International Journal on Engineering, Science and Technology
2022, Volume 4, Issue 4 ISSN: 2642-4088 (Online)

Probability-Based Two-Stage CNN for Pulmonary Embolism Detection in Computed Tomography Pulmonary Angiography

Jinn-Yi Yeh
National Chiayi University, Taiwan

Sheng-You Wu
National Chiayi University, Taiwan

Abstract: Pulmonary embolism (PE) can impede the normal flow of blood, which can result in reduced oxygen levels in vital organs and be life-threatening. Therefore, early detection and treatment of PE can effectively reduce mortality. Computed tomography pulmonary angiography (CTPA) is the clearest and most immediate tool available for PE diagnosis. With an increase in the CTPA image quality and number of slices, physician fatigue associated with interpreting these images during the diagnostic process has also increased. Therefore, this study proposes a modified two-stage convolutional neural network-based approach to automatically detect areas suspected of containing PE in CTPA images to reduce the burden of diagnosis on doctors. The main functions for eliminating false positives include principal component analysis, enlarged feature map (EFM), and probability-based anchor point extraction (PAE). Experimental results showed that although EFM can slightly improve the accuracy for small-object detection, it significantly increases the time required for training. PAE can reduce the training time and slightly improve detection accuracy. But the simultaneous use of EFM and PAE can significantly increase sensitivity (10.57%) with an increase of only 29 min in the training time.

Keywords: Pulmonary embolism, Detection, Convolutional neural network

Introduction

Pulmonary embolism (PE) is a phenomenon wherein the pulmonary artery or its small blood vessels are blocked by an embolus comprising air, fat, or blood clots. In the United States, an average of 1 in every 500 people annually suffers from PE, and 11% of them die within one hour of its onset. Without treatment, the average mortality rate of PE is approximately 30%. Hence, PE is a common condition with high morbidity and mortality that requires early and accurate diagnosis, and treatment with thrombolytics or anticoagulants, which can effectively reduce mortality by approximately 2–8% (Goldhaber and Bounameaux, 2012). Therefore, early and accurate PE diagnosis is an extremely crucial and challenging task for physicians.

Computed tomography pulmonary angiography (CTPA) is one of the most important and detailed techniques for examining the blood vessels inside the lungs. It can show whether the blood vessel itself is abnormal (stenosis, blockage, rupture, etc.). CTPA can also show the relative position of the lesion and blood vessels, and the

distribution of blood vessels inside the lesion. A CTPA image can serve as an important basis for helping doctors diagnose and treat PE cases (Figure 1).

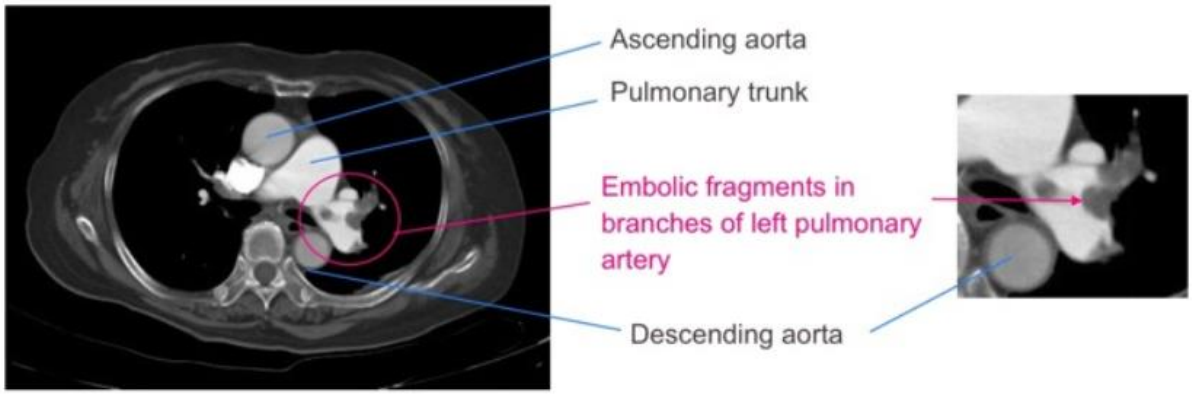


Figure 1. Computed Tomography Pulmonary Angiography Image (González et al., 2020)

Radiologists typically detect dark areas of PE in CTPA images by manually circling contours; however, this process is time consuming. Moreover, it is difficult to determine the blood vessel boundary owing to noise or artifact interference, and experienced experts may be required to interpret them clearly. Therefore, a computer-aided detection (CAD) system is required to assist doctors during interpretation. This study proposes a modified two-stage convolutional neural network (CNN) approach from Yang et al. (2019) to automatically detect areas suspected of containing PE in CTPA images. In the proposed method, the candidate cubes that may belong to the embolism region are first extracted from the original 3D image based on the image features.

A distribution model is then established for the original image based on the PE probability. Thereafter, the enlarged feature map (EFM) corresponding to more positive anchor points with better quality is extracted from the candidate region to facilitate subsequent detection. The entire process can improve the PE detection accuracy without increasing the time required. The remainder of this paper is organized as follows: Section 2 contains a literature review, Section 3 details the research methods used in this study, Section 4 presents an analysis of the experimental results, and Section 5 concludes the paper.

Literature Review

Most traditional PE detection methods rely on the professional judgment of doctors. Generally, the blood flow in the peripheral lobe during vascular embolism is lower than that in the normal area. Therefore, if a detailed and accurate description of the pulmonary lobe blood flow distribution can be obtained, PE occurrence and the scope of its impact can be identified. Qanadi et al. (2000) used traditional double-helix computed tomography (CT) to analyze whether embolism occurred in major blood vessels, such as the pulmonary artery. Among 158 patients, 147 were correctly diagnosed (93%). But the diagnostic accuracy rate of CT for microvascular embolisms is not ideal.

A study by Michiels et al. (2005) conducted on 53 patients pointed out that the accuracy rate was only 60% (32 patients), while pulmonary artery computed tomography angiography can increase the accuracy to 94% (50 patients). Nicholas and Harvey (2003) stated that there are two main reasons for the low sensitivity of traditional CT for diagnosing PE. First, there are numerous thin blood vessels in the pulmonary lobes, and their volume is too low to be easily observed. Therefore, the diagnosis must be performed carefully, and the visual fatigue and misjudgment of a radiologist can easily increase.

Second, the information corresponding to each CT image contains the description of the actual thickness of the lungs from 2–5 mm, which may be offset by breathing or heartbeats, causing the finer branch vessels to produce shadows with low grayscale values similar to emboli in the image and resulting in misjudgment. Therefore, after a comprehensive analysis of the research results of these scholars, we decided to use pulmonary artery computed tomography angiography images as the original image data source.

There are several unwanted areas in CTPA images that can easily affect the CAD training process, such as the outer ring of the ribs and spine. No blood vessels are present in these skeletal areas, and it is unlikely that emboli will be present in these areas. Therefore, it is necessary to perform image segmentation to remove unnecessary areas and reduce the amount of calculation required. Several methods exist for segmenting lung blocks using image processing technology. Some utilize features around the lungs or diseases located in the lungs to determine the location of the lungs. For example, Park et al. (1998) proposed using the distribution characteristics of the trachea as the basis for segmentation. Dajnowiex and Alirezaie (2004) utilized the state of air circulation in the lungs as their characteristic appearance.

Blechschmidt et al. (2001) used emphysema appearance as a characteristic of the lungs of patients. Yongbum et al. (2001) and Augusto et al. (2001) first used the detection of small pulmonary nodules as features to determine the location of the lungs in an image. They then used other image processing and artificial intelligence techniques, such as contour search and neural networks, to segment the lungs in the images. Regarding automatic segmentation, methods have been proposed by Margarida et al. (2006) and Xu et al. (2005) to automatically segment lung regions using multiple active contours and outlier models. After comparing the various methods proposed by these scholars, we decided to use the active contour model (ACM) to segment the lung area and remove unnecessary areas from the image.

Traditional CAD systems for detecting PE usually include four stages: (1) extracting regions of interest (ROI) from the original image by performing lung or blood vessel segmentation; (2) using algorithms to select probable PE candidates in these ROIs; and (3) manually extracting features from each probable PE candidate and using a rule-based classifier. Subsequently, a neural network or k-nearest neighbor classifier computes a confidence score for each candidate.

Several studies have proposed different CAD systems for PE diagnosis; however, they all have serious limitations. For example, although the systems proposed by Masutani et al. (2002) and Pichon et al. (2004) achieved sensitivities of 100% and 86%, respectively, they were trained with only a small number of PE images (21 and

22, respectively); therefore, it was impossible to determine the degree of generalization. Another system proposed by Das et al. (2003) achieved a sensitivity of 88%, but it is unclear how well it performs on all PE images because the image samples used in their study only considered embolization images with vessel endings.

