Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2019) 27: 710 – 722

doi:10.3906/elk-1710-157

Lung segmentation in chest radiographs using fully convolutional networks

Rahul HOODA[1], **Ajay MITTAL**[2,*], **Sanjeev SOFAT**[1]
[1]Department of Computer Science and Egineering, Punjab Engineering College, Chandigarh, India
[2]University Institute of Engineering and Technology, Panjab University, Chandigarh, India

Received: 16.10.2017 • **Accepted/Published Online:** 22.08.2018 • **Final Version:** 22.03.2019

Abstract: Automated segmentation of medical images that aims at extracting anatomical boundaries is a fundamental step in any computer-aided diagnosis (CAD) system. Chest radiographic CAD systems, which are used to detect pulmonary diseases, first segment the lung field to precisely define the region-of-interest from which radiographic patterns are sought. In this paper, a deep learning-based method for segmenting lung fields from chest radiographs has been proposed. Several modifications in the fully convolutional network, which is used for segmenting natural images to date, have been attempted and evaluated to finally evolve a network fine-tuned for segmenting lung fields. The testing accuracy and overlap of the evolved network are 98.75% and 96.10%, respectively, which exceeds the state-of-the-art results.

Key words: Chest X-rays, deep learning, lung Segmentation, medical imaging

1. Introduction

Medical image analysis is an immensely active and fast-growing area that has evolved into an established discipline. Medical images are acquired using different modalities such as radiography, computed tomography (CT), and magnetic resonance imaging (MRI). Among all these, radiography is the most commonly used modality for diagnosing pulmonary and abdominal abnormalities. Chest radiography, colloquially known as chest X-ray (CXR), is used for screening various pulmonary diseases such as lung cancer, tuberculosis (TB), pneumoconiosis, and emphysema. While CXR is easy to acquire, its interpretation is extremely challenging and heavily depends on the expertise of the person interpreting it. It has been observed that there are substantial interobserver and intraobserver variations in the interpretation of CXRs. Since triaging and clinical decisions heavily depend upon CXR interpretation, it becomes essential to develop a computer-aided diagnosis (CAD) system that can automatically interpret CXRs and assist clinicians in decision making.

As stated, lung field segmentation (LFS) is a preliminary step in any chest radiographic CAD system. The LFS problem has been extensively studied since the 1970s, and Section 2 succinctly presents the related work. Despite enormous research effort, the problem has not yet been satisfactorily solved, and active research is still being pursued to develop a robust LFS method. The recent research trends have seen a paradigm shift, and deep learning techniques are now being applied to solve the LFS problem.

The resurrection of deep learning, after its conceptualization in the 1990s, is accredited to the easy availability of graphics processing units (GPUs) and large image datasets. Nowadays, most of the research in the field of medical image analysis has shifted towards it. The convolutional neural network (CNN) is the most popular deep learning technique. In recent years, various CNN-based architectures such as LeNet [1], Alexnet [2],

*Correspondence: ajaymittal@pu.ac.in

VGGnet [3], GoogleNet [4], and ZFNet [5] have been proposed to perform image classification. Researchers have also customized these architectures to perform semantic segmentation. However, their segmentation performance is not satisfactory. Thus, networks specifically designed for semantic segmentation have also been developed. These networks include the fully convolutional network (FCN) [6], Segnet [7], and U-Net [8]. Among these, FCN is the best performing architecture and is thus chosen for this study. In this paper, FCN is customized for segmenting lung fields from gray-scale CXRs.The significant contribution of this paper is in reengineering the FCN architecture, which includes several modifications to the original architecture such as augmentation of skip layers, removal of pooling layers, and addition of dropout layers. The effect of modifications when applied in isolation and conjunction on the original architecture has been evaluated to retain the effective modifications and to report the best performing architecture.

The rest of the paper is organized as follows. Section 2 briefly presents literature related to LFS methods and semantic segmentation using CNNs. Section 3 discusses the modifications in the FCN architecture to make it suitable for LFS. Performances of modified architectures for LFS are reported and compared to state-of-the-art LFS methods in Section 4. Finally, the conclusion is drawn in Section 5.

2. Related work

This section succinctly presents the relevant literature along two separate threads: lung field segmentation and semantic segmentation using a CNN.

2.1. Lung field segmentation

LFS methods presented in the literature can be broadly categorized into three categories, namely rule-based methods, machine learning-based methods, and deformable model-based methods.

i *Rule-based methods*: These methods employ heuristic rules based on lungs' characteristics such as position and texture to segment the lung field. The rules are formulated from the prior knowledge of the lung anatomy and are implemented using low-level image processing operations. These methods are flexible as the rules can be applied in different permutations and combinations to achieve the desired results. Some of the popular rule-based LFS methods were presented in [9–11].

These methods do not require annotated datasets for training and are thus unsupervised. However, these methods are fragile when the lung portion is missing or its shape is highly deformed.

ii *Machine learning-based methods*: These methods classify each pixel of the CXR image as either lung or background region using a binary classifier. The classifier is trained using the features extracted from the training dataset. Depending on how the features are extracted, these methods are further categorized as shallow learning-based methods and deep learning-based methods. In shallow learning-based methods, feature extraction is performed manually and requires extensive domain knowledge. The most significant challenge in these methods is to decide the appropriate class of features to be extracted. Some of the popular shallow learning-based LFS methods were presented in [12–15].

