Local Characteristic Features for Computer Aided Detection of Pulmonary Embolism in CT Angiography

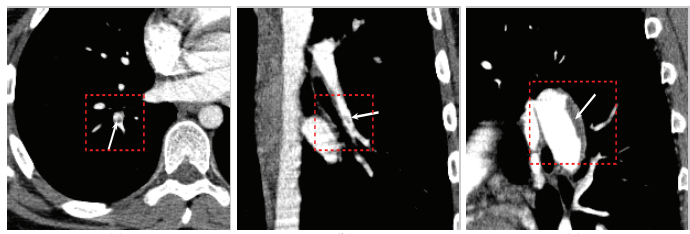
Jianming Liang and Jinbo Bi
Siemens Medical Solutions, CKS IKM
Malvern, PA 19355, USA

Abstract. An automated detection system is constructed for detecting pulmonary embolism from computed tomographic pulmonary angiographic images. Our previous work has presented novel effective algorithms to identify suspicious PE regions from images and reduce false detections by designing powerful classifiers. However, these techniques have to take effects in conjunction with discriminative features used to characterize each identified PE candidate. This paper investigates three sets of novel features: 1) features based on local candidate co-occurrence matrices to remove false detections induced by noise and poorly mixed contrast; 2) features characterizing vessel properties to eliminate candidates outside of vessel; 3) features discriminating between arteries and veins to remove candidates from veins. We tested these features in our multiple instance learning classification setting, and they constantly improved the detection accuracy when the 3 sets of features are included sequentially. The resulted PE CAD system has been deployed in clinical settings with capabilities of incrementally reporting any detection immediately once becoming evident during searching, offering real-time support and achieving 85% sensitivity at 5 false positives.

1 Introduction

Computer aided detection of pulmonary embolism (PE) in computed tomographic pulmonary angiographic (CTPA) images has received growing attention in recent years. Pulmonary embolism (PE) is the third most common cause of death in the US, with at least 650,000 cases occurring annually, although treatment with anti-clotting medications is highly effective. Each year more than 400,000 PE cases are not correctly diagnosed, and approximately 100,000 patients die who would have survived with the proper diagnosis and treatment. CTPA has emerged as the first-line diagnostic tool for PE. In CTPA, an embolus appears as dark regions residing in bright vessel lumen as illustrated in Fig. 1. A CTPA volume consists of hundreds slices of the lung. The accurate and efficient interpretation of such a large image volume is complicated by various PE look-alikes and also limited by human factors, such as attention span and eye fatigue. It is highly desirable to have an automated detection system to assist radiologists in detecting and characterizing emboli in an accurate, efficient and reproducible way. Such a CAD system must satisfy stringent real-time requirement since PE cases often occur in emergency room. It must achieve high detection sensitivity with as few false positives as possible to acquire clinical acceptance, because inaccurate use of anti-clotting medications can lead to subsequent hemorrhage and bleeding.

PE detection has been attacked by several groups [1–4]. All the existing methods in the literature require sophisticated vessel segmentation. Vessel segmentation is computationally time-consuming, has been problematic in small vasculature where



(a) (b) (c)

Fig. 1. Acute (a,b) and chronic (c) pulmonary emboli.

subsegmental PEs often occur, and further complicated by PE if exists. None of these approaches can meet the real-time requirement. To circumvent the problem, our early work developed an effective system to search for any suspicious PE regions in the entire lung without segmenting the vessels, and then reduce non-PE regions by distinguishing them from true PE candidates with local characteristic features. Since the number of suspicious PE regions (called candidates) is rather limited (about 50) in each case, and it is much more efficient to compute the features descriptive to the local vascular structure of the candidates than segmenting the whole pulmonary vascular tree.