The sensitivity of the system proposed by Digumarthy et al. (2006) was 92%. However, their study only used images with good shading and no obvious motion artifacts or other lung diseases as the original image sources; hence, the system performance for other more representative imaging data remains unknown. Maizlin et al. (2007) performed a test using the same system with more representative imaging data, which showed 58% less sensitivity. Unlike previous studies, the study by Buhmann et al. (2006) was the only one to use a large dataset (352 images) containing respiratory artifacts and parenchymal lung disease imaging data. They did not exclude images based on the size or location of the thrombus and considered all types of thrombi. Their experimental results were poor, with a sensitivity of only 47% and 52% and an average number of false positives being 3.9 and 11.4 in the training and testing datasets, respectively. The visualization method proposed by Kiraly et al. (2006) also achieved only 50% sensitivity, which is not significantly different from the former. The above-mentioned manually produced feature representations have limited capabilities and usually result in a high number of false positives to achieve an acceptable sensitivity.

Recently, new algorithms for PE detection have been proposed. For example, Wang et al. (2012) proposed a pulmonary vascular tree segmentation algorithm for lung segmentation to improve the CAD performance for PE detection. Their method achieved a sensitivity of 62.5%, while reducing the false-positive rate by 16.2%. Liang and Bi (2007) applied the sled algorithm to mark the pixels belonging to PE in the image and used multi-instance classification to assist CAD in PE detection. They achieved a sensitivity of 80% in the best case, with four false positives in each dataset. Ozkan et al. (2014) proposed a novel method for detecting PE in CTPA images that uses vessel segmentation and a rule-based classifier. Their method yielded eight false positives for each dataset and the best case sensitivity of approximately 62%.

Although the aforementioned scholars used different feature representation methods that are superior to manual ones, they may still generate numerous false positives, which imposes a heavy burden on subsequent steps. In this study, we combined all steps, including the PE region extraction method, image processing to align the blood vessels, and elimination of false positive candidates, through a modified two-stage CNN to achieve high accuracy while avoiding additional time and space consumption.

Research Methods

This study established an automated detection method to automatically detect areas suspected of containing PE in CTPA images using a modified two-stage CNN. A flowchart of the process is shown in Figure 2. First, the CTPA images are read from the PE challenge database. The system then pre-processes the expert-interpreted images and executes a modified two-stage CNN. The performance of the system was also evaluated against other PE detection methods.

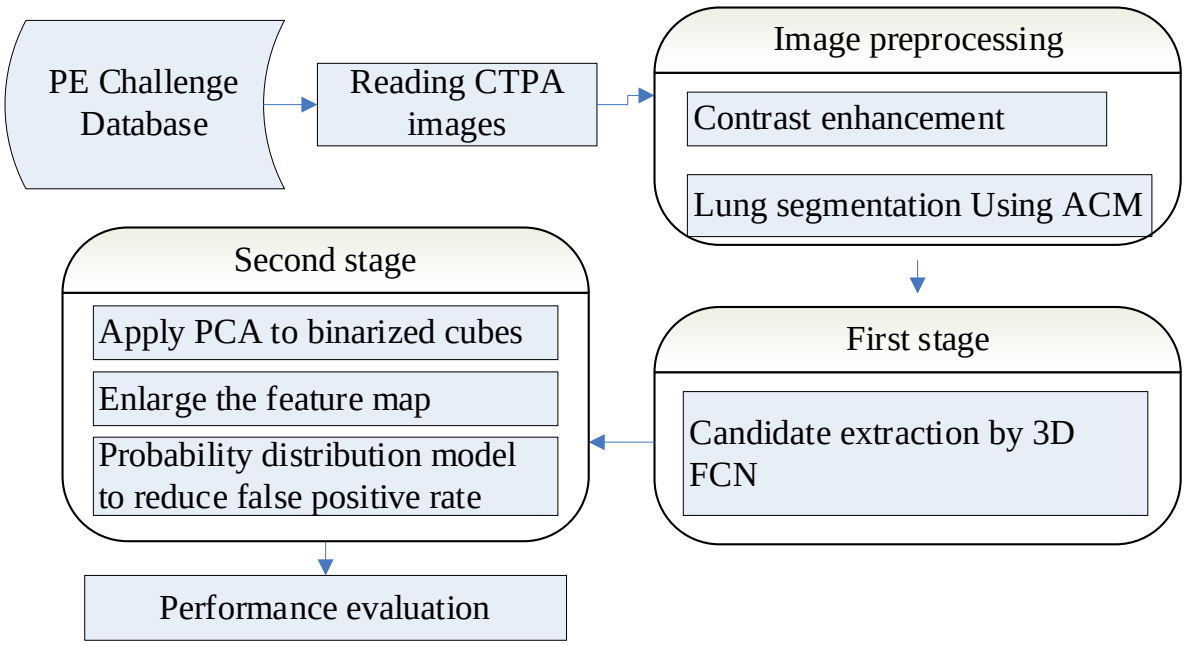


Figure 2. Flowchart of the Modified Two-Stage PE Detection Network

Image Preprocessing

In an original CTPA image, the CT value is generally used to represent the intensity of a pixel in Hounsfield units (HU). However, general computer monitors are incompatible with this range; therefore, the image must be converted to a 256 grayscale resolution. The higher the CT value of the image, the higher the tissue density. If the CT value is less than 0 HU, it appears black in the image. When the CT value is greater than 1000 HU, it appears bright white in the image. The bright white parts indicate the surrounding ribs and vertebrae with high-density values and the arteries near the center. At this stage, image processing methods are used to enhance the contrast and segment the lung area. Image segmentation not only reduces unnecessary computation, but also improves the identification of suspicious areas in the image.

Contrast Enhancement

Because the thin blood vessels around the lung lobes are not obvious in an original image, this study used a cubic curve to enhance the image (Domingo et al., 2000), which is shown in Figure 3. It can be observed that the curve contains an inflection point TP3. By controlling the position of TP3, the curve can be bent to different degrees to adjust the image. Equation (1) was used to calculate the cubic adjustment curve, where x is the pixel value in the original image and y is the pixel value of the image after adjusting the curve. Because the curve must pass through two points (0, 0) and (255, 255), the value of d can be deduced as 0.

$$y = f(x) = ax^3 + bx^2 + cx + d \quad (1)$$

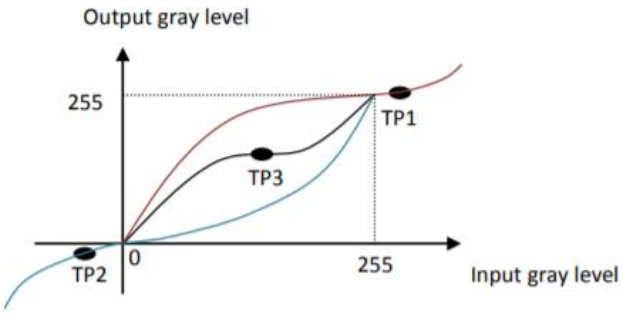


Figure 3. Appropriate Compensation Curve

Next, the following formulas were used to deduce the cubic adjustment curve required to adjust the image: Equation (5) defines the inflection point in the cubic curve, where I is an image and x is the pixel value of any point in I. Point A was used to find other unknown variables to derive the cubic compensation curve required to adjust the image. The calculation methods for all variables in the curve are shown in Equations (2)–(5), where $min_{x\in I}$ and $max_{x\in I}$ represent the smallest and largest pixel values in the image, respectively.

$$c = 1 - a \times (255)^2 - b \times 255 \tag{2}$$

$$b^2 = 3 \times a - (255)^2 \times 3a^2 - 255 \times 3 \times s \times b \tag{3}$$

$$a = \frac{1}{(225)^2 - 3\times 255\times A + 3\times A^2} \tag{4}$$

$$A = min_{x\in I}\{x\} + 0.7\{max_{x\in I}\{x\} - min_{x\in I}\{x\}\} \tag{5}$$

Lung Region Segmentation

The ACM is a framework used for extracting object contour lines from 2D images that may contain noise. It involves establishing an initial curve on the image first; the shape is not restricted, but the target object must be wrapped inside the contour line. Several equations must then be established, including those for the standard curve and contour line of the target object. During this process, finding the minimum value of the formula for the standard curve shape can cause the curve to continue shrinking towards the inside of the outline and maintain a smooth state. Finding the minimum value of the formula for the closeness of the standard curve to the contour of the target object should be continued until the two overlapping objects can maintain the curve close to the target.