In deep learning-based methods, the feature extraction process is automatic and hierarchical and has multiple levels of abstraction. It more closely resembles the way the human brain does it. These methods have a deep architecture with multiple processing layers consisting of linear and nonlinear transformations. Recently, these methods have efficiently replaced shallow learning-based methods in different medical

image segmentation studies. However, the usage of deep learning techniques in the area of LFS remains relatively unexplored. LFS methods based on deep learning were presented in [16–18].

iii *Deformable model-based methods*: These methods use a model that can modify its shape according to desired objects (lungs, in this case) using internal forces and external forces. While the internal forces ensure the shape to be smooth and stretchable, the external forces enrich the model with desired image characteristics such as terminations, edges, and lines. Deformable models that have been used for segmentation of lungs are further classified as parametric models and geometric models. Some of the popular deformable model-based LFS methods were presented in [19–22].

The registration-based LFS method presented by Candemir et al. [22] is the best performing method to date. It belongs to the category of deformable model-based methods and involves a registration-driven lung boundary detection technique. A content-based image retrieval approach is used to obtain similar training images and then scale-invariant feature transform (SIFT)-flow nonrigid registration is applied to create an initial lung model. After that, a graph cut-based method is used for deformation to obtain the final segmented result.

2.2. Semantic segmentation using CNNs

Semantic segmentation is defined as understanding the image at the pixel level, which means to assign each pixel of the image to an object class. This requires high-level visual understanding of the image. Initially, approaches such as random forest-based classifiers were used for semantic segmentation. However, after deep learning took over, CNNs attained enormous success in solving segmentation problems. Current state-of-the-art methods for semantic segmentation include SegNet [7], FCNs [6], and U-Net [8].

i *SegNet*: It is a deep-layered architecture proposed by Badrinarayanan et al. for semantic pixel-wise labeling. The architecture consists of a stack of encoders and decoders. It is among the first few architectures specifically designed for semantic pixel-wise segmentation as the initial deep-learning approaches for segmentation have tweaked classification architectures such as AlexNet and VGGNet to perform segmentation. The changes in classification-based architectures include removal of fully connected layers of the architecture and enhancement in the resolution in the last layer to obtain output with the same dimensions as the input image. However, in SegNet architecture, the resolution is enhanced in a stepwise manner by transferring the max-pooling indices from the encoder part to the corresponding decoder part. This transferred information helps the decoder in mapping features from subsampled layers to the final layer.

ii *U-Net*: It is another popular architecture that makes use of an encoder-decoder network like SegNet. It is called U-Net because the architecture is shaped similarly to the letter U. The first few layers (i.e. encoder layers) perform downsampling while the later layers (i.e. decoder layers) perform upsampling. The architecture performs upsampling by passing a complete feature map to the corresponding decoder layer. It is different from SegNet, in which only max-pool indices are passed. It is specifically proposed for biomedical images where the number of annotated images is limited, and thus this architecture uses excessive data augmentation techniques. The U-Net architecture takes as input images of size 572×572 and produces an output of size 388×388. This means that it produces output of smaller size as compared to input, which is a disadvantage of this method.

iii *Fully-convolutional networks (FCNs)*: FCNs [6] are dense prediction networks that perform semantic pixel-wise segmentation. These networks are an extension of CNNs and prediction is made for each pixel individually. They apply a skip architecture that combines and takes advantage of coarse as well as fine information. In these networks, classifiers are modified to obtain dense prediction by converting fully connected (FC) layers into convolutional layers, and thus a heat map or classification map is received as output. This modification helps FCNs to take the input of different sizes and produce classification maps of varying size. These maps are then combined, and their resolution is upsampled using deconvolutional operation to receive output of the same size as input. In addition to that, the most significant difference between CNN and FCN is that FCN keeps on learning filters even in the last layer as all the layers are convolutional.

