Front. Comput. Sci., 2021, 15(2): 152701
https://doi.org/10.1007/s11704-020-9050-z

RESEARCH ARTICLE

DFD-Net: lung cancer detection from denoised CT scan image using deep learning

Worku J. SORI ((✉))[1,2], Jiang FENG [3,4], Arero W. GODANA [3], Shaohui LIU [3], Demissie J. GELMECHA [5]

1 School of Electrical Engineering and Computing, Department of Computer Science and Engineering, Adama Science and Technology University, Adama 1888, Ethiopia
2 Faculty of Natural Computational Science, Department of Mathematics Bule Hora University, Bule Hora 144, Ethiopia
3 School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China
4 Peng Cheng Laboratory, Shenzhen 518052, China
5 School of Electrical Engineering and Computing, Department of Electronics and Communications Engineering, Adama Science and Technology University, Adama 1888, Ethiopia

Abstract The availability of pulmonary nodules in CT scan image of lung does not completely specify cancer. The noise in an image and morphology of nodules, like shape and size has an implicit and complex association with cancer, and thus, a careful analysis should be mandatory on every suspected nodules and the combination of information of every nodule. In this paper, we introduce a "denoising first" two-path convolutional neural network (DFD-Net) to address this complexity. The introduced model is composed of denoising and detection part in an end to end manner. First, a residual learning denoising model (DR-Net) is employed to remove noise during the preprocessing stage. Then, a two-path convolutional neural network which takes the denoised image by DR-Net as an input to detect lung cancer is employed. The two paths focus on the joint integration of local and global features. To this end, each path employs different receptive field size which aids to model local and global dependencies. To further polish our model performance, in different way from the conventional feature concatenation approaches which directly concatenate two sets of features from different CNN layers, we introduce discriminant correlation analysis to concatenate more representative features. Finally, we also propose a retraining technique that allows us to overcome difficulties associated to the image labels imbalance. We found that this type of model easily first reduce noise in an image, balances the receptive field size effect, affords more representative features, and easily adaptable to the inconsistency among nodule shape and size. Our intensive experimental results achieved competitive results.

Keywords medical image, discriminant correlation analysis, features fusion, image detection, denoising.

Received May 27, 2019; accepted February 10, 2020

E-mail: worku.jifara@gmail.com

1 Introduction

Globally, as the 2018 World Health Organization (WHO) evidence, 1.76 million deaths were estimated as a result of lung cancer. In the United States of America alone, 83,550 deaths were predicted where lung cancer is considered as a reason [1]. Figure 1 reveals the major cancer varieties for the estimated deaths in the USA, where lung cancer ranked first in both males and females. This figure is likely bigger in developing countries, especially in African countries. Overall, lung cancer is a serious cause of death among the death related to cancer.

With timely diagnosis and treatment, the probability that one can cure from lung cancer is much higher. Therefore, lung cancer screening is the first and crucial step, with better identification approaches leading to the improved patient result. According to [2], in high-risk demographics, early detection of lung cancer from computed tomography image of lung decreased mortality rate by 20% compared to single view radiography. But, detecting cancer from lung image prone to false positive and true negative, which vulnerable a patient for additional investigation, cost, and tension and a doctor for additional burden [3]. The recent involvement of computer-aided diagnosis (CAD) in lung cancer detection reveals that there have been observed a great achievement that CAD is improving lung cancer detection accuracy than human, timely.

The immense success of deep learning methods in natural image detection and recognition are also transformed to different medical imaging problems and modalities. Deep convolutional neural network model is a deep learning model that provides enormous success over the other state of the art methods in various medical image challenges [4–6]. For instance, in the area of medical image, convolutional network achieved better sensitivity and accuracy result than human expert and other methods in skin cancer metastases classification [5], gaining noticeably better sensitivity performance than human. Further-

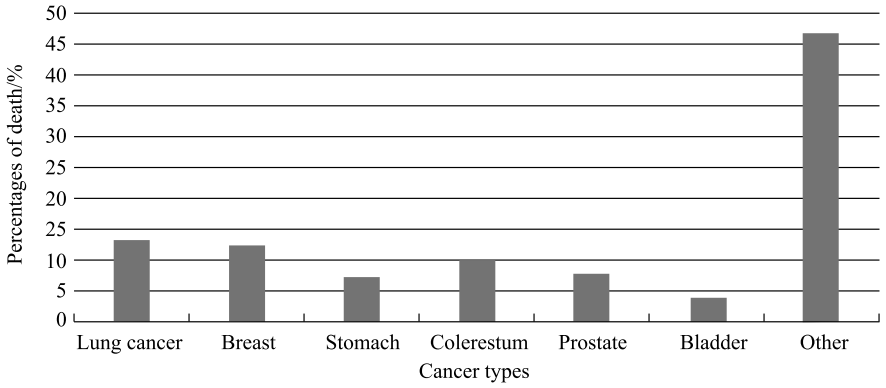

Fig. 1 An estimated percentages of death due to cancer in USA, 2018

more, [6] introduced a three dimensional convolutional network to detect lung cancer and achieved encouraging results, where this model however limited when the issue of memory and time complexity is taking into consideration. However, due to several motives explained in this work, the existing lung cancer detection CNN based models have not revealed sufficiently accurate and robust outcome for real time employment. Considering a computed tomography image of lung, we identified three lung cancer detection challenges. These are:

1) The intricacy information's that nodule is made of. Nodule is found anywhere in a lung. They might be normal or abnormal. It is hard to model or characterize their nearby pixel, because their nearby pixel's is highly correlated. Thus, because of their highly correlated nature of nearby pixel nodule, the existing detection methods encounter the problem of accuracy, and usually hard for experts.
2) The innate morphology of nodules across a computed tomography image of a lung. Nodules shape are inconsistent. Some nodules are circular, some are elliptical, and some are shapeless. Their shape is not absolutely representing the normality or abnormality of a nodule. Similarly, nodules sizes are not constant; some are extremely small, some are medium sized, and some are larger. Their size does not absolutely reveal the normality or abnormality of a nodule. Most of the time, this create ambiguity among expert, and thus, further investigations are recommended like intensive care, blood examination, and surgery, and as a result, it becomes a load for patients. It is not only difficult for an expert alone, CAD faces similar situations. Figure 2 demonstrates the pointed inconsistency among nodules.
3) The predominance of noise. It is known that medical imaging quality getting better and better through time, however, with the development of medical imaging technology that produces high resolution medical images, noise remains the predominant issue. This makes an image quality somehow deteriorated, create misunderstanding, and complicate the overall subsequent medical image analysis task such as the detection task. That is why "denoising first" detection (DFD) strategy is considered in this paper.

Detecting lung cancer via deep learning model is depending on how much the nearby pixel values are designed, i.e., the better local and global range dependencies are well designed, the better the model achieves better performance. Majority of researches that were devoted for lung cancer detection used 3×3, 5×5, or 7×7 kernel size to design a nearby pixel and follow one path, through the model [6–8]. With one path and fixed kernel size, even if those methods noticeably achieved better performance compared to the conventional lung cancer detection methods, none of those methods were focused on the integration of local (nearby pixels) and global range dependencies together, instead they independently emphases on local or global dependencies. Due to this, during designing the nearby pixels or global pixels, the fixed kernel size puts some impact on both dependencies. This implies, those aforesaid difficulties are not jointly reflected, as the size, shape, global, and local dependencies are all about how better designing nearby pixels. From Fig. 3, if we see what would R_i, R_j, and R_k outputs using convolution is somehow different, because of their receptive field size utilized. The best size depends on the problem at hand. Thus, it is significant to design a model which balance the impact of receptive field size.

The other challenge that lead cancer detection result to false positive via convolutional neural network model is: (1) a shortage of dataset — it has been proved by researchers that deep learning models are benefitted from large number of dataset than small dataset during training [7,8]. However, when we come to medical image field, one of the greatest challenge is to find large number of dataset and (2) the labelled dataset class imbalance- even if we found certain medical image dataset, we usually face the problem of labelled data class imbalance. For

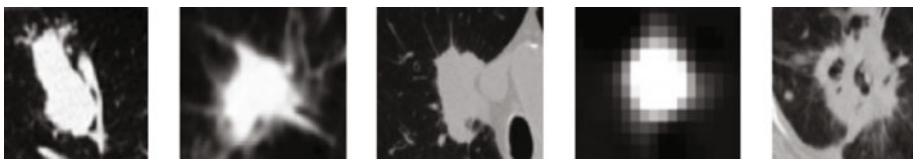

Fig. 2 Nodules with different morphology (zoomed from the original images)

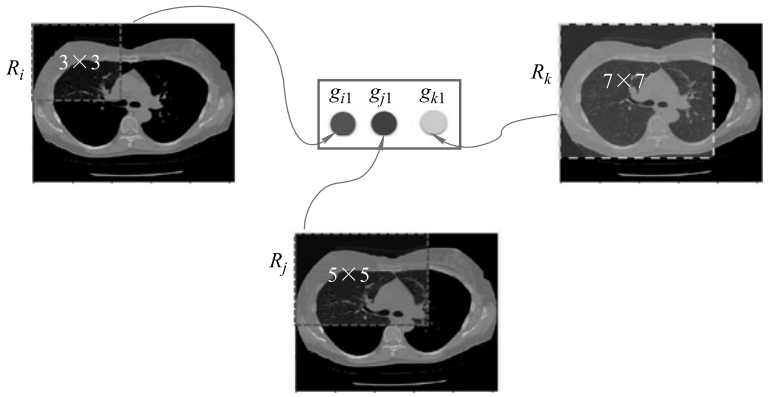

Fig. 3 Demonstration of different receptive field size: R_i, R_j, and R_k are distinct receptive field size and g_{i1}, g_{j1}, and g_{k1} are their corresponding output through convolution

instance, in the collection of Kaggle Data science Bowl Challenge 2017, the number of labelled data has a ratio of 1 to 7, i.e., the ratio of normal (cancer) to abnormal (No-cancer). Clearly this class imbalance benefits the larger target while disfavouring the lower target during learning process which in turn reduce model accuracy.

To address these aforesaid encounters, taking the denoised computed tomography lung image as an input, this paper introduces a two-path convolutional neural network with efficient features fusion approach. Contested to the three dimensional convolutional neural networks that cost much time and memory during training, we used the two dimensional convolutional network that move forward in two-path, which lead the cancer detection task one step forward. What make our proposed convolutional network difflerent from the convensional model are:

One, our model is differ in structure. The model has two paths, where each path is designed in such a way that the effect of receptive field sizes are to be minimized during modelling short and long range dependencies. This is very important since the morphology of a nodule is highly determined by the integration of nearby pixels. In this two-path model, the aim of the first path is to model local dependencies, whilst, the aim of the second path is to model long-range dependencies.

Two, our model is the first to use feature fusion mechanism in two path convolutional networks. Instead of directly concatenating features from different layers like the conventional feature fusion approach, we introduce discriminative correlation analysis to fuse more important features extracted from the images by the two paths together. The features resulted from the discriminative correlation mechanism is more discriminative and informative than the initially estimated features and it has immense advantage for a CNN model.

Three, our method is focused on denoised computed tomography image of lung. Contrary to the usual lung cancer detection approach which directly leverage cancer detection from raw images, we first add noise to a computed tomography lung image with additive White Gaussian noise (noise level $\sigma = 10$) at some stage. To achieve the intended objective, we denoise these images first. For this purpose, we introduced DR-Net model which removes noise first. This model is proposed by our research group first in [9]). Here, one could ask why we add noise while detection is possible from raw CT scan images. Since noise is the most predominant in medical images and no medical image is free of noise, our goal is that: (1) to show the research communities that to what extent one could detect lung cancer from denoised images and (2) practically, denoising is the primary step for detection which most of the time left from the detection task and hence in this work, we motivated to combine and introduce the two tasks together. So, this method is intended to introduce lung Cancer detection model from denoised computed tomography image of lung and at the same time focuses on designing robust detection model.

Four, the use of retraining strategy makes our model differ. Several researchers have proposed different CNN training strategy when there is a shortage in dataset and labelled dataset class imbalance, such as transfer learning [6] and fine tuning [7] the whole model. However, these methods do not provide intended detection accuracy. Thus, in this work, we used a retraining strategy where the first training is similar with the conventional one, but the retraining is based on fine-tuning only some part of the model. In general, our network is designed in a unified form to circumvent the effect of distant dependencies, and provide competitive performance. We summarize our method as follows:

1) We introduce a two-path convolutional network that is trained with denoised Computed Tomography image of lung to detect lung cancer. The aim of this model is to characterize and learn different morphology of nodule features contextual information. To the best of our knowledge, this is the first work that integrate denoising and detection tasks. Details of the contribution are found in Section 3.
2) To better polish our model accuracy, we also deal with discriminant correlation analysis strategy. It concatenates more representative features than features estimated by the two paths, where we found better accuracy with this approach. The details contribution is found in Section 3.3.1.
3) To deal with the critical class difference problem, we propose two stage training where all model layers' parameters are fine-tuned after the first training stage, except the output layer in which we employ retraining strategy. Details of the contribution are found in Section 3.4.2

The rest of this work is arranged as: Section 2 briefly review related works. In Section 3, details of the introduced two-path

convolutional network model (descriptions of the two paths, dataset used and the conducted preprocessing) is presented. We present the experimental results and discussions in Section 4 and finally, Section 5 gives a conclusion.

2 Related work

The primary goal of automatic lung cancer detection methods is to aid a professional to give better decision during diagnosis. It minimizes energy and time of an expert, and reduces cost of a patient. The investigation made by [10], indicates that the number of research conducted on lung cancer detection has rapidly increasing in the recent time. This observation highlights automatic lung cancer detection methods are very significant and it is still a task in progress.

Several medical image detection approaches have been proposed by researchers in general and for lung cancer detection in particular. We categorize them into two major approaches. These are discriminative models and handcrafted models. Discriminative models construct detail information about lung anatomy. It digs out low level features and automatically models the relation of original features and its corresponding label. On the other side, different from the discriminative methods, the hand crafted methods are immensely focus on domain specific understanding about the identification of labels. Nodule presence is challenging to describe, and previous handcrafted methods commonly misscreening the true labels, i.e., the predicted label is not akin with the ground truth [11,12].

Since a comprehensive analysis is away from the range of our work, we only give some current discriminative methods. Majority of an image detection models are focused on two dimentional image. For instance, faster-RCNN was introduced by [13], where the proposed model suggested some bounding boxes during an initial stage and the class decision estimated during second stage. Moreover, the current methods extend to a single stage, where class probabilities as well as bounding boxes are forecasted immediately [14] or without generating the proposal, the class probabilities can be predictable for default boxes [15]. Generally, single stage methods are faster nonetheless two-step methods are better perfect. Recently, [16] proposed a new partially supervised training paradigm, together with a novel weight transfer function, that enables training instance segmentation models on a large set of categories all of which have box annotations, but only a small fraction of which have mask annotations. Following this, the new batchnormalization technoque called group normalization [17] is incorporated in deep learning and for example [18] employed for image detection in cascaded way.

Very sooner, medical image detection model based on three dimensional convolutional networks were also introduced. Some of these work are found at [3, 7, 8, 19]. Considering small sample size, [19] proposed fast CapsNet for lung cancer screening. The method achieved better result than the tradition CNN and three times faster as well. Taking a computed tomography scan lung image, [3, 5, 6] proposed three dimensional convolutional network and [20] proposed a multi path CNN for lung cancer detection to detect Cancer and have obtained reasonably encouraging results. However, the time needed to train this model is very expensive and the memory need to process is very huge. Assume how it is so challenging if the task is to be daily work. In addition, these approaches do not take into consideration those previously elaborated challenges and never perform the detection task from the denoised image, but in this paper, we address those challenges.

3 Methodology

3.1 The preprocessing stage

3.1.1 Dataset

The datasets used are taken from Kaggle Data Science Bowl 2017 challenge (KDSB) and LUNA 16. These datasets are a collection of Computed Tomography scan images, in DICOM format. For each individual patient, in both datasets, there are about 100–400 axial slice of size 512 height and 512 width. The two datasets are not similarly labelled. The Kaggle Data Science Bowl dataset contains 2,101 labeled data. It is labelled as 0 and 1. The label 0 represents the negative outcome (normal or cancer free) and 1 represents the label of positive outcome (abnormal or with cancer).