To reduce the time required for the removing unnecessary image regions, this study used ACM to circle the lung regions in the CTPA images. Before using the ACM method, a circle of control points must be set outside the lung area. Each segment of the control points is then connected in series with a cloud-shaped curve (B-spline curve) to form the most primitive outline. The image used in this study comprised 4096 grayscale values, and the parts with brighter grayscale values in the image indicate structures with higher density, such as the ribs and vertebrae. The darker parts of the image indicate structures with lower density, such as the lungs, and the gray

level of fat, which lies between the bone and air. Therefore, this study used these features to set the contour points of the lungs by combining the projections of the grayscale value in the x- and y-axes directions. After the contours were generated, the angles obtained during the edge detection process were used to gradually converge the lung region in the image based on the ellipse shape.

Probability-based two-stage CNN

This study used a modified two-stage CNN proposed by Yang et al. (2019) as the framework for detecting PE, which is illustrated in Figure 4. The first stage uses a 3D fully convolutional network (FCN) for extracting candidate regions (which may be PE targets). The second stage extracts 3D candidates after the blood vessel alignment. The 2D cross sections of these cubes were fed into a probability-based ResNet-18 classifier to eliminate false positives.

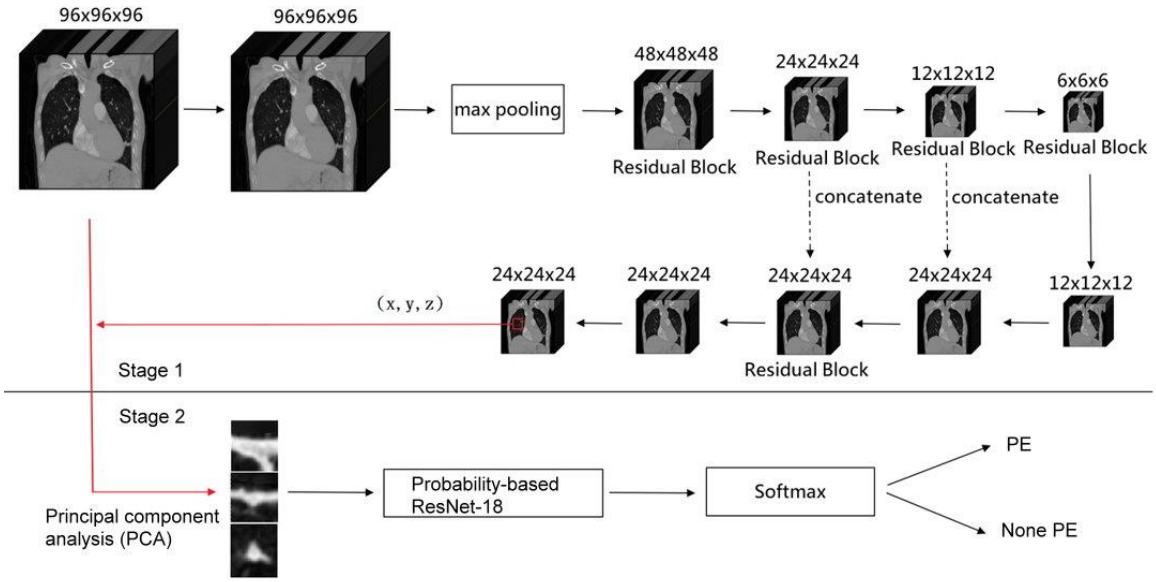


Figure 4. Framework of the probability-based two-stage CNN

Stage 1: Extracting Candidate Regions

The goal of the first stage is to obtain high sensitivity with a reasonable number of false positives. To fully use the 3D image information from the CTPA images, a 3D FCN was used to extract the feature hierarchy from the 3D images to obtain the candidate regions. As shown in the upper part of Figure 5, the 3D FCNs are connected to each other through the encoder and decoder with residual connections. The encoder begins with a 3D convolutional layer, followed by a max-pooling layer and four other residual modules in sequence to encode feature maps hierarchically. The decoder then up-samples the feature maps using two deconvolutional layers, a residual module, and two convolutional layers. Residual connections are used to connect the last two residual modules in the encoder and the corresponding residual modules in the decoder.

In addition to visual features, location information is an important indicator for identifying PE. Because PE is usually located in fixed areas, such as the main pulmonary artery or lobar branches, it is also necessary to input location information and combine it with decoding. Additionally, the FCN feature maps in the generator are combined. Specifically, a 3-channel location map is created in the second deconvolution layer (i.e., 24 × 24 × 24), having the same size as the FCN feature map. Each voxel of the position map is a 3D vector responsible for displaying the x-, y-, and z-coordinates in the entire 3D space. Next, the network directly connects the 3-channel location map with the 64-channel FCN feature map, which passes from the residual connection to create a 131-channel feature map. The residual module is then used to connect the feature maps for information fusion.

To select candidate regions from the fused 3D feature map, cubes with anchor points (reference points when the convolution kernel slides) were input into the network responsible for extracting candidate regions for accurate detection. Note that the lesions may vary in size. Specifically, a cube with anchor points is a predefined multiscale 3D window whose center is located at the position of each voxel in the feature map. This study assigned three cubes with anchor points for each voxel location.

Each cube had a distinct size ratio s (10, 30, and 60 mm). Five regression variables were set for each cube with anchor points representing the position of the cube $(\Delta x_s, \Delta y_s, \Delta z_s)$ and size $(\Delta d_s, p_s)$, to calculate these five values. They were responsible for defining the location and space occupied by the candidate region in the entire lung volume with the probability that the candidate region contains PE. Subsequently, based on s, the position and size offset values of the candidate regions with anchor points were regressed and trained. To achieve this, the entire network architecture first applied a 3D convolutional layer with 64 kernels of size 1 × 1 × 1 each to the fused feature maps, followed by another 3D convolutional layer with 15 kernels (each kernel of size 1 × 1 × 1) to output feature maps of size 24 × 24 × 24. Each voxel of the feature map output during the research process was a 5N-dimensional vector, expressed as $(\Delta x_s, \Delta y_s, \Delta z_s, \Delta d_s, p_s)$, where N = 3.

Stage 2: Eliminating False Positives

The second stage aims to eliminate as many false positives as possible using the classifier while maintaining high sensitivity. This is a very challenging task, as the first stage may select several false-positive candidate regions, resulting in a severe imbalance between the positive and negative samples. Moreover, owing to the differences in the orientation, size, and shape of PEs, the 2D cross-sectional images of all possible candidate regions may noticeably differ in appearance.

Using 3D classifiers can overcome the problem of appearance variation to some extent; however, owing to the lack of sufficient 3D samples to train 3D classifiers, the entire research process is severely limited by the amount of training data. To solve these problems, this study referred to an image representation method for aligning blood vessels proposed by Tajbakhsh et al. (2015). This method uses principal component analysis (PCA) of 3D image information to create the candidate and affected areas. The orientation of the vessels is aligned to reduce the appearance of PE in 2D cross-sectional images.

Principal Component Analysis

The image processing process first crops a small ROI cube containing PE (PE-ROI) from the candidate region. The cube is then binarized with intensity thresholding. According to the experience of radiologists, normal blood vessel wall image intensity values are higher than 100 HU, while that of other tissues are lower than 100 HU. Because PE introduces a filling defect (a dark area surrounded by bright vascular lumen) in CTPA images, the image intensity value of PE will be slightly lower than that of blood vessels. In this study, 70 HU was used as the threshold value for binarization.

During the binarization process, voxels with image intensities higher than 70 HU were marked as blood vessels, whereas the other pixels were set to 0 to indicate non-vessel parts. PCA was then applied to the binarized cubes to calculate the orientation of vessel segments in them. The specific method involves extracting three eigenvectors (v_1, v_2, v_3) and their corresponding eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$, where $\lambda_1 \geq \lambda_2 \geq \lambda_3$. According to the physical significance of the eigenvectors, v_1 represents the direction in which the blood vessel extends, and v_2 and v_3 represent two orthogonal directions perpendicular to the v_1 plane. Then, according to $(\lambda_1, \lambda_2, \lambda_3)$, a 3D affine transformation matrix A_θ was defined to apply the 3D rotation transformation to the cube as follows:

$$\begin{bmatrix} x^s \\ y^s \\ z^s \end{bmatrix} = A_\theta \begin{bmatrix} x^t \\ y^t \\ z^t \\ 1 \end{bmatrix} = \begin{bmatrix} s_x e_1^T v_1 & s_y e_1^T v_2 & s_z e_1^T v_3 & t_x \\ s_x e_2^T v_1 & s_y e_2^T v_2 & s_z e_2^T v_3 & t_y \\ s_x e_3^T v_1 & s_y e_3^T v_2 & s_z e_3^T v_3 & t_z \end{bmatrix} \begin{bmatrix} x^t \\ y^t \\ z^t \\ 1 \end{bmatrix} \quad (6)$$

where (x^t, y^t, z^t) denote the cube coordinates after vessel alignment, and (x^s, y^s, z^s) denote the original coordinates of the cube. During the study, all coordinates were normalized to [-1, 1], that is, $-1 < x^t, y^t, z^t < 1$ and $-1 < x^s, y^s, z^s < 1$. t_x, t_y, and t_z represent the offset values of the cube relative to the center of the entire blood vessel; $s_x, s_y,$ and s_z represent the scaling ratios between the cube and blood vessel; and $e_1, e_2,$ and e_3 represent the identity matrices.