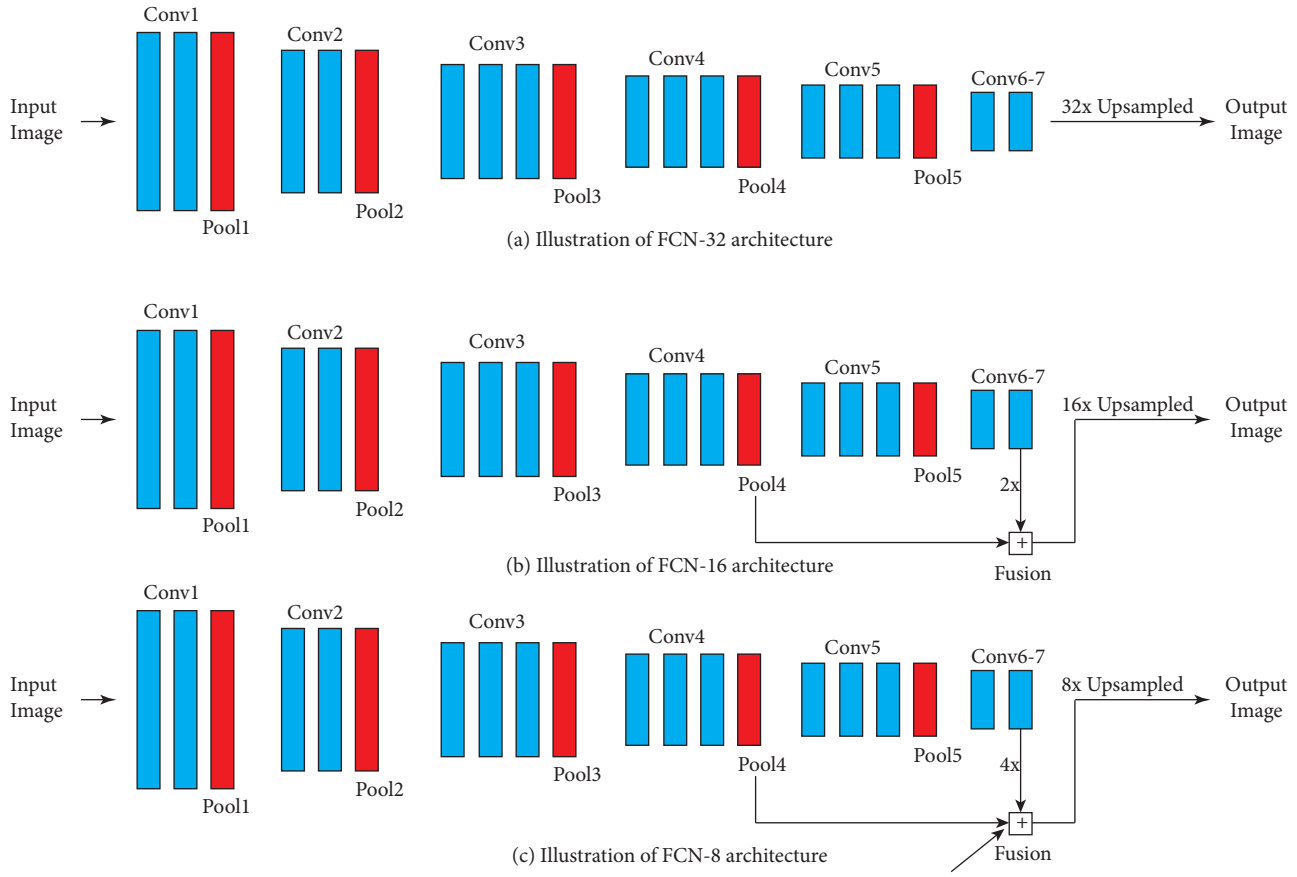

Figure 1. Illustration of different FCN architectures.

FCNs are based on VGG-16 networks and are obtained by swapping FC layers with convolutional layers having 1×1 filters for dense prediction. In the original paper, the performances of three types of FCN architectures, i.e. FCN-32, FCN-16, and FCN-8, were evaluated and compared with each other with the PASCAL VOC 2011 dataset. In FCN-32, the output of the final prediction layer needs to be upsampled 32 times and thus omits the fine details of the scene. The other two versions of FCN resolve this issue by adding skip layers. In skip connections, the output of lower layers (with finer details) is added to the final prediction layer. It turns a straight line topology into directed acyclic graph (DAG) topology,

in which edges from lower layers directly jump to higher layers while skipping intermediate layers. In FCN-16, the output of the Pool4 layer is fused with the Conv7 layer (upsampled two times). The fusion is further upsampled 16 times using deconvolutional operation to obtain the final output. Similarly, in FCN-8, the output of the Pool3 layer is fused with the Pool4 layer (upsampled two times) and Conv7 layer (upsampled four times). The fusion is further upsampled 8 times to obtain the final output. Out of all three variations, FCN-8 gives the best output with finer details and is used to achieve segmentation. The architecture details of three variations of FCNs (i.e. FCN-32, FCN-16, and FCN-8) are shown in Figure 1.

3. Proposed modifications

The study proposes the following modified architectures based on standard FCN architecture to perform LFS.

i ***FCN-4 architecture:*** The implementation of FCN-4 architecture has not been done to date. FCN-4 implementation extends the standard FCN architecture and combines the output of the Pool2 layer with the output of three more layers, i.e. the Pool3 layer (upsampled two times), Pool4 layer (upsampled four times), and Conv7 layer (upsampled eight times). The output of the fusion is further upsampled 4 times to obtain the final output of the same size as that of the input image. This addition does not increase the complexity of the network and also helps in reducing the upsampling of the fusion to half. Due to the reduced upsampling, the final output has fewer pixel-level predictions and thus constitutes fine details. The architectural detail of FCN-4 is shown in Figure 2.

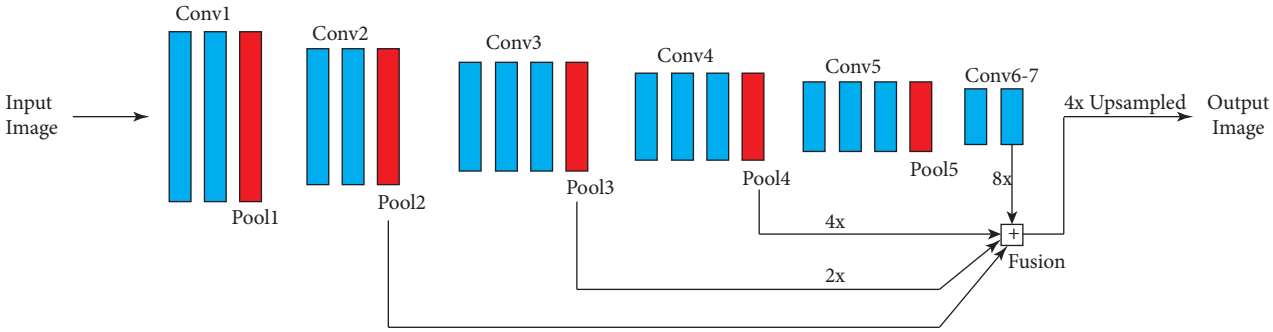

Figure 2. Illustration of FCN-4 architecture.

ii ***Architecture with dropout layers:*** There is a considerable number of parameters used in the FCN architecture. Since medical images are scarce, training a large network with no regularization would often lead to overfitting. In this architecture, a dropout layer has been added after each convolutional layer for better optimization. Addition of dropout layers is commonly used in modern deep architectures for better regularization [18]. The details of this modified architecture are shown in Figure 3.

iii ***Architecture with only conv layers:*** It is evident from the literature that usage of convolutional layers to perform downsampling instead of pooling layers can improve the performance of an architecture [23]. These changes enhance the network by introducing new parameters. In this modification, all the pooling layers have been removed and their operation is performed by the preceding convolutional layer using strides. The modified architecture is shown in Figure 4.

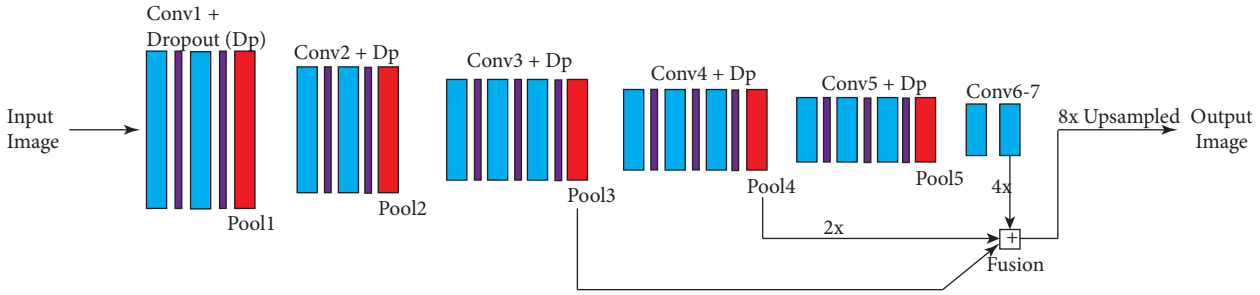

Figure 3. Illustration of modified FCN-8 architecture with added dropout layers.

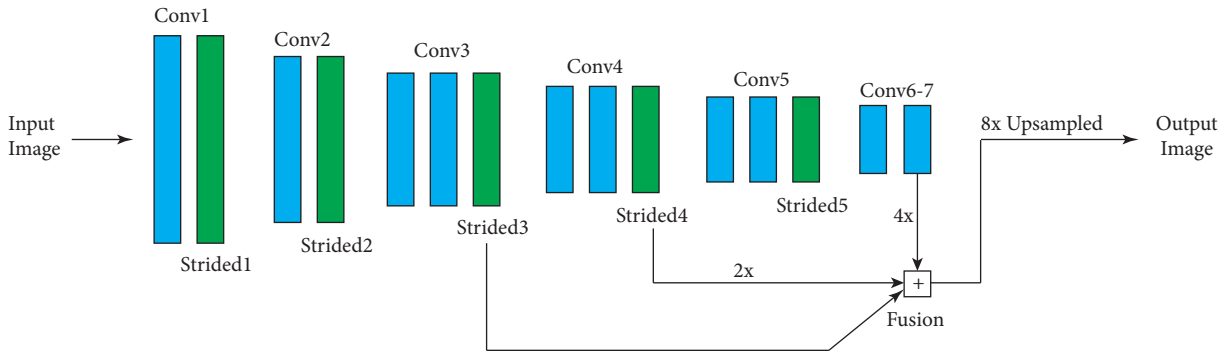

Figure 4. Illustration of modified FCN-8 architecture with only convolutional layers.

4. Experimental results

4.1. Datasets and evaluation metrics

Datasets: In this study, two datasets, namely JSRT [24] and the Montgomery dataset [25], have been used. These are the standard LFS datasets that are used in various LFS studies and are publicly available. The details of these datasets are as follows.

i *JSRT dataset:* It was created by the Japanese Society of Radiological Technology (JSRT) in collaboration with the Japanese Radiological Society. It consists of 247 posteroanterior (PA) CXRs collected from different institutions in Japan and the United States. Of all 247, 154 CXR images have lung nodules, while 93 have none. The CXR images have a size of 2048×2048 pixels and gray-scale depth of 12 bits. The dataset also provides the labeled annotations of different anatomic structures including lung.

ii *Montgomery dataset:* The dataset was created by the US National Library of Medicine (USNLM) in collaboration with the Department of Health and Human Services, Montgomery County (MC), USA. It consists of 138 PA CXRs collected via MC's tuberculosis screening program. Of all 138, 80 CXRs have been classified as normal, while the remaining 58 have manifestations of TB. The CXR images were acquired at two different spatial resolutions with gray-scale depth of 12 bits. It also includes labeled annotations of lung regions.

Evaluation metrics: LFS is a binary classification task, including the classes L,B, which stand for lung and background event, respectively, and the predicted classes l, b, which denotes predicted lung and predicted background event. There can be four possible outcomes from the classifier and they can be displayed in a 2×2 confusion matrix, as shown in Figure 5.

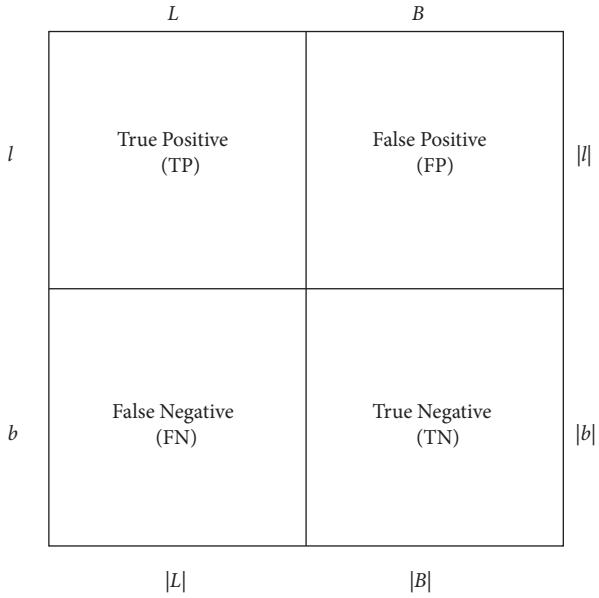

Figure 5. 2×2 confusion matrix for a lung classifier.

True positive (TP), i.e. correct prediction, is a lung pixel correctly identified as a lung pixel; false positive (FP), i.e. false alarm, is a background pixel falsely classified as a lung pixel; false negative (FN), i.e. a miss, is a lung pixel falsely classified as a background pixel; and true negative (TN), i.e. correct rejection, is a background pixel correctly classified as a background pixel.

To evaluate and compare the performance of the proposed architectures with other algorithms, two commonly applied metrics, i.e. accuracy and overlap, are used.

i *Accuracy:* It is defined as the ratio of correct predictions to the total number of predictions made by the classifier. It can be obtained by using Eq. 1.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

ii *Overlap:* Also known as the Jaccard similarity coefficient, it is the ratio of the area of intersection to the area of union between the ground truth image (G) and the segmented image (O), and it can be determined by using Eq. 2.

$$Overlap = \frac{|G \cap O|}{|G \cup O|} = \frac{TP}{TP + FP + FN} \tag{2}$$

The accuracy measure is to be cautiously used while analyzing the quality of a predictive model as it suffers from the accuracy paradox. This measure is used with other evaluation metrics such as precision, overlap, and recall to determine the quality of a predictive model. Therefore, in this study, both accuracy and overlap have been used as the evaluation metrics.

4.2. Training model

To perform training, the datasets have been divided into training and testing portions. In this study, two sets of experiments have been performed. Experiment 1 is conducted using the JSRT dataset. In this experiment, 186 images out of the total 247 images are used for training and the remaining 61 images are used for testing. Since the dataset has a smaller number of images, the chances of overfitting are high. To avoid overfitting, data augmentation is performed on the training images. For each training image, eight images are added to the training set. Out of these eight augmented images, three have been obtained by rotating the image by 90°, 180°, and 270°. Horizontal and vertical flips are performed to obtain another two images. The rest of the images are obtained by performing random cropping. Experiment 2 is performed using both the datasets. In this experiment, the complete JSRT dataset with augmentation is used for training and the Montgomery dataset is used for testing to determine the generalizing capabilities of the architectures.

Loss function: Let the input training dataset be denoted by $Tr = \{(X_n, Y_n), n = 1, ..., N\}$, where X_n denotes the raw input image, Y_n denotes the corresponding binary ground truth, and N denotes the number of images in the training dataset. Each layer of the architecture has a set of weights or parameters. We denote the complete set of parameters as W. The parameters are initially set to random values. In each iteration, the predicted output, PRY_n, is obtained using the parameters. Based on Y_n and PRY_n, the loss function of the architecture L is calculated using the following formula:

$$L = \frac{1}{N}\left(\sum_{n=1}^{N}(Y_n * log(PRY_n))\right) \tag{3}$$

The objective is to minimize the loss function L. The evaluated value is used by the optimization algorithm to update the parameters W in the next iteration. This process happens in each iteration and finally optimized parameters are obtained.

4.3. Results and discussion

In this section, the performances of default FCN architectures and modified architectures are reported and compared with each other. The results of Experiment 1 are as follows.

Default FCN architecture: Out of all three default architectures, FCN-8 performs the best as it obtains auxiliary outputs from two different pooling layers and requires small upsampling after the fusion. On the other hand, FCN-32 performs worst as it does not perform any fusion using the output from previous layers and upsampling at the end of the network is also very high. Thus, the FCN-32 network's segmentation results are coarse while FCN-8 includes the fine details around the lung boundary, which enhances its performance. The training performance and loss convergence on the JSRT dataset are shown in Figure 6. The performances of different networks, as shown in Figure 6, are quite close to each other and hence clear distinction can be made only by the testing performance achieved by the architectures. The testing accuracy of FCN-32, FCN-16, and FCN-8 networks are 97.19%, 97.48%, and 98.51%, respectively, while the testing overlaps of these networks are 91.37%, 93.92%, and 94.76%, respectively. These results indicate that inclusion of skip layers improves the performance of the network.

FCN-4 architecture: Since the fusion of different layers improves the performance and incorporates fine details in the segmented output, a network in which the output of an additional previous layer is included at the final layer is created. The output of the second pooling layer is added to the fusion, which reduces the final upsampling to 4 times only. Due to this, the performance of the modified architecture is significantly improved

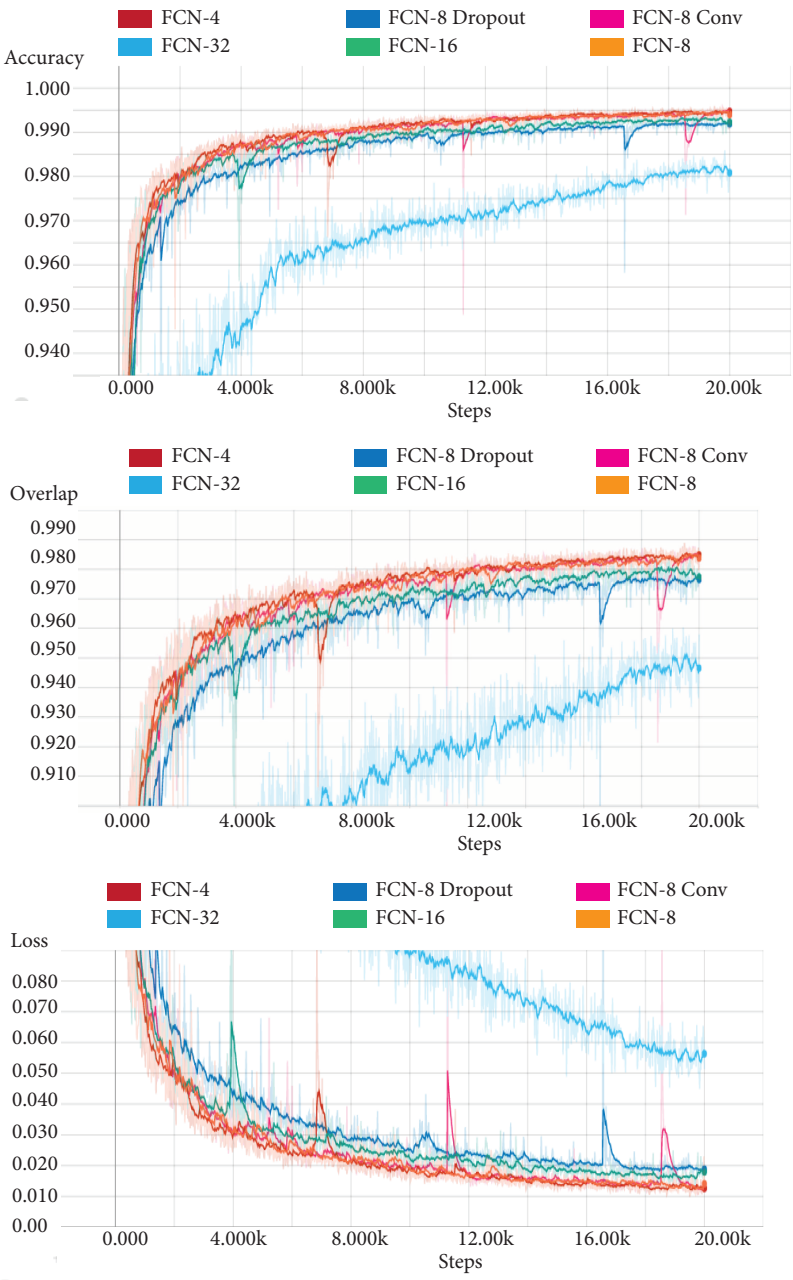

Figure 6. Training performance of the different architectures' (a) accuracy, (b) overlap, and (c) loss.

and crossed the state-of-the-art performance for LFS. The training performance and loss convergence on the JSRT dataset are shown in Figure 6. The testing accuracy and overlap of this network are 98.75% and 96.10%, respectively.

Architecture with dropout layers: As each FCN network has a large number of parameters, regularization is needed in the network to avoid overfitting. Dropout layers are thus sandwiched between all pairs of convolutional layers in the FCN-8 network to provide better optimization. The addition of dropout layers lowered the training performance; however, as observed, it slightly improves the testing performance. This observation suggests that the introduction of dropout layers regularizes the network and enhances its perfor-

mance. The training performance and loss convergence on the JSRT dataset are shown in Figure 6. The testing accuracy and overlap of this network are 98.61% and 95.07%, respectively.

Architecture with only conv layers: In this architecture, all the pooling layers have been removed from the FCN-8 architecture and downsampling operation is performed using strided convolution. The performance of this network is slightly better as compared to the FCN-8 architecture. It happens because the strided convolution introduces new parameters in the network. The training performance and loss convergence on the JSRT dataset are shown in Figure 6. The testing accuracy and overlap of this network are 98.54% and 95.44%, respectively.

Two different architectures in which all the modifications are applied in conjunction have also been evaluated. One such architecture includes the FCN-4 network with dropout layers and pooling layers removed, and another architecture includes the same changes applied to the FCN-8 network. As listed in the Table, the testing performance of these networks is slightly lower than the performance of the FCN-4 network but better than the performance of the standard FCN-8 network.

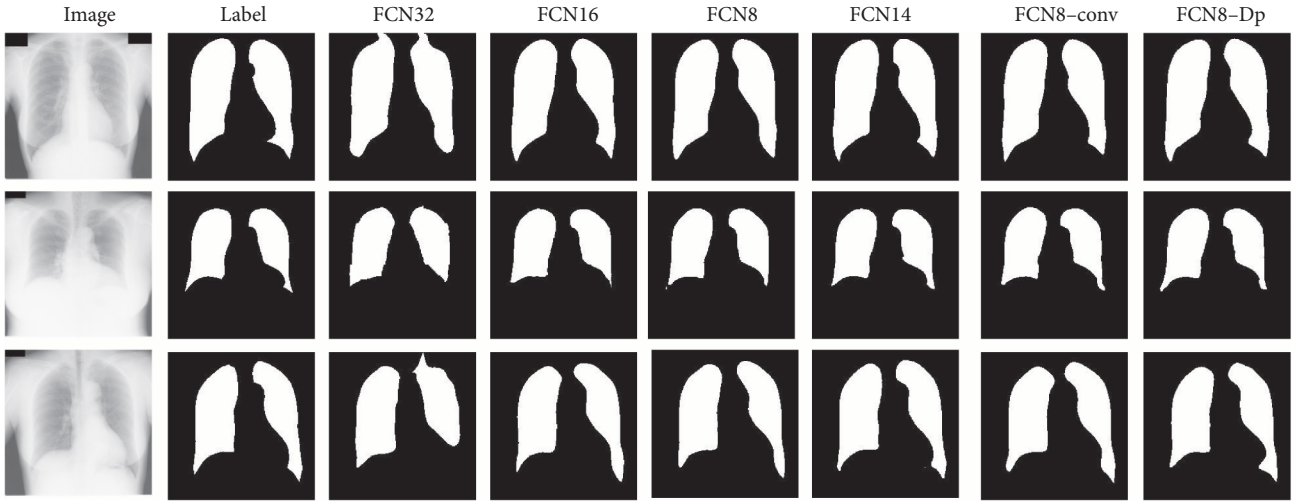

Figure 7. Output obtained on testing dataset on some images.

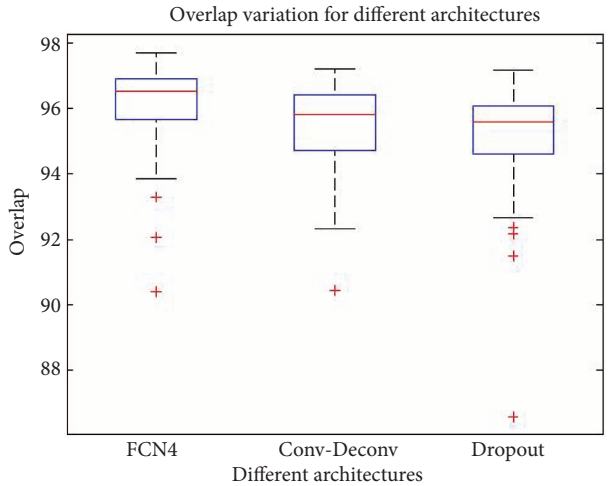

Figure 8. Comparison of the performance of different proposed architectures.

4.4. Comparison with other methods

Figures 7 and 8 show the output and performance of all modified architectures, respectively. The FCN-4 network performs the best and outperforms the human observer as well as other state-of-the-art methods in the literature. The performance of other modifications is on par with other methods reported in the literature. The Table shows the comparison of the performance of different architectures evaluated with the state-of-the-art techniques. In Figure 9, sample output obtained from FCN-4 architecture and Candemir's state-of-the-art method [22] is shown. The performance of these methods is evaluated on the same set of 61 test images. The overlap obtained by Candemir's method is 94.40%, which is significantly less than the 96.10% overlap attained by the proposed architecture.

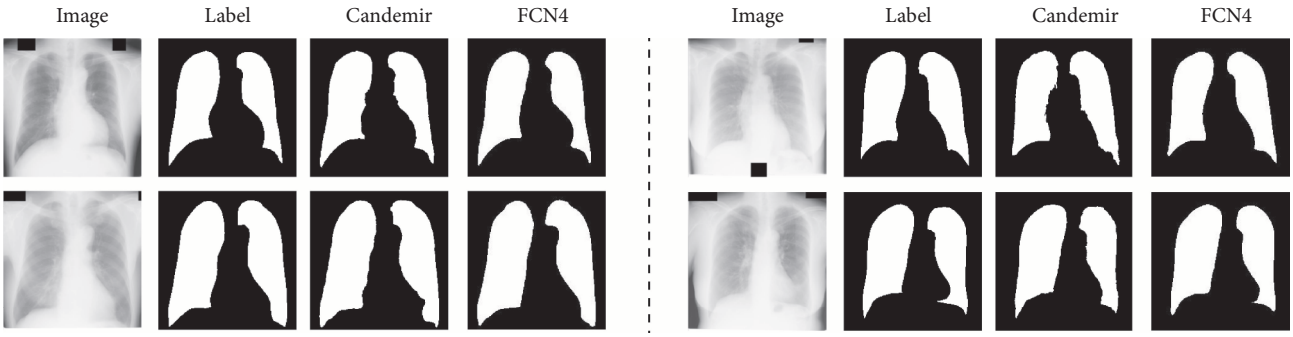

Figure 9. Comparison of output of FCN-4 architecture and Candemir's method [22].

Table. Performance comparison of different proposed architectures with LFS algorithms reported in the literature.

Method	Overlap score(%)	Method	Overlap score(%)
FCN-4	96.10	SIFT + graph cut [22]	95.40
FCN-4 (all modifications combined)	95.93	Novikov et al. [18]	95.00
FCN-8 (all modifications combined)	95.81	Hybrid voting [14]	94.90
FCN-8 with conv layers only	95.44	PC postprocessed [14]	94.50
FCN-8 with dropout	95.07	ASM optimal feature [19]	92.70
FCN-8	94.76	Ahmad et al. [11]	87.00

Experiment 2 evaluates the performance of the best architecture, i.e. FCN-4. It uses an augmented JSRT dataset for training and a completely different Montgomery dataset for testing. The architecture on the Montgomery dataset gives testing accuracy of 97.36% and testing overlap of 90.57%. The performance is acceptable and indicates that the method can be used to segment entirely new CXR images as well.

5. Conclusion

In this study, different architectures have been proposed to improve the LFS performance on CXRs. These architectures are evaluated on two standard datasets, and their performance shows improvement over the standard FCN architectures. The FCN-4 architecture achieves the best performance and surpasses the performance of state-of-the-art LFS methods. It is also concluded that application of each modification has increased the performance as compared to the standard FCN-8 network. The proposed method also gives a satisfactory performance when an entirely different dataset is used for testing purpose.

References

[1] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE 1998; 86: 2278-2324.

[2] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems; 2012; Lake Tahoe, NV, USA. pp. 1097-1105.

[3] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint, arXiv:1409.1556, 2014.

[4] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2015. pp. 1-9.

[5] Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. In: European Conference on Computer Vision; 6 September 2014. pp. 818-833.

[6] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2015. pp. 3431-3440.

[7] Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE T Pattern Anal 2017; 39: 2481-2495.

[8] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention; 5 October 2015. pp. 234-241.

[9] Li L, Zheng Y, Kallergi M, Clark RA. Improved method for automatic identification of lung regions on chest radiographs. Acad Radiol 2001; 8: 629-38.

[10] Duryea J, Boone JM. A fully automated algorithm for the segmentation of lung fields on digital chest radiographic images. Med Phys 1995; 22: 183-91.

[11] Ahmad WS, Zaki WM, Fauzi MF. Lung segmentation on standard and mobile chest radiographs using oriented Gaussian derivatives filter. Biomed Eng Online 2015; 14: 1-26.

[12] McNitt-Gray MF, Sayre JW, Huang HK, Razavi M. Pattern classification approach to segmentation of chest radiographs. In: Proceedings of SPIE Medical Imaging, International Society for Optics and Photonics; 1993. pp. 160-171.

[13] Tsujii O, Freedman MT, Mun SK. Automated segmentation of anatomic regions in chest radiographs using an adaptive sized hybrid neural network. Med Phys 1998; 25: 998-1007.

[14] Van Ginneken B, Stegmann MB, Loog M. Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database. Med Image Anal 2006; 10: 19-40.

[15] Shi Z, Zhou P, He L, Nakamura T, