hospital of Qingdao University, The study was conducted in accordance with the relevant guidelines and regulations.

The overall framework of the model proposed in this paper is shown in Fig. 2. The input 3D image data is first processed through four coding layers for feature extraction, and then passed to four decoding layers for reconstruction and segmentation. Specifically, the first convolutional layer performs initial feature extraction on the input data to obtain key image features. The second coding layer introduces the CSSA-C module to focus on segmentation and fusion of channel features, while the third coding layer applies the CSSA-A module to further extract the combined features of channels and spatial information. After these feature extractions, the feature maps are passed to the fourth coding layer for processing using the GT-AFM module. This module integrates geometric and topological information from the 3D axial structure, enhancing the perception of axial information and spatial features.

In the decoding stage, the feature maps are up-sampled and reconstructed through the four decoding layers, ultimately generating probabilistic predictive maps for vessels and thrombi. These probability maps are then analyzed by the threshold adjustment module to determine the optimal thresholds for achieving accurate

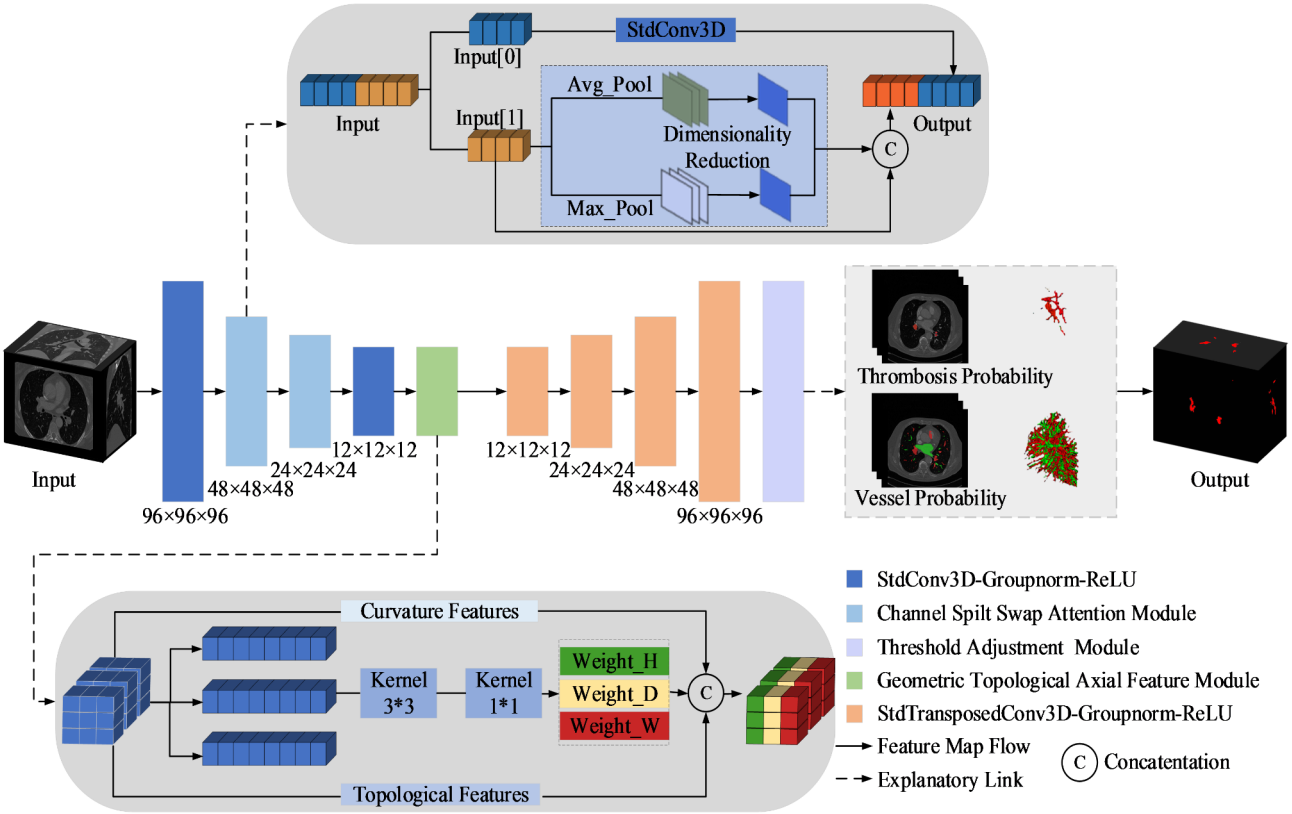

Fig. 2. Overall model architecture.

segmentation results. Finally, the TAD module fine-tunes the predicted probabilities for vessels and thrombi, resulting in accurate segmentation maps for thrombi.

Threshold adjustment module

The TAD module is an adaptive parameter adjustment mechanism for 3D CTPA based pulmonary embolism detection. It combines thrombus and vessel probabilistic information for accurate results. The model in this paper predicts vascular and thrombus probabilities simultaneously, leveraging the physiological property that thrombi are confined within blood vessels.

To optimize the prediction, the module identifies regions with high thrombus probability from the thrombus probability map using a predefined threshold. These regions are then cross-validated with the vascular probability map to ensure their physiological plausibility. Regions aligned with vascular areas are retained as valid, while others are adjusted to reduce false positives. A dynamic adjustment mechanism based on local consistency between thrombus and vascular probabilities is used to ensure the model's learnable parameters adapt to the data characteristics.

The optimization incorporates the logarithmic difference between thrombus and vascular probability maps, a crucial step in ensuring their consistency and enhancing the model's performance in thrombus detection. This approach dynamically adjusts based on the probability differences between the two maps. In scenarios where the thrombus prediction probability is high while the corresponding vascular probability is low, it is likely an incorrect thrombus prediction. The logarithmic difference effectively amplifies this significant disparity, enabling the model to identify such potential errors.

Conversely, when the thrombus prediction and the vascular probability in the corresponding area and its neighborhood are both high, indicating a likely correct identification of a thrombus within a vessel, there can still be differences due to the size difference between the thrombus and the vessel (the vessel's volume being larger, leading to a higher probability in its neighborhood). Here, the logarithmic difference reduces these differences, maintaining the model's stability and ensuring accurate judgment.

In addition to the role of the logarithmic difference, L_p-norm regularization promotes smooth and consistent predictions. By penalizing excessive fluctuations in probability values, especially in regions where probabilities are similar, it improves the model's generalization ability and robustness, which is particularly important in areas where thrombus and vessel overlap or are in close proximity.

A nonlinear transformation further refines the interaction between thrombus and vascular probabilities. It magnifies relevant differences to better distinguish thrombus from surrounding tissue, while suppressing noise in low-probability regions. Additionally, scaling and normalization help balance disparities, preventing overemphasis on extreme values and improving detection accuracy.