Once the vessel-aligned cube was obtained, two 2D cross-sectional images of the cube were extracted, including longitudinal and cross-sectional images, as input images for the probabilistic ResNet-18 classification network. Figure 5 briefly illustrates the entire vessel alignment procedure, where (a) shows a voxel around an embolus, (b) shows a PE-ROI cube, (c) shows the beginning of PE-ROI PCA to determine vessel axis v_1 (blue arrow) and two orthogonal directions (v_2 and v_3), (d) depicts the use of v_1, v_2, and v_3 to form a planar image of the cross-section (shown in blue) and a planar image of the longitudinal section (shown in green), (e) shows the multiple longitudinal and cross-sectional planes obtained by rotating v_2 and v_3 around v_1, and (f) shows the division of the rotated planar images into two structures. Finally, the emboli are generated by randomly selecting one planar image from the transverse section and two planar images from the transverse section, as shown in (g).

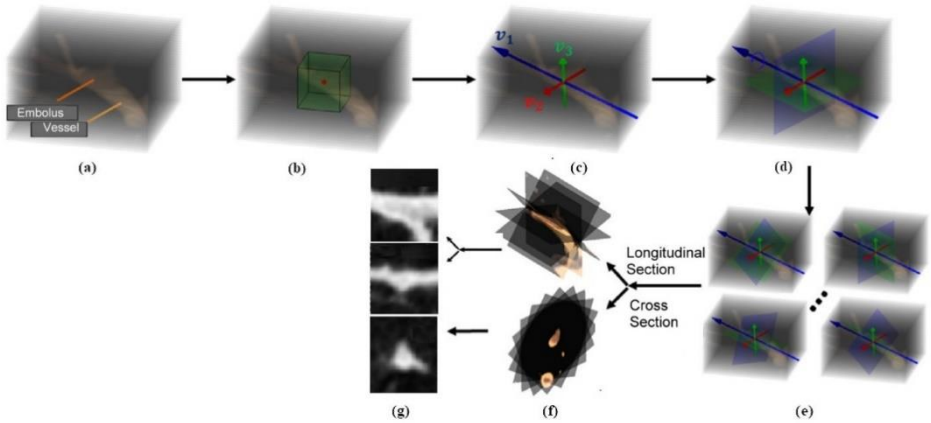


Figure 5. Vessel Alignment Process

Enlarging the Feature Map

In the process of selecting candidate regions, unless the intersection over union (IoU) between the anchor and ground truth is the highest, or the IoU between the anchor and ground truth is greater than a threshold t, it is considered a negative anchor. A positive anchor is responsible for participating in the loss calculation. However, during the research process, it was discovered that the number of positive anchor points was very small and their quality was poor, resulting in a high number of false positives in the extracted candidate regions. Therefore, the sampling stride must be adjusted to obtain more positive anchor points.

As shown in Figure 6, assuming that both the anchor point and ground truth are square, the side lengths of the anchor point and ground truth are S_a and S_g, respectively. To make the IoU value between the anchor point and ground truth greater than the threshold t, the sampling stride length d should satisfy

$$\frac{(S_a+S_g-d)^2/4}{S_a^2+S_g^2-(S_a+S_g-d)^2/4} \geq t \tag{7}$$

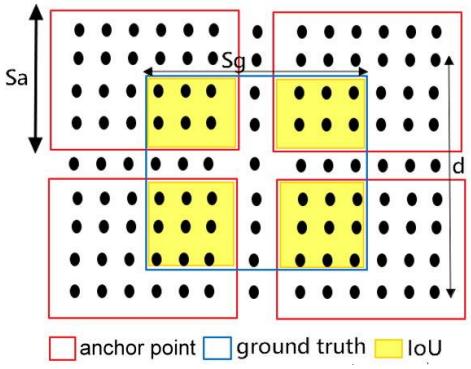


Figure 6. Minimum IoU

In Equation (7), the minimum value on the left-hand side of the inequality is satisfied only when the ground truth lies in the middle of the four adjacent anchors. Moving the ground truth in any direction increases the IoU value.

However, considering the irregular shape of the PE as an example, most positive anchors and ground truths fail to satisfy Equation (7). Therefore, the sampling stride must be reduced by enlarging the enlarging feature map through up-sampling. The relationship between the sampling stride length and the up-sampling scale should satisfy the following:

$$d = w_{image}/(w_{feature} * S) \tag{8}$$

where d is the sampling stride length, w_{image} is the input image width, $w_{feature}$ is the feature map width, and S is the up-sampling scale. Up-sampling the feature map not only makes the extracted anchors denser but also enriches local details, which is beneficial for small object detection such as PE.

Modeling Probability Distributions

By up-sampling the feature map, the matching success rate of the anchor points and ground truths can be improved. However, the EFM occupies significant memory space, which increases the amount of subsequent computation. Therefore, this study used an anchor-point extraction method based on the probability of PE vascular distribution features, which can reduce the time and space consumption required by the EFM and improve the detection performance for small objects, such as PE.

Generally, PE-ROIs rarely appear randomly in the entire image. Most PE-ROIs are concentrated on both sides of the mid-axis, where the pulmonary artery and its branches are located. Therefore, extracting anchors with different densities can easily correspond to the ground truth. Hall et al. (2020) established a probability model to describe the probability of PE-ROI features in each voxel of the image. In their model, the probability distribution was modeled using a Gaussian mixture model (GMM), which is defined in Equation (9). Given a random variable X, GMM can be expressed as

$$p(X) = \sum_{k=1}^{k} \pi_k \mathcal{N}(x|\mu_k, \sum_k) \tag{9}$$

where $\mathcal{N}(x|\mu_k, \sum_k)$ is the Gaussian distribution, that is, the k-th component in the mixture model, and π_k is the mixture coefficient that satisfies the following conditions:

$$\sum_{k=1}^{k} \pi_k = 1 \text{ and } 0 \leq \pi_k \leq 1 \tag{10}$$

In this study, X is a two-dimensional random variable representing the center position of the PE-ROI. The probability of a PE-ROI feature appearing in each pixel of the image can be calculated using Equation (9). The PE-ROI distribution can be expressed as a combination of two Gaussian distributions, that is, k = 2. The expectation maximization algorithm is then used to calculate the unknown parameters $(\pi_k, \mu_k, \sum_k)$ in Equation (9). Finally, the probability distribution of the PE-ROI in each coordinate is obtained.

Using the probability distribution, a candidate region can be sampled from the image to extract anchor points. First, the k-th Gaussian distribution is selected based on the mixture coefficient value. Subsequently, a coordinate is randomly selected based on the selected Gaussian distribution. If the selected coordinates are duplicates, they are rejected. Next, the position information is extracted from the obtained candidate anchor points, and then numerous samplings are performed.

By combining the EFM and probability-based anchor extraction strategies, the network framework can effectively improve the detection performance for small objects, such as PE. Specifically, the feature map is first up-sampled to enrich the local details of small objects and then anchor points with a higher density are extracted to enlarge the feature map. The feature map establishes a candidate region and performs anchor-point extraction on it. This approach can not only extract anchor points closer to the ground truth but also reduce the time and space required owing to the feature map enlargement.

Training Process

In this study, the objective function of the network for extracting candidate regions is defined as

$$L(\{p_i\},\{t_i\}) = \frac{1}{N_{cls}}\sum_i L_{cls}(p_i, p_i^*) + \lambda\frac{1}{N_{reg}}\sum_i p_i^* L_{reg}(t_i, t_i^*) \tag{11}$$

where the classification loss L_{cls} is the binary cross-entropy loss, and the regression loss L_{reg} is the smoothing of the L_1 loss. Both are normalized by the minimum batch size N_{cls} and number of anchor positions N_{reg}, weighted by λ. It is important to note that L_{reg} works only for positive anchors. i represents the i-th anchor in the mini-batch; p_i and p_i^* represent the probability of being predicted as PE and the ground truth label (p_i = {0,1}), respectively; t_i and t_i^* represent the predicted and ground truth position with the association of a positive anchor, respectively. This position comprises the following four parameters:

$$\Delta x = (x - x_a)/d_a, \Delta y = (y - y_a)/d_a, \tag{12}$$

$$\Delta z = (z - z_a)/d_a \text{ and } \Delta d = \log(d/d_a), \tag{13}$$

where (x, y, z, d) are the center coordinates of the predicted or ground truth cube and its side length, and (x_a, y_a, z_a, d_a) are the center coordinates for the cube with anchor points and its side length.

To collect training samples for the first-stage network, a binary class label was assigned to each anchor. An anchor was marked as positive if it overlapped with a ground truth and its IoU was greater than 0.5, whereas it was marked as negative if its IoU was less than 0.02 compared to all ground truths. Anchors that were neither marked as positive nor negative were not used for training. This study used hard example mining during training by randomly selecting M negative samples in each mini-batch and ranking them in descending powers based on their

classification scores. Calculating the objective function by selecting the top k samples in the ranking as samples for hard mining is helpful.

The first-stage 3D FCN was pretrained on the LUNA16 dataset. In this study, the first stage of the model was trained for 100 epochs using a stochastic gradient descent (SGD) optimizer with a learning rate of 1e-3, momentum of 0.9, and weight decay of 1e-4. To train the second-stage ResNet-18, the SGD optimizer was used with a momentum of 0.99 and an initial learning rate of 1e-4. The learning rate was reduced by a factor of 10 after every 30 epochs during training. ResNet-18 was trained for approximately 100 epochs until convergence.

The objective function of the second-stage network was a cross-entropy loss function with softmax. False positive anchors based on the output of the first stage were collected as training data for the second stage. If the center of the candidate cube generated by the first-stage network did not fall on any ground-truth mask, the candidate cube was marked as a negative sample for training. Conversely, if the center of a candidate cube fell on the ground-truth mask, the candidate cube was marked was a positive sample for training. To alleviate the problem of severe imbalance in the numbers of positive and negative samples, the positive samples were augmented by scaling, randomly translating, and rotating the original cubes, and then extracting candidate cubes that aligned with the blood vessels from these cubes. Specifically, a random scaling of N_s times was performed in the range of 15–35 mm, a random translation of N_t times was performed in the range of -5–5 mm, and a random rotation of N_r times was performed around the v_1 axis. During random translation, it was ensured that the center of the candidate cube that was moved still lay on the ground-truth mask. Therefore, each positive training sample could be eventually increased by $N_p = N_s \times N_t \times N_r$ in the second stage.

During data augmentation, N_s was set to 3 (i.e., 15, 25, and 35 mm) and the cross-sections of all candidate cubes were resized to 32 × 32 to enable easy input into ResNet-18. During the translation operation, the candidate cube was moved N_t = 4 times in a random direction within a 5 mm range. For the rotation operation, N_r was set to 5. Therefore, the positive training samples of the second-stage network increased by a factor of 60. In the entire architecture, two stages of training were conducted: training the first-stage network until convergence, and then training the second-stage network based on the output of the trained first-stage network model.

Performance Evaluation

The performance of the probability-based two-stage CNN was evaluated using free-response operating characteristic curves. The detection result was considered positive if the detected position was within the three localization error ranges of the ground-truth mask (i.e., 0, 2, and 5 mm). Figure 7 illustrates the assessment of the localization error using a raw CTPA image containing a PE. The upper-left image shows a saddle embolus in the pulmonary artery, the upper right image shows the reference standard when the positioning error is 0 mm (i.e., the ground truth), the lower left image shows the reference standard when the positioning error is 2 mm, and the lower right image shows the reference standard when the positioning error is 5 mm. The masks for these positioning errors were manually delineated by two radiologists having more than ten years of interpretation

experience on all CTPA images in the PE Challenge database, and were used as a reference standard during training and testing.

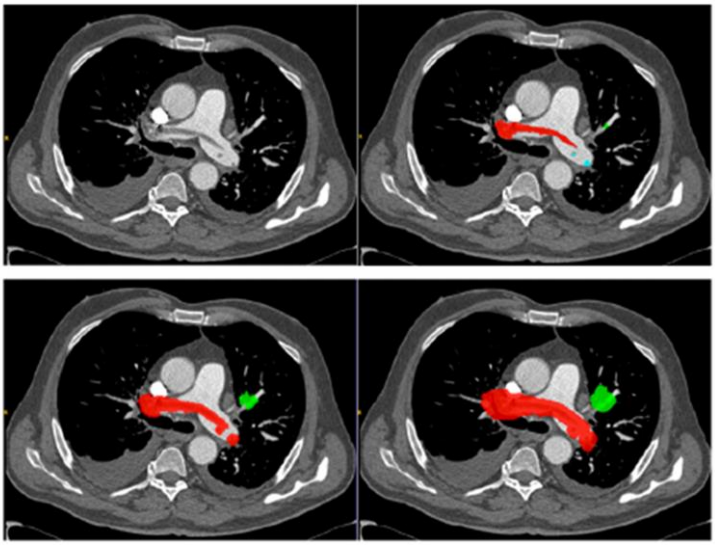


Figure 7. Three Positioning Error Ranges for Ground Truth Masks (González et al., 2020)

Sensitivity was used as the evaluation index. Sensitivity, also known as the true positive or recall rate, refers to the proportion of actual positive samples that are predicted to be positive. It can be calculated as follows:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{14}$$

where TP and NP are the numbers of true positives and false negatives, respectively. To further examine the effectiveness of each part of the probability-based two-stage CNN, the following four controlled experiments were performed using the strategy proposed by the research method through the analysis of control variables: (1) using only the original two-stage CNN; (2) using a modified two-stage CNN, including two strategies of EFM and PAE; (3) using a modified two-stage CNN and only the EFM strategy; and (4) using a modified two-stage CNN and only the PAE strategy.

Experimental Design and Analysis

The hardware used for this study included a personal computer with an Intel Core(TM) i7-10750H 2.60 GHz CPU, 16GB of RAM, and an NVIDIA GeForce GTX 1650M GPU. Ubuntu 14.04 was used as the operating system and Python 3.7 as the programming language, which included the Pytorch 0.4.0, TensorFlow 1.10.0, Keras 2.0.8, scikit-image, and OpenCV packages.

Database

The CTPA images used in this study were sourced from the PE Challenge database, which contains clinical data obtained from six different local hospitals. The radiology research center used was Unidad Central de

Radiodiagnóstico, Madrid, Spain. The CTPA image dataset contains 235 emboli, with a mean size of 3.43 ×□103 ± 8.6 ×□103 mm3 (minimum size = 8.63 mm3, maximum size = 6.34 ×□104 mm3). The minimum and maximum numbers of emboli in each case were 1 and 21, respectively (González et al., 2020).

The database comprised 3,438 CTPA images with PE, which were sorted according to the medical record number of the patient (ID) and slice order. Moreover, this database is accompanied by annotations, such as the embolism position and size in each image, which can be used to evaluate the effect of image processing. These images were divided into two parts: training set (1,978 images from 20 patients) and test set (1,460 images from another 20 patients). During the experiment, using the same patient data for training and testing was avoided.

Preprocessing

The grayscale value range of the images used in the experiment was between -2048 and 2048. First, the grayscale value range was adjusted to be between 0 and 4095, and the resolution of the image was 512 × 512. The slice thickness was set to 1 mm. CTPA images from 40 patients were used in the experiment. Each patient had a dataset, and each dataset contained approximately 250–280 images, but not every image contained a PE. After screening each dataset, only approximately 30–50 images containing PE were obtained. Figure 8 (a) shows the normalized CTPA image, wherein the overall contrast is still low. A cubic curve was used to enhance the contrast of the entire image, the result of which is shown in Figure 8 (b). In addition to increasing the brightness difference between the general pulmonary artery and PE, thin blood vessels in the pulmonary lobe can also be highlighted, thereby reducing the difficulty of detecting PE.

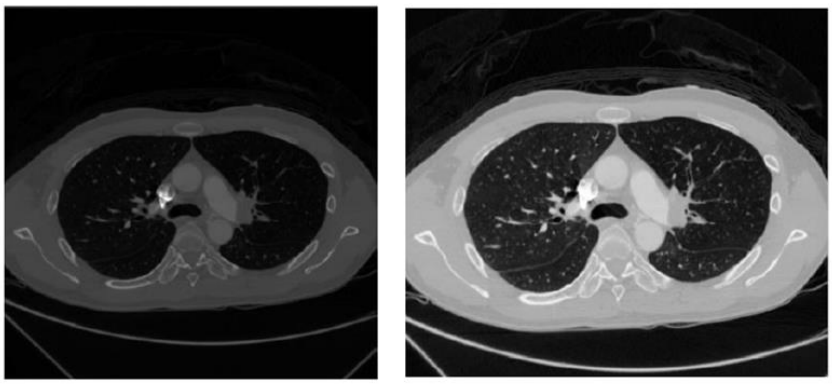
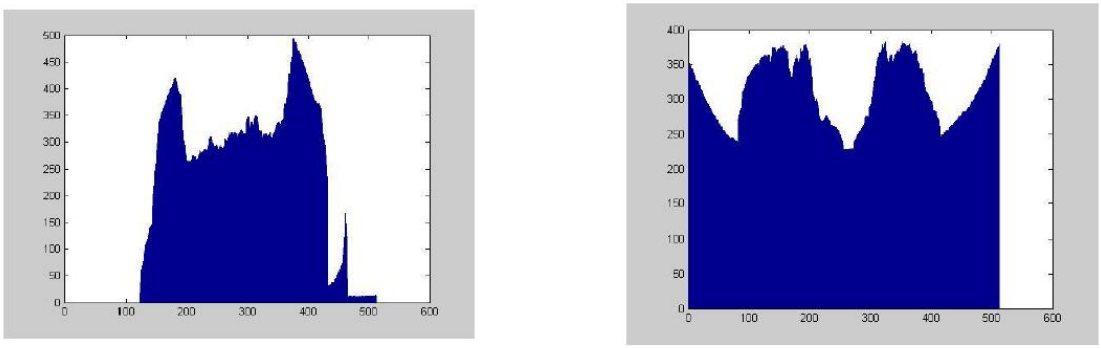


Figure 8. Contrast enhancement

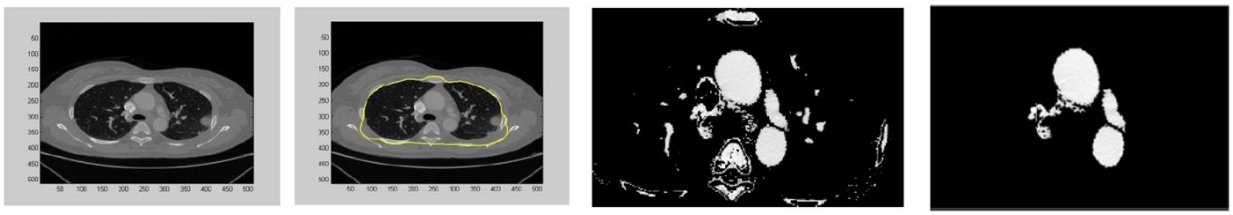
Figures 9 (a) and (b) show the projection results of the grayscale value of the image in the x- and y-axes directions, respectively. The ACM then uses these distribution features in combination with grayscale values to determine the contours of the lung region.



(a) (b)

Figure 9. Projection Results of Grayscale Value Distribution

After the contour was generated, the angle obtained in the edge detection process was used to gradually converge the lung block of the image based on the ellipse shape. Figure 10 (a) shows the original CTPA image, and (b) shows the ACM using the angle obtained in the edge detection process to gradually converge the image of the lung block based on the ellipse shape. After performing ACM, the area included in the range could be segmented, such as in the image of the main blood vessel area shown in Figure 10 (c). After obtaining the lung area framed by the ACM, the connected element algorithm was used to obtain the required area, as shown in Figure 10 (d).



(a) (b) (c) (d)

Figure 10. Working of the Active Contour Model

Probability-Based Two-Stage CNN

In the testing phase, the first stage of the probability-based two-stage CNN generated 2,687 PE-ROIs, of which 672 were true PEs and 2,015 were false positives. For each patient, the average number of false positives in the first phase of the network was 39.44. After aligning the blood vessels in the cube covering the PE through PCA, the emboli appearance was interpreted more consistently than the original irregular shape by using two types of images as representations. As can be observed in Figure 11, in the case wherein the appearance of almost all emboli is different, the PCA results represent the PE covered by the blood vessel as both elongated longitudinal and cross-sectional structures. The orientation of the vessels is different, and the original transverse appearance does not exhibit this characteristic. Consistent image appearance is key for accurately training CAD systems. Figure 11 (a) shows the transverse (left), coronal (middle), and sagittal (right) slice images of the original PE cubes in the three cases. It can be observed that they all appear different. Figure 11 (b) shows the longitudinal (left, middle) and cross-sectional (right) slices after vessel alignment.

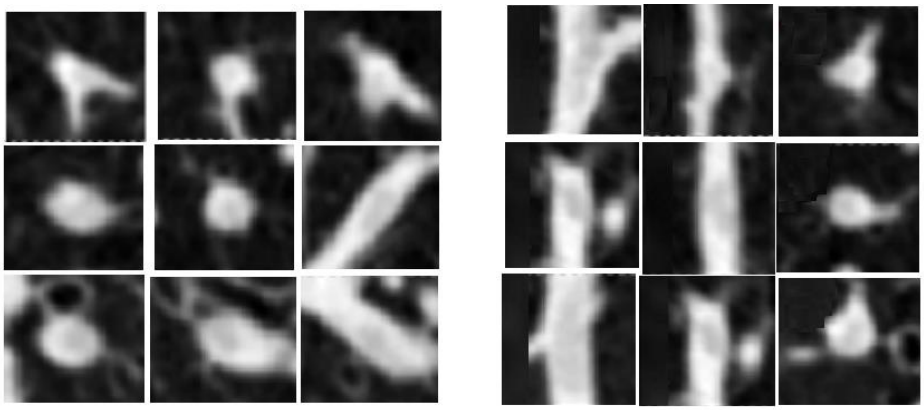


Figure 11. Vessel Alignment

In the second stage, the three feature maps generated through PCA were input into ResNet-18 to extract the anchors, and the size of the feature maps was 32 ×16. Therefore, according to the sampling stride length definition in Equation (8), the sampling stride length of the three feature maps is (8, 16, 32). However, at this sampling stride length, the quality and quantity of positive anchors are scarce. As presented in Tables 1 and 2, in ResNet-18, when t in Equation (7) is equal to 0.7, the average number of positive anchor points is 2.01. When the IoU of all the positive anchor points is greater than 0.7, the proportion is only 12.58%. This is because numerous anchors are identified as positive because they have the largest IoUs (but $t < 0.7$) between them and the ground truth, and not because the condition of $IoU > 0.7$ is met.

Table 1. Average Number of Positive Anchors at Different Sampling Stride Lengths and IoU Thresholds

	Average number of positive anchors				
Sampling stride	t = 0.5	t = 0.6	t = 0.7	t = 0.8	t = 0.9
ResNet-18	3.41	2.67	2.01	1.65	1.52
(1, 2, 4)	20.32	11.71	6.94	6.05	5.56
(1, 4, 4)	29.75	15.34	8.02	7.57	6.98
(1, 4, 8)	45.27	25.96	12.39	10.24	9.65
(2, 2, 4)	32.59	19.24	7.22	6.41	5.84
(2, 4, 4)	12.33	7.43	4.58	3.92	3.46
(2, 4, 8)	11.89	6.01	3.72	3.43	3.17

Excessive distance between positive anchors and ground truth adversely affects the detection results. The error between the anchor and ground truth can be considered a linear error only when they are sufficiently close; otherwise, it is considered to be a complex nonlinear problem. While enlarging the feature map, as the sampling stride decreases, the number of positive anchors gradually increases. As can be observed from Tables 1 and 2, when the sampling stride of the three feature maps is (1, 4, 8), the average number of positive anchor points is 12.39, which is nearly six times that of the original, and the number of positive anchors is 12.39. The proportion of anchor points with an IoU greater than t also increased to 35.70%; however, the sampling stride is not as small

as possible. As the sampling step shrinks, the magnification of the feature map increases exponentially, which may result in memory overflow.

Table 2. The Ratio of Positive Anchor Points Greater than t under Different Sampling Stride Lengths and IoU Thresholds

	Ratio of positive anchors greater than t				
Sampling stride	t = 0.5	t = 0.6	t = 0.7	t = 0.8	t = 0.9
ResNet-18	81.31%	26.88%	12.58%	0.92%	0.19%
(1, 2, 4)	93.53%	54.57%	28.15%	1.43%	0.47%
(1, 4, 4)	95.07%	49.19%	32.72%	2.97%	0.51%
(1, 4, 8)	98.69%	47.58%	35.70%	3.67%	0.79%
(2, 2, 4)	98.49%	52.65%	31.26%	2.73%	0.64%
(2, 4, 4)	95.38%	49.14%	26.84%	2.83%	0.43%
(2, 4, 8)	95.63%	45.58%	25.09%	3.53%	0.31%

To obtain the correct number of positive anchors, the threshold t was set to 0.5. As can be observed from Tables 1, 2, and 3, when t is equal to 0.5, and the sampling stride length is (1, 4, 8), the average number of positive anchors is 45.27 and the proportion of points with an IoU greater than 0.5 is 98.69%. Compared with ResNet-18, the EFM method achieved a higher proportion of positive anchors with IoU values greater than 0.5. This proves that the anchor points extracted using the EFM method are closer to the ground truth.

Table 3. Average IoUs of Positive Anchors at Different Sampling Stride Lengths and IoU Thresholds

	Average IoU of positive anchors				
Sampling stride	t = 0.5	t = 0.6	t = 0.7	t = 0.8	t = 0.9
ResNet-18	0.52	0.53	0.55	0.56	0.56
(1, 2, 4)	0.55	0.56	0.57	0.57	0.55
(1, 4, 4)	0.56	0.57	0.57	0.56	0.56
(1, 4, 8)	0.59	0.58	0.58	0.57	0.57
(2, 2, 4)	0.57	0.59	0.62	0.59	0.58
(2, 4, 4)	0.55	0.61	0.58	0.58	0.57
(2, 4, 8)	0.56	0.59	0.61	0.58	0.58

By obtaining the probability distribution, the false-positive anchor points can be removed from a region in the image. First, the k-th Gaussian distribution is selected based on the value of the mixture coefficient, and then a coordinate is randomly selected based on the selected Gaussian distribution. If the selected coordinates are duplicates, they are excluded. Subsequently, the position information is extracted from the anchor points in the obtained image, and then a significant number of samplings are performed. This makes it easy to extract anchors that match the ground truth. Figure 12 (a) shows the location distribution of the center points of all PE-ROIs in the training set, and (b) shows the Gaussian mixture model established from these center points.

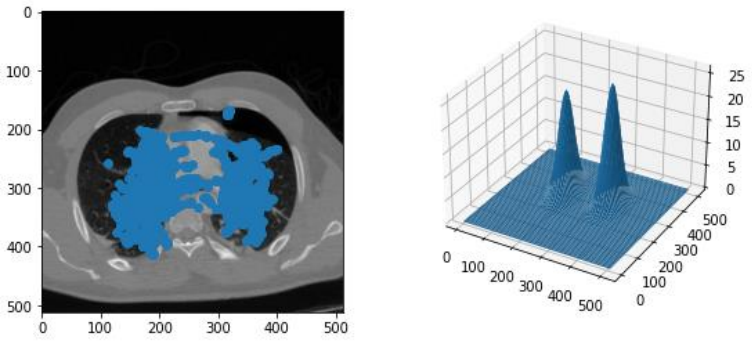


Figure 12. (a) Location Distribution of PE-ROIs and (b) the Established Gaussian Mixture Model

Figures 13 (a) and (b) show the results of successful detection under a positioning error of 0 mm; the area in the red frame indicates embolism, and Figures 13 (c) and (d) show the results of failed detection. In this study, failure occurred owing to closeness of the pixel brightness and plug, which caused misjudgment.

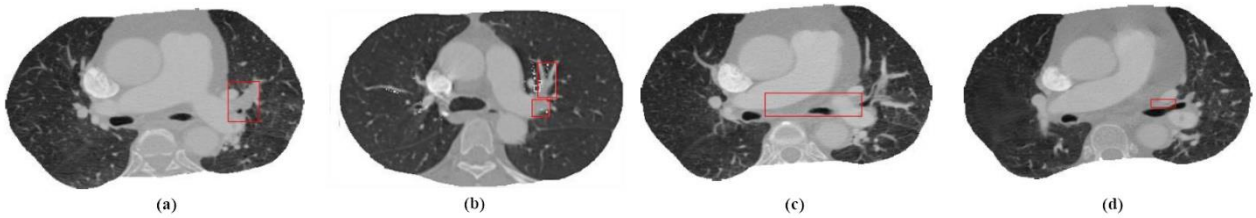


Figure 13. Detection Results

Figure 14 shows the performance of the probability-based two-stage CNN for localization errors of 0, 2, and 5 mm. Radiologists require a system that can maintain a relatively low number of false positives (i.e., 1–5 per CTPA image) while maintaining high sensitivity. The graph depicts the sensitivity when two false positives are observed per scan. The probability-based two-stage CNN achieved 85.97%, 87.31%, and 88.71% sensitivity in the best case of scanning two false positives under the 0, 2, and 5 mm positioning errors of the second stage, respectively.

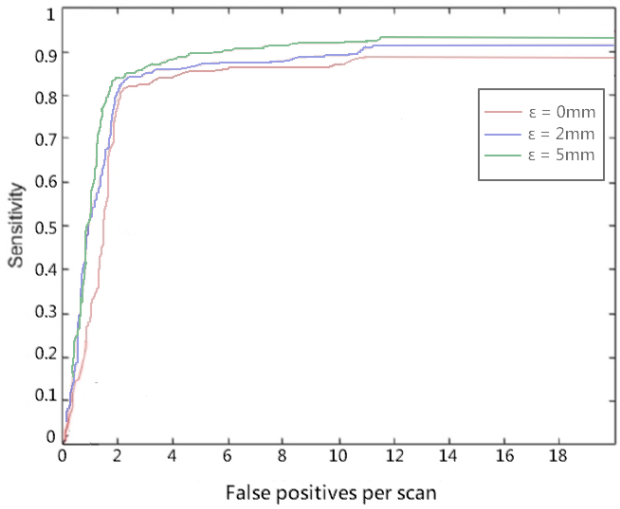


Figure 14. Stage-Two Performance with 0, 2, and 5 mm Positioning Errors

Table 4 lists the results of this study and those of other research teams that participated in the CAD-PE Challenge competition. This competition is organized by the creators of the PE Challenge database, and all participating research teams must continue to use the same materials and performance evaluation criteria to ensure fairness. The purpose is to compare the teams that can develop the best algorithm to detect PE. This study also evaluated a probability-based two-stage CNN using the same performance evaluation criteria.

The results in Table 4 are based on the case of a 0 mm localization error; that is, the detected PE is considered a true positive only when the detection result falls entirely within the ground truth mask. This study argues that optimizing performance under a positioning error of 0 mm can be more helpful for clinical applications and is more challenging than having a higher performance advantage under 2 and 5 mm positioning errors. As can be observed from Table 4, although the related research work started after the competition, the results ranked first compared with other teams that have previously participated in the CAD-PE Challenge. The proposed method achieved a sensitivity of 85.97%, with two false positives per scan. The teams marked in blue, including the one that performed this study, used deep learning methods, and those marked in green used traditional machine learning methods, such as support vector machines that support vector learning or Adaboost. Overall, deep learning methods outperformed traditional machine learning methods in PE detection, and are more suitable for developing CAD systems for PE detection.

Table 4. Results of This Study and Other Teams that have Previously Participated in the CAD-PE Challenge

Teams	Number of training sets	Sensitivity	
		1FP	2FP
UA-2D	20	37%	43%
UA-2.5D	71	50%	58%
ASU-Mayo	20	28%	33%
Lin et.al	20	N/A	75%
This Study	20	N/A	85%
Mevis Inc.	20	N/A	28%
UPM	20	15%	22%
FUM-Mvlab	20	7%	22%
BWH	20	1%	2%
LMNIT	20	0.3%	0.7%

Table 5 lists the experimental results based on different factors. It can be observed that when tested on the test set, the original two-stage CNN has a sensitivity of 75.40% in the best case, wherein the localization error is 0 mm and two false positives are scanned. Compared with the original two-stage CNN, the EFM improves the sensitivity by 7.92%, but also increases the training time by 1.59 h. In contrast, PAE improves the sensitivity by 5.69% while reducing training time by approximately 56 min. Using EFM and PAE simultaneously can significantly improve the sensitivity by 10.57%, whereas increasing the training time by only 29 min. Therefore, combining EFM and PAE with a two-stage CNN can improve the PE detection performance with acceptable extra time consumption.

Table 5. Experimental Results Based on Different Factors

Network architecture	Factor	0 mm	2 mm	5 mm	Training time (h)
ResNet-18		75.40%	75.40%	75.40%	4.45
ResNet-18	EFM	83.32%	86.37%	87.22%	6.04
ResNet-18	PAE	81.09%	84.12%	84.63%	3.51
ResNet-18	EFM+PAE	85.97%	87.31%	88.71%	4.94

Conclusion

In this study, a deep learning approach was used to extract candidate regions from raw CTPA images by using a 3D FCN. Through the PCA process of aligning the blood vessels and calculating the regularity of PE lesion distributions in the image, the false-positive categories in the candidate area could be conveniently eliminated. During the experiment, the limitations of the 3D FCN in detecting small objects, such as PE, in medical images were discovered. Because the extraction stride of the anchor points was too large, most of the anchor points extracted by the 3D FCN could not match the ground truth, resulting in poor detection performance. Increasing the density of samples by enlarging the feature map resulted in additional memory and time consumption. Most PEs only appear in some specific areas of the image, rather than evenly distributed over the entire image. Therefore, by calculating the frequency of all PEs in the training set at different positions in the image, a Gaussian mixture model could be used to establish the probability distribution of PE at different positions in the image. The extraction of candidate regions facilitated the extraction of positive anchors from the probability distribution through sampling. By combining the EFMs and candidate targets, a series of anchor points that were easier to match with the ground truth could be extracted with higher accuracy, while avoiding additional time and space consumption. It was concluded that combining EFMs and PAE with a two-stage CNN can improve the PE detection performance with acceptable extra time consumption.

This study mainly contributes to the field in two aspects: (1) it proposes a modified probabilistic-based two-stage CNN framework. By combining image processing for vessel alignment, proper feature map up-sampling, and a probability-based anchor extraction strategy, the PE detection performance can be significantly improved. (2) After modification, the probability-based two-stage CNN improved the PE detection sensitivity from 75.4% to 85.97%.

A limitation of the PE Challenge database is that all selected cases are PE-positive; therefore, the predictive value of this study architecture for PE-negative cases could not be assessed. Although other public datasets of CT images exist, especially those related to lung cancer screening, to the best of our knowledge, there is no public dataset of pulmonary computed tomography angiography images with negative cases. The second limitation is the amount of training data; therefore, the network architecture used in this study includes data augmentation in the second stage. It is believed that adding new cases to the PE Challenge public database can further help in developing new algorithms and improving the detection performance.

During the experiment, it was also discovered that there is still room for growth in the study of pulmonary artery computed tomography angiography. Misjudgment as a candidate target to reduce the number of false positives that must be eliminated in the second stage is an issue that can be further studied. Additionally, this study also discovered that some other medical image recognition problems can use a similar framework, such as calcified plaque detection in cervical angiography and tumor detection in other CT images; in these cases, the distribution of lesions may appear in other forms, rather than the probability distribution calculated in this study. For example, the probability of plaque distribution in the carotid or coronary arteries can also be calculated using a Gaussian distribution. Based on these extensions, it is expected that the probability distribution of other lesions in medical images can be established in the future, and a corresponding CAD system can be developed to ensure radiologists concentrate more on the locations in the images that have a high probability of lesions.

Acknowledgements

We are grateful to the Ministry of Science and Technology of Taiwan for providing a research grant that supported this study (MOST 110-2221-E-415-014).

References

Augusto, S., Jose, S. S., Beatriz, S. S., & Carlos, F. (2001). Fast Pulmonary Contour Extraction in X-ray CT images: A Methodology and Quality Assessment. *Medical Imaging 2001: Physiology and Function from Multidimensional Images*, 4321, 216–224.

Blechschmidt, R.A., Werthschutzky, R., & Lorcher, U. (2001). Automated CT image evaluation of the lung: a morphology-based concept. *IEEE Transactions on Medical Imaging*, 20(5), 434–442.

Buhmann, S., Herzog, P., Stoeckel, J., Salganicoff, M., Wolf, M., & Reiser, M. F. (2006). Clinical evaluation of a CAD prototype for the detection of PE. In *European Congress of Radiology*, (pp. 585).

Dajnowiex, M., & Alirezaie, J. (2004). Computer simulation for segmentation of lung nodules in CT images. In *IEEE International Conference on Systems Man and Cybernetics*, (Vol. 5, pp. 4491–4496).

Das, M., Schneider, A., Schoepf, U., Cheema, J., Wood, S., & Costello, P. (2003, November). Computer aided diagnosis of peripheral pulmonary emboli. In *Radiological Society of North America*, (pp. 351–352). RSNA.

Digumarthy, S. R., Kagay, C. R., Legasto, A. C., Muse, V. V., Wittram, C., & Shepard, J. O. (2006). *Computer-aided detection (CAD) of acute pulmonary emboli: evaluation in patients without significant pulmonary disease*, Chicago, IL: SSC.

Domingo, M., Martí-Bonmatí, L., Dosdá, R., & Pallardó, Y. (2000). Interobserver agreement in the diagnosis of PE with helical CT. *Eur. J. Radiol*, 34, 136–140.

Goldhaber, S. Z., & Bounameaux, H. (2012). PE and deep vein thrombosis. *Lancet*, 1835(46), 379.

González, G., Jimenez-Carretero, D., Rodríguez-López, S., Cano-Espinosa, C., Cazorla, M., Agarwal, T., Agarwal, V., Tajbakhsh, N., Gotway, M. B., & Liang, J. et al. (2020). Computer Aided Detection for PE Challenge (CAD-PE). Available online: https://arxiv.org/abs/2003.13440

Hall, D., Dayoub, F., Skinner, J., Zang, H., Miller, D., Corke, P., Carneiro, G., Angelova, A., & Suenderhauf, N. (2020). Probabilistic object detection: Definition and evaluation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, (pp. 1031–1040). WCAC.

Kiraly, A. P., Novak, C. L., Naidich, D., Vlahos, P. I., Ko, J. P., & Brusca-Augello, G. T. (2006, March). A comparison of 2D and 3D evaluation methods for PE detection in CT images, *Proceedings of the SPIE 6146, Medical Imaging 2006: Image Perception, Observer Performance, and Technology Assessment*, 61460H.

Liang, J., & Bi, J. (2007). Computer aided detection of PE with tobogganing and multiple instance classification in CT pulmonary angiography. In *Karssemeijer, N., Lelieveldt, B. (eds.)*, (Vol. 4584, pp. 630–641). IPMI.

Maizlin, Z., Vos, P., Gody, M., & Cooperberg, P. (2007). Computer-aided detection of PE on CT angiography: initial experience. *Journal of Thoracic Imaging*, 22(4), 324–329.

Margarida, S., & Jorge, M. (2006). Automatic segmentation of the lungs using multiple active contours and outlier model. In *IEEE, 2006. Engineering in Medicine and Biology Society Annual International Conference*, (pp. 3122–3125).

Masutani, Y., MacMahon, H., & Doi, K. (2002). Computerized detection of PE in spiral CT angiography based on volumetric image analysis. I*EEE Transactions on Medical Imaging*, 21, 1517–1523.

Michiels, J. J., Hoogsteden, H., & Pattynama, P. M. (2005). Non-invasive diagnosis of PE. *Acta Chir Belg*, 1, 26–34.

Nicholas, J. S., & Harvey, O. C. (2003). Detection of lung perfusion abnormalities using computed tomography in a porcine model of PE. *Journal of Thoracic Imaging*, 18, 14–20.

Ozkan, H., Osman, O., Sahin, S., & Boz, A. F. (2014). A novel method for PE detection in CTA images. *Computer Methods and Programs in Biomedicine*, 113(3), 757–766.

Park, W., Hoffman, E. A., & Sonka, M. (1998, Aug). Segmentation of intrathoracic airway trees: a fuzzy logic approach. *IEEE Transaction on Medical Imaging*, 17(8), 489–197.

Pichon, E., Novak, C. L., Kiraly, A. P., & Naidich, D. P. (2004). A novel method for pulmonary emboli visualization from high-resolution CT images. *Proceedings of the SPIE Medical Imaging*, 5367, 161–170.

Qanadi, S. D., Hajjam, M. E., & Mesurolle, B. (2000). PE detection: Prospective evaluation of dual section helical CT versus selective pulmonary arteriography. *Radioloay*, 217, 447–455.

Tajbakhsh, N., Gotway, M. B., & Liang, J. (2015, November). Computer-aided pulmonary embolism detection using a novel vessel-aligned multi-planar image representation and convolutional neural networks. In *Proc. Int. Conf. Med. Image Comput*, (pp. 62–69). Comput-Asist. Intervent.

Wang, X., Song, X., Chapman, B. E., & Zheng, B. (2012). Improving performance of computer-aided detection of PEs by incorporating a new pulmonary vascular-tree segmentation algorithm In. *International Society for Optics and Photonics*, (83152U). SPIE Medical Imaging.

Xu, D., J. Lee., Raicu, D. S., Furst, J. D., & Channin, D. (2005, June). Texture Classification of Normal Tissues in Computed Tomography. In *The 2005 Annual Meeting of the Society for Computer Applications in Radiology*, (pp. 2–5). Orlando, Florida.

Yang, X., Lin, Y., Su, J., Wang, X., Li, X., Lin, J., & Cheng, K. T. (2019). A two-stage convolutional neural network for pulmonary embolism detection from CTPA images. *IEEE Access*, 7, 84849–84857.

Yongbum, L., Takeshi H., Hiroshi F., Shigeki I., & Takeo I. (2001). Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique. *IEEE Transactions on Medical Imaging*, 20(7), 595–604.