DOI: 10.21917/ijivp.2025.0513

LUNG NODULE SEGMENTATION BASED ON LUNG-RANGE-STANDARDIZATION

O Chung-Hyok, Ri Jong-Hyok, Om Chol-Nam

Institute of Information Technology, High-Tech Research and Development Center, Kim Il Sung University, Democratic People's Republic of Korea

Abstract

Checking radiological image is a very toilsome work for radiologists because it requires long time practice and experienced skill. Therefore, many computer-aided diagnosis (CAD) systems have been introduced to cooperate with radiologists and nowadays many CAD systems based on deep learning exceed human experts in diagnosing accuracy. Nowadays, the much of progress has been made in designing architectures. However, peculiar pre-processing method customized for a certain problem can also increase the model accuracy. After checking the LIDC dataset [44], it has been realized that the locations and sizes of lungs were not regularized. Therefore, in this paper, a new pre-processing method (lung-range-standardization) is proposed in order to improve the general accuracy of lung-related diagnosis systems. And the efficiency of the proposed pre-processing method is validated through comparison between the nodule segmentation model trained using our proposed pre-processing method and the nodule segmentation model, which is trained using the prior pre-processing methods. By using lung-range-standardization we could reduce the difference between train loss and test loss in a great deal (from 0.337 to 0.119).

Keywords:

Network Lung Nodule, Convolutional Neural Networks, Lung Cancer

1. INTRODUCTION

Lung cancer is one of the most common and lethal cancers in the world. So far It has been propagated around more than 180 countries and it brought 1.6 million deaths in 2012 and 1.76 million deaths in 2018 [1] [2]. Among men and women, it will be the second cause of death next to the prostate cancer for men and the breast cancer for women [3]. So, lung cancer which accounts more than a quarter of all cancer death is major threats to human health on a worldwide scale [4].

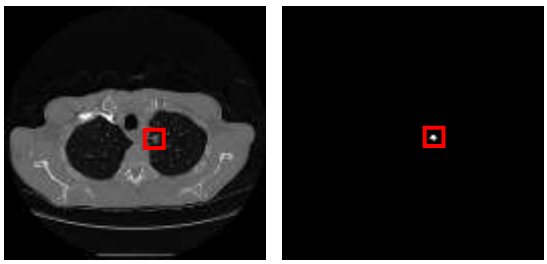
Lung cancer, an abnormal development of cells, is often called pulmonary nodule inside the lungs and the diameter of nodules ranges from 3mm to 30mm. Although, this cancer threats to human life severely, people can control it sufficiently with correct and timely diagnosis. In fact, nowadays, histopathology and molecular biology are widely used tumours pathology diagnosis but it's very annoying and stressful because of such needle biopsy or surgical resection [5].

However, early detection of lung cancer from CT (computed tomography) image of lung decreased the mortality rate by 20% compared to single view radiography [6]. Even if screening programs have provided the appealing chance for better diagnosis, actual nodule detection from radiograph is another challenge.

Therefore, CAD systems have been developed to assist radiologist in reading process and thereby potentially making lung cancer screening more effective. By combining a CAD system to ordinary diagnostic, accuracy of diagnostic performance can be increased [7,8]. And the rapid progress in deep learning and its wide applications has powered the CAD systems [9,10,11]. Especially CNN based deep learning systems have been widely applied in computer vision tasks and the results from them have proven their reliance in object detection and localization in variant images. These successes have been also transformed to different medical imaging problems and showed the state-of-the-art results in various challenges of image processing [12, 13].

Fig.1 shows the examples of a CT image and the segmentation label mask of it. The area highlighted by red rectangles in Fig.1 represents the corresponding areas in both images.

The CNN based CAD systems for lung cancer can be mainly categorized as automatic cancer diagnosis [15, 16], lung segmentation [17], lung nodule segmentation [18], lung nodule detection [19], cancer classification [20] or nodule malignancy assessment [21, 22]. In order to train more smart deep learning model, one must consider about the pre-processing data, designing architecture and establishing loss function.



(a) Lung slice image (b) Nodule segmentation

Fig.1. An example of lung CT image and its nodule segmentation.

Recent CAD systems pay more attention on designing architecture: using 3D convolution [39, 43], multi-views [37], residual block [40] and so on. However, the pre-processing remains the mere elimination of useless information from image such as noise and artefacts. Therefore, in this paper we focus on improving the pre-processing method and establishing loss function. Image pre-processing step is not only the first step but also an important step for CAD systems. In fact, for the visual recognition systems, in the era of deep learning, it seems that image pre-processing step has been neglected a little even though it was regarded as an important step in the era of hand-engineered features because of the learning power of deep networks and the large amount of available training data.

However, dataset size of lung CT images is relatively small compared with other image dataset like Pascal or COCO. The Table.1 represents the analysis of different lung CT image dataset. So significant image pre-processing may be very useful for lung nodule detection systems in order to avoid over-fitting. Generally, in machine learning, data standardization is regarded as very important pre-processing method. In the case of image processing, the standardization often contains histogram and geometrical standardization. Geometrical standardization is very

important because, in general, the sizes and placements of target objects change in query images. Therefore, some face recognition systems use face alignments methods (one of the geometrical standardizations) in order to train more robust and accurate models. However, applying geometrical standardizations requires base things. In the case of face, these base things can be eyes, nose, mouth and ears (often called landmarks) and the bounding box of face. For the lung CT image standardization, we used the minimal rectangle of universal lung region as the base thing.

Table.1. Analysis of datasets

Dataset	The number of CT scans	The number of nodules	Annotation
LIDC-IDRI	1018	36378	Ok
LUNA16	888	13799	Ok
Ali Tianchi	1000	1000	Ok
NSCLS	211	-	Ok
ELCAP	50	-	Ok

The loss function is also important because the minima of it encode solutions to the real-world problem. Sometimes, practical machine learning boils down to understanding the different types of loss functions and which loss function should be applied to which problem. So, much of machine learning is simply the art of turning complicated real-world systems into suitable loss function. In the case of nodules, not only the sizes of lung nodules are relatively very small but also number of them is small. Thus, we used multi-levels loss by generating the labels of different levels using max-pooling in order to achieve our aim to segment both small and large nodules.

This paper consists of three main section: Sect. 2 provides an overview of basic knowledge and concepts that is helpful to understand CAD systems for lung nodule diagnosis and reviews some related works. Sect. 3 describes the proposed pre-processing method, multi-levels loss and a main architecture of the nodule segmentation model. Sect. 4 represents the training process and evaluate effectiveness of proposed pre-processing method on the LIDC.

2. METHOD

2.1 COMPUTED TOMOGRAPHY IMAGING

Computer Tomography is a useful method for visualizing inner substance of an object, which x-ray can pass through. As a result, CT makes most sense for visual diagnosis because it can produce several cross-sections of the body. The X-ray source and detector, facing each other, rotate around the body. During the rotation, several snapshots are taken and are then processed to produce an image. So, sometimes you can see the particular interference figure on CT scan images like Fig.2. Each pixel of CT scan images represents the HU (Hounsfield Unit), which is the quantitative measurement of Radio density of substances. The Table.2 shows the relation between the HU values and the associated substances.