Finally, the TAD module utilizes the learned threshold parameters to adaptively threshold the predicted thrombus and vessel probabilities to ensure the accuracy and consistency of the prediction results. The logarithmic interpolation method used to calculate the vessel and thrombus probabilities is shown in Eq. (1).

$$adj = \sum_{i=1}^{n} \left(\frac{\alpha_i \cdot tp_i + \varepsilon}{\beta_i \cdot vp_i + \varepsilon} \right) \omega_i + \lambda \cdot \|t_P - v_p\|_p^p \tag{1}$$

where α_i and β_i are learnable parameters, and tp_i and vp_i represent the thrombus probability tensor and the vascular probability tensor for the ith voxel, respectively. ω_i is a weighting factor for the logarithmic difference, which adjusts its effect on each voxel. λ is a parameter used to control the degree of influence of the regularization term. $\|t_P - v_p\|_p^p$ denotes the p power of the p-paradigm between tp and vp and is used as a measure of the difference between the thrombus and vascular probabilities. After the adjustment of the logarithmic difference, the thrombus probability values are scaled nonlinearly, as shown in Eq. (2).

$$adj_{res} = \left(tp^{\gamma_1} + \eta \cdot cos\left(\omega \cdot tp\right) \right) \cdot exp\left(-\gamma \cdot adj + \Phi \cdot \frac{\partial\, adj}{\partial\, tp} \right) \tag{2}$$

The formulation combines a nonlinear scaling of thrombus probability tp^{γ_1} with a periodic adjustment term $\eta \cdot \cos(\omega \cdot \mathrm{tp})$, achieving precise control over thrombus detection. The exponential function is adjusted according to the logarithmic difference adj. By incorporating the periodic term, the model introduces oscillatory behavior, enhancing sensitivity to periodic thrombus patterns. Changes in thrombus probability are captured by the differential term $\Phi \cdot \frac{\partial\, adj}{\partial\, tp}$, allowing the scaling factor γ to better adapt to data characteristics. This design ensures the adjustment process captures both global trends and local variations effectively. Further nonlinear transformations and normalization steps, which are essential for improving the accuracy of the model, are described in Eq. (3).

$$adj_{res} = \sigma \left(\frac{adj_{res} \cdot \left((1+\delta)\right)}{1 + \|adj_{res}\|_q^q} \right) + \varepsilon \cdot \log\left(\frac{1}{1 + \exp\left(-adj_{res}\right)} \right) \tag{3}$$

The σ stands for the Sigmoid function, which is used to constrain the results to the range [0, 1]. It ensures that the output aligns with the requirements of probability values and prevents results from exceeding a reasonable range. The nonlinear adjustment factor δ and the normalization factor $\|adj_{res}\|_q^q$ are utilized to enhance the optimization of the results. By combining with the nonlinear adjustment factor δ, the Sigmoid function provides finer control over the results, offering greater flexibility to adapt to different image features. The logarithmic function log and the correction factor ε are applied to smooth the adjustment results and prevent excessive fluctuations. Additionally, the regularization term $\|adj_{res}\|_q^q$ in the denominator helps suppress large oscillations during the adjustment process, ensuring more stable results and mitigating overfitting. Finally, adaptive threshold control is employed to manage the thrombus probability. See Eq. (4) for details.

$$final_{res} = \begin{cases} \min\left(\frac{\sum_{j=1}^{m} \omega_j \cdot tanh\left(\Phi_j \cdot vp_j\right)}{m}, 1 \right), if\ Softplus\,(vp) \\ 0, \qquad\qquad otherwise \end{cases} \tag{4}$$

The impact of vascular information is assessed using the weighted average of the $\frac{\sum_{j=1}^{m} \omega_j \cdot \tanh\left(\Phi_j \cdot vp_j\right)}{m}$, and the final adjustment is combined with the Softplus function.

Channel split swap attention module

The CSSA module enhances the representation of 3D feature maps by dividing the input feature map into two subsets, each processed with distinct operations. Based on the principle of channel blending[21], one subset extracts local spatial structure and texture information via a standard 3D convolution module, while the other subset evaluates the importance of each region in the feature map using the Adaptive Attention Unit (AAU) and adjusts feature responses to suppress irrelevant regions. The CSSA module then exchanges the processing order between the two subsets, providing two configuration modes: CSSA-C and CSSA-A. In CSSA-C mode, the subset is first subjected to the 3D convolution operation, while in CSSA-A mode, it is processed by the AAU first. This design explores different feature space combinations[22], supporting subsequent feature fusion and segmentation operations.

Upon completing the convolution and attention operations, the two subsets are fused along the channel dimension, producing a feature representation that integrates local details with global context[23,24]. The channel attention mechanism in the AAU module captures global context via global pooling, generating a weight map to emphasize specific channels, while the spatial attention refines the feature representation across the spatial dimension using a larger receptive field[25,26]. In contrast to the traditional Convolutional Block Attention Module (CBAM)[27], the AAU module combines depthwise separable convolution with dilated convolution to enhance feature extraction efficiency and reduce computational complexity. To mitigate overfitting, the AAU module incorporates a regularization layer, bolstering the model's robustness across diverse datasets[28].

Geometric-topological axial feature module

The GT-AFM module improves the spatial processing of 3D CTPA images by enhancing axial features within 3D convolutional neural networks, specifically in pulmonary embolism detection and segmentation tasks. This module refines the representation of features along the height, width, and depth axes, offering more detailed spatial guidance, which aids in the precise localization and identification of pulmonary emboli. By fusing geometric and topological features, the module further enhances the recognition of details and the understanding of structures in medical image segmentation.

Geometric features capture local structural information through curvature and edge detection, enabling the model to accurately delineate boundaries and shapes, By calculating image curvature, they focus on local geometric characteristics of thrombus locations, where vessel deformation or curvature changes often occur, facilitating accurate thrombus localization. In contrast, topological features reflect structural persistence across scales using persistence maps, emphasizing global structures like blood vessel connectivity. These features identify disruptions or anomalies in vascular networks caused by thrombus, aiding in the recognition of large-scale structural information and reducing confusion with other tissues.

Figure 3 illustrates key features of the GT-AFM module. Figure 3a highlights enhanced contrast in image structure and detail. Figure 3b shows a curvature map from Gaussian filtering, capturing edge features and geometric information. Figure 3c enhances structural edges using the Sobel operator. Figure 3d presents persistence maps: dimension 0 indicates stability of connected regions at smaller scales, while dimension 1 shows points clustering near the diagonal and others above it. This reflects feature merging across scales, where diagonal points represent convergence of small-scale regions, and off-diagonal points highlight independent features persisting at smaller scales. These maps reveal the hierarchical and multiscale structural evolution, aligning with the GT-AFM module's functionality.

Incorporating geometric and topological features into the axial attention mechanism enhances the representation of complex structures in 3D medical images. Geometric features capture image details by identifying boundaries, curvatures, and local shapes, revealing spatial relationships and shape variations between regions, which is crucial for identifying and segmenting key structures. Topological features, in contrast, emphasize the structural properties of enclosed regions, effectively reflecting structural persistence and hierarchical relationships across different scales, particularly in complex shapes like tumors. After separately computing the geometric and topological feature maps, they are fused through element-wise feature fusion, where corresponding elements from both feature maps are combined. This fused representation strengthens the model's ability to detect subtle and intricate structures, thereby improving the accuracy and reliability of medical image detection and segmentation.

During feature processing, an adaptive mean pooling layer compresses features along the height, width, and depth axes independently, ensuring comprehensive capture of multi-scale information in each direction. A fully connected layer then encodes and recalibrates this global information. The channel attention mechanism further enhances this by automatically adjusting feature importance through learned weights. Specifically, the

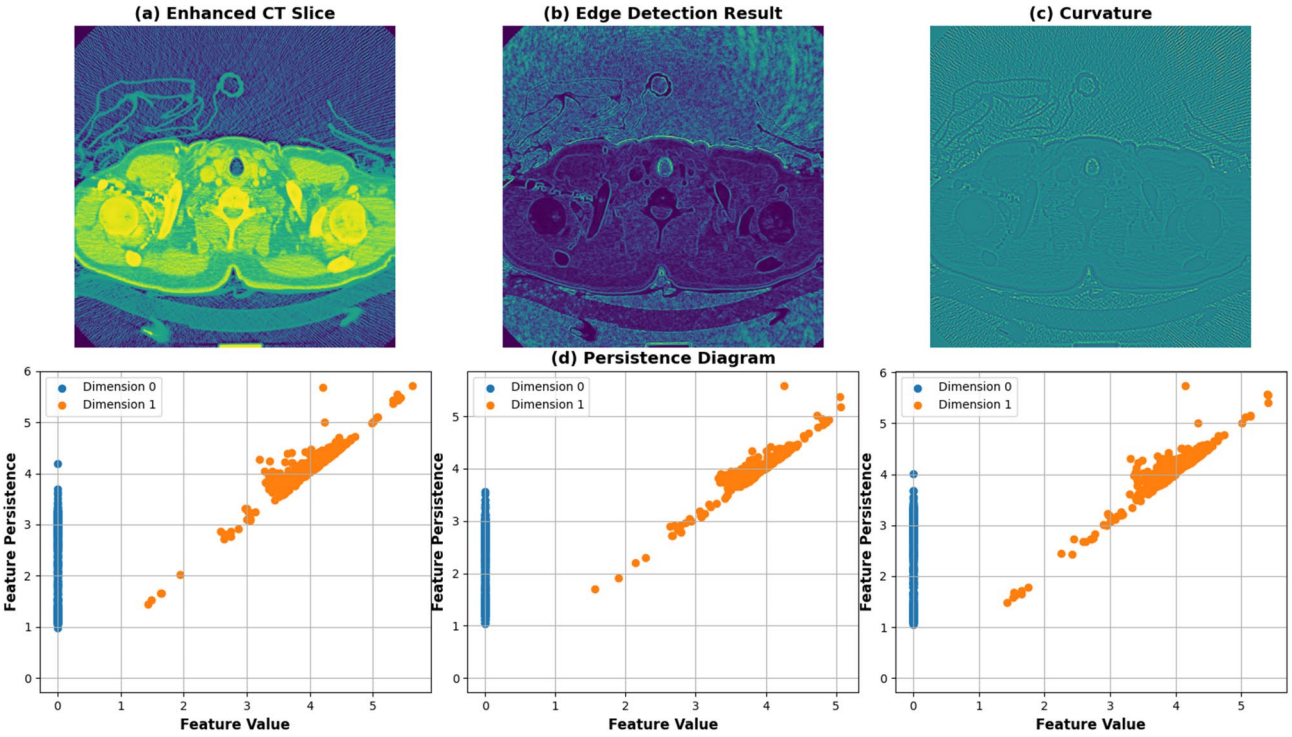

Fig. 3. Combined visualization of contrast-enhanced CT slices, edge detection results, curvature maps, and persistence diagrams.

mechanism computes feature importance weights for each channel by calculating the global average feature value, amplifying key information relevant to the task while reducing interference from irrelevant data. This adjustment optimizes the model's sensitivity to critical features, as reflected in Eqs. (5), (6), and (7), thereby improving overall model performance.

$$G_\sigma(x) = \frac{1}{(2\pi\sigma^2)^{3/2}} \exp\left(-\frac{||x-\mu||^2}{2\sigma^2}\right) \tag{5}$$

$$PH_k(X) = \{(b_i, d_i) \,|i \in I\} \tag{6}$$

$$Y = \odot(w_h + w_w + w_d) + G_\sigma(x) + PH_k(X) \tag{7}$$

where σ denotes the standard deviation; w_h, w_w, w_d represents the feature weights extracted along each of the three dimensions; G_σ (x) is the result of geometric feature extraction; $PH_k(X)$ is the result of topological feature extraction (persistent homotopy); and $\odot$ denotes element-wise multiplication.

Loss function

This paper designs the Focal Tversky Loss to enhance the thrombus detection model's performance in complex images. Based on the Tversky Loss[29], it incorporates a focusing mechanism to boost the model's attention on hard - to - classify samples.

The Tversky Loss gauges error sensitivity via the weighted relation of TP, FP, and FN. In thrombus detection, it balances FN and FP, helping the model recognize rare thrombus areas and complex backgrounds. The added focusing mechanism, with an exponential term in the Tversky Score, reduces weight on easy - to - classify regions and intensifies focus on difficult ones. This significantly improves the model's ability to detect hard - to - distinguish thrombi.

where TP, FP, and FN denote true positives, false positives, and false negatives, respectively, γ controls the strength of the focal mechanism, and ϵ is a constant that prevents the division by zero error.

With this design, the loss function can effectively improve the model's ability to detect thrombus regions, especially when dealing with rare and indistinguishable thrombi, which improves the accuracy and reliability of detection.

$$L_{Tversky} = 1 - \frac{\sum_i TP_i + \varepsilon}{\sum_i TP_i + \sum_i FP_i + \sum_i FN_i + \varepsilon} \tag{8}$$

$$L_{Focal\ Tversky} = (1 - T_{Tversky})^\gamma \tag{9}$$

Experiments

Datasets

The dataset is derived from the Pulmonary Embolism Challenge[30] and includes 91 CTPA pulmonary artery scans obtained from six hospitals in the Madrid region. These images were diagnosed and annotated by experienced radiologists to confirm the presence of.

pulmonary embolism. For this study, the dataset comprises 71 images for model training and 20 images for model evaluation. The images are represented as 3D volumetric data with dimensions of (512, 512, 452) and a resolution of (0.583984375, 0.583984375, 0.5), reflecting the resolution in each spatial dimension. This spatial information is crucial for understanding the relationships and scales among different regions and serves as a foundation for further data analysis and model optimization. A limitation of this dataset is that 12 pulmonary embolism labels are absent from the officially provided data. Given that the detection and segmentation of thrombi are inherently challenging, this limitation contributed to the model's poor performance in sensitivity during subsequent evaluations.

Data preprocessing

In this paper, a detailed design approach is adopted during the data preprocessing and training process. Data preprocessing involves several steps. First, fixed - size image blocks are extracted from 3D medical imaging data. Then, cropping and spatial transformations such as random scaling, translation, rotation, and flipping are performed. These operations are used to simulate different patient poses and scanning conditions. As a result, the diversity of the training data is increased, and the model's generalization ability is enhanced. Additionally, density window adjustments were applied to ensure consistency among image blocks. A goal - directed sampling strategy was implemented. This strategy prioritizes image regions containing pathological features. By doing so, the model's ability to recognize lesions is improved, and its robustness against noise and variations in image quality is enhanced.

During the training process, to further improve the model's generalization ability, various data enhancement techniques such as padding, random cropping, and noise injection were applied. Additionally, by adjusting the cropping strategy and enhancement intensity, a balanced training set containing lesion and non - lesion image blocks was generated. The test set was only subjected to the necessary density window adjustment and standard cropping, and no random spatial transformation was applied to ensure the objectivity and accuracy of the evaluation. The preprocessing and training strategies were finely designed to address the challenges in medical image analysis. These strategies aim to optimize model performance, ensure the reliability of study results, and

enhance the practical applicability of the models. Focusing on common problems in medical image analysis, the data enhancement techniques are designed to improve the coping ability of deep learning models. This, in turn, is expected to enable the models to demonstrate better generalization capabilities in new patient data and unseen samples.

Experimental results

Table 1 provides a detailed comparison of the performance of each model in the task of pulmonary embolism detection and segmentation, The methods Mevis, UPM, and LMNIIT are derived from the results published in the ISBI Challenge[36,37], including sensitivity, false alarm rate (FP/S), and positive predictive value (PPV) under different tolerances ($\varepsilon = 0$ mm, 2 mm, and 5 mm). At $\varepsilon = 0$ mm, the model proposed in this paper achieved a sensitivity of 0.761, outperforming the other models in terms of pulmonary embolism detection accuracy. The FP/S was 1.273, indicating the model's effectiveness in reducing false alarms. Additionally, the PPV of 0.600 validated the model's accuracy in positive prediction. When the tolerance was expanded to 2 mm, the sensitivity increased to 0.850, which remained higher than that of most other models, indicating stable performance with a relaxed detection threshold. The FP/S decreased significantly to 0.565, reflecting fewer false alarms, while the PPV remained stable at 0.618, demonstrating reliable predictive ability. At $\varepsilon = 5$ mm, the sensitivity further improved to 0.878, illustrating the model's ability to detect pulmonary embolism under relaxed criteria. The FP/S dropped to 0.515, showing continued reduction in false positives. Although the PPV slightly decreased to 0.585, it remained at a relatively high level, suggesting the model's potential for practical applications. Meanwhile in Table 2 we compare the DSC (%) of thrombus segmentation as well as the HD95 metrics, our DSC metrics are better than the rest of the model, but the HD95 values show poor performance because the TAD module excludes a large number of thrombi with low potential, which can lead to poor segmentation of the borders as well as missed detection of some positive thrombi.

The overall analysis indicates that the proposed model demonstrated maintained excellent and stable performance across all tolerance conditions, maintaining a balance between sensitivity and false alarm rate. This outcome highlights the model's effective detection capability and suggests its potential utility in clinical applications.

The comparison of the number of true positives (TP) and false positives (FP) between the model proposed in this paper and better-performing models, as shown in Table 1, is visualized in Fig. 4. The heatmap on the left side of Fig. 4 illustrates that the proposed model consistently achieved a higher number of true positives across the three tolerance levels (0 mm, 2 mm, and 5 mm). Compared to the other models under the same conditions, the proposed model significantly reduced the false positive rate, especially at the 5 mm tolerance, where it yielded only 56 false positives, in contrast to Dual-Hop (112), nnUNet (94), AANet (98), and UNet-3D (59). These results suggest that the proposed model effectively lowered the false alarm rate while maintaining high detection accuracy, further validating its reliability and stability in pulmonary embolism detection tasks.

Figure 5 provides a detailed analysis of the pulmonary embolism segmentation performance of the proposed model at three tolerance levels ($\varepsilon = 0$ mm, 2 mm, and 5 mm). The evaluation, conducted using the Free-Response Receiver Operating Characteristic Curve (FROC), offers a comprehensive view of the model's sensitivity at different levels of false positives. In this figure, the y-axis represents sensitivity, defined as the ratio of correctly identified positive cases to all true positive cases, while the x-axis denotes the number of false positives, referring to cases incorrectly reported as positive by the model. Unlike the previously discussed sensitivity, typically assessed at a specific threshold or false positive count, the FROC curves illustrate the variation in the model's sensitivity across different false positive numbers, reflecting its performance under varying detection conditions.

At $\varepsilon = 0$ mm, the model's sensitivity reaches a maximum value of 0.76 when the number of false positives is 3. At $\varepsilon = 2$ mm, sensitivity increases to 0.85 with the same false positive count of 3. Furthermore, sensitivity rises to 0.88 when the number of false positives decreases to 2 at $\varepsilon = 5$ mm. These results demonstrate that the model achieves high sensitivity with fewer false positives, significantly outperforming other models.

This performance trend illustrates the robustness of the model and its ability to balance true positive detection with false positive control. Sensitivity continues to improve with increasing tolerance levels, highlighting the model's adaptability to varying degrees of detection stringency. Although false positives increase at higher tolerance levels, the model consistently maintains high sensitivity, which is crucial for accurate diagnosis and reliable clinical application.

In conclusion, the results underscore the model's utility in providing high sensitivity for pulmonary embolism detection while effectively managing the false alarm rate, positioning it as an important tool for clinical diagnosis.

Figure 6 presents the pulmonary embolism segmentation results, with each row containing four images. These images are categorized into three groups based on the sectioning planes: axial, sagittal, and coronal. Each group includes the original image, a ground truth map manually annotated by a medical expert, and the model-generated segmentation map. These visualizations demonstrate the model's segmentation performance across different planes. The original images depict the anatomical structure and morphological features of the pulmonary embolism region, while the ground truth maps provide a reference standard for evaluation against the model's output.

Ablation experiments

Systematic ablation experiments were conducted to evaluate the contribution of different modules to the performance of the model in.

pulmonary embolism detection and segmentation. The results compare the overall performance of the Baseline Model (BL) with the FROC curve scores of its modules at various boundary tolerances and false - positive thresholds. These results are summarized in Table 3. The BL exhibited mean sensitivity values of 0.6405 ($\varepsilon = 0$ mm), 0.7026 ($\varepsilon = 2$ mm), and 0.7436 ($\varepsilon = 5$ mm), reflecting its baseline performance in detecting pulmonary

Tolerance		UA-2D [31]	UA-2.5D [31]	UA-3D [31]	Mevis	ASU-Mayo [14]	FUM-Mvlab [32]	AA Net [18]	UPM	LMNIIT	AG-CNN[33]	Dual-Hop[34]	nnUNet[35]	Ours
ε = 0 mm	Sens	0.515	0.392	0.700	0.278	0.362	0.269	0.753	0.292	0.098	0.717	**0.769**	0.75	*0.761*
	FP/S	3.500	**0.250**	1.613	2.350	3.330	2.700	2.597	4.550	26.45	2.434	*0.523*	–	1.273
	PPV	0.489	**0.910**	0.562	0.440	0.416	0.393	0.473	0.295	0.024	0.324	0.446	–	*0.600*
ε = 2 mm	Sens	0.596	0.477	0.789	0.368	0.395	0.342	**0.859**	0.316	0.140	0.793	0.847	0.80	*0.850*
	FP/S	3.350	**0.200**	0.620	1.650	3.250	2.350	0.687	3.900	24.90	2.231	0.590	–	*0.565*
	PPV	0.503	**0.927**	0.584	0.560	0.409	0.453	0.492	0.316	0.031	0.458	0.451	–	*0.618*
ε = 5 mm	Sens	0.633	0.511	0.822	0.400	0.456	0.411	*0.871*	0.400	0.178	0.804	0.866	0.81	**0.878**
	FP/S	3.300	**0.200**	0.574	1.350	3.200	2.100	1.015	3.300	23.80	1.895	0.554	–	*0.515*
	PPV	0.463	**0.920**	0.556	0.571	0.390	0.468	0.446	0.353	0.033	0.467	0.411	–	*0.585*

Table 1. Evaluation values of different models at different tolerances, where bold and italic fonts represent the top two in the overall evaluation, respectively.

Method	DSC(%)	HD95
Swin-UNet[38]	67.49	16.56 ± 36.13
UNet++[39]	71.83	6.67 ± 9.81
SCUNet++[40]	80.42	**3.54 ± 7.91**
Ours	**82.57**	19.69 ± 25.54

Table 2. Comparison of DSC (%) and HD95 metrics across different network models. Significant values are in bold.

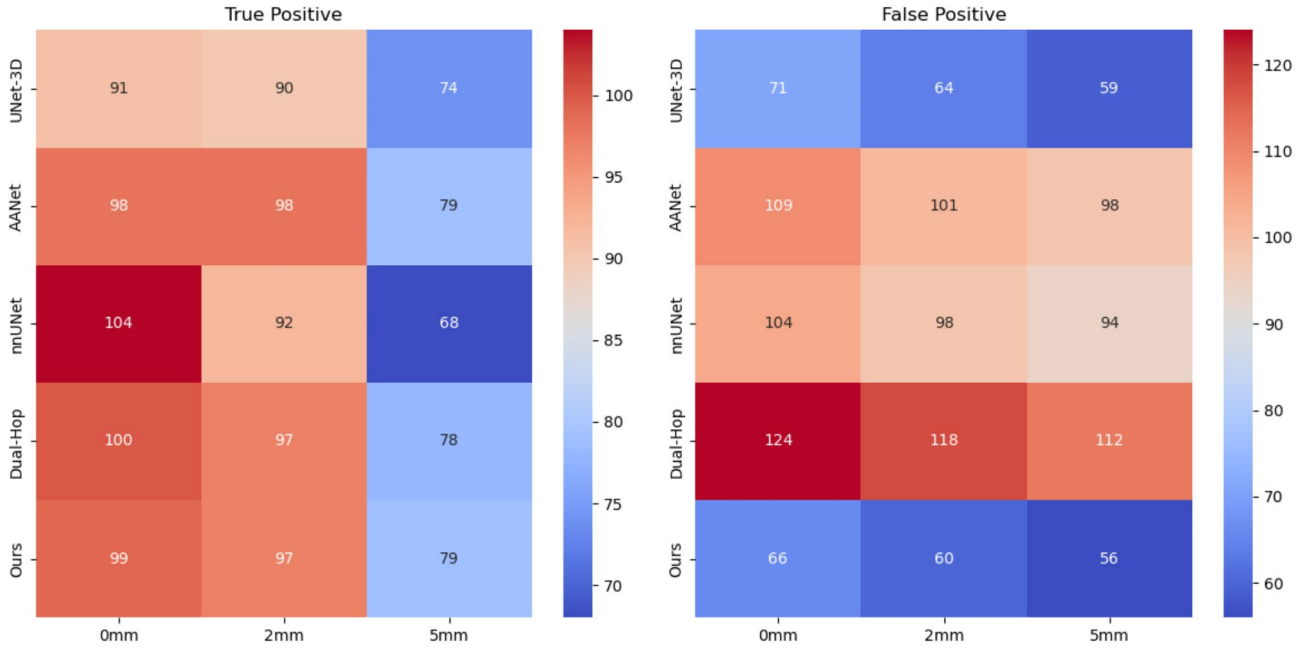

Fig. 4. Heat map of true positives and false positives based on the model with better overall detection performance in Table 1: Higher color values in the red areas indicate higher confidence in the true positive areas, and lower color values in the blue areas indicate a lower number of false positives and better model performance.

embolism. However, with the gradual integration of additional modules, the model's performance improved significantly.

After introducing the GT-AFM module, the mean sensitivity values increased to 0.6475 (ε = 0 mm), 0.7134 (ε = 2 mm), and 0.7459 (ε = 5 mm). The GT-AFM module enhances spatial information processing for 3D CTPA images, particularly in tasks involving the identification and segmentation of pulmonary embolism. This module optimizes the representation of features along the height, width, and depth axes by incorporating geometric and topological features, improving the model's capacity to recognize local structures.

Following the addition of the CSSA module, the model's mean sensitivity decreased slightly to 0.6432 (ε = 0 mm) and 0.7113 (ε = 2 mm), but increased to 0.7519 for ε = 5 mm. The CSSA module boosts 3D feature representation by effectively extracting channel features from the input feature map. By dividing the feature map into two subsets and applying different strategies to each—standard 3D convolution for general features and adaptive attention for detailed, task - relevant features, this module enhances the model's ability to emphasize critical information in complex 3D medical images.

Ultimately, the full model, with all modules combined, achieved the highest sensitivity across boundary tolerances, with mean values of 0.6538 (ε = 0 mm), 0.7478 (ε = 2 mm), and 0.7825 (ε = 5 mm). The TAD module contributes to the model by optimizing the prediction of thrombus - vessel relationships through a learnable threshold - tuning parameter. This optimization significantly improves the accuracy of critical region predictions, reduces false alarms, and enhances segmentation precision. As a result, the model ensures high sensitivity even in the presence of complex anatomical structures.

The results of the ablation experiments collectively validate the unique contribution of each module and its synergistic effect, emphasizing the key role of modular design in enhancing the performance of medical image detection and segmentation. These findings have provided an important theoretical foundation for further optimization of biomedical image analysis tools and have laid a solid groundwork for future research directions.

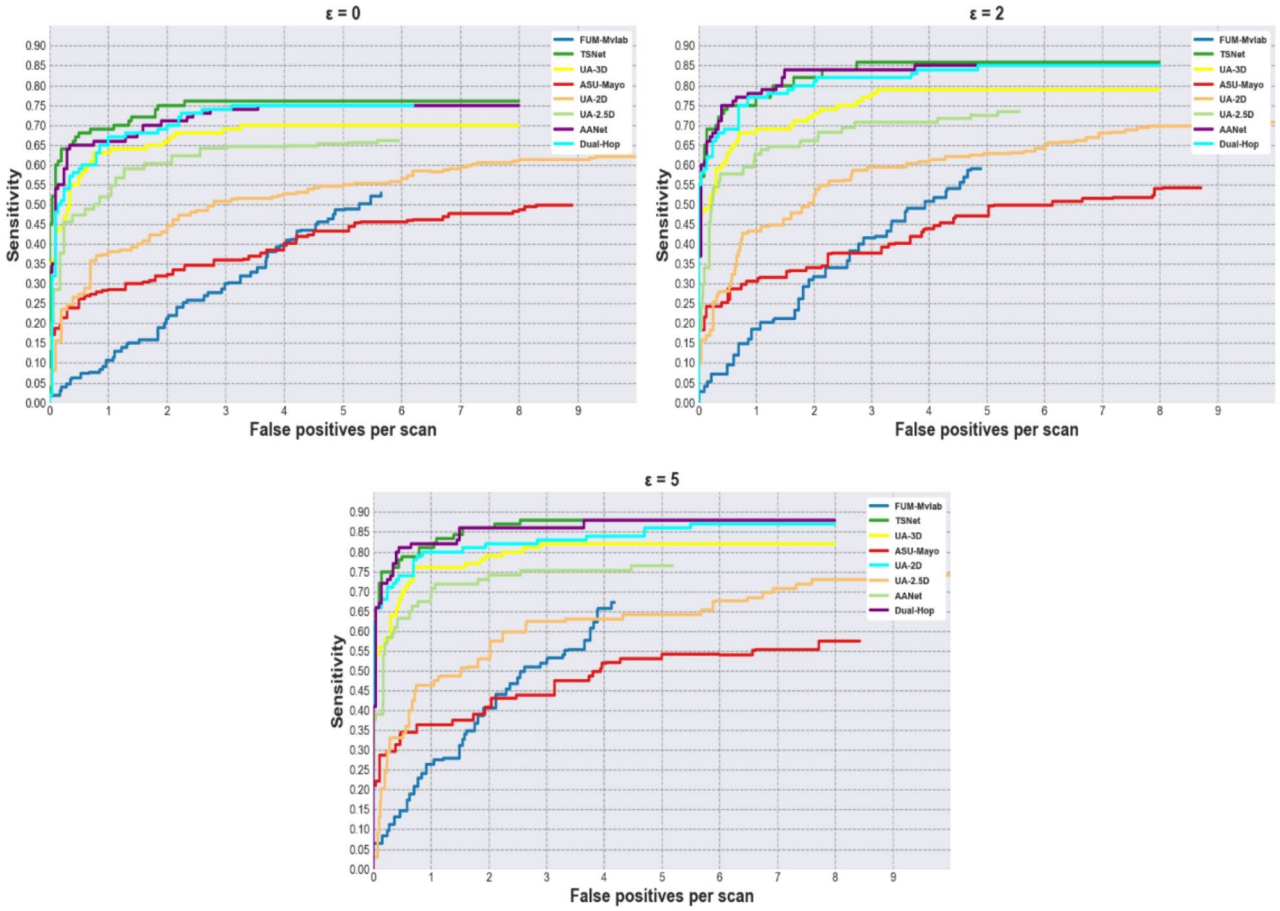

Fig. 5. Sensitivity curves for different boundary tolerance levels.

Conclusion

The proposed model tackles the challenge of automatically detecting and segmenting pulmonary embolism in CTPA images while significantly reducing the false positive rate. It enhances vessel and thrombus feature identification via a multi - task learning strategy and an innovative feature extraction module. By deeply understanding the interaction between thrombus and vessel structures, the model improves its ability to distinguish thrombus from surrounding tissues, ensuring that the outputs are highly consistent with physiological structures, thus enhancing overall performance. Experimental results demonstrate that this model exhibits excellent accuracy and robustness across various image qualities and lesion patterns. In clinical applications, it provides physicians with reliable diagnostic information, improving diagnostic efficiency and reducing the risk of misdiagnosis, thus showing significant potential. In conclusion, through deep learning and biological feature analysis, the model exhibits strong performance and practical clinical value in pulmonary embolism detection and segmentation.

The proposed model shows some limitations in its performance. While it excels at detecting pulmonary embolism with high accuracy under high false alarm rate conditions, its sensitivity remains suboptimal in low false alarm rate scenarios, potentially due to the model's complexity, which can impair the detection of subtle lesions. Additionally, the segmentation of thrombus, particularly for small or irregularly shaped lesions, presents challenges. The model struggles to accurately delineate thrombus boundaries, especially in regions with subtle or diffuse lesions.

Future research will focus on addressing these limitations by improving sensitivity and segmentation performance. Specifically, incorporating advanced techniques such as diffusion models and semi-supervised learning can enhance feature extraction and the ability to generalize from limited labeled data. Diffusion models, with their ability to model complex spatial relationships, could improve the segmentation of small and intricate thrombus structures. Semi-supervised learning, on the other hand, could leverage large amounts of unlabeled data to refine the model's generalization capabilities, especially in detecting subtle lesions. These strategies aim to improve the model's robustness and accuracy in clinical settings, providing more reliable and precise detection and segmentation of pulmonary embolism.

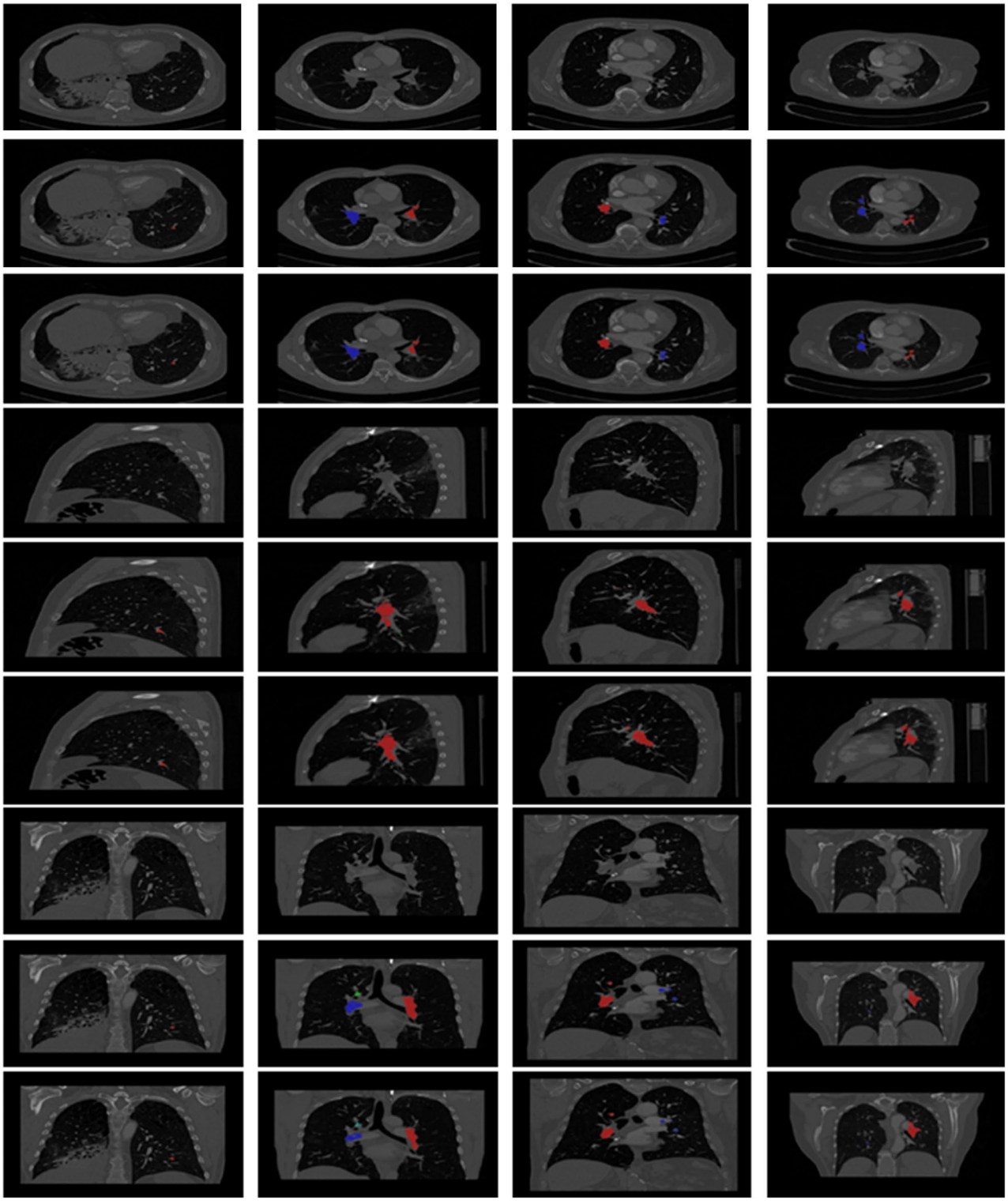

Fig. 6. Comparison of the results of the TSNet model for segmentation of pulmonary embolisms.

Model components	Overall sensitivity performance at different false positive thresholds		
	ε = 0 mm	ε = 2 mm	ε = 5 mm
BL	0.6405	0.7026	0.7436
BL + GT-AFM	0.6475	0.7134	0.7459
BL + GT-AFM + CSSA	0.6432	0.7113	0.7519
BL + GT-AFM + CSSA + TAD	0.6538	0.7478	0.7825

Table 3. Sensitivity values of different component modules at various boundary tolerance levels respectively.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 20 October 2024; Accepted: 24 February 2025
Published online: 01 March 2025

References

1. None. "The clinical course of pulmonary embolism". *N. Engl. J. Med.* **327**, 957–958 (1992).
2. Velmahos, G. C. et al. Spiral computed tomography for the diagnosis of pulmonary embolism in critically ill surgical patients: a comparison with pulmonary angiography. *Arch. Surg.* **136**, 505–511 (2001).
3. Sood, S. et al. Role of CT angiography in pulmonary embolism and its comparative evaluation with conventional pulmonary angiography. *Indian J. Radiol. Imaging* **16**, 215–220 (2006).
4. Meaney, J. F. M. et al. Diagnosis of pulmonary embolism with magnetic resonance angiography. *N. Engl. J. Med.* **336**, 1422–1427 (1997).
5. Safriel, Y. & Zinn Harry. CT pulmonary angiography in the detection of pulmonary emboli. *Clin. Imaging* **26**, 101–105 (2002).
6. Ageno, W. et al. Factors associated with the timing of diagnosis of venous thromboembolism: results from the MASTER registry. *Thromb. Res.* **121**, 751–756 (2008).
7. Ronneberger, O., Fischer, P. & Brox, T. U-Net: convolutional networks for biomedical image segmentation. *Med Image Comput. Comput. Assist. Interv. MICCAI 2015* **18**, 234–241 (2015).
8. Pu, J. et al. Automated detection and segmentation of pulmonary embolisms on computed tomography pulmonary angiography (CTPA) using deep learning but without manual outlining. *Med. Image Anal.* **89**, 102882 (2023).
9. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **60**, 84–90 (2017).
10. Masutani, Y., Macmahon, H. & Doi, K. Computerized detection of pulmonary embolism in spiral CT angiography based on volumetric image analysis. *IEEE Trans. Med. Imaging* **21**, 1517–1523 (2002).
11. Yuan, C. et al. Pulmonary arteries segmentation from CT images using PA - Net with attention module and contour loss. *Med. Phys.* **50**, 4887–4898 (2023).
12. Zhou, C. et al. Preliminary investigation of computer-aided detection of pulmonary embolism in three-dimensional computed tomography pulmonary angiography images. *Acad. Radiol.* **12**, 782–792 (2005).
13. Soffer, S. et al. Deep learning for pulmonary embolism detection on computed tomography pulmonary angiogram: a systematic review and meta-analysis. *Sci. Rep.* **11**, 15814 (2021).
14. Tajbakhsh, N., Gotway, M. B. & Liang, J. Computer-aided pulmonary embolism detection using a novel vessel-aligned multi-planar image representation and convolutional neural networks. *Med. Image Comput. Comput. Assist. Interv.* **9350**, 62–69 (2015).
15. Huhtanen, H. et al. Automated detection of pulmonary embolism from CT-angiograms using deep learning. *BMC Med. Imaging* **22**, 43 (2022).
16. Long, K., Tang, L., Pu, X., Ren, Y. & Deng, F. Probability - based mask r - cnn for pulmonary embolism detection. *Neurocomputing* **422**, 345–353 (2021).
17. Huang, S. et al. Penet - a scalable deep - learning model for automated diagnosis of pulmonary embolism using volumetric Ct imaging. *NPJ Digit. Med.* **3**, 61 (2020).
18. Guo, J. et al. Artery-Aware network for pulmonary embolism detection in CTPA images. *Med. Image Comput. Comput. Assist. Interv.* **13431**, 473–483 (2022).
19. Islam, N. U., Zhou, Z., Gehlot, S., Gotway, M. B. & Liang, J. Seeking an optimal approach for computer-aided diagnosis of pulmonary embolism. *Med. Image Anal.* **91**, 102988 (2024).
20. Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K. & Yuille, A. L. Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected Crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* **40**, 834–848 (2018).
21. Zhang, X., Zhou, X., Lin, M. & Sun, J. ShuffleNet: an extremely efficient convolutional neural network for mobile devices. *IEEE/CVF Conf. Comput. Vis. Pattern Recognit.* 6848–6856 (2018).
22. Zhang, J., Xie, Y., Wang, Y. & Xia, Y. Inter-slice context residual learning for 3D medical image segmentation. *IEEE Trans. Med. Imaging* **40**, 661–672 (2021).
23. Dai, Y. et al. *IEEE Winter Conf. Appl. Comput. Vis.* 3559–3568 (2020). 3559–3568 (2020). (2021).
24. Chollet, F. & Xception Deep learning with depthwise separable convolutions. *IEEE Conf. Comput. Vis. Pattern Recognit.* 1800–1807 (2017).
25. Hu, J., Shen, L. & Sun, G. Squeeze - and - excitation networks. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7132–7141 (2018).
26. Jaderberg, M., Simonyan, K., Zisserman, A. & Kavukcuoglu, K. Spatial transformer networks. *Proc. 29th Int. Conf. Neural Inf. Process. Syst.* 2017–2025 (2015).
27. Woo, S., Park, J., Lee, J. Y. & Kweon, I. S. CBAM: convolutional block attention module. *Comput.Vis. ECCV* **2018**, 3–19 (2018).
28. Qin, X. et al. Boundary-aware segmentation network for mobile and web applications. *ArXiv* (2021).
29. Darooei, R. et al. Loss-Modified Transformer-Based U-Net for accurate segmentation of fluids in optical coherence tomography images of retinal diseases. *J. Med. Signal. Sens.* **13**, 253–260 (2023).
30. González, G. et al. Computer aided detection for pulmonary embolism challenge (CAD-PE). *ArXiv* (2020).
31. Cano-Espinosa, C., Cazorla, M. & González, G. Computer aided detection of pulmonary embolism using multi-slice multi-axial segmentation. *Appl. Sci.* **10**, 2945 (2020).

32. Masoudi, M. et al. A new dataset of computed-tomography angiography images for computer-aided detection of pulmonary embolism. *Sci. Data* **5**, 1–9 (2018).
33. Bushra, F. et al. Deep learning in computed tomography pulmonary angiography imaging: A dual-pronged approach for pulmonary embolism detection. *Expert Syst. Appl.* **245**, 34–53 (2024).
34. Condrea, F. et al. Anatomically aware dual-hop learning for pulmonary embolism detection in CT pulmonary angiograms. *Comput. Biol. Med.* **174**, 108464 (2024).
35. González, C. et al. Lifelong nnU-Net: a framework for standardized medical continual learning. *Sci. Rep.* **13**, 1–13 (2023).
36. Tharwat, A. Classification assessment methods: a detailed tutorial. *Appl Comput. Inform.* (2018).
37. Gonzalez, G., Ledesma-Carbayo, M. J., Estepar, S. J., Fraile Moreno, R. & Kassarjian, E. A. & Jimenez-Carretero, D. CAD-PE challenge: Developing a baseline for CAD-PE development. www.cad-pe.org
38. Cao, H. et al. Swin-Unet: Unet-Like pure transformer for medical image segmentation. *Comput Vis. ECCV 2022 Workshops* **13803**, 205–218 (2023).
39. Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., Liang, J. & UNet++ A nested U-net architecture for medical image segmentation. *Deep Learn. Med. Image Anal. Multimodal Learn. Clin. Decis. Support* **11045**, 3–11 (2018).
40. Chen, Y. et al. SCUNet++: Swin-UNet and CNN bottleneck hybrid architecture with Multi-Fusion dense skip connection for pulmonary embolism CT image segmentation. *IEEE/CVF Winter Conf. Appl. Comput. Vis.* 7744–7752 (2024).

Author contributions
H.Y.L. wrote the main manuscript and Codes. J.C.F. and Y.L. reviewed the manuscript and evaluated the results. Y.Q.Q and Y.D.R checked the grammar of the manuscript and the datasets.

Funding
This work is supported by Shandong Provincial Natural Science Foundation of China under Grant ZR2018MF009, the State Key Research Development Program of China under Grant 2017YFC0804406, National Natural Science Foundation of China under Grant 42472324, the Special Funds of Taishan Scholars Construction Project, and the foun-dation of Key Laboratory of Mining Disaster Prevention and Control (Shandong University of Science and Technol-ogy).

Declarations

Competing interests
The authors declare no competing interests.

Ethics approval
The study was approved by the Medical Ethics Committee of the Affiliated Hospital of Qingdao University (No. QYFYWZLL29335), informed consent was waived by the Medical Ethics Committee of the Affiliated Hospital of Qingdao University, and the study was conducted in accordance with the relevant guidelines and regulations. The approval document from the Medical Ethics Committee of the Affiliated Hospital of Qingdao University has been uploaded as an attachment.
Correspondence and requests for materials should be addressed to Yang Li.

Additional information
Correspondence and requests for materials should be addressed to Y.L. or Y.R.

Reprints and permissions information is available at www.nature.com/reprints.

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