We investigate three sets of novel local characteristic features in this paper. First, features based on local candidate co-occurrence matrices are calculated to remove false detections due to noise and poorly mixed contrast. Second, Hessian-based features are designed to characterize vessel properties to eliminate candidates outside of vessel. Last, image features discriminating between arteries and veins are computed to remove candidates from veins. These features were tested in the multiple instance classification setting. Together with our candidate generation and classification algorithms, the overall approach forms a comprehensive and clinically usable PE CAD system which is capable of reporting the first detection if any within 30 seconds and achieving 85% sensitivity under 5 false positives on average across a large number of cases. For completion, we first review our concentration oriented tobogganing algorithm for quickly generating suspicious regions for PE incrementally (candidate generation) and our multiple instance learning (MIL) framework for false positive reduction (classification) in Section 2 before presenting our three sets of novel features in Secs. 3, 4 and 5, respectively.

2 Retrospect of candidate generation and classification

2.1 Candidate generation

When the vascular structure is not segmented, a major challenge for automatic PE detection is to effectively separate the emboli from the vessel wall and to quickly remove partial volume effects around the vessel boundaries while correctly preserving the PE pixels, since all the voxels in those areas have the same original CT values as those in the PE regions. To this end, we first reported an approach based on basic tobogganing in [5] with a simple operation called "sliding": A pixel v with intensity $P(v)$ and

Table 1. False positive distributions based a meticulous analysis of training cases.

FP types	in vein	in artery	on lymph	on artery wall	between vessels	vein wall	as Nodule	at bifurcation	others
percentage	39.06%	20.24%	12.00%	9.88%	4.71%	4.47%	3.29%	2.82%	3.53%

neighbors $N(v)$ slides down to pixel $g = \arg\min_{t \in N(v) \cup \{v\}} P(t)$. A pixel that cannot slide to any of its neighbors is called a *concentration*. All the pixels that slide down to the same concentration form a toboggan *cluster*. The basic idea is to treat the image as a landscape and let pixels slide within the landscape. Almost all the pixels along the vessel wall in the same intensity range as true PEs will merge into air regions and the true PE regions will stand out automatically. However, this basic tobogganing approach requires to scan the whole volume two times and cannot meet the real-time requirement for PE detection. To accelerate the tobogganing process, we developed a concentration-oriented tobogganing algorithm, achieving the same sensitivity as basic tobogganing but reducing computation time from originally 2 minutes to 27 seconds.

Readers can consult with [6] for a complete description of our concentration-oriented tobogganing algorithm. Briefly, the algorithm has two steps. It first searches for concentration c from the given pixel s and then expands from the found concentration c to extract the whole toboggan cluster C. The expansion includes a base step and an iterative step. In the base step, it includes the concentration c as the first pixel in the cluster and pushes all its neighbors with CT values between -50 HU and 100 HU into an active list A. In the iterative step, it selects pixel q with the minimal CT value from the active list A, if the selected pixel toboggans to an already clustered pixel, then conditionally pushes its neighbors to the active list A to ensure the uniqueness of the pixels in the active list, otherwise, the selected pixel belongs to the cluster's external boundary B. The iterative step is repeated till the active list A is empty. This concentration oriented tobogganing algorithm is repeatedly applied on all those external boundary pixels, until a whole PE candidate has been extracted. This process is illustrated in Fig. 2.

The concentration oriented tobogganing algorithm is then iteratively applied on each of the external boundary pixels with CT value in the PE HU range. Any additional extracted toboggan cluster is merged into the previously extracted toboggan clusters, and any additional external boundary pixels are also merged. Once no external boundary pixel is left, all the toboggan clusters are extracted and merged, automatically forming a connected component – a PE candidate.

We validated the tobogganing algorithm through a clinical study of 177 cases (collected from multiple medical institutions) with 872 clots marked by expert chest radiologists. These cases were divided into two sets: training (45 cases with 156 clots) and test (132 cases with 716 clots). This algorithm successfully detected 90.38% (141/156) of the PE in the training set and 90.1%(645/716) of the PE in the test set. However, it also produces false positives—candidates that do not intersect with any PEs. On average, 47.5 and 40.3 false positives for each case were generated for the training set and the test set, respectively. Our false positive distribution is shown in Tab. 1.

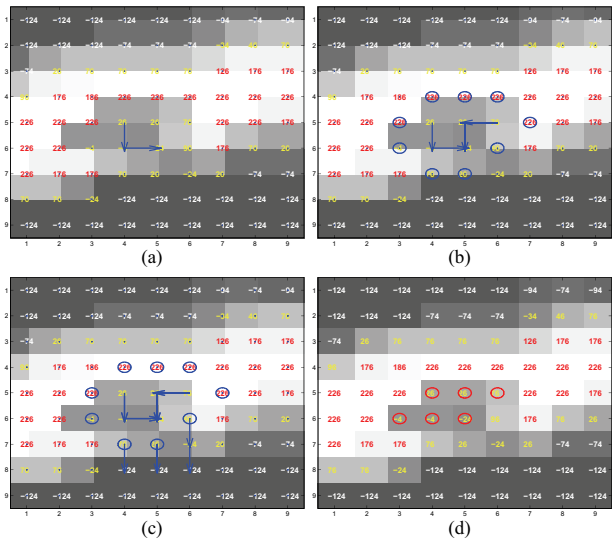


Fig. 2. Using the concentration oriented toboggan algorithm for detecting PE in a small vessel. During the scan in row by row, the first pixel which does merge into dark regions (< -50 HU) is (4,5) and regarded as a PE pixel, from which our concentration oriented toboggan algorithm is applied to extract a whole PE candidate. (a) Step A of the algorithm: Finding the concentration. (b) Step B: Expanding from the concentration to cover a whole toboggan cluster and providing all the external boundary pixels (circled). (c) Repeatedly apply the algorithm on all those external boundary pixels with CT value between -50 HU and 100 HU to form a PE candidate (d).

2.2 Classification

A system that "cries wolf" too often will be rejected out of hand by radiologists. Therefore, to reduce the false positives, an effective multiple instance learning algorithm was designed [7] based on the 1-norm support vector machine (SVM). The 1-norm SVM constructs a linear decision function $f(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + b$ by minimizing the regularized empirical error $\gamma||\mathbf{w}||_1 + \sum_{i\in S_+}\xi_i + \sum_{i\in S_-}\xi_i$ where γ is a tuning parameter and $\xi = \max\{0, 1-y(\mathbf{w}^T\mathbf{x}+b)\}$ and $y = 1$ for candidates in S_+ (PEs), $y = -1$ for candidates in S_- (non-PEs). The multiple-instance learning approach makes use of the fact that some candidates can belong to the same PE, so that a PE is missed only if all the candidates belonging to it are classified as non-PE. Let the i^{th} PE contain ℓ_i candidates, represented as a set of feature vectors $\{\mathbf{x}_{ij}\}_{j=1}^{\ell_i}$. Let $S_i, i = 1, \cdots, m$ be the index set of all candidates that belong to the i^{th} PE assuming there exist totally m PEs. The goal of our classification algorithm is to determine a decision boundary that separates, with high accuracy, at least one candidate from each set S_i, $i = 1, \cdots, m$ on one side and as many as possible negative detections on the other side. It implies that the candidate in S_i with the smallest loss ξ should be correctly classified for each i. Mathematically, it means we

merely need to minimize the smallest error occurred on each set S_i and hence the objective function is to $\min_{\mathbf{w},\xi} \gamma||\mathbf{w}||_1 + \sum_{i=1}^{m} \min\{\xi_{ij},\ j \in S_i\} + \sum_{i \in S_-} \xi_i$ subject to constraints (1) $\mathbf{w}^T\mathbf{x}_i + b \geq 1 - \xi_i,\ i \in U_{i=1,\cdots,m} S_i$, (2) $\mathbf{w}^T\mathbf{x}_i + b \leq -1 + \xi_i,\ i \in S_-$, and (3) $\xi_i \geq 0,\ i = 1, \cdots, \ell$.

We resort to aggregation of multiple classifiers trained over various sample patient sets to reduce the variance of the constructed classifier. We carry out T trials, and in each trial, two third of the training cases are randomly sampled and used for training. Linear decision functions are constructed in different trials. The final classifier is obtained by averaging the weight vectors of these decision functions. Features with very small weights in the final model will be removed to reduce noise level. This classification algorithm can reduce much more false positives naturally due to its design. However, the power of the classification algorithm relies on the discriminativeness of the features. We therefore develop three sets of features described in the following sections.