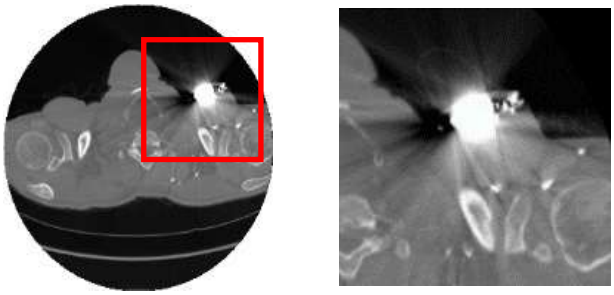


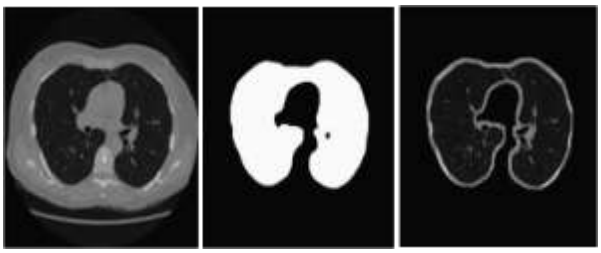
Fig.2. Interference figure of CT image

Table 2. HU values for different substances

Substance	HU
Air	-1000
Lung	-500
Water	0
Blood	30~45
Soft tissue	100~300
bone	700~3000

2.2 COMMON PRE-PROCESSING METHODS AND LOSS FUNCTIONS

Most of pre-processing methods for lung CAD systems include two parts: lung segmentation and HD normalization. In terms of the lung segmentation, most common methods used by scholars are k-means [23], thresholding [24], watershed [25] and clustering [26]. Lung slice image described in Fig.3(a). Lung region is so dark that it can be easily segmented. Therefore, lung region is expressed as mask image (Fig.3(b)). By choosing the part of lung image associated with the lung region mask, the segmented lung image can be obtained as Fig.3(c). In the case of HD normalization, two common methods are dominative: one is to transforms relevant HD range into target range [27] [28], and another method uses mean and standard deviation of HD [29].



(a) Lung image (b) Lung region mask (c) Lung segmentation

Fig.3. An example of lung image pre-processing

And with respect to loss function, the way to set up the loss function depends on the types of proposed problem. In supervised machine learning there are two sub problems of classification and regression. The L_2 loss is commonly used for regression problems. Suppose that x is a collection of data and y is the associated labels then the L_2 loss is expressed as follows.

$$L(x,y) = \| f(x) - y \|_2 = \sum_{i=1}^{n} \sqrt{(f(x_i) - y_i)^2} \quad (1)$$

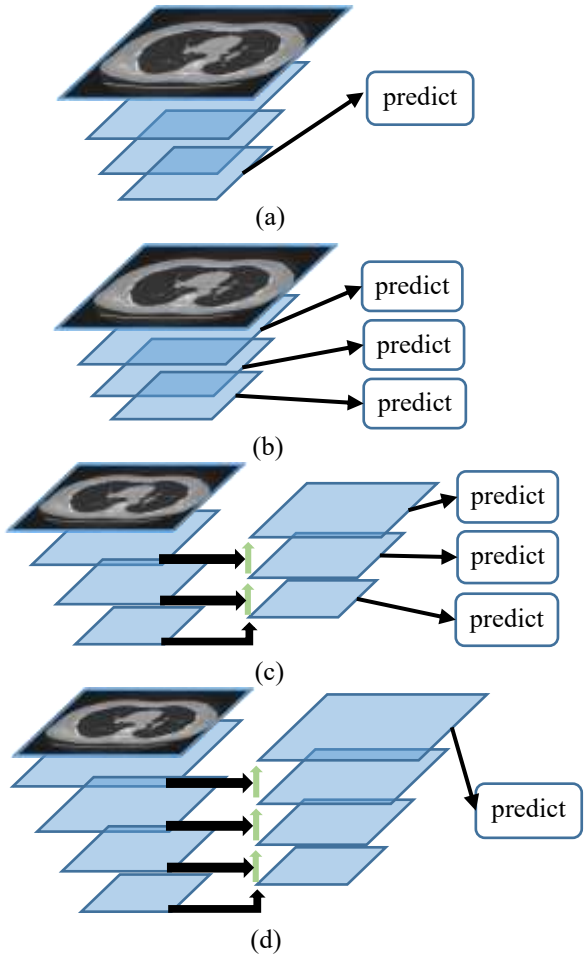


Fig.4. Different kinds of architectures

On the other hand, cross-entropy loss is useful for classification. Cross-entropy is a mathematical method for gauging the distance between two probability distributions. Suppose that $p=(y,1-y)$ is the true data distribution and $q=(y_{\text{pred}},1-y_{\text{pred}})$ is the prediction from machine learning system then the cross-entropy loss is described as follows.

$$L(p,q)=y\log y_{\text{pred}}+(1-y)\log(1-y_{\text{pred}}) \tag{2}$$

These two loss functions were widely used to solve segmentation problems before they alternated with the dice similarity coefficient (DSC) loss function, which is regarded as a suitable loss function for in image segmentation tasks using a U-Net CNN. By minimizing the negative DSC, the model attains the maximal overlap of the predict mask with the ground truth mask [30].

$$\text{DSC}=\frac{2TP}{FN+2TP+FP} \tag{3}$$

where, TP means true positive, FN means false negative and TP means false positive.

2.3 COMMON ARCHITECTURES

Feature pyramids are a basic component in recognition systems for detecting objects at different scales which is a fundamental challenge in computer vision. Featured image pyramid were widely used in the traditional object detection systems that used hand-engineered features [31, 32].

The Fig.4 shows the different kinds of architectures and their prediction methods.

After advent of ConvNet, feature extraction tasks were alternated by ConvNets. However, pyramids were still useful to get more accurate results. Featuring each level of an image pyramids has an advantage of producing a multi-scale feature representation, but it had limitations of increasing process time. For that reason, Fast and Faster- RCNN [33, 34] did not use featured image pyramids (Fig.4(a)).

Deep ConvNets can use feature hierarchies to compute a multi-scale feature representation instead of image pyramids. The SSD (Single Shot Detector) [35] is one of the first attempts at using a ConvNet's pyramidal feature hierarchy (Fig.4(b)). But in this case, SSD cannot use higher-resolution maps of the feature hierarchy so it's hard to detect small objects.

In order to overcome this shortcoming of pyramidal feature hierarchy, FPN (Feature Pyramid Networks) was introduced [36] (Fig.4(c)). FPN architecture combines low-resolution, semantically strong features with high-resolution, semantically weak features. Especially, improved FPN architectures like PRB-FPN [45] can detect both small and large objects very accurately.

U-net [41] is also similar to FPN however, the output of U-net is a mask because it is designed for segmentation systems. Therefor U-net, fully composed with convolutional neural networks (include up-convolution) takes an input image to the output predict mask (Fig.4(d)).

3. PROPOSED METHOD

3.1 PRE-PROCESSING

In this paper, proposed pre-processing pipeline of CT image involves three main steps: lung segmentation, HD normalization and lung-range standardization. For the earlier two steps, we have mentioned briefly about them in the previous section and the lung-range-standardization is our proposed new pre-processing method.

In a word, lung-range is the universal range of all lung regions of a person. In order to help your understanding, we depict the meaning of lung rang in Fig.5. The CT scans of a person's thorax consist of a number of lung slice images (Fig.5(a)). The corresponding lung region masks are represented in Fig.5(b). By joining those lung region masks together, we get the lung-range mask (Fig.5(c)) which can cover all lung regions adequately.

The example images in Fig.6 can express the necessity of lung-range-standardization. The images of the second raw in Fig.6 are the lung slice images in which the bronchus is separated.

The Fig.6 shows the different kinds of placements of lungs in LIDC dataset: a) represents the typical case, b) represents the small-size case, c) represents the rotated case, d) represents the ratio-variant case and e) represents the upside-down case. Because LIDC dataset does not contain plenty of lung images in different cases, a new additional pre-processing that maps the region of lung-range onto the specific region is needed. We call this pre-processing ''lung-range-standardization''.

