



Computer Aided Diagnosis for Pulmonary Embolism Nodule Detection

¹Pragati D. Pawar, ²Sanjay L. Badjate

¹Assistant Professor, ²Principal

¹Electronics & Telecommunication Engineering Department,

¹Jawaharlal Darda Institute of Engineering and Technology, Yavatmal, Maharashtra, India

Abstract: Computer-aided detection and diagnosis systems have been adopted widely to improve the diagnosis performance by detecting and analyzing the lung diseases. The pulmonary embolism is considered as a fatal condition related to lung where the blood clot cause blockage to the lung arteries and this condition can cause death to the patient. Early detection of blood clot can help to diagnose the pulmonary embolism. In order to detect the PE, lung segmentation and nodule detection is the main task for any CAD system. Several approaches have been introduced to perform the segmentation but the accuracy and false positives of segmentation remains a challenging task in this field. Thus, we focus on the lung segmentation and nodule detection using computer vision approach for PE detection and developed Maximum Likelihood Probabilistic model for Pulmonary Embolism nodule detection (ML-PPED). According to the proposed approach, first of all we extract the lungs regions i.e. left and right lung regions followed by segmentation and finally a maximum likelihood based probabilistic model is developed to detect the lung nodules. The experimental analysis shows that the proposed approach improved the segmentation performance when compared with the existing techniques.

Index Terms - Pulmonary embolism, nodule detection, computer vision, probabilistic likelihood maximization

I. INTRODUCTION

Pulmonary embolism (PE) is a life-threatening fatal disease and also a leading cause of human death. This condition occurs due to the extrication of blood clot (this blood clot is formed in leg and known as thrombus or embolus) from deep vein thrombosis (DVT) and blocks the pulmonary artery circulation. The smaller blood clot can cause moderate bleeding inside the lung, moreover excessive blood clot can cause the more blockage of the artery and patient can die due to the no respirational activity. The main symptoms of PE are manifested as chest pain, shortness of breath, orthopnoea, and coughing etc [1]. PE is one of the most chronic disease, a study shows that this disease affects 500,000 to 600,000 people per year and causes 200,000 to 300,000 deaths per year in United States [2]. Another study presented in [3] shows that 100 new cases/100,000 are estimated per year in Italy [3]. According to the study presented in [4], the PE disease will become the third most chronic disease by 2030.

The American Heart Association categorizes the PE into three different categories such as: low-risk, intermediate risk, and high risk. High risk PE is described as the presence of continuous hypotension whereas intermediate-risk PE is defined in normotensive patients where strain is caused in right ventricular (RV) by echocardiogram, identified using computed tomography (CT) scan or cardiac biomarkers. The high risk PE is also known as Acute Pulmonary Embolism which has the higher risk and threats to the human. Thus, the early-identification of PE can be useful for efficient diagnosis.

In the medical field, Computer based applications has gained huge attraction from research community. These applications provide algorithmic solution to interpret the analysis from medical data which is acquired in the form of medical images or bio-medical signals. In this case of PE, the detection of PE and finding the location is considered as a challenging task. In order to overcome this issue, computer vision based schemes are introduced to solve the detection and diagnosis problem. The computer vision based solutions are divided into two groups as computer-aided detection (CAD) and the computer-aided diagnosis (CAD-x) which are based on the principle of image processing. CAD/CAD-x based applications are widely adopted for detection and diagnosis of several health related issues such as pulmonary nodule detection [5].

In computer vision based medical application, Computed tomography imaging is considered as a promising technique which is widely adopted for diagnosis purpose. The CT X-ray images are widely used in PE diagnosis, these images are called as computed tomography angiography (CTA) images. These images efficiently display the lesions in blood vessels which helps to extract several information. In CT images of PE, the color is dissolved in blood and improves the brightness at the particular location however, the PE region remains dark because the color does not dissolve to embolism. Thus, it becomes a challenging task for radiologists to identify the dark spots and their corresponding embolism, moreover, it consumes more time and also results are unreliable because of false recognition of different masses as PE. In order to improve the PE image analysis, several techniques based on CAD systems are introduced using CTA images. Recently, numerous techniques are present to identify and detect the PE based on CAD systems. Generally, the PE and tissues has the same intensity hence, identifying the PE from tissue is

a difficult task. Limited works are presented to discriminate the tissue and PE. In order to achieve the accurate detection of PE, Zkan et al. [6] developed a new approach for lung vessel segmentation.