3 Local candidate characteristic features

As shown in Tab. 1, there are many false positives generated from vessels which are often due to noise or poorly mixed contrast material. One way to distinguish these false positives from true PE candidates is to design features based on the spatial distributions of intensity values computed within the candidate cluster. Therefore, in addition to the basic intensity distribution features such as mean, variance, skewness and kurtosis, we compute a number of features according to gray level co-occurrence matrices.

The co-occurrence matrices are widely used in texture analysis to capture the spatial dependence of intensity values within an image using second-order statistics. For a given 3D volume I of $n_1 \times n_2 \times n_3$ and a number of intensity levels N, a co-occurrence matrix $\mathbf{C}$ is an $N \times N$ matrix and acts as an accumulator over the image I. Specifically, a co-occurrence matrix along a direction (d_x, d_y, d_z) is defined as follows:

$$C_d(i,j) = \sum_{x=1}^{n1} \sum_{y=1}^{n2} \sum_{z=1}^{n3} \begin{cases} 1, & \text{if} I(x,y,z) = i \ \&\ I(x+d_x, y+d_y, z+d_z) = j, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Haralick [8] proposed fourteen statistical features computed from the co-occurrence matrices. Based on our experiments, the following four features have the most discriminative power in capturing the spatial dependence of intensity values within a PE candidate:

- **Entropy**, $\mathcal{E} = -\sum_i^N \sum_j^N C_d(i,j) \log C_d(i,j)$, measures the randomness of a gray-level distribution and is expected to be high if the gray levels are distributed randomly within a PE candidate.
- **Energy**, $\mathcal{G} = -\sum_i^N \sum_j^N C_d^2(i,j)$, measures the number of repeated pairs of gray levels, and is expected to be high if co-occurrence concentrates on certain pairs (i,j).
- **Contrast**, $\mathcal{C} = -\sum_i^N \sum_j^N (i-j)^2 C_d(i,j)$, measures the amount of local variations within a PE candidate and is expected to be low if the gray levels of each pixel pair are similar.

– **Homogeneity**, $\mathcal{O} = -\sum_i^N \sum_j^N C_d(i,j)/(1+|i-j|)$, measures the smoothness of a PE candidate and is expected to be large if the gray levels of all pixel pairs are similar.

Based on these definitions, we expect that true PE candidates have lower entropy, higher energy, lower contrast and larger homogeneity, so as to be distinguished from those false positives due to noise or poorly mixed contrast material. These features serve as the basic set of image features. We validate these features by evaluating the PE detection performance using a recently-developed classification algorithm [7]. Figure 3 shows test performance of 60% sensitivity at 5 false positives per volume with these features, and hence this set of features is not yet sufficient to meet our goal.

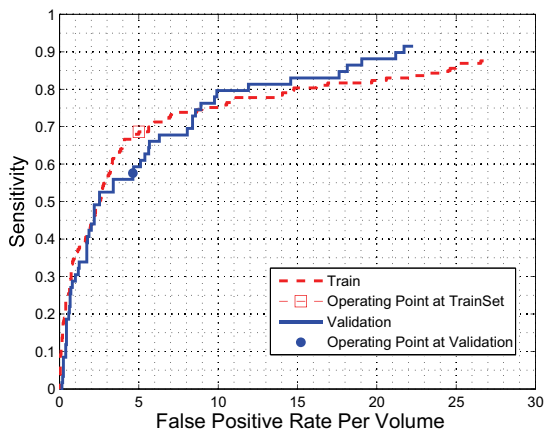


Fig. 3. ROC plot using local candidate features.

4 Vesselness characteristic features

Again as shown in Tab. 1, we have 12% of false positives generated from the connective (lymph) tissues located outside of the arteries. To eliminate this type of false positives, we design a set of features to determine if a candidate is located inside of a vessel based on the multi-scale Hessian vessel-likelihoods. Ideally, if a voxel is within a vessel, the Hessian matrix **H** formed from the second derivatives of intensity in the neighborhood of this voxel will have two negative eigenvalues and a third one is close to zero. The eigenvectors corresponding to the first two eigenvalues lie in a plane orthogonal to the central axis of the vessel, and the eigenvector corresponding to the third eigenvalue is in a line with this central axis. Based on the eigenvalues, vessel-likelihoods can be

computed based on the Frangi [9], Sato [10] and Lorenz [11] methods. However, this idea works only if the vessel does not have PEs. So, we precede this procedure by filling the detected regions (candidates) with high intensity values, to simulate the inside of a PE-ridden artery. To accommodate the different vessel radii, three vessel-likelihoods are computed at five different scales and the maximal response for each vessel-likelihood is taken across the five scales. In summary, our Hessian-based vesselness features are computed as follows:

1. Fill the detected candidates with high intensity values
2. Compute vessel-likelihood features for each voxel $\mathbf{x}$ in the candidate by:
   (a) Constructing the Hessian matrix $\mathbf{H}(\mathbf{x}, s)$ based on the second derivatives in scale s at voxel $\mathbf{x}$
   (b) Decomposing $\mathbf{H}$ into eigenvalues λ_1, λ_2, and λ_3 with $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$ in scale s
   (c) Computing three vessel-likelihoods in scale s at voxel $\mathbf{x}$:

$$\mathcal{F}(\mathbf{x}, s) = \left[1 - e^{-\frac{1}{2}(\frac{\lambda_2}{\alpha\lambda_3})^2}\right] e^{-\frac{1}{2|\lambda_2\lambda_3|}(\frac{\lambda_1}{\beta})^2} \left[1 - e^{-\frac{\lambda_1^2+\lambda_2^2+\lambda_3^2}{2\gamma^2}}\right] \tag{2}$$

$$\mathcal{S}(\mathbf{x}, s) = \begin{cases} \sigma^2|\lambda_3| \left(\frac{\lambda_2}{\lambda_3}\right)^{\xi} \left(1 + \frac{\lambda_1}{|\lambda_2|}\right)^{\tau}, & \lambda_3 \leq \lambda_2 \leq \lambda_1 \leq 0 \\ \sigma^2|\lambda_3| \left(\frac{\lambda_2}{\lambda_3}\right)^{\xi} \left(1 - \rho\frac{\lambda_1}{|\lambda_2|}\right)^{\tau}, & \lambda_3 \leq \lambda_2 \leq 0 \leq \lambda_1 \leq \frac{|\lambda_2|}{\rho} \end{cases} \tag{3}$$

$$\mathcal{L}(\mathbf{x}, s) = \sigma^{\eta} \left|\frac{\lambda_2 + \lambda_3}{2}\right| \tag{4}$$

   (d) Maximizing vessel-likelihoods at voxel $\mathbf{x}$ across scale s:

$$\mathcal{F}(\mathbf{x}) = \max_s \mathcal{F}(\mathbf{x}, s), \quad \mathcal{S}(\mathbf{x}) = \max_s \mathcal{S}(\mathbf{x}, s), \quad \mathcal{L}(\mathbf{x}) = \max_s \mathcal{L}(\mathbf{x}, s) \tag{5}$$

3. Compute statistical features (*i.e.*, maximum, minimum, mean, median and standard deviation) based on the computed three vessel-likelihoods for all the voxels in the candidate.

True PE candidates have much higher vessel-likelihoods than those generated outside of vessels (*e.g.*, lymph false positives). We plot the training and test detection performance of our classification algorithm in Fig. 4 with these vesselness features in addition to the basic features. It clearly shows the improvement of around 20% increase on sensitivity at 5 false positives when vesselness features are included.

5 Vein/Artery characteristic features

We have nearly 40% false positives generated in veins. In CTPA images, we found that the veins have different intensity distribution and texture patterns from arteries. Furthermore, those false positives within arteries were generated mainly due to poorly mixed contrast material, demonstrating distinct intensity and texture patterns. Therefore, we design a set of features to capture the intensity and texture patterns of veins and arteries to eliminate false detections from veins. To do so, we extract a segment of vessel from a given candidate and then compute a set of features based on the extracted vessel segment. While taking advantage of the benefits of vessel segmentation, our approach is to only segment the relevant areas around the candidate, thereby avoiding an expensive complete segmentation of the complete pulmonary vessel tree.