3.1.2 Segmentation of the lung

In addition to a lung tissue, a lung image obtained by CT modality contains other substances like water, air, bone and blood. Their presence is irrelevant for cancer detection, but their existence disturb the capability that the model labels the nodules accurately, and hence we omit those substances.

First, we convert pixel values of each CT scan into Hounsfield Unit (HU). HU is a quantitative measurement which is used to represent Radio density of substances that are found in CT scan images of lung. For instance, lung has –500, bone has 700, Blood has 0, Kidney has 30, and Water has 0 Radio density in HU [21]. After HU conversion, each CT scan comprising several slices, where each pixel value is corresponding to HU and found between [–1024, 3071].

Next, we omit those tissues. To omit those tissues, a common method used by scholars are k-means [21], thresholding [22], watershed [23] and clustering [24]. In this work, thresholding is employed. To this end, Gaussian filter is used and then we normalize the pixel value is to be between [0, 1]. We used –600 as a threshold. Figure 4 shows a slice of CT scan image of a patient and its segmentation result based on thresholding.

Finally, in order to use the proposed network, we transform the HU values of each slice into UINT8, i.e., the raw data that were initially transformed within [–1024, 3071] are linearly transformed within [0, 255]. And then, the mask used to segment the lung tissue is multiplied with these values. Substances that are out of the mask is set to 170, representing a common tissue of luminance.

Fig. 4 Some segmentation results by thresholding

3.1.3 Suspicious nodule detection

To effectively and efficiently segment some suspicious nodules from CT scan image of a lung, we used U-Net initially proposed by [23]. It is a two dimensional network that has been extensively functioned for many Bio-medical image segmentation problems like for brain nodules localization and segmentation. By modifying some parameters, like depth of the model, filter size, the size of an input patch, we train a U-Net. To this end, we employed LUNA 16 dataset. The detail of U-Net is found at [23].

After training U-Net with LUNA 16 dataset, we then used the network to segment the KDSB dataset, i.e., LUNA 16 is used during training and KDSB dataset is used for testing. In such a way, we determine suspicious nodule region which increase the degree that our model will predict Cancer nodule from the normal one. We then used these images (image having suspicious nodules) to train our proposed DFD-Net by dividing them into training set, validation set, and test set. These suspicious candidates have variable size and shape. Figure 5 shows some samples of an image generated by U-Net.

3.2 DR-Net

In this section, we give brief explanation of the denoising model, DR-Net. The purpose of this model is to denoise lung CT scan images before these images are used to be processed by the detection part of the model. First, the model is trained with LUNA 16 dataset, and then the model is evaluated with KDSB dataset. Both datasets are CT scan images of lung, they have similar characteristics except their Cancer location and Cancer status, where LUNA 16 has Cancer location but no Cancer status and KDSB has Cancer status but no Cancer location. The one we used for detection is the denoised one, i.e., the denoised KDSB images.

To train the model, we follow the residual learning approach proposed by [9]. The model is illustrated in Fig. 6 and the parameters of the model's are given in Table 1. To obtain the trainable parameters of the DR-Net, the loss function between the desired residual image $\{(z_k - x_k)\}$ and predicted one $\{\Im(z_k;\kappa)\}$ is formulated as

$$l(\kappa) = \frac{1}{2M}\sum_{k=1}^{M} \parallel \Im(z_k;\kappa) - (z_k - x_k) \parallel_F^2, \tag{1}$$

where κ denotes the parameters to be trained, $\{z_k, x_k\}_{k=1}^{M}$ are M noise-clean image pairs, respectively and $\|.\|_F^2$ is Frobenius norm. We initialize the weights as in [25] and use Adam [26] algorithm with $\epsilon = 10^{-8}$, $\alpha = 0.01, \beta_1 = 0.9$, and $\beta_2 = 0.999$. The size of image batch is set to 10. The model is trained for 50 epochs. For the 50 epoches the learning rate is decayed exponentially from 0.01 to 0.0001. To conduct this experiment MatConvNet framework [27] is employed to train the network. Since the data is small in number, we also used data augmentation strategy. This part of the experiment is carried out in the Matlab (R2015b) language on a computer with Intel(R) Xeon(R) CPU E3-1230v3 3.30GHz and NVIDIA Tesla K40c GPU.

While the training detail of this method is found at [9], in our detection model, we used the reconstruction part, i.e., given a noisy CT scan image of lung (particularly after preprocessing), we used the reconstruction phase of DR-Net to reconstruct the noise free images. These reconstructed image are then directly used in the detection part. So, for the detection part, we can assume DR-as a pretrained model which denoise a noisy images.

3.3 Two-path CNN with feature fusion

In the introduction section, we presented that the inconsistency of nodule morphology, and relationship of the nearby pixels are among the crucial factors when detecting lung Cancer using deep learning methods. Because of these factors, a lung cancer detection models using deep learning is not providing sufficiently greater accuracy. One of the motives is the effect of receptive field size during modeling local context and global context and the fixed number of path corresponding to these receptive field sizes. For instance, most CNN centered detection networks are so limited to fixed receptive field size throughout

Fig. 5 Segmented and detected sample of suspecious nodules by U-Net

Table 1 DR-Net parameters

Layers	1	2–14	15
Types	conv+relu	conv+BN+relu	conv+relu
#kernel/Kern.	64/3×3×1	64/3×3×64	1/3×3×64
stride	1	1	1

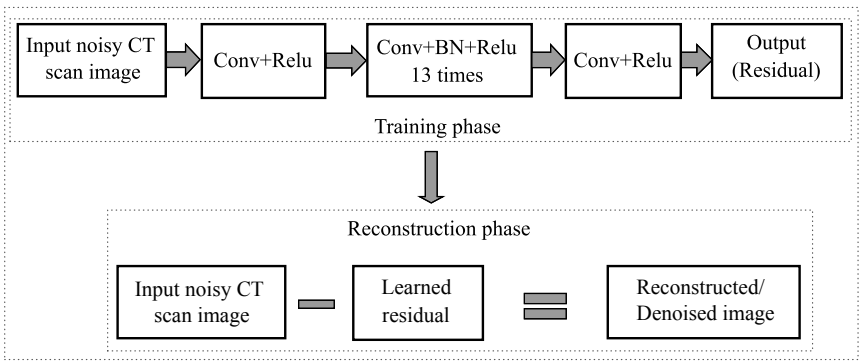

Fig. 6 DR-Net framework

all layers (for instance, the use of 3×3, 5×5 or 7×7 size in the entire model) and one-way path (path that only rely on fixed receptive field size). This influence the nearby pixels and high level information during feature extraction.

To overcome the challenges, we have introduced a two-path CNN that consider different receptive field sizes correspondingly. Each path is with unique kernel size, where we named these two-paths as, first path and second path. The first path used a receptive filed of size 3×3 and that of the second path is 5×5. At this position, one could raise a question as if 5×5 size is larger enough or 3×3 is suitable enough? Reply may vary; it might rely on a specific challenge. To investigate the stated challenges, we just emphasis on the mentioned kernel sizes and two-paths. The model parameters are given in Table 2.

Furthermore, in the proposed model we incorporate various feature fusion strategy. For both paths (first and second path), apart from the traditional concatenation approach, we have introduced an efficient way to concatenate their feature map from the fourth and third convolutional layer respectively. First, we transform the fourth convolutional layer of the first path and the third convolutional layer of the second path features and then, after this transformation, we concatenate them together. For three reasons, we inspired for such type of model. These are, the prediction of the label is clearly determined by how nearby pixel values are well modeled, the use of concatenation benefits in modelling more representative features and modeling long and short-term dependencies, and the selection of more important features from the last CNN layer boost the prediction accuracy, where we used DCA strategy in our case. Inspired by these, we relatively overcome the lung Cancer detection challenges. Figure 7 shows the network of the proposed model, where the overall algorithm of our model is illustrated in Table 3. We name this network as "denoising first" detection with two-path CNN (DFD-Net).

Table 2 DFD-Net parameters setup, ReLu activation function is used after each convolution except the conv. after concatenation layer (where soft-max is used)

First path			Second path		
Layers	w.size/ #weights	Input	Layers	w.size/ #weights	Input
conv.	3×3/64	128×128×1	conv.	5×5/64	128×128×1
maxp.	2×2	64×126×126	maxp.	2×2	64×124×124
conv.	3×3/64	64×125×125	conv.	5×5/64	64×123×123
maxp.	2×2	64×123×123	maxp.	2×2	64×119×119
conv.	3×3/64	64×122×122	conv.	5×5/64	64×118×118
maxp.	2×2	64×120×120	maxp.	2×2	64×115×115
conv.	3×3/64	64×119×119			
maxp.	4×4	64×117×117			
Concatenation (primary or secondary) layer					
conv.	114×114	192×114×114			
Output		2×1×1			

3.3.1 Feature fusion in two-path CNN

In this sub-section, since the concatenation layer is mainly focuses on discriminant correlation analysis (DCA), we provide the details of DCA. Several feature fusion mechanism has been exploited in convolutional network aiming at obtaining more

Table 3 The overall algorithm of DFD-Net

Input: KDSB, LUNA16.

Step 1: Convert the input data to HU.

Step 2: Segment the lung using thresholding as proposed in Section 3.1.2.

Step 3: Train U-Net with LUNA16 dataset for segmentation.

Step 4: Test the trained U-Net with KDSB dataset to determine the suspicious nodule region.

Step 5: Add Gaussian noise to LUNA16 dataset with noise level $\sigma = 10$.

Step 6: Train DR-Net with the data in Step 5.

Step 7:Add Gaussian noise to the outputs of Step 4 with noise level $\sigma = 10$.

Step 8: Test DR-Net with Step 7 outcomes.

Step 9: Train DFD-Net with some outputs of Step 8 (as some are used in Step 10).

Step 10: Test DFD-Net with data that are not included during performing Step 9.

Output: Prediction (as Cancer or No-Cancer).

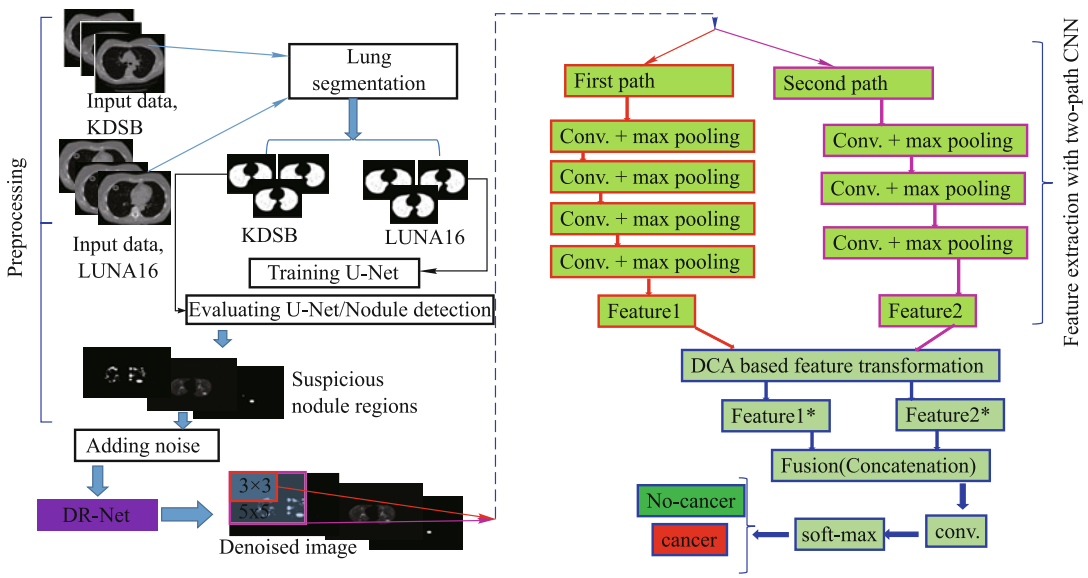

Fig. 7 DFD-Net framework. The figure reveals the input patch undergoes some preprocessing steps, then through two paths of convolutional layers

helpful descriptor. For instance, serial features fusion strategy [28] and parallel features fusion strategy [29, 30] mechanism are immensely employed.

Serial features fusion strategy easily concatenates two independent features into one feature. Assumes that f_1 and f_2 are two set of features having a, b dimension, respectively. Using serial feature fusion mechanism, this two features are transformed to a single feature l having size $(a + b)$. Similarly, parallel strategy transforms the two sets of features into a single feature, but the transformed feature became a complex vector.

Canonical correlation analysis (CCA) feature mechanism has been introduced by [31] to better provide better informative features than the initially estimated features. We briefly explain the idea of this method as follows.

Canonical feature fusion: Assume that $F_1 \in \mathbb{R}^{a\times n}$ and $F_2 \in \mathbb{R}^{b\times n}$ represents set of features, with n representing total number of feature set and a and b denotes the dimension of F_1 and F_2, respectively. Again, assume $C_{f_1f_1} \in \mathbb{R}^{a\times a}$ and $C_{f_2f_2} \in \mathbb{R}^{b\times b}$ represent the covariance matrices (κm) of F_1 and F_2, respectively, and $C_{f_1f_2} \in \mathbb{R}^{a\times b}$ is the intermediate covariance, where $C_{f_2f_1} = C^{\mathrm{T}}_{f_1f_2}$. The general covariance matrix $C \in \mathbb{R}^{(a+b)\times(a+b)}$ is then calculated as

$$C = \begin{pmatrix} \kappa m(F_1) & \kappa m(f_1, f_2) \\ \kappa m(f_2, F_1) & \kappa m(f_2) \end{pmatrix} = \begin{pmatrix} C_{f_1f_1} & C_{f_1f_2} \\ C_{f_2f_1} & C_{f_2f_2} \end{pmatrix}, \tag{2}$$

As pointed in [32], it is tough to apprehend the relationship between these two features from matrix C. The main objective of canonical correlation analysis is that to introduce two linear combination $F^*_1 = X^{\mathrm{T}}_{f_1}F_1$ and $F^*_2 = X^{\mathrm{T}}_{f_2}F_2$, which optimizes the correlation (cr) between the two features F_1 and F_2.

$$cr(F^*_1, F^*_2) = \frac{\kappa m(F^*_1, F^*_2)}{\sigma^2(F^*_1).\sigma^2(F^*_2)}, \tag{3}$$

where $\kappa m(F^*_1, F^*_2) = X^{\mathrm{T}}_{f_1}C_{f_1f_2}X_{f_2}$ is a correlation between (F^*_1, F^*_2), $\sigma^2(F^*_1) = X^{\mathrm{T}}_{f_1}C_{f_1f_1}X_{f_1}$ is variance of (F^*_1), and $\sigma^2(Y^*) = X^{\mathrm{T}}_{f_2}C_{f_2f_2}X_{f_1f_2}$ is a variance of (Y^*). To satisfy the constraint $\sigma^2(F^*_1) = \sigma^2(F^*_2) = 1$, the optimization problem of the $c(F^*_1)$ and $c(F^*_2)$ is given in [32] via Lagrange multipliers.

At the end, as introduced in [31], the transformed features are computed by concatenation as

$$Z = \begin{pmatrix} F^*_1 \\ F^*_2 \end{pmatrix} = \begin{pmatrix} X_{F_1} & 0 \\ 0 & X_{F_2} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}, \tag{4}$$

or summation as

$$Z = F^*_1 + F^*_2 = \begin{pmatrix} X_{F_1} \\ X_{F_2} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}. \tag{5}$$

This type of concatenating features is termed as canonical correlation discriminant (CCD). The CCA fuse different layer features to construct representative features. However, CCA approach was not considered the relation or the correlation between various features and thus in this paper, we interested to optimize the correlation between features before concatenation. Recently, to circumvent CCA weakness, [32] introduced the discriminant correlation analysis (DCA). We give the brief idea of DCA as follows:

Discriminant feature fusion: Assume that E represents an input data, k is the number of classes in E, F_1 denotes features obtained from the data E and $f_{1_{ij}} \in F_1$ represents the features extracted from the ith image of the jth class. Also, let $\hat{f}1_i = \frac{1}{n_i}\sum_{j=1}^{n} f_{1_{ij}}$ and $\hat{f}_1 = \frac{1}{n_i}\sum_{i=1}^{k} n_i f_{1_i}$ represent the mean of the $f_{1_{ij}}$ under the ith category and the whole feature set, respectively, with n denotes the number of data that belong to the ith class. In [33], the intermediate-class scatter matrix is formulated as:

$$C_{xf_{1(a\times b)}} = \Omega_{xf_1}\Omega^{\mathrm{T}}_{xf_1}, \tag{6}$$

where

$$\Omega_{xf_{1(a\times k)}} = [\sqrt{n_1}(\hat{f}_{11} - \hat{f}_1), \sqrt{n_2}(\hat{f}_{12} - \hat{f}_1), ..., \sqrt{n_k}(\hat{f}_{1k} - \hat{f}_1)]. \tag{7}$$

According to [32], the substantial eigenvectors of $(\Omega_{xf_1}\Omega^{\mathrm{T}}_{xf_1})_{a\times a}$ can be calculated by mapping the eigenvector of $(\Omega^{\mathrm{T}}_{xf_1}\Omega_{xf_1})_{k\times k}$. To greatly isolate the categories, matrix $(\Omega^{\mathrm{T}}_{xf_1}\Omega_{xf_1})$ should be a diagonal matrix [33] as:

$$L^{\mathrm{T}}(\Omega^{\mathrm{T}}_{xf_1}\Omega_{xf_1})L = \Gamma, \tag{8}$$

where L is positive definite eigenvalues which are in the main diagonal matrix, Γ is the eigenvectors matrix and is orthogonal, and m is the non-zero eigenvectors matrix that is denoted by $\omega_{(k\times m)}$ and is the highest values. Given a matrix L, we can have

$$\omega^{\mathrm{T}}(\Omega^{\mathrm{T}}_{xf_1}\Omega_{xf_1})\omega = \Gamma_{(m\times m)}. \tag{9}$$

As suggested in [32], the substantial eigenvectors of C_{xf_1} can be computed by mapping $\Omega \to \Omega_{xf_1}\omega$ as:

$$(\Omega_{xf_1}\omega)^{\mathrm{T}}C_{xf_1}(\Omega_{xf_1}\omega) = \Gamma_{(m\times m)}, \tag{10}$$

$y_{xf_1} = \Omega_{xf_1}\omega\Gamma^{-1/2}$ is the transformation of the matrix F_1 that unitizes C_{xf_1} and decreases the dimensionality of F_1 from a to m where

$$X^{\mathrm{T}}_{xf_1}C_{xf_1}X_{xf_1} = Id, \tag{11}$$

$$F'_{1(m\times n)} = X^{\mathrm{T}}_{xf_{1(m\times a)}}F_{1_{(a\times n)}}, \tag{12}$$

similarly for the second feature F_2, the transformation matrix X_{xf_2} is calculated, which unitizes C_{xf_2} and decreases the dimensionality of F_2 from b to r.

$$X^{\mathrm{T}}_{xf_2}C_{xf_2}X_{xf_2} = Id, \tag{13}$$

or

$$XF'_{2(m\times n)} = X^{\mathrm{T}}_{xf_{2_{m\times b}}}F_{2_{(b\times n)}}, \tag{14}$$

where m is the feature length of the changed features

$$m \leqslant \min(k - 1, rank(F_1), rank(F_2)). \tag{15}$$

Now, at this stage our goal in this work is to optimize (clearly maximization) the correlation between features throughout the two feature sets. To this end, we need to diagonalize the intermediate-set covariance matrix of changed feature, $C'_{f_1f_2} = F'_1F'^{\mathrm{T}}_2$. According to [32], it can be achieved by singular value decomposition

$$C'_{f_1f_{2(m\times m)}} = V\Phi U^{\mathrm{T}}V^{\mathrm{T}}C'_{f_1f_2}U = \Phi, \tag{16}$$

setting $X_{kf_1}V\Phi^{-1/2}$ and $X_{kf_2} = U\Phi^{-1/2}$, we have

$$V(\Phi^{-1/2})^{\mathrm{T}}C'_{f_1f_2}(U\Phi^{-1/2}) = Id, \tag{17}$$

which unitizes matrix $C'_{f_1f_2}$ and this equation has the same form with 14. Thus, similar to Eq. (15), the transformation of the feature set can be written as,

$$F_1^* = X_{kf_1}^{\mathrm{T}}F'_1 = X_{kf_1}^{\mathrm{T}}X_{xf_1}^{\mathrm{T}}F_1 = X_{f_1}F_1, \tag{18}$$

$$F_2^* = X_{kf_2}^{\mathrm{T}}F'_2 = X_{kf_2}^{\mathrm{T}}X_{xf_2}^{\mathrm{T}}F_2 = X_{f_2}F_2. \tag{19}$$

Finally, after applying this transformation, using desecrate correlation analysis approach, fusing of features can be executed by either summation or concatenation. Though the use of summation approach provides features with lower dimensions and also reduces computational time, we found that this approach provides less accuracy than the concatenation approach, and thus we used concatenation approach instead of summation in this work. That is, first, we used feature sets of the last convolutional layers of the two paths and compute the transformation using Eqs. (18) and (19). Then, we apply concatenation approach on the transformed features. We named the concatenation layer considering the traditional approach (without the introduced transformation) as primary concatenation layer and the one considering DCA approach as secondary concatenation layer, and they are discussed as follows.

3.3.2 Primary concatenation layer

Basically the proposed network constitutes primary and secondary concatenation layer where both concatenation layer is independent of one another. The secondary concatenation layer is employed with discriminant correlation analysis and explained in the next section. To explicitly average the effect of receptive field size, the primary concatenation layer concatenates the output of the fourth convolutional layer of the first path and the output of the third convolutional layer of second path together. Next, convolution operation is once applied on the concatenated features with 192 filters where each filter is of size 114×114. The output of this last convolution is then followed by soft max classifier for the final Cancer nodule estimation as shown in Table 2. Compared to the traditional CNN detection models that have only one path, the computational complexity of this model is higher, however, it achieved better detection accuracy. In our future work, we will focus toward this challenge.

3.3.3 Secondary concatenation

As we have explained earlier, the secondary concatenation is based on discriminative correlation analysis presented above. This layer helps us to select more representative features from the first and the second path. Contrary to the primary concatenation layer which directly concatenate features of the fourth convolutional layer of the first path and features of the third convolution layer of second path, the secondary concatenation layer concatenates the transformed features where the transformation is based on discriminant correlation analysis as in Eqs. (18) and (19). Similar to the primary concatenation layer, convolution operation is once applied on the concatenated features with 192 filters where each filter is of size 114×114. The output of this last convolution is then followed by soft max classifier for the final Cancer nodule estimation. In general, the employment of discriminant correlation analysis in two path model improves the efficiency of convolutional networks and at the same time it effectively models nearby pixels in the Cancer detection challenges.

3.4 Training stage

In this section, first we give the general training procedure and then, to further deal with the improvement of the proposed model, we consider two training stage. We call the first training phase, first-training stage, and the second training phase, re-training stage. We give the details of both training stages one by one as follows.

3.4.1 General training rule

To train a convolutional network for a detection task, a common optimization rule used by many researchers is either maximizing the probability of all labels in training set or minimizing the probability of the negative logarithm.

$$-\log p(M|N) = \sum_{i,j} -\log(M_{i,j}|N), \tag{20}$$

for every labeled lung slices.

Through frequent selection of labels $M_{i,j}$ at an arbitrary subset of patches within each lung, we employed stochastic gradient method. Instead of processing the whole patches for parameter updates, we made an updates only on some subset of patches. This method makes the training process faster; at the same time, it provides adequate updates for learning. Above all, this approach is executed through creating mini batches dataset of smaller slices of lungs image patches, matching with the similar detection label as the objective.

It has been proved by researchers that the momentum method has been fruitful [21], we used it to improve our model. We used this method as

$$v_{i+1} = \eta * v_i - \gamma * \nabla W_i, \tag{21}$$

$$W_{i+1} = W_i + v_{i+1}, \tag{22}$$

where W_i is the CNN parameters at step i, η is the momentum, ∇W_i is the gradient at W_i, v_i is the combined velocity and γ is the learning rate. The value of initial and final momentum η is initialized to 0.05 and 0.009, respectively. Finally, in our model, we set γ=0.0005 and it is decayed exponentially with decay factor 0.1.

3.4.2 First and re-training stage

Earlier in the introduction part, we pointed that intra-dataset label imbalance in medical images field is one of the great challenges. The reason is that deep learning models are highly benefitted from large volume of data. As a result, in imbalanced class label, a model highly favors the higher target during training, while disfavoring a small target. For example, in KDSB 2017 challenge, about 26 % of the patient are labelled as 1 (Cancer) and about 74% of the patients are labelled as 0 (free of cancer). Figure 8 shows this class imbalance problem. During training, a patch having 0 label is highly favored because of their size in the dataset, while the patch having label 1 is disfavored by the

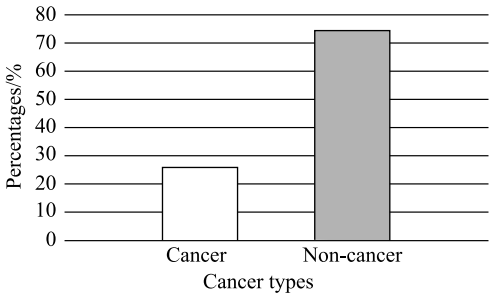

Fig. 8 The number of cancer to non-cancer in KDSB 2017 challenge

model. However, the overall result disturbs the training accuracy of CNN models.

To solve this, we increase the lower target by data augmentation technique until the number of lower target patches (those labelled as one) are equivalent to higher target patches (those labelled as zero). Data augmentation is first proposed by [34] and is highly used technique in natural image segmentation. We call this training stage first training stage.

Then, considering the original dataset (i.e., the imbalanced dataset), we apply a fine tuning strategy and train the output layer only. That is, by maintaining the optimized parameters of the first training phase (except the parameters of the output layer), we only train the output layer. We refer to this training stage as re-training stage. In this manner, we obtain the best outcome.

4 Experimental results and discussion

4.1 Experimental setup

We used Tensor flow [35] with Kera backend for the implementation of our proposed method. Tensor flow backend with Kera is one of the deep learning library that support graphical processing unit (GPU). The use of Tensor flow backend with GPU is to speed up a machine learning algorithm.

The number convolutional layer of the first path was set to four and the number of convolutional layer of the second path was set to three. By conducting various experiments (by altering number of hidden layer sizes), we fixed the depth of our model to the existing size. During experimentation, we identified that appending additional layers to both first and second path was not improved our results. Also, we identified that reducing the number of layers from both first and second path was not improved our results. In this way, we fixed the size of the proposed model to the existing size.

We also fix the parameters of each layer as shown in Table 2. The parameters shown in this table is the one for which our model achieved best result on validation set. For both paths, each pooling and convolutional layers' stride were set to 1. It benefits us to maintain the per-pixel accuracy. We initialize the filters of all layers randomly from the uniform distribution U(–0.005, 0.005) except the parameters of the softmax layer. For the softmax layer, we initialize the filter to the log of the label. Finally, the bias of our network was set to zero.

4.2 Performance evaluation

Many medical image detection performance evaluation has been proposed by researchers. Among those metrics, accuracy, recall, and specificity are commonly employed. We also employed these metrics to see how our model well perform on the testing set. These metrics are defined as follows

$$S = \frac{TP}{TN + FP}, \tag{23}$$

$$R = \frac{TP}{TP + FN}, \tag{24}$$

$$A = \frac{TP + TN}{TP + FP + FN + TN}, \tag{25}$$

where S, R, and A shows specificity, recall, and accuracy, respectively and TP, FP, FN, and TN, denotes the number of true positives, false positives, false negatives, and true negative, respectively.

4.3 Two-path CNN

Most convolutional neural networks model that are designed to detect lung cancer is in feed forward manner, which we refer as one-way path. Apart from those methods, the proposed DFD-Net has two pathways. These paths aimed to better estimate local and global dependencies of the nearby pixels, where the first path aimed to model the details and the second path aimed to model the contextual information. Experimental results reveal that our motivation on both context and detail greatly achieved better accuracy compared to the recent lung cancer detection work.

To reveal how the combination of the first path and second path is so relevant in detecting lung cancer, we provide the performance results of each path and results of the two path with both primary concatenation layer and secondary concatenation layer, independently. As the label class imbalance challenge is also considered in this work with first and retraining technique presented earlier in Section 3.4.2, we provide the results of our model with both first and retraining stage. To make our work more visible, we denote our proposed convolutional neural network made of first path alone as DFD-f and the one made of the second path alone as DFD-s. Also, we denote our convolutional network model with primary concatenation layer as DFD-tp and the overall two path CNN with secondary concatenation layer as DFD-Net. The superscript "s" append to those notation indicates the retraining stage.

Table 4 illustrates the results of various version of our models. As one can observe from the table, the results of the DFD-Net, DFD-tp, and DFD-f are given with both single stage and retraining stage, and the results of DFD-s with only retraining phase is presented. Accordingly, the first path with retraining stage (DFD-f^s) ranked as third performer model, however, using the first training stage, the model (DFD-f) ranked as seventh. Moreover, with retraining stage, DFD-f^s has gained 18.2% recall, 16.8% specificity and 17.8% accuracy values over DFD-f. On the other hand, if we see the results of DFD-tp and DFD-Net, it achieved less accuracy, specificity and recall values than DFD-f^s, but after the retraining stage, DFD-Net and DFD-tp ranked first and second, respectively. This shows how the introduced retraining stage tactic is so relevant for the class imbalance problem, and the combination of the two path is so significant to alleviate lung Cancer detection challenges. If we consider the effect of the primary (i.e., without DCA) and secondary (i.e., with DCA) concatenation layers, DFD-Net^s provides better performance. More specifically, without DCA feature fusion strategy, i.e., with only primary concatenation layer,

Table 4 DFD-Net model and its variation performance

Methods	R	S	A
DFD-Net^s	**0.874**	**0.891**	**0.878**
DFD-tp^s	0.865	**0.891**	0.871
DFD-f^s	0.855	0.851	0.854
DFD-s^s	0.830	0.821	0.828
DFD-Net	0.789	0.782	0.788
DFD-tp	0.768	0.762	0.766
DFD-f	0.673	0.683	0.676

our model has still shown better performance, however with DCA strategy, the overall performance of the model is boosted. As shown in Table 4, DFD-Net^s has gained about 0.9% recall and 0.7%accuracy values over DFD-tp^s. Overall, DFD-Net^s is the best performer model. This shows how the proposed DCA feature strategy is important in our model.

To make our best performer model more visible, we have shown the Confusion Matrix results in Table 5. The proposed method is also compared with some other recent state of the art lung cancer detection methods. Table 6 shows the results of our best performer model compared to other state of the art lung cancer detection approaches. As shown in the table, DFD-Net^s achieved better accuracy, specificity and recall values. It achieved 0.014 accuracy, 0.016 recall, and 0.010 specificity values over [6]. For the rest mentioned approaches in the table, it achieved higher accuracy, recall, and specificity values. The second performer model DFD-tp^s, even achieved better recall, accuracy, and specificity values over [22, 24, 36]. Those methods are different compared to our two path network in many ways, for example, those models follows one way convolutional networks. To make more feasible comparison, we have implemented DR-Net and DCA method in [22, 24, 36] and compared its results with our proposed method. Figure 9 shows the results achieved after DR-Net and DCA strategy implementation. The result shows that still our proposed network outperformed those methods. In some methods, (e.g., [6, 36]), there is some improvement with DR-Net and DCA, and in some of the methods, (e.g., [22, 24]), the performance of the model reduced with DR-Net and DCA strategy. We believe this is for various reasons (e.g., the difference of the network structure, the retraining strategy employed, the employed preprocessing technique, and many other).

Table 5 Confusion matrix results of DFD-Net using test image

Actual class	Predicted	
	Cancer	Non-cancer
Cancer	**0.8746**	0.1254
Non-cancer	0.1089	**0.8911**

Table 6 Contestation of our best performer model DFD-Net^s with other methods

Methods	R	S	A
DFD-Net^s	**0.874**	**0.891**	**0.878**
Kuan et al. [6]	0.858	0.881	0.864
Chon et al. [22]	0.840	0.841	0.840
Rao et al. [24]	0.815	0.801	0.811
Huang et al. [36]	0.724	0.742	0.728

4.4 The preprocess stage

As we have mentioned earlier, a Computed Tomography image of lung is composed of many substances. These substances are not important to detect lung cancer from lung image. To see its effect, firstly, we experiment the proposed model with KDSB without conducting any preprocessing. However, the result obtained was very low. Secondly, we removed non relevant substances such as bone, blood, water, and trained our best performer model (DFD-Net^s), again the achieved results were not satisfactory. Thirdly, as most lung cancer detection models, we perform lung segmentation and nodules detection from raw lung images. We applied thresholding to segment the lung, and then, we used U-Net to detect suspicious nodules.

Table 7 shows the results of the proposed DFD-Net^s when preprocessing (lung segmentation and suspicious nodule detection) are applied and preprocessing are not performed. From the table, one can observe that the performance of DFD-Net^s boosted when lung segmentation and suspicious nodule detection is considered. The proposed DFD-Net^s trained with the preprocessing stage has 0.101 accuracy, 0.218 specificity, and 0.105 recall values than DFD-Net^s that is not trained with the aforesaid preprocessing. Similar result has been achieved with DFD-tp^s. In general, we observed that the use of segmentation via thresholding and suspicious nodule detection via U-Net as a preprocessing strategy were enhanced the DFD-Net^s accuracy.

4.5 Analysis of DR-Net impact

The denoising network used (DR-Net) is the recent model in-

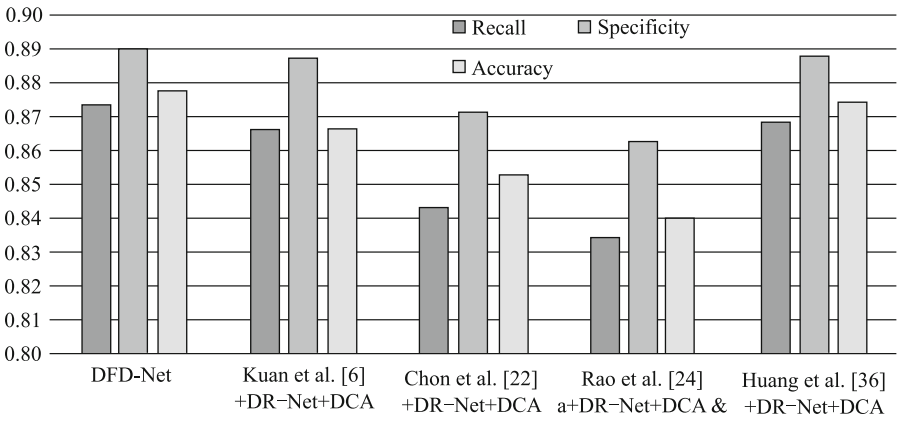

Fig. 9 Comparison of our method with other methods (when DR-Net and DCA strategy is employed in [6, 22, 24, 36])

Table 7 DFD-Net and DFD-f^s performances with preprocessing (wpp) and without preprocessing (wtpp) stage

Model	A		S		R	
	wpp	wtpp	wpp	wtpp	wpp	wtpp
DFD-Net^s	**0.874**	0.773	**0.891**	0.673	**0.878**	0.773
DFD-f^s	**0.865**	0.643	**0.891**	0.663	**0.871**	0.623

troduced by [9]. It used residual learning approach using CNN to denoise medical images. Particularly, it was designed for lung images denoising. According to this paper, it outperformed the well know image denoising methods like BM3D and KSVD. Thus, we prefer this model to use in our DFD-Net to first denoise lung CT scan image for cancer detection purpose.

In medical image analysis, interpretation, and detection task, the first step to be considered is that noise removal or reduction. It makes the result of those tasks more realistic, otherwise, those tasks are complicated [13]. Therefore, denoising strategy is very crucial step for the other tasks and also irrespective of those tasks. To this end, we used DR-Net to first denoise lung CT scan image of lung before the detection task. However, during our experimentation, we observed that the overall result of DFD-Net is affected by the use of DR-Net compared to the results of DFD-Net without DR-Net. The reason is that during denoising CT scan image of lung with DR-Net, we lost some details about an images. More or less, the result obtained is promising. Without using DR-Net, we obtained better accuracy results. It has about 0.125 accuracy higher. In our future work, we work on it to obtain the same or greater accuracy value with the denoised images.

4.6 The subject of morphology

To see how our model is robust enough toward a nodule size and shape variations, we identified 300 images and test our best performer model. The experimental results obtained are shown in Table 8. Among the 300 images, 150 images has relatively bigger sizes (in diameter) and has multi shapes, and the rest 150 images has relatively smaller nodules size (in diameter) with inconsistent shape. All these images are taken from KDSB 2017 challenge.

From Table 8, one can observe that of 150 larger nodule nominated, only 3.33% were wrongly predicted with our proposed DFD-Net^s, and of 150 smaller nodules nominated, only 4.66% were wrongly predicted. From the experiment, we observed that DFD-Net^s best perform on larger nodules than on smaller nodules. This observation is also valid when the nodules is screened by an expert, i.e., when nodule screening is made by an expert, the accuracy of larger nodules is higher than that of smaller nodules. Contested to the other state of the art, our best performer model DFD-Net^s has achieved better performance results. For example, while [6] predicts 95.33% cancer nodules correctly, DFD-Net^s predicts 96.66% cancer nodules correctly, which is about 1.33% larger. Overall, the proposed DFD-Net better address the contextual and morphological challenge during lung cancer detection is conducted.

5 Conclusion

This paper presented the detection of lung cancer from the denoised Computed Tomography images of lung with two path CNN. Different architecture has been considered aiming to obtain best lung cancer detection results. The conducted Experiment on Kaggle Data Science bowl 2017 challenge with our proposed DFD-Net revealed better lung cancer detection results than the recently introduced approach on accuracy, sensitivity and specificity.

The first training and retraining strategy proposed has improved the model accuracy. We also suggest this strategy can be used in other detection tasks where there is a problem of class imbalance within the data. Since we used GPU technology, we found the time the model cost during training is reasonable, where the test time is so faster even with CPU.

Table 8 Morphological context performance of our model and its contestation with other approaches (No. S: number of samples, No.cp: number of correctly predicted, No. wp: number of wrongly predicted, %cp: percentage of correctly predicted, and %wp: percentage of wrongly predicted)

Approaches	No. S	No. cp	%cp	No. wp	%wp
DFD-Net^s	Bigger(150)	145	96.66	5	3.33
	Smaller(150)	143	95.33	7	4.66
DFD-tp^s	Bigger(150)	144	96	6	4
	Smaller(150)	143	95.33	7	4.66
Kuan et al. [6]	Bigger(150)	143	95.33	7	4.66
	Smaller(150)	141	94	9	6
Chon et al. [22]	Bigger(150)	140	93.33	10	6.66
	Smaller(150)	141	94	9	6
Rao et al. [24]	Bigger(150)	139	92.66	11	7.33
	Smaller(150)	136	90.66	14	9.33
Huang et al. [36]	Bigger(150)	134	89.33	16	10.66
	Smaller(150)	132	88	18	12
Kuan et al. [6] +DR-Net+DCA	Bigger(150)	144	96	6	4
	Smaller(150)	143	95.33	7	4.66
Chon et al. [22] +DR-Net+DCA	Bigger(150)	139	92.66	11	7.33
	Smaller(150)	139	92.66	11	7.33
Rao et al. [24] +DR-Net+DCA	Bigger(150)	134	89.33	16	10.66
	Smaller(150)	132	88	18	12
Huang et al. [36] +DR-Net+DCA	Bigger(150)	139	92.66	11	7.33
	Smaller(150)	136	90.66	14	9.33

Feature fusion strategy called DCA is introduced to combine more important features. We examined the new transformation with concatenation and this strategy improved the model performance greatly.

Overall, contested to the recently proposed lung Cancer detection approaches with convolutional neural networks, our proposed DFD-Net can be easily adopted to variation among nodules size and also it can be easily adopted for other medical image detection task even with the denoised data that loses some details during denoising process.

Acknowledgements This work was partially funded by the national Key research and development program of China (2018YFC0806802 and 2018YFC0832105) and Bule Hora University of Ethiopia. We would like to acknowledge the editors and the anonymous reviewers whose important comments and suggestions led to greatly improved the manuscript.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel R L, Torre L A, Jemal A. Global cancer statistics 2018. A Cancer Journal for Clinicians, 2018, 68(6): 394–424
2. National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. New England Journal of Medicine, 2011, 365(5): 395–409
3. Patz E F, Pinsky P, Gatsonis C, Sicks J D, Kramer B S, Tammemagi M C, Chiles C, Black W C, Aberle D R. Over diagnosis in low-dose computed tomography screening for lung Cancer. JAMA Internal Medicine, 2014, 174(2): 269–274
4. Alvarez J M, Gevers T, LeCun Y, Lopez A M. Road scene segmentation from a single image. In: Proceedings of the 12th European Conference on Computer Vision. 2012, 376–389
5. Liu Y, Gadepalli K, Norouzi M, Dahl G E, Kohlberger T, Boyko A, Venugopalan S, Timofeev A, Nelson P Q, Corrado G S, Hipp J D. Detecting cancer metastases on giga pixel pathology images. 2017, arXiv preprint arXiv: 1703. 02442
6. Kuan K, Ravaut M, Manek G, Chen H, Lin J, Nazir B, Chen C, Howe T C, Zeng Z, Chandrasekhar V. Deep learning for lung cancer detection: tackling the kaggle data science bowl 2017 challenge. 2017, arXiv preprint arXiv: 1705. 09435
7. Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio Y, Pal C, Jodoin P M, Larochelle H. Brain tumor segmentation with deep neural networks. Medical Image Analysis, 2017, 35:18–31
8. Pereira S, Pinto A, Alves V, Silva C A. Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Transactions on Medical Imaging, 2016, 35(5): 1240–1251
9. Jifara W, Jiang F, Rho S, Cheng M, Liu S. Medical image denoising using convolutional neural netwok: a residual learning approach. Journal of Super Computing, 2019, 75(2): 704–718
10. Razzak M I, Naz S, Zaib A. Deep learning for medical image processing: overview, challenges and future. Classification in BioApps: Automation of Decision Making, 2017, 26: 323
11. Clark M C, Hall L O, Goldgof D B, Velthuizen R, Murtagh F R, Silbiger M S. Automatic tumor segmentation using knowledge-based clustering. IEEE Transaction on Medical Imaging, 1998, 17(2): 187–201
12. Lin D T, Yan C R. Lung nodules identification rules extraction with neural fuzzy network. In: Proceedings of the 9th International Conference on Neural Information Processing. 2002, 2049–2053
13. Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. In: Proceedings of Advances in Neural Information Processing Systems. 2015, 91–99
14. Redmon J, Farhadi A. Yolo: better, faster, stronger. 2016, arXiv preprint arXiv:1612.08242
15. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C Y, Berg A C. SSD: single shot multi box detector. In: Proceedings of European Conference on Computer Vision. 2016, 21–37
16. Ronghang H, Piotr D, Kaiming H, Trevor D, Ross G. Learning to segment everything. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018, 4233–4241
17. Wu Y, He K. Group normalization. In: Proceedings of the European Conference on Computer Vision. 2018, 3–19
18. Jiang X, Pang Y, Sun M, Li X. Cascaded sub patch networks for effective cnns. IEEE Transactions on Neural Networks and Learning Systems, 2017, 29(7): 2684–2694
19. Mobiny A, Van Nguyen H. Fast capsnet for lung cancer screening. In: Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention. 2018, 741–749
20. Sori W J, Feng J, Liu S. Multi-path convolutional neural network for lung cancer detection. Multidimensional Systems and Signal Processing, 2019, 30(4): 1749–1768
21. Gurcan M N, Sahiner B, Petrick N, Chan H P, Kazerooni E A, Cascade P N, Hadjiiski L. Lung nodule detection on thoracic computed tomography images: preliminary evaluation of a computer-aided diagnosis system. Medical Physics, 2002, 29(11): 2552–2558
22. Chon A, Balachandar N, Lu P. Deep convolutional neural networks for lung cancer detection. Standford University, 2017
23. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention. 2015, 234–241
24. Rao P, Pereira N A, Srinivasan R. Convolutional neural networks for lung cancer screening in computed tomography (CT) scans. In: Proceedings of International Conference on Contemporary Computing and Informatics. 2016, 489–493
25. He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: surpassing human-level performance on imageNet classification. In: Proceedings of IEEE International Conference on Computer Vision. 2015, 1026–1034
26. Kingma D P, Ba J. Adam: a method for stochastic optimization. 2014, arXiv preprint arXiv: 1412. 6980
27. Vedaldi A, Lenc K. Matconvnet: convolutional neural networks for matlab. In: Proceedings of the 23rd ACM International Conference on Multimedia. 2015 689–692
28. Liu C, Wechsler H. A shape-and texture-based enhanced fisher classifier for face recognition. IEEE Transaction on Image Process, 2001, 10(4): 598–608
29. Yang J, Yang J Y. Generalized K-L transform based combined feature extraction. Pattern Recognition, 2002, 35(1): 295–297
30. Yang J, Yang J Y, Zhang D, Lu J F. Feature fusion: parallel strategy vs. serial strategy. Pattern Recognition, 2003, 36(6): 1369–1381
31. Sun Q S, Zeng S G, Liu Y, Heng P A, Xia D S. A new method of feature fusion and its application in image recognition. Pattern Recognition, 2005, 38(12): 2437–2448
32. Schott J R. Principles of multivariate analysis: a user's perspective. Journal of the American Statistical Association, 2002, 97(458): 657–659
33. Haghighat M, Abdel-Mottaleb M, Alhalabi W. Discriminant correlation analysis: real-time feature level fusion for multi-modal bio-metric recognition. IEEE Transaction on Information Forensics Security, 2016, 11(9): 1984–1996
34. Krizhevsky A, Sutskever I, Hinton G E. Image net classification with deep convolutional neural networks. In: Proceedings of Advances in Neural Information Processing Systems. 2012, 1097–1105
35. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado G S, Davis A, Dean J, Devin M, Ghemawat S. TensorFlow: large-scale machine learning on heterogeneous distributed systems. 2016, arXiv Preprint

arXiv: 1603. 04467

36. Huang X, Shan J, Vaidya V. Lung nodules detection in CT using 3D Convolutional neural networks. In: Proceedings of the 14th IEEE International Symposium on Biomedical Imaging. 2017, 379–383

Worku J. Sori received the BEd degrees in Mathematics from Mada Walabu University, Ethiopia in 2009, the Master degree in Applied Mathematics from Addis Ababa University, Ethiopia in 2011 and PhD degree in computer science and technology from Harbin Institute of Technology, China in 2019. He is now an assistant professor at the School of Electrical Engineering and Computing, department of Computer science and Engineering, at Adama Science and Technology University (ASTU), Ethiopia. His research interests include medical data processing, pattern recognition, image and video processing, and large data compression.

Jiang Feng received the BS, MS, and PhD degrees in computer science from Harbin Institute of Technology (HIT), China in 2001, 2003, and 2008, respectively. He is now a full professor in the Department of Computer Science, Harbin Institute of Technology and a visiting scholar in the School of Electrical Engineering, Princeton University. His research interests include computer vision, pattern recognition and image and video processing.

Arero W. Godana received the BSc degree from Bule Hora University, Ethiopia, and MSc degrees in computer science from Harbin Institute of Technology, China in 2016 and 2019, respectively. He is now a Lecturer at Arsi University, Ethiopia. His research interests include Satellite image processing, pattern recognition and image and video processing.

Shaohui Liu received the BS. MS. and PhD degrees in computer science from Harbin Institute of Technology (HIT), China in 2000, 2002, and 2007, respectively. He is now an Associated Professor in the Department of Computer Science, Harbin Institute of Technology, China. His research interests include data compression, pattern recognition and image and video processing.

Demissie J. Gelmecha received the BSc degree in physics from Haramaya University, Ethiopia in 2006, the MSc degree in physical electronics engineering from Central China Normal University, Institute of Technology, China in 2011 and PhD degree in optical Engineering from Harbin Institute of Technology, China in 2018. Currently, he is an Assistant Professor of Optical Engineering at the Department of Electronics and Communication Engineering, School of Electrical Engineering and Computing at Adama Science and Technology University (ASTU), Ethiopia. He is also serving the University as Dean of Academic Staff Affairs. His current research interests include fiber-optic communications, non-linear chiral fibers for optical communications and new optical devices, and full-duplex communication systems.