The Fig.7 shows the examples of lung-range-standardization. As you can see from this Figure, even if the lung region placements are different (right side images of Fig.7(a) and c)), their placements become similar after standardization (right side images of Fig.7(c) and d)).

Lung-range-standardization starts with finding minimum rectangle that spreads over the lung-range (the red rectangles of Fig.7(a) and Fig.7(b). Luckily Open-CV package offers a proper function (implemented as minAreaRect), which evaluate the minimum area rectangle of proposed contour. And then calculate the affine transformation matric that transforms the rectangle points into specific points.

Let's suppose that specific region is a rectangle with size of (n,n) and that the size of minimum rectangle is (w,h), then the specific points are calculated as follows.

$$\begin{pmatrix} SP_{\text{top-left}} \\ SP_{\text{bottom-right}} \end{pmatrix} = \frac{n}{m} \cdot \left(\begin{pmatrix} 0 & 0 \\ w & h \end{pmatrix} + \begin{pmatrix} \frac{m-w}{2} & \frac{m-h}{2} \\ \frac{m-w}{2} & \frac{m-h}{2} \end{pmatrix} \right)$$

$$m = \max(w,h) \quad (4)$$

After that, the transformation matric $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ can also be easily obtained by using Open-CV function (implemented as getAffineTransform). By using the Open-CV function (implemented as warpAffine), which uses A, proposed lung image can be mapped to the standardized lung image.

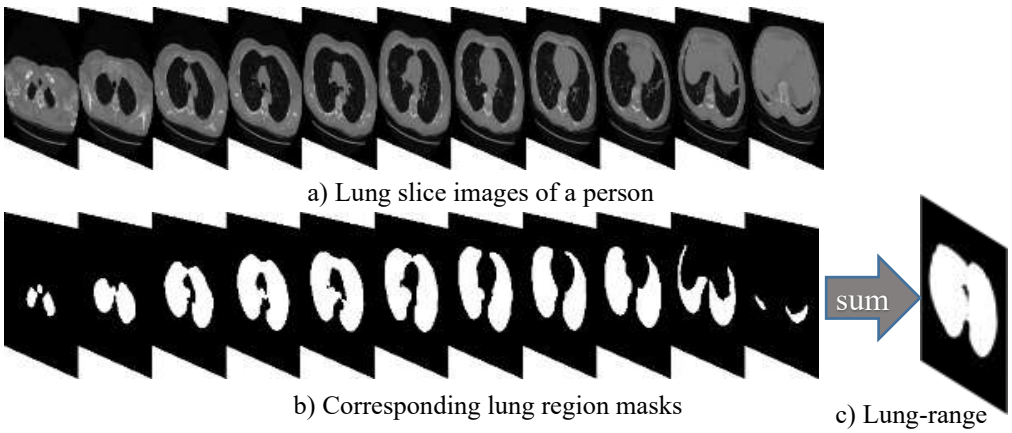
Then how can we get the label of the transformed lung image. One easy way is to transform the original label mask image with A. However, this method was used in our study. The labels of lung nodules are given as the form of the coordinates of nodule border points.

$$B = \{p_1(x_1, y_1), p_2(x_2, y_2), \ldots, p_n(x_n, y_n)\} \quad (5)$$

Then the transformed border points can be calculated as follows.

$$B' = \{p_1'(x_1', y_1'), p_2'(x_2', y_2'), \ldots, p_n'(x_n', y_n')\} \quad (6)$$

$$\begin{pmatrix} x_1' & x_2' & \cdots & x_n' \\ y_1' & y_2' & \cdots & y_n' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \end{pmatrix} \quad (7)$$

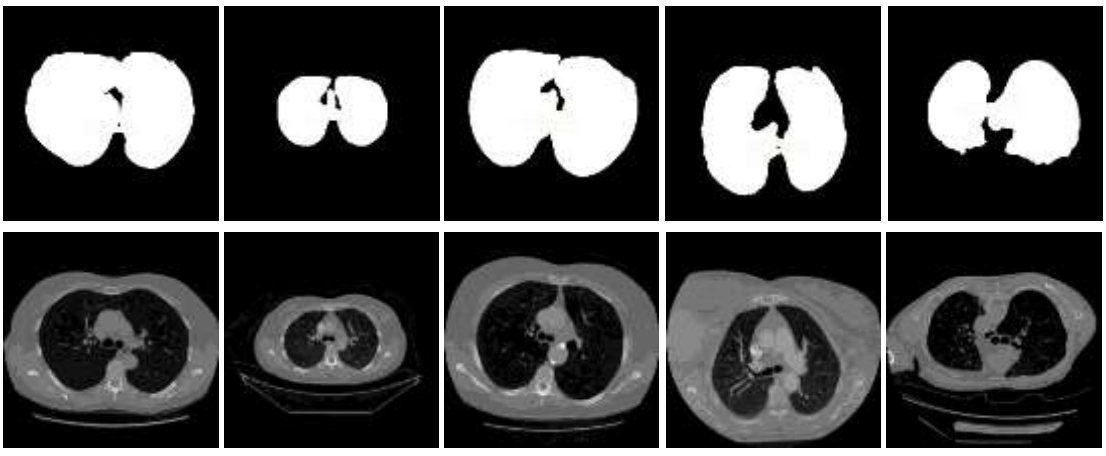


a) Lung slice images of a person

b) Corresponding lung region masks

c) Lung-range

Fig.5. Pictorial definition of lung-range



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

Fig.6. Different lung placements of LIDC dataset

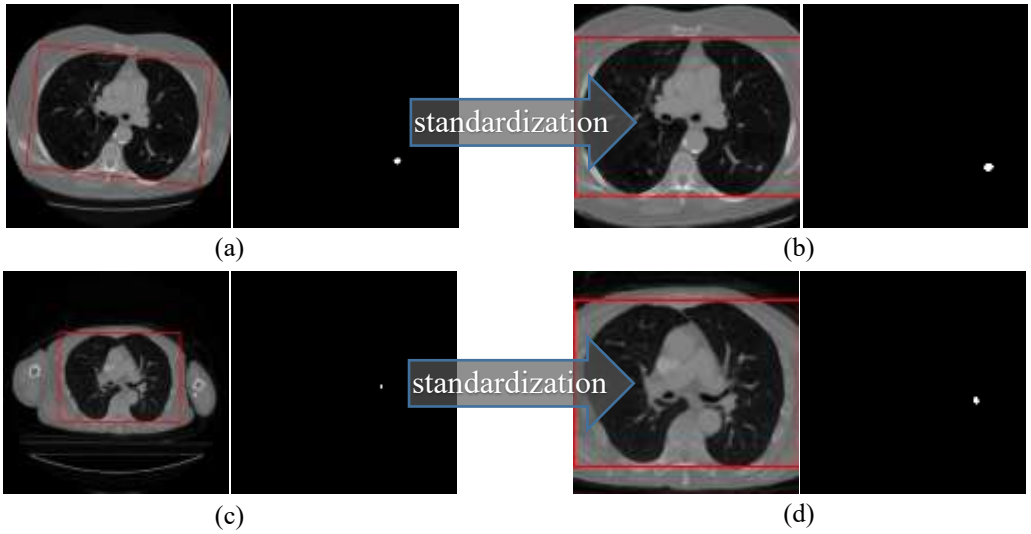
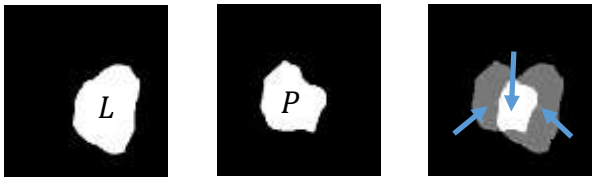


Fig.7. Example results of lung-range-standardization

3.2 MULTI-LEVELS LOSS AND MAIN ARCHITECTURE

Although the DSC loss is suitable for training a segmentation model has a serious problem. That is the model trained using DSC loss is not able to separate objects in an image individually. Therefore, if an image contains both large and small objects, small objects may not be segmented frequently. It can be explained theoretically with the deep analysis of DSC loss. The Fig.8 shows the meaning of DSC.



a) Label b) Prediction c)

Fig.8. Meaning of Dice similarity coefficient

The white area of Fig.8(a) and Fig.8(b) represents the positive points of label and prediction result, and those are expressed as the set L and the set P. And the white area of Fig.8(c) represents the intersection of L and P. It can be denoted as either $P \cap L$ or TP. The grey regions beside the intersection, represents FN and FP (right side is FP and left side is FN). Then, DSC is denoted as following.

$$\mathrm{DSC} = \frac{2TP}{FN + 2TP + FP} = \frac{2TP}{(FN + TP) + (TP + FP)} = \frac{2S(P \cap L)}{S(P) + S(L)} \tag{8}$$

In practice, DSC loss is commonly used as the form of following.

$$\mathrm{DCL} = 1 - \frac{2S(P \cap L) + \mathrm{smooth}}{S(P) + S(L) + \mathrm{smooth}} \tag{9}$$

Here smooth is added to poise the loss value and to avoid dividing by zero, and $S(U)$ denote the size or number of elements of set U. If the predict mask is identical with the label mask, the dice coefficient loss is zero and the larger the intersection is, the smaller the DCL is, under the same size of predicted mask.

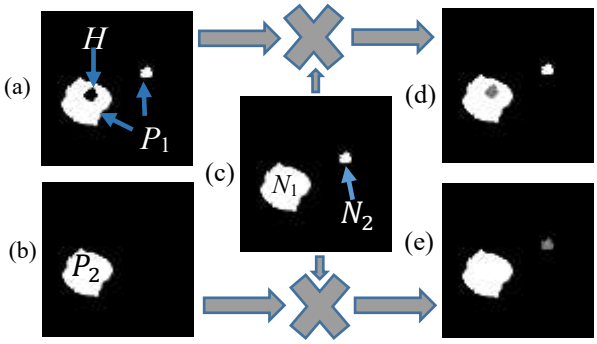


Fig.9. DSCs for different predictions

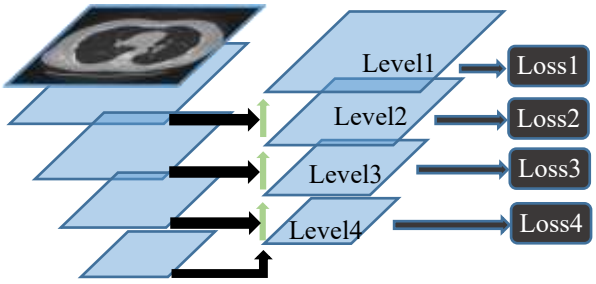


Fig.10. The illustration of multi-levels loss

Then let's suppose that the label mask (Fig.9(c)) contains large and small nodule and two predicted masks are given as Fig.9(a) and Fig.9(b), where $S(H) = S(N_2)$ then $S(P_1) = S(P_2)$. The label mask can be denoted as $N_1 + N_2$ and the first prediction (Fig.8 a)) is denoted as $P_1 = N_1 \cup N_2 - H$ and the second prediction is denoted as P_2 (P_2=N_1). The grey regions of Fig.9(d) and Fig.9(e) represent FP. Thus, the DCL of first prediction is expressed as follows:

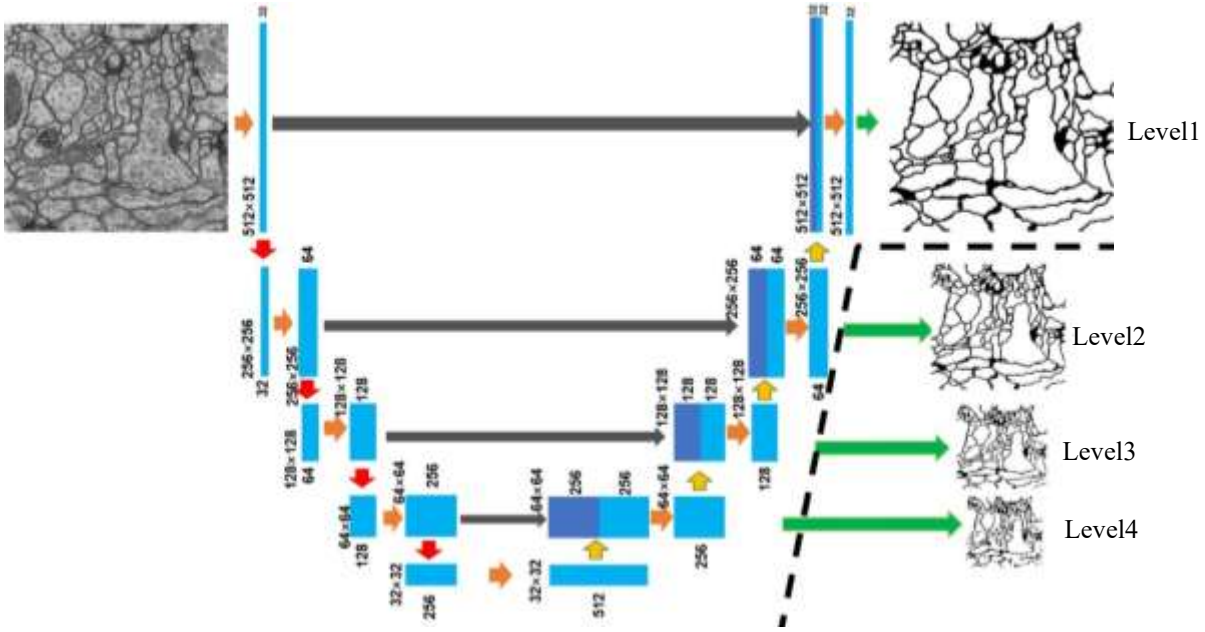


Fig.11. Main architecture of [42]

$$DCL_1 = 1 - \frac{2(S(N_1) - S(H) + S(N_2)) + \text{smooth}}{S(P_1) + S(L) + \text{smooth}} \tag{10}$$

while the of second prediction is expressed as follows.

$$DCL_1 = 1 - \frac{2S(N_1) + \text{smooth}}{S(P_1) + S(L) + \text{smooth}} \tag{11}$$

As can be seen from the above, although, the predictions are different, in both cases the DCL is the same. However, the meaning of both predictions is largely different from each other because first prediction can be regarded to find two nodules but second prediction finds only one nodule. So, this may lead the model to find only large nodules.

In order to overcome this shortcoming, in this paper, multi-levels loss is used. The goal of this idea is to calculate losses at several levels not only one level and to add them up. In fact, SSD [35] also calculate losses at several levels however those losses are not added up because different levels make predicts. More detail, in SSD, deeper levels are responsible for detecting larger objects and shallower levels are responsible for detecting smaller objects. Therefore, it can be regarded that SSD uses multi-levels prediction because the final prediction is made by merging predictions of all levels. The multi-levels loss is illustrated in Fig.10.

In fact, the prepared nodule segmentation labels can be used for the labels of the final level. Then how can get the labels for different levels. One very simple and easy way is to resize label image using interpolation methods(resize-labelling) step by step [42]. Fig.11 and Fig.12 shows the main architecture of [42] and the labels at different levels. As it can see from Fig.12, the label image of level4 has many broken parts because the narrow parts are attenuated after a number of resizing.

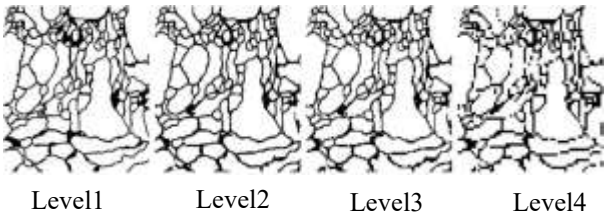


Fig.12. Generated label images of different levels of [42]

In the case of nodule segmentation, by using resize-labelling, small nodules (some nodules are smaller than 5pixels in CT scans) can be lost at deeper levels. The Fig.13 shows the difference between resize-labelling and max-pool-labelling. Original label image (on the left end of Fig.13) is resized by different ways step by step.

In every step, the image size decreases as half of its size. Resized label images, over the blue arrow, are generated by bilinear interpolation and other label images, over the green arrow, are generated by max-pooling method. As you can see from Fig.13, if max-pool-labelling is used, the small nodules still remain even at the deepest level.

Therefore, it may be useful for generating the labels of different levels by using max-pooling. The efficiency of max-pool-labelling has been demonstrated through, the following toy experiments.

First, by using max-pool-labelling, risk of overfitting can be decreased. In order to train two models (using max-pool-labelling and resize-labelling) a data generator, which generates query images and labels like Fig.14 is created. Right side images of Fig.14(a) and Fig.14(b) of represent the labels, and left side images are the query images. The white regions in label images, which have the random shapes, can be regarded as nodule segmentations.

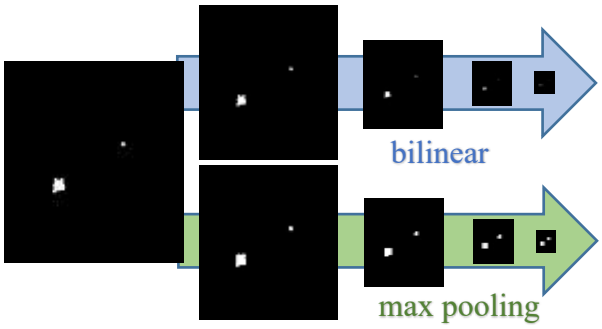
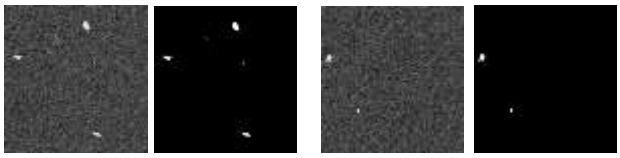


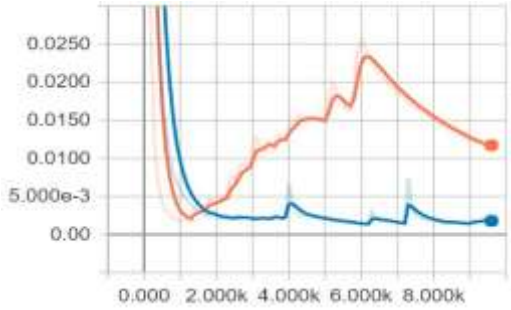
Fig.13. Resizing results by bilinear and max pooling

Here, we limit the sizes of them to range from 2 to 8 pixels. The query images are created by adding random backgrounds to label images.

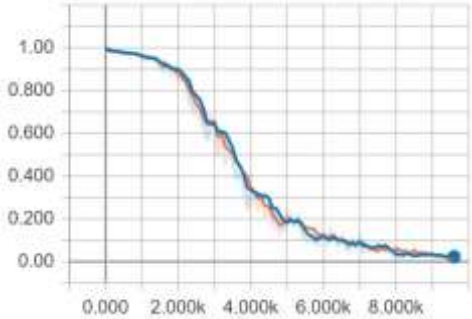


(a) (b)

Fig.14. Training samples of the toy experiment



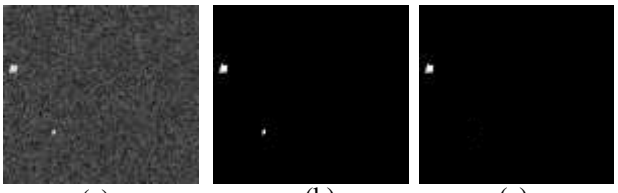
(a) regression losses



(b) dice coefficient losses

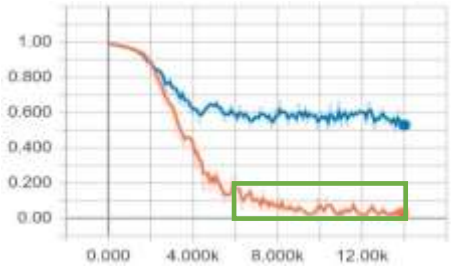
Fig.15. Training course of max pooling and resizing model

Then two models are trained on generated data in different ways: using resize-labelling and using max-pool-labelling. The Fig.15 shows the variance of regression losses and dice coefficient losses respected to the step during the training course. The blue curves of a) and b) represent the losses of max-pool-labelling model while red curves are for the resize-labelling model. The Fig.15(a) shows that the regression losses of max-pool-labelling model converges more steadily than resize-labelling model although the dice coefficient losses converge similarly each other. From this it can be assumed that the max-pool-labelling model is more generalized.

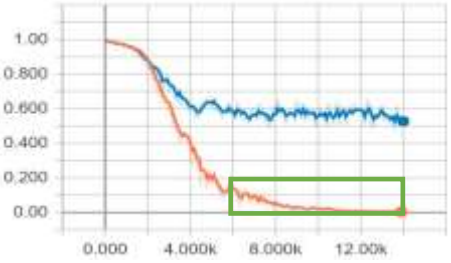


(a) (b) (c)

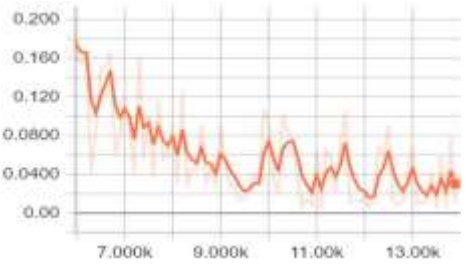
Fig.16. Training samples of the toy experiment



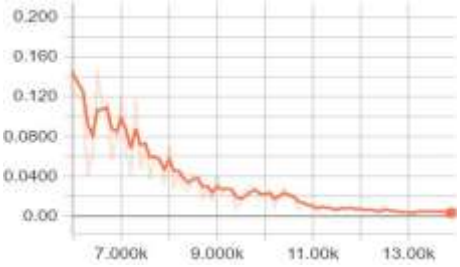
(a) Training losses of models



(b) Testing losses of models



(c) Training losses of max pooling mode



(d) Testing losses of max pooling mode

Fig.17. Training course on error containing data.

Second, by using max-pool-labelling, the model becomes more robust against errors. In order to generate error containing data make errors to label images. In detail, “nodules”, whose size

is smaller than 4 pixels are removed with a probability of 20%. Then the training samples change like Fig.16. The true label image (Fig.13(b)) changes into error label image (Fig.16(c)).

The Fig.17(a) shows the variance of training losses of both models. As you can see from the Fig.17, the training loss of max-pool-labelling model reached near zero point after 10000 steps, but the training loss of resize-labelling model remains around 0.6. It is not because of the errors that the training samples contain. The testing losses is also similar to the training losses (Fig.17(b), Fig.17(c) and Fig.17(d)) is the losses of max-pool-labelling model after 6000 steps (highlighted by green boxes on a) and b)). The training loss vibrates after 11000 steps because of errors, but the testing loss is fixed stably almost zero. This means that the max-pool-labelling model is very robust against errors. The Fig.18 shows the main architecture of our proposed model.

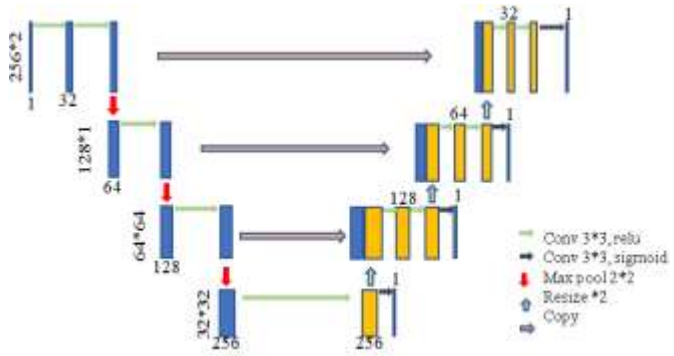


Fig.18. Proposed architecture

4. TRAINING AND RESULTS

4.1 DATA AUGMENTATION

The Fig.19 shows the different data augmentation methods and their results. Here, red box represents the original minimum rectangle and the blue box represents interesting region.

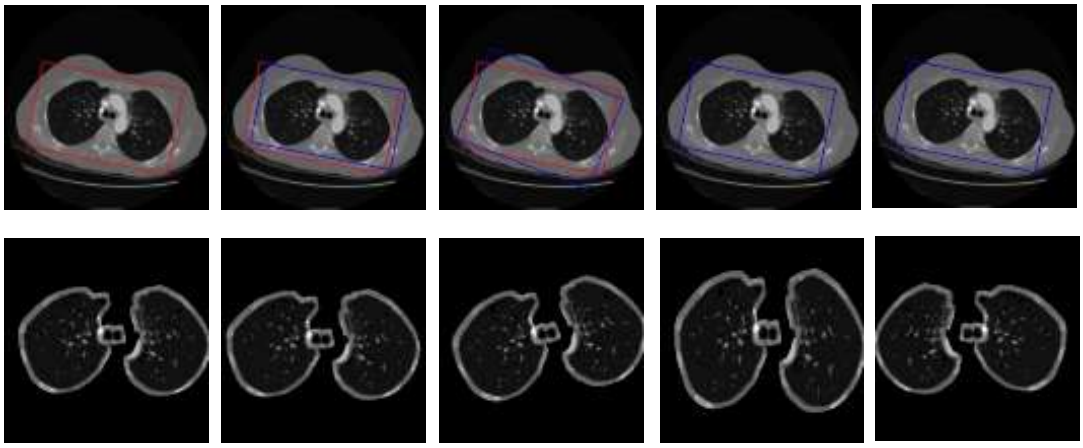
In order to increase the general accuracy of nodule segmentation system, we use several data augmentation methods: random crop, random rotate, random ratio, random flip (both vertical and horizontal). These methods (commonly called affine transformations) were well known already but applying them to our system requires some trick because of the speciality of proposed pre-processing method. Lung-range-standardization fully depends on finding a transformation matric, which is decided by associated points. So, the data augmentation is merely finding associated points.

In order to increase the general accuracy of nodule segmentation system, we use several data augmentation methods: random crop, random rotate, random ratio, random flip (both vertical and horizontal). These methods (commonly called affine transformations) were well known already but applying them to our system requires some trick because of the speciality of proposed pre-processing method. Lung-range-standardization fully depends on finding a transformation matric, which is decided by associated points. So, the data augmentation is merely finding associated points.

For the first let's check about the random crop case. Suppose that the red rectangle of Fig.19(b) is represented by $\{bp_1(bx_1, by_1), bp_2(bx_2, by_2)\}$ and that α denotes the leaning angle of the red rectangle then BP_{random_crop} (blue rectangle of Fig.19(b)-top) is calculated follows.

$$BP_{\text{random\_crop}} = \begin{pmatrix} bx_1' & bx_2' \\ by_1' & by_2' \end{pmatrix} = \begin{pmatrix} bx_1 & bx_2 \\ by_1 & by_2 \end{pmatrix} + \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} vx_1 & vx_2 \\ vy_1 & vy_2 \end{pmatrix} \quad (12)$$

(vx_1, vy_1) and (vx_2, vy_2) is the randomly selected variation vectors of (bp_1, bp_2). Then the crop result is obtained as Fig.19(b) bottom. Then, for the random rotate case, if the rotate angle and border points of lung-range are given as β and $LP = \{lp_1(lx_1, ly_1), lp_2(lx_2, ly_2), \ldots, lp_n(lx_n, ly_n)\}$, then the BP_{random_rotate} (blue rectangle of Fig.19(c)) is expressed as follows. And the rotated result is Fig.19(c)-bottom.



(a) normal (b) crop (c) rotate (d) re-ratio (e) flip

Fig.19. Data augmentation examples

$$\begin{pmatrix} lx_1', lx_2', \ldots, lx_n' \\ ly_1', ly_2', \ldots, ly_n' \end{pmatrix} = \begin{pmatrix} \cos(\alpha+\beta) & \sin(\alpha+\beta) \\ -\sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix} \cdot \begin{pmatrix} lx_1, lx_2, \ldots, lx_n \\ ly_1, ly_2, \ldots, ly_n \end{pmatrix} \quad (13)$$

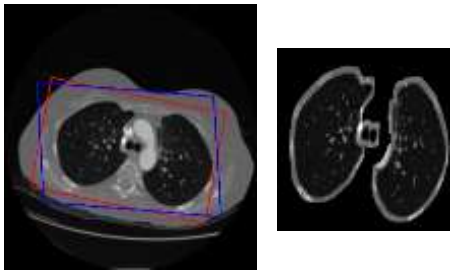
$$BP_{\text{random\_rotate}} = \begin{pmatrix} bx_1' & bx_2' \\ by_1' & by_2' \end{pmatrix} = \begin{pmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix} \cdot \begin{pmatrix} \min_i(lx_i') & \max_i(lx_i') \\ \min_i(ly_i') & \max_i(ly_i') \end{pmatrix} \quad (14)$$

But one must be careful about the applying order of transformations. In fact, complex of these transformations can be expressed with the multiplication of each transformation matrix. In general, matrix multiplication dose not satisfy the commutative law. Thus, the different order of transformations generates different outcomes. It is depicted as Fig.20.

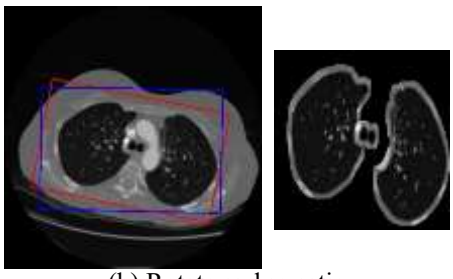
The images on the left side of Fig.20 a) and Fig.20 b) represent the original minimal rectangles (the red rectangle) and their associated rectangles (the blue rectangles). Because the associated points are different from each other the transformation results are also different (the images on the right side of a) and b) of Fig.20). In our case we choose the following order. 1-Random ratio (RRa), 2- random rotate (RRo), 3- random crop (RC) and 4-random flip (RF). And the data augmentation parameters are tabulated in Table.3.

4.2 COMPARISON

We train and test our model on subset of LIDC, which consists of samples that contain at least a nodule (13123 slice images of 800 persons are for training and 1195 slice image of 72 persons are for testing). In order to demonstrate the efficiency of lung-range-standardization we train three models of two architectures. The first one (architecture-1) is our proposed architecture and the second one (architecture-2) is made by adding convolution layers and a pooling layer to the proposed architecture. the Fig.21 shows the second architecture.



(a) Re-ratio and rotate



(b) Rotate and re-ratio

Fig.20. Transformation results in different orders

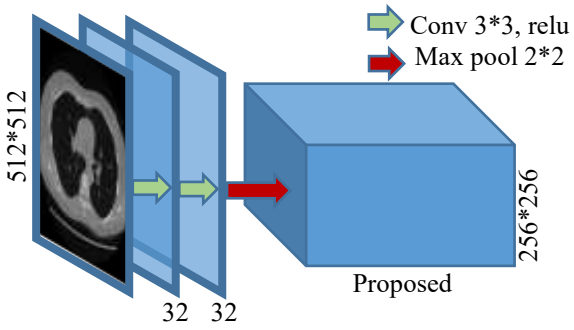


Fig.21. The second architecture

Table.3. Data increasing parameters

Parameters	Value
Random rotate	-5~5
Random ratio	1.2~1.7
Random crop	-10%~10%

The second architecture is used to compare the models with different input size. In fact, if lung slice images (original size is 512*512) are resized to (256*256), the lung regions of some images become so small that the relevant texture information of small nodules can be diluted. The properties of different models are summarized in Table.4.

One of the important things about the LIDC dataset is that the lung nodules are annotated by four professors, respectively. This means the labels of training samples are different for different professors. Thus, the model accuracy mainly depends upon the training label selection. The Fig.22 shows the examples of labels. In this paper, those nodules, which are annotated by at least one professor (not black region of Fig.22(f)), are used as training labels.

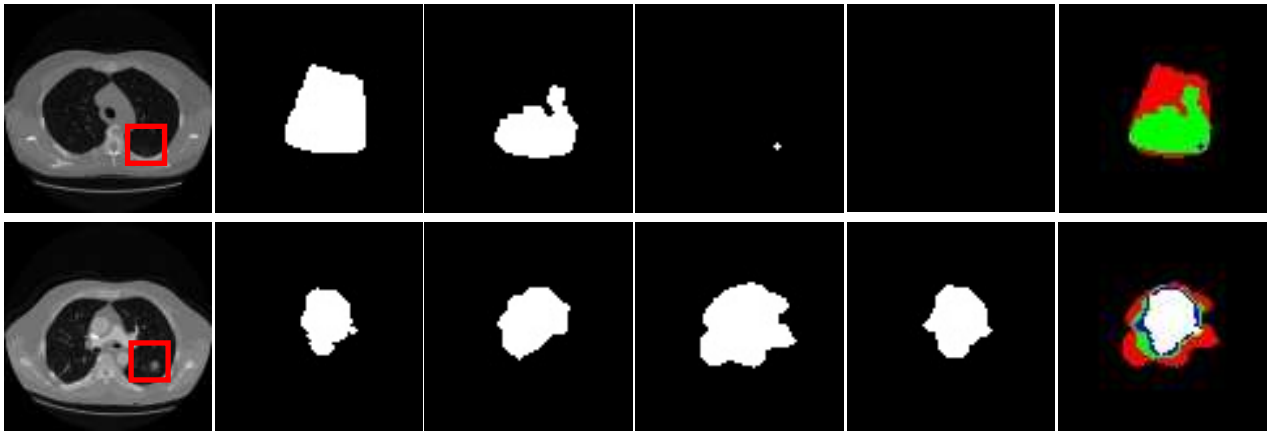
The models are compared using DSC, SEN and PPV.

$$\text{SEN} = \frac{TP}{TP+FN};\ \text{PPV} = \frac{TP}{TP+FP} \quad (16)$$

The Table.5 shows the estimation results of model-1, model-2 and model-3. Here "professors" represents how many professors annotated the label. For example, "professors" <= 1 means the nodules that are annotated at least a professor. From the table, it is easy to know that the model-2 is much better than the model-1 in all measurements and is similar to model-3. And also, the processing speed of model-3 is much faster than model-2 because the input image size of model-3 is a quarter of model-2.

Table.4. Models for comparison

Model name	**Pre-processing**	**Data Augmentation**	**Architecture**	**Loss function**
Model-1	Lung segmentation (LS)	RRa+ RRo+ RC+ RF	Architecture-2	Multi-levels loss
Model-2	LS + lung-range-standardization (LRS)		Architecture-2	
Model-3	LS + LRS		Architecture-1	



(a) (b) (c) (d) (e) (f)

Fig.22. Label examples according to professors

However, these testing result cannot explain the efficiency of lung-range-standardization precisely because there are some problems in the preciseness of labels of testing data as it is mentioned before.

Table.5. Testing results of model-1, model-2 and model-3

Professors	Model-1			Model-2			Model-3		
	DSC	SEN	PPV	DSC	SEN	PPV	DSC	SEN	PPV
≤ 1	0.584	0.615	0.587	0.633	0.655	0.646	0.623	0.644	0.628
≤ 2	0.688	0.669	0.744	0.723	0.692	0.795	0.715	0.682	0.778
≤ 3	0.705	0.647	0.808	0.735	0.664	0.857	0.741	0.665	0.861
≤ 4	0.646	0.565	0.803	0.679	0.582	0.863	0.688	0.584	0.871

For example, the DSC scores of models change a lot according to the professors. Therefore, we checked DSC scores of training and testing. The Table.6 shows the DSC scores of model-1and model-3. The training DSC of model-1 is much greater than model-3 but the testing DSC of model-1 is lower than testing DSC of model 3. Especially the training DSC of model-1is very high (0.92) while the testing DSC is very low (0.58). That is because model-1 is overfitted for the training dataset. And then, this result shows that lung-range-standardization prevent over-fitting properly although the best DSC score of model-3 (0.741) is lower than other state-of-the-art results (more than 0.82) include [14].

Table.6. Training and testing DSC scores

Models	DSC (training)	DSC (test)	DSC (test)
Model-1	0.921	0.584	0.337
Model-3	0.742	0.623	0.119

5. CONCLUSION

In this paper we have discussed about the lung-range-standardization method and multi-levels loss. In fact, the idea of lung-range-standardization comes from the deep analysis of the LIDC dataset. Therefore, by using the proposed methods we can increase the general accuracy of the lung segmentation model. Especially, lung-range-standardization can be useful for other lung cancer relation systems (include detection and classification) and multi-levels loss can be used for training small object segmentation models.

Our future work will focus on predict minimum rectangle of lung-range from one slice image. Now the lung-range is obtained from a person's full lung slice images. But, sometimes, it can be necessary to detect nodules from one or few slice images. Therefore, in this case, to apply lung-range-standardization it is necessary to predict minimum rectangle from one slice image.

REFERENCES

[1] J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D.M. Parkin, D. Forman and F. Bray, "Cancer Incidence and Mortality Worldwide: Sources, Methods and Major Patterns in Globocan 2012", *International Journal of Cancer*, Vol. 136, No. 5, pp. 359-386, 2015.

[2] H. MacMahon, J.H. Austin, G. Gamsu, C.J. Herold, J.R. Jett, D.P. Naidich, E.F. Patz Jr and S.J. Swensen, "Guidelines for Management of Small Pulmonary Nodules Detected on CT Scans: A Statement from the Fleischner Society", *Radiological Society of North America*, Vol. 237, No. 2, pp. 395-400, 2005.

[3] R.L. Siegel, K.D. Miller, S.A. Fedewa, D.J. Ahnen, R.G. Meester, A. Barzi and A. Jemal, "Colorectal Cancer Statistics", *CA Cancer Journal for Clinicians*, Vol. 67, No. 3, pp. 177-193, 2017.

[4] P.C. Hoffman, A.M. Mauer and E.E. Vokes, "Lung Cancer", *Lancet*, Vol. 355, pp. 479-485, 2000.

[5] T. Song, Alfonso Rodríguez-Patón, Z. Pan and X. Zeng, "Spiking Neural P Systems with Colored Spikes", *IEEE Transactions on Cognitive and Developmental Systems*, Vol. 10, No. 4, pp. 1106-1115, 2018.

[6] D.S. Alberts, "National Lung Screening Trial Research Team, Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening, *New England Journal of Medicine*, Vol. 365, No. 5, pp. 395-409, 2011.

[7] D.E. Midthun, "Early Detection of Lung Cancer", *F1000Research*, Vol. 5, pp. 1-9, 2016.

[8] G. Veronesi, P. Maisonneuve, L. Spaggiari, C. Rampinelli, A. Pardolesi, R. Bertolotti, N. Filippi and M. Bellomi, "Diagnostic Performance of Low-Dose Computed Tomography Screening for Lung Cancer Over Five Years",

Journal of Thoracic Oncology, Vol. 9, No. 7, pp. 935-939, 2014.

[9] M.J. DCruz, M.A. Jadhav, M.A. Dighe, M.V. Chavan and J. Chaudhari, "Detection of Lung Cancer using Backpropagation Neural Networks and Genetic Algorithm", *International Journal of Advanced Research in Computer and Communication Engineering*, Vol. 5, No. 4, pp. 1-5, 2016.

[10] K. Kuan, M. Ravaut, G. Manek, H. Chen, J. Lin, B. Nazir, C. Chen, T.C. Howe, Z. Zeng and V. Chandrasekhar, "Deep Learning for Lung Cancer Detection: Tackling the Kaggle Data Science Bowl 2017 Challenge", *Computer Vision and Pattern Recognition*, pp. 1-9, 2017.

[11] J. Kuruvilla and K. Gunavathi, "Lung Cancer Classification using Neural Networks for CT Images", *Computer Methods and Programs in Biomedicine*, Vol. 113, No. 1, pp. 202-209, 2014.

[12] J.M. Alvarez, T. Gevers, Y. LeCun and A.M. Lopez, "Road Scene Segmentation from a Single Image", *Proceedings of International Conference on Computer Vision*, pp. 376-389, 2012.

[13] Y. Liu, K. Gadepalli, M. Norouzi, G.E. Dahl, T. Kohlberger, A. Boyko, S. Venugopalan, A. Timofeev, P.Q. Nelson, G.S. Corrado and J.D. Hipp, "Detecting Cancer Metastases on Gigapixel Pathology Images", *Computer Vision and Pattern Recognition*, pp. 1-13, 2017.

[14] V. Hernández-Solis, A. Tellez-Velázquez, A. Orantes-Molina and R. Cruz-Barbosa, "Lung-Nodule Segmentation using a Convolutional Neural Network with the U-Net Architecture", *Proceedings of International Conference on Pattern Recognition*, pp. 335-344, 2021.

[15] H. Polat and H.D. Mehr, "Classification of Pulmonary CT Images by using Hybrid 3D-Deep Convolutional Neural Network Architecture", *Applied Sciences*, Vol. 9, No. 5, pp. 1-15, 2019.

[16] E. Simie and M. Kaur, "Lung Cancer Detection using Convolutional Neural Network", *Proceedings of International Conference on Issues and Challenges in Intelligent Computing Techniques*, Vol. 5, No. 4, pp. 284-292, 2019.

[17] J. Zhu, J. Zhang, B. Qiu, Y. Liu, X. Liu and L. Chen, "Comparison of the Automatic Segmentation of Multiple Organs at Risk in CT Images of Lung Cancer between Deep Convolutional Neural Network based and Atlas-based Techniques", *Acta Oncologica*, Vol. 58, No. 2, pp. 257-264, 2019.

[18] X. Huang, W. Sun, T. Tseng, C. Li and W. Qian, "Fast and Fully-Automated Detection and Segmentation of Pulmonary Nodules in Thoracic CT Scans using Deep Convolutional Neural Networks", *Computerized Medical Imaging and Graphics*, Vol. 74, pp. 25-36, 2019.

[19] H. Xie, D. Yang, N. Sun, Z. Chen and Y. Zhang, "Automated Pulmonary Nodule Detection in CT Images using Deep Convolutional Neural Networks", *Pattern Recognition*, Vol. 85, pp. 109-119, 2019.

[20] W. Alakwaa, M. Nassef and A. Badr, "Lung Cancer Detection and Classification with 3D Convolutional Neural Network (3D-CNN)", *International Journal of Advanced Computer Science and Applications*, Vol. 8, No. 8, pp. 99-110, 2017.

[21] M.B. Zia, Z.J. Juan, Z.U. Rehman, K. Javed, S.A. Rauf and A. Khan, "The Utilization of Consignable Multi-Model in Detection and Classification of Pulmonary Nodules", *International Journal of Computer Applications*, Vol. 177, No. 27, pp. 1-6, 2019.

[22] Y. Onishi, "Automated Pulmonary Nodule Classification in Computed Tomography Images using a Deep Convolutional Neural Network Trained by Generative Adversarial Networks", *Biomed Research International*, Vol. 2, No. 5, pp. 99-110, 2019.

[23] M.N. Gurcan, B. Sahiner, N. Petrick, H.P. Chan, E.A. Kazerooni, P.N. Cascade and L. Hadjiiski, "Lung Nodule Detection on Thoracic Computed Tomography Images: Preliminary Evaluation of a Computer-Aided Diagnosis System", *Medical Physics*, Vol. 29, No. 11, pp. 2552-2558, 2002.

[24] A. Chon, N. Balachandar and P. Lu, "Deep Convolutional Neural Networks for Lung Cancer Detection", Available at https://www.academia.edu/84820317/Deep_Convolutional_Neural_Networks_for_Lung_Cancer_Detection, Accessed in 2017.

[25] O. Ronneberger, P. Fischer and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation", *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234-241, 2015.

[26] P. Rao, N.A. Pereira and R. Srinivasan, "Convolutional Neural Networks for Lung Cancer Screening in Computed Tomography (CT) Scans", *Proceedings of International Conference on Contemporary Computing and Informatics*, pp. 489-493, 2016.

[27] J. Worku Sori, Jiang Feng, W. Arero Godana, Shaohui Liu and J. Demissie Gelmecha, "DFD-Net: Lung Cancer Detection from Denoised CT Scan Image using Deep Learning", *Computer Science*, Vol. 15, No. 2, pp. 1-9, 2021.

[28] RuoXi Qin, Zhenzhen Wang, LingYun Jiang, Kai Qiao and Jinjin Hai, "Fine-Grained Lung Cancer Classification from PET and CT Images based on Multidimensional Attention Mechanism", *Hindawi Complexity*, Vol. 2020, pp. 1-12, 2020.

[29] Jia Ding, Aoxue Li, Zhiqiang Hu and Liwei Wang, "Accurate Pulmonary Nodule Detection in Computed Tomography Images using Deep Convolutional Neural Networks", *Medical Image Computing and Computer Assisted Intervention*, pp. 559-567, 2017.

[30] Mike Huang, "Automated Extraction and Classification of Pulmonary Lung Nodules from CT Scans", Available at https://github.com/NoahApthorpe/CellMagicWand, Accessed in 2018.