Pichon et al. [7] also presented PE detection approach where lung vessel segmentation is introduced using morphology scheme and later, a vessel tree is constructed to distinguish the contrast-enhanced blood. Choi et al. [8] developed feature extraction based scheme for pulmonary nodule detection where lung vessel segmentation is presented and later feature extraction process is implemented and finally, SVM (Support Vector Machine) classifier is used for classifying the nodules. This method of Computed tomography pulmonary angiography (CTPA) shows a significant improvement detection of the occurrence of PE [9]. With the help of CTPA, a study was conducted known as PLOPED (Prospective Investigation of Pulmonary Embolism Diagnosis) which shows that the specificity and sensitivity is obtained as 96% and 83% [10]. Dogan et al. [9] presented a study which shows that the sensitivity is obtained as 83-100% and specificity of CTPA scans is obtained as 89-96%. This study shows that the existing techniques may cause misdiagnoses which can be potentially life-threatening.

Wu et al. [11] developed computed tomography imaging approach for PE detection. CTPA imaging is acquired and iodine distribution is also identified which helps to detect the perfusion defects. Later, iodine densities of perfusion defect and normal lung are measured to identify the blood flow and detection of perfusion defects. Tajbakhsh et al. [12] discussed about the PE detection using CAD based CTPA. Based on the previous studies, authors demonstrated that conventional CAD based PE detection systems suffer from the false positive detection which causes unnecessary burden for radiologists. In order to mitigate the issue of false positive detection, this work introduces convolution neural network based approach where a new model is developed to represent the 3D vessel-aligned image as a multi-planer image of emboli. According to this work, the aligned images are processed through the CNN training process for detection. This representation provides summarized information about emboli in 2 image channels, it helps to align the embolus according to the orientation of affected vessel and also, it helps for data augmentation for training.

The perfusion defect plays important role in PE detection. Recently, Cid et al. [13] presented graph-model approach for to extract the significant information about perfusion from the several region of lung. This study shows that the previous works focus on the lung region analysis whereas this approach focused on the complete pulmonary circulatory and pulmonary hypertension. To apply this approach, Dual Energy Computed Tomography (DECT) images are considered where weighted graphs are constructed for encoding the network energy distribution.

Generally, computed tomography based approaches are well-studied in this field of PE detection, similarly, Nagle et al. [14] presented magnetic resonance angiography (MRA) for PE diagnosis as an alternative of computed tomography. However, the performance of these methods depends on several characteristics such as data acquisition, and artifacts during acquisition. In this study, authors identify that most of the radiologists are comfortable with the CT imaging when compared with the MRA. Hence, in this work authors presented data acquisition details, image analysis and some other key aspects to improve the MRA image analysis which can be helpful for radiologists to improve the PE detection.

According to the study presented in [15], the Computed tomography pulmonary angiography (CTPA) imaging systems have significant impact on the PE analysis and detection. The CTPA helps to evaluate the functionality of right ventricular (RV) dysfunction and right-to-left ventricular (RV/LV) diameter ratio which are related to PE mortality. Moreover, CTPA helps to identify the abnormal lung parenchymal features such as lung disease, infraction and effusions etc. Thus authors have suggested a postulate which shows the combining aforementioned ancillary finding with PE detection can help to improve the PE detection. In order to overcome these issue, we present a novel approach for PE detection using CT images.

II. PROPOSED MODEL

In this section we present the proposed solution for PE detection from CT images. According to proposed **ML-PPED** model, first of all we present lung segmentation approach to extract the left and right parts of the lung from 3D image. The proposed solution of lung extraction is applied into three main stages which are as follows: (a) lung localization, (b) discriminating the left and right lung and (c) boundary detection and enhancement. Later, in the second phase, we present the lung vessel segmentation to detect the PE. Figure 1 shows the process of nodule lung region extraction and segmentation. Initially, the input images are considered in the DICOM format where we apply shape identification approach based on the edge detection where we consider the identified shape for the lung segmentation but due to poor shape identification, the segmentation performance degrades. Thus we apply adaptive thresholding approach which provides the complete shape of the lung. Based on this shape, we extract left and right region and the final segmented output is achieved which is further used for detecting the nodules.

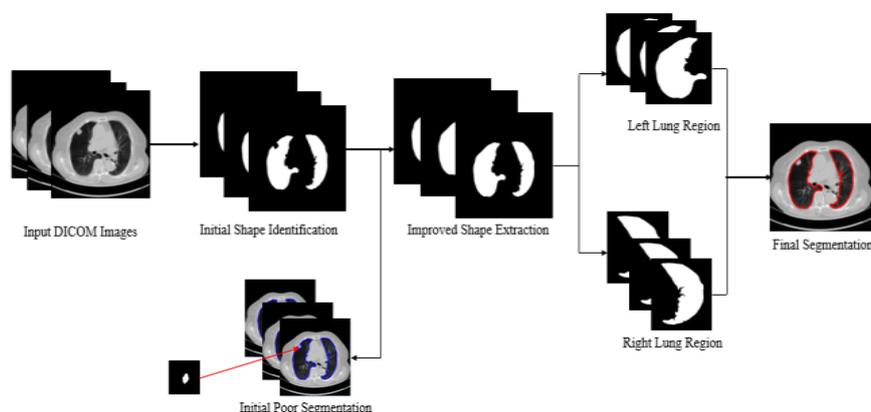


Fig.1. Lung region detection and segmentation process

III. LUNG SEGMENTATION

In this subsection, we describe the proposed solution for efficient lung segmentation to improve the PE detection process.

3.1 Lung Extraction

The conventional methods of lung segmentation use threshold based approach for detection and localization the lung but due to complex structure and artifacts during image acquisition, these methods fails to provide the desired solution for the localization. Thus we present the adaptive threshold selection method. This threshold helps to identify the small variation in the tissue density. Here, it is assumed that the acquired image has only two voxels such as: *body voxel* these voxels are located in the dense body wall and chest structure, another is called as *non-body voxels* which are surrounding the subject's body. The proposed adaptive threshold separates the body and non-body voxels to localize the lung. This process is performed using iterative process where T_i is the threshold at i^{th} iteration. In order to select the new threshold, we apply T_i and separate the body and non-body voxels. Let φ_{bv} denotes the *body voxel* and φ_{nb} denotes the *non-body voxel*. With the help of these voxels, the threshold for the current iteration can be computed as:

$$T_{i+1} = \frac{\varphi_{bv} + \varphi_{nb}}{2} \quad (1)$$

This process of threshold selection is repeated until there is not much changes observed in threshold value. However, the initial threshold is selected based on CT number and for pure air and the number total number of voxels within the chest.

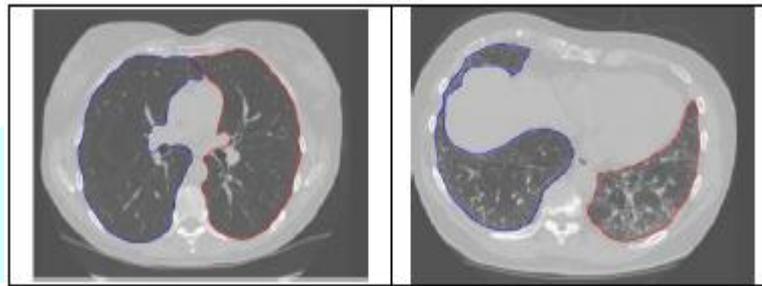


Fig.2. Sample images of lung segmentation using proposed PPED approach.

Once the voxels are identified, we apply 3D connected component analysis where background region is removed by identifying the connected components which are connected to the boundary of the still image. Moreover, the smaller and disconnected regions are also discarded and finally, we keep the biggest region in the considered volume. Thus the unwanted smaller volumes are removed from the processed image. However, at this phase, some internal cavities may remain untouched hence image filling is applied to obtain the complete shape of the lung.

After localizing the lung regions, the left and right lung discrimination is the main task where anterior and posterior junction can be obtained between left and right lung. However, this junction can be very thin and the conventional thresholding methods fail to achieve the junction positions. Here, our main aim is to find the junction between lung parts to separate the right and left lungs. In order to find the junction position, we apply a dynamic programming where a search region is considered on a 2-D slice and it is propagated through all the slices. Here, we apply morphological erosion to distinguish the region of right and left lungs and later dilation is applied to restore the original shape of lung without connecting the removed parts. This process is performed using n -fold erosion, given as:

$$S = Z \ominus nB_4 \quad (2)$$

Where Z denotes the set of lung slices, \ominus represents the binary morphological operation, B_4 denotes the connected binary elements and n is used for scaling term.

After achieving the left and right region, erosion is applied to reconstruct the S using conditional dilation which is performed iteratively. This conditional dilation process is given as:

$$C^{i+1} = C^i \cup \{ \{p\} \oplus B_4 \} \quad (3)$$

Where \oplus is the morphological dilation operation, $p \in C^i \cap Z$, $C^0 = S$ used as initial condition for dilation and later it is updated iteratively. This dilation process provides a reliable solution to discriminate two lungs. At this stage, we achieve left and right lung parts. Figure 2 shows a sample outcome of this process of lung localization. In next phase we focus on the segmentation of pulmonary embolism.

3.2 Segmentation

In the field of medical imaging, the 3D filtering approaches play an important role to enhance the structure of medical data such as vessels, nodules and tissue boundaries. However, due to complex structure of lung, the lung segmentation and nodule detection becomes a challenging task in this field. Let us consider that a pixel value of lung is corresponding to the i^{th} segment (These segments are obtained using fuzzy clustering and each segment represents a class) which has the probability density as $p(x|\theta_i)$ where θ_i denotes the set of parameters for density function. The probability density of a pixel can be expressed as:

$$p(x) = \sum_{i=1}^{\mathbb{S}} p(x|\theta_i)\alpha_i \quad (4)$$

Where \mathbb{S} denotes the total number of obtained segments, α_i denotes the prior probability of i^{th} segment. Let us consider a set of image pixel is given as $X = \{x_k, k = 1, \dots, n\}$. The segmentation task can be achieved inverse problem of parameter estimation as $\Theta = \{\theta_i, i = 1, \dots, \mathbb{S}\}$ i.e. the parameters estimation of each segment can provide us the segmented outcome of the image.

We assume that the pixels values are independent and the maximization likelihood can helps to obtain the set of parameters of the current density function. This function can be written as:

$$p(X|\Theta) = \prod_{k=1}^n \sum_{i=1}^{\mathbb{S}} p(x_k|\theta_i)\alpha_i \quad (5)$$

In this process, we represent each image segment with the help of Gaussian distribution given as $(p(x|\theta_i) - G(\mu_i, \sigma_i))$. Based on this distribution, we can compute the expected probability distribution of new pixel of the next segment of cluster. This can be computed as:

$$p(i|x_k, \Theta) = \frac{\alpha_i p(x_k|\theta_i)}{\sum_{j=1}^{\mathbb{S}} \alpha_j p(x_k|\theta_j)} \quad (6)$$

The eq. (5) gives the probability of next pixel, similarly, we compute the maximized likelihood of the complete data to estimate the parameters as:

$$\begin{aligned} \alpha_i &= \frac{1}{n} \sum_{k=1}^n p(i|x_k, \Theta) \\ \mu_i &= \frac{1}{n} \frac{\sum_{k=1}^n p(i|x_k, \Theta)}{\alpha_i} \\ \sigma_i^2 &= \frac{1}{n} \frac{\sum_{k=1}^n p(i|x_k, \Theta)(x_k - \mu_i)^2}{\alpha_i} \end{aligned} \quad (7)$$

In this process, the parameters are computed for each iteration and thee likelihood also increases for eachiteaartion. This process is repeated until the overall process reaches to the convergence. Finally, *maximum a posterior* is applied for all pixel values to obtain the segment labels as:

$$L_k = \underset{i}{\operatorname{argmax}} p(i|x_k, \Theta), \forall x_k \quad (8)$$

In order to achieve the optimal filter, we consider the lung vessel size and similar size filter response is considered for analysis. This can be obtained by computing the second order derivative of image at scale s which is given as:

$$\nabla^2 I(\vec{r}, s) = s^\tau \begin{bmatrix} I_{xx}(\vec{r}) & I_{xy}(\vec{r}) & I_{xz}(\vec{r}) \\ I_{yx}(\vec{r}) & I_{yy}(\vec{r}) & I_{yz}(\vec{r}) \\ I_{zx}(\vec{r}) & I_{zy}(\vec{r}) & I_{zz}(\vec{r}) \end{bmatrix} \quad (9)$$

τ denotes the normalization parameters.

In order to perform the PE segmentation, we integrate all the segmented vessels along with their size information. To integrate the segments, we consider that the largest vessels are \mathbb{S}_{max} and smallest vessels are extracted as \mathbb{S}_{min} and the process of integration also performed from largest to smallest segment. The segment integration can be represented into two stages as:

(a) Initial stage integration at \mathbb{S}_{max} :

$$F(x, y, z) = \begin{cases} \mathbb{S}_{max}, & \text{if } \mathcal{S}(x, y, z; \mathbb{S}_{max}) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

Where $\mathcal{S}(x, y, z)$ denotes the voxel structure at (x, y, z) ,

(b) if $F(x, y, z) = 0$, then for $\mathbb{S}_k = \mathbb{S}_{max}$ to \mathbb{S}_{mi} the integration can be expressed as:

$$F(x, y, z) = \begin{cases} \mathbb{S}_k, & \text{if } \mathcal{S}(x, y, z; \mathbb{S}_k) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

IV. RESULTS AND DISCUSSION

In this section we present the performance analysis using proposed ML-PPED approach of pulmonary nodule segmentation. The complete experimental analysis is carried out using MATLAB tool using DICOM images. In this work, we have obtained publically amiable data and their corresponding ground truth [16]. For each patient, the CTA image slices are stored in DICOM format and ground truth is stored in .MAT format. The ground truth contains foreground and background voxels. This dataset include, 17 male and 18 female patients for data acquisition where these patients are between 24-82 years. Total 8792 slices are present in this dataset where Patient24 and Patient32 have no PE-clots. Figure 2 shows some sample images obtained from this dataset.

Similarly, we have considered LIDC standard image dataset to evaluate the performance of proposed ML-PPED lung nodule segmentation approach. This dataset contains both low and high does CT images and these images are breath—held CT images with the size of 512x512. In this dataset, the in-plane pixel size varies between 0.5 and 0.8 mm/pixel. During acquisition, the kVp range of these data was considered as 120-140 with 120 and the mA range was 30-634 with the 145.1 standard deviation. The complete process of proposed approach is depicted in figure 3 where lung nodule region identification and the lung nodule detection outcomes are depicted.

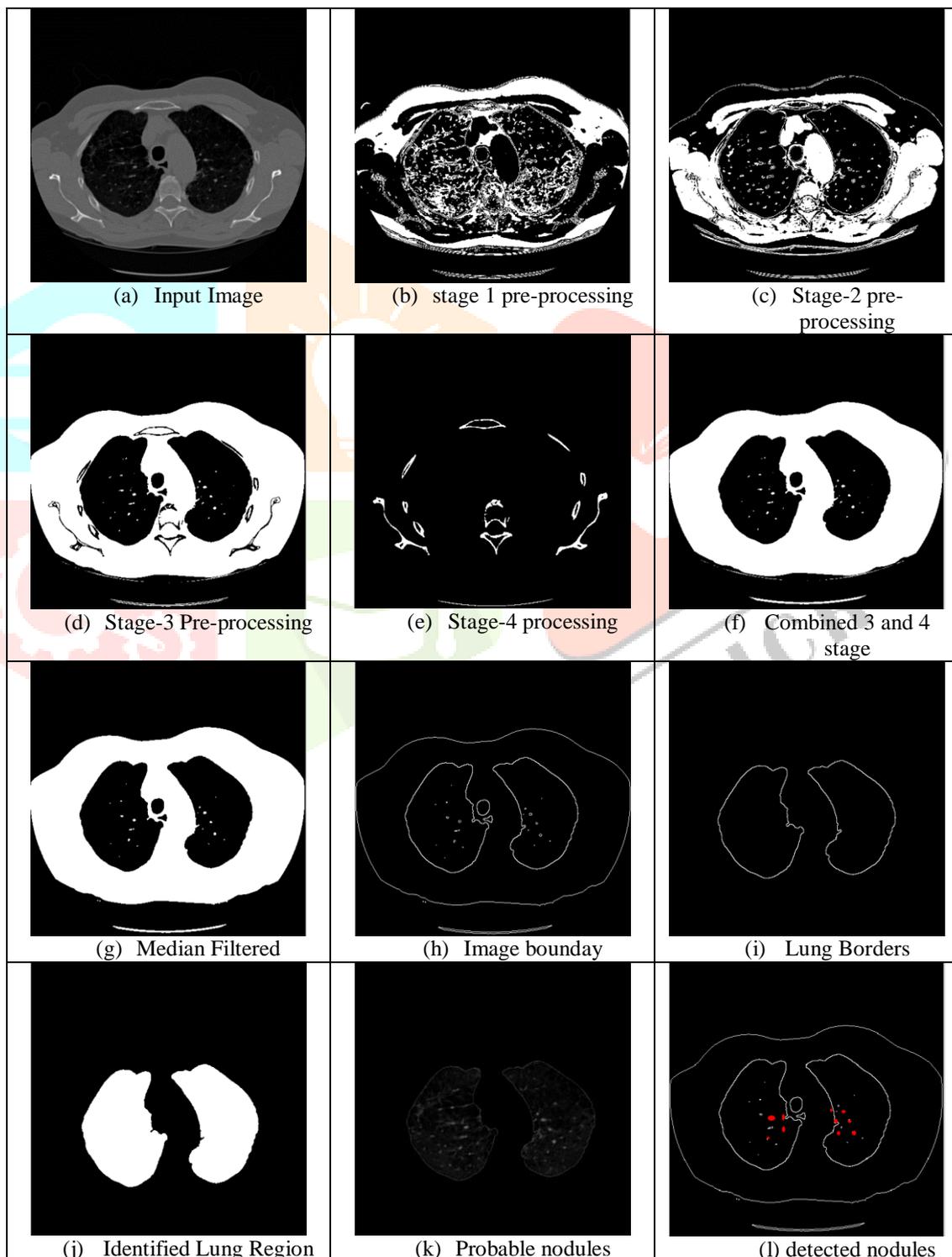


Fig.3. Lung nodule detection process.

V. PERFORMANCE MEASUREMENT METRICS

The performance of proposed approach is compared with other approaches like distance regularized level set (DRLSE) [17] and a weight based level set method [18]. Dice similarity co-efficient (DSC) of two level set based algorithms DRLSE [17], WLSE [18] and proposed approach (ML-PPED) are shown in Table 1. It can be stated from Table 2 that the proposed method outperforms other level set techniques even without using any learning.

Table 1. DSC performance comparison

Technique	Isolated Nodules		Juxtaplural Nodule	
	Min. DSC	Max. DSC	Min. DSC	Max. DSC
DRLSE [17]	0.33	0.93	0.71	0.91
WLSE [18]	0.78	0.94	0.79	0.87
SDLS [19]	0.87	0.94	0.81	0.93
Proposed (ML-PPED)	0.89	0.95	0.83	0.94

VI. CONCLUSION

In this work, we have focused on pulmonary embolism detection and segmentation. Generally, this occurs in the lung artery where a blood clot blocks the pulmonary circulatory. This issue of blood clot can lead towards the death of the patient, hence, early detection of this disease is an important task for medical researchers. However, due to complex structure of lung, the identification of these types of nodules is a crucial task. In this work, we present a novel approach for pulmonary nodule detection and segmentation using computer vision approach. According to the proposed approach, first of all, we identify the region of left and right lungs and their positions where we also remove the small connected components and background connected components to achieve the complete structure of lung. Later, we present probabilistic likelihood maximization approach with Gaussian distribution model where pixel probabilities are computed and updated in an iterative process where we achieve different types of vessels which are later integrated from larger segment to smaller segment to obtain the segmented region. This approach is implemented on publically available CTA image database and the performance of proposed approach is compared with the ground truth of the database.

REFERENCES

- [1] Stein, P. D., Beemath, A., Matta, F., Weg, J. G., Yusen, R. D., Hales, C. A., & Buckley, J. D. (2007). Clinical characteristics of patients with acute pulmonary embolism: data from PIOPED II. *The American journal of medicine*, 120(10), 871-879.
- [2] Marshall PS, Mathews KS, Siegel MD. Diagnosis and management of life-threatening pulmonary embolism. *J Intensive Care Med* 2011;26:275–294
- [3] Celi, A., Marconi, L., Villari, L., & Palla, A. (2016). The diagnosis of pulmonary embolism. *Monaldi Archives for Chest Disease*, 71(2).
- [4] Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Medicine*. 2006;3(11):e442:2011-2030. DOI: 10.1371/journal.pmed.0030442.
- [5] Setio, A. A. A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C., Van Riel, S. J., ... & van Ginneken, B. (2016). Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. *IEEE transactions on medical imaging*, 35(5), 1160-1169.
- [6] H. zkan, O. Osman, S. ahin, and A. F. Boz, "A novel method for pulmonary embolism detection in CTA images," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 3, pp. 757–766, Dec 2014.
- [7] E. Pichon, C.L. Novak, A.P. Kiraly, D.P. Naidich, A novel method for pulmonary emboli visualization from high-resolution CT images, *Proceedings of the SPIE Medical Imaging 5367* (2004) 161–170
- [8] Choi, W. J., & Choi, T. S. (2014). Automated pulmonary nodule detection based on three-dimensional shape-based feature descriptor. *Computer methods and programs in biomedicine*, 113(1), 37-54.
- [9] Dogan H, de-Roos A, Geleijns J, Huisman MV, Kroft LJM. The role of computed tomography in the diagnosis of acute and chronic pulmonary embolism. *DiagnIntervRadiol*. 2015;21:307–16. community. *ArteriosclerThrombVasc Biol*. 2008;28:370–2.
- [10] Stein PD, Fowler SE, Goodman LR, Gottschalk A, Hales CA, Hull RD, Leeper KV, Popovich J, Quinn DA, Sos TA, Sostman HD, Tapson VF, Wakefield TW, Weg GJ, Woodard PK. Multidetector computed tomography for ac
- [11] Wu, H. W., Cheng, J. J., Li, J. Y., Yin, Y., Hua, J., & Xu, J. R. (2012). Pulmonary embolism detection and characterization through quantitative iodine-based material decomposition images with spectral computed tomography imaging. *Investigative radiology*, 47(1), 85-91.
- [12] Tajbakhsh, N., Gotway, M. B., & Liang, J. (2015, October). Computer-aided pulmonary embolism detection using a novel vessel-aligned multi-planar image representation and convolutional neural networks. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 62-69). Springer, Cham.
- [13] Cid, Y. D., Müller, H., Platon, A., Janssens, J. P., Lador, F., Poletti, P. A., & Depeursinge, A. (2016). A lung graph-model for pulmonary hypertension and pulmonary embolism detection on DECT images. In *Medical Computer Vision and Bayesian and Graphical Models for Biomedical Imaging* (pp. 58-68). Springer, Cham.
- [14] Nagle, S. K., Schiebler, M. L., Replinger, M. D., François, C. J., Vigen, K. K., Yarlaga, R., ... & Reeder, S. B. (2016). Contrast enhanced pulmonary magnetic resonance angiography for pulmonary embolism: building a successful program. *European journal of radiology*, 85(3), 553-563.
- [15] Kumamaru, K. K., Saboo, S. S., Aghayev, A., Cai, P., Quesada, C. G., George, E., ... & Rybicki, F. J. (2016). CT pulmonary angiography-based scoring system to predict the prognosis of acute pulmonary embolism. *Journal of cardiovascular computed tomography*, 10(6), 473-479.

- [16] Masoudi, M., Pourreza, H. R., Saadatmand-Tarzjan, M., Eftekhari, N., Zargar, F. S., & Rad, M. P. (2018). A new dataset of computed-tomography angiography images for computer-aided detection of pulmonary embolism. *Scientific data*, 5.
- [17] C. Li, C. Xu, C. Gui, M. D. Fox, Distance regularized level set evolution and its application to image segmentation, *IEEE Transactions on Image Processing* 19 (12) (2010) 3243.
- [18] A. Khadidos, V. Sanchez, C.-T. Li, Weighted level set evolution based on local edge features for medical image segmentation, *IEEE Transactions on Image Processing* 26 (4) (2017) 1979–1991.
- [19] Roy, R., Chakraborti, T., & Chowdhury, A. S. (2019). A Deep Learning-Shape Driven Level Set Synergism for Pulmonary Nodule Segmentation. *Pattern Recognition Letters*.
- [20] R. Tachibana, S. Kido, Automatic segmentation of pulmonary nodules on ct images by use of nci lung image database consortium, in: *Medical Imaging 2006: Image Processing*, Vol. 6144, International Society for Optics and Photonics, 2006, p. 61440M.
- [21] V. V. Kishore, R. Satyanarayana, Performance evaluation of edge detectors-morphology based roi segmentation and nodule detection from dicom lung images in the noisy environment, in: *Advance Computing Conference (IACC), 2013 IEEE 3rd International*, IEEE, 2013, pp. 1131–1137
- [22] J. Dehmeshki, H. Amin, M. Valdivieso, X. Ye, et al., Segmentation of pulmonary nodules in thoracic ct scans: a region growing approach, *IEEE transactions on medical imaging* 27 (4) (2008) 467–480
- [23] Q. Li, F. Li, K. Doi, Computerized detection of lung nodules in thinsectionct images by use of selective enhancement filters and an automated rule-based classifier, *Academic Radiology* 15 (2) (2008) 165–175.