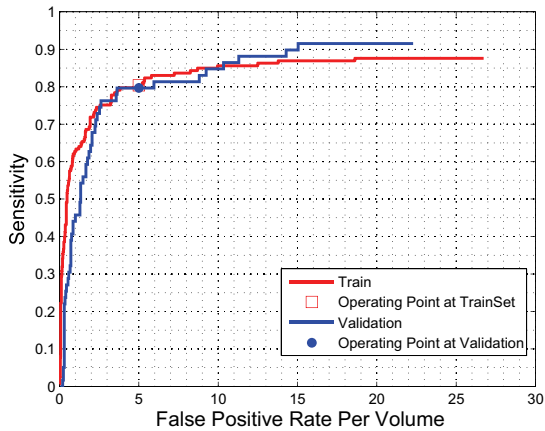


Fig. 4. ROC plot using vesselness features in addition to local candidate features.

5.1 Vessel segment extraction

Each PE candidate generated by our CG is a connected component, consisting of a set of voxels. In this step, we first extract the candidate boundary surface and then obtain the vessel segment with a modified Dijkstra's algorithm. The PE candidate boundary surface is extracted by checking the neighbors of each voxel of the candidate. If any of its neighbors does not belong to the candidate, then the voxel is on the boundary surface and is recorded. The vessel segment containing the candidate component is extracted by a graph-searching based iterative region-growing process by taking the extracted boundary points of the PE candidate as initial seeds. This process creates a minimum cumulative cost path map. The cost is set to the Euclidean distance from the boundary of the candidate component. The growth occurs from the voxel with the lowest cumulative distance cost on the boundary of the growing region. In the growing process, only those voxels with intensity value above a threshold (100 HU) are considered, so that the growth is along the contrast enhanced vessels. As a result, the extracted region grows in layers (equal distance to the boundary of the candidate component) and stops if the vessel length is larger than a given value (30 mm in our case).

5.2 Vessel feature computation

Once the vessel segment is extracted, similarly we can compute the basic intensity distribution features (mean, variance, skewness and kurtosis) and texture features based on gray level co-occurrence matrices from the extracted vessel segment as described in Section 3. In addition, we compute wavelet-based texture features. In the case of volumetric images, the discrete wavelet decomposition is obtained by applying a pair of

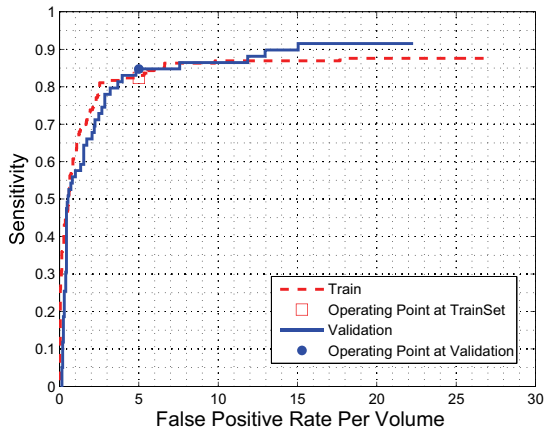


Fig. 5. ROC plot using the entire set of features.

wavelet filters: a lowpass filter h and a highpass filter g, along the x, y, and z direction of the volumetric image. In a one-level decomposition of the 3D discrete wavelet transform (DWT), the particular pair of filters h and g corresponds to a particular type of wavelet used and $\downarrow_x 2$ stands for downsampling by 2 along x direction, generating nine subvolumes from the input volumetric image. For example, subvolume LLL corresponds to the lowest frequencies and subvolume LHH gives the low frequency along x direction and high frequencies along y and z directions. A multi-level DWT decomposition is obtained by repeating the same procedure to the subvolume LLL until the desired level is reached. In our application, we used Daubechies 8-tap filters and a 2-level decomposition of 3D DWT to compute the wavelet subvolumes. Then for each subvolume, the energy feature is calculated from its wavelet coefficients,

$$\mathcal{W} = \frac{1}{N_x N_y N_z} \sum_x \sum_y \sum_z |w(x,y,z)| \tag{6}$$

where N_x, N_y, and N_z are dimensions of the subvolume and $w(...)$ is a wavelet coefficient within the subvolume. For each candidate, we have nine energy features computed from the nine subvolumes.

We validate the third set of features by incrementally including them into our training process. Hence the final set of features include co-occurrence based features in Section 3 and features described in Sections 4 and 5. The final system achieves 85% test sensitivity at 5 false positives per volume.

6 Conclusion

We have developed three sets of new local characteristic features for eliminating false positives in automated PE detection. The developed features have been fully integrated

into our PE CAD system and validated on over 130 cases, showing that the PE CAD system performance has been incrementally improved by more than 20% at the ROC operating point with our second and third set of the features, respectively, in our multiple instance learning framework. Our approach has a set of distinguished features, requiring no vessel segmentation, reporting any detection incrementally in real time, and detecting both acute and chronic pulmonary emboli, achieving a sensitivity of 85% at 5 false positives, resulting in a clinically usable PE CAD system, which has been deployed in many clinical sites around the world.

References

1. Y. Masutani, H. MacMahon, and K. Doi, "Computerized detection of pulmonary embolism in spiral CT angiography based on volumetric image analysis," *IEEE Transactions on Medical Imaging*, vol. 21, no. 12, pp. 1517–1523, 2002.
2. M. Quist, H. Bouma, C. van Kuijk, O. van Delden, and F. Gerritsen, "Computer aided detection of pulmonary embolism on multi-detector CT," in *RSNA*, Chicago, USA, Nov 2004.
3. C. Zhou, H. Chan, S. Patel, P. Cascade, B. Sahiner, L. Hadjiiski, and E. Kazerooni, "Preliminary investigation of computer-aided detection of pulmonary embolism in three-dimensional computed tomography pulmonary angiography images," *Academic Radiology*, vol. 12, no. 6, pp. 782–792, 2005.
4. S. Park, C. Bajaj, G. Gladish, and D. Cody, "Automatic pulmonary embolos detection and visualization," poster at http://www.ices.utexas.edu/∼smpark, 2004.
5. J. Liang, M. Wolf, and M. Salganicoff, "A fast toboggan-based method for automatic detection and segmentation of pulmonary embolism in ct angiography," in *The Eighth International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2005)*.
6. J. Liang and J. Bi, "Computer aided detection of pulmonary embolism with tobogganing and multiple instance classification in CT pulmonary angiography," in *Proceedings of the 20th International Conference on Information Processing in Medical Imaging, 20th International Conference, IPMI 2007*, Kerkrade, The Netherlands, 2007, pp. 630–641.
7. J. Bi and J. Liang, "Multiple instance learning of pulmonary embolism detection with geodesic distance along vascular structure," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'07),*, Minneapolis, Minnesota, USA, 2007.
8. R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3, no. 6, pp. 610–621, 1973.
9. A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *Medical Image Computing and Computer-Assisted Intervention*, 1998, pp. 130–137.
10. Y. Sato, S. Nakajima, H. Atsumi, T. Koller, G. Gerig, S. Yoshida, and R. Kikinis, "3d multiscale line filter for segmentation and visualization of curvilinear structures in medical images," in *CVRMed-MRCAS'97: Proceedings of the First Joint Conference on Computer Vision, Virtual Reality and Robotics in Medicine and Medial Robotics and Computer-Assisted Surgery*. London, UK: Springer-Verlag, 1997, pp. 213–222.
11. C. Lorenz, I.-C. Carlsen, T. M. Buzug, C. Fassnacht, and J. Weese, "Multi-scale line segmentation with automatic estimation of width, contrast and tangential direction in 2d and 3d medical images," in *CVRMed-MRCAS'97: Proceedings of the First Joint Conference on Computer Vision, Virtual Reality and Robotics in Medicine and Medial Robotics and Computer-Assisted Surgery*. London, UK: Springer-Verlag, 1997, pp. 233–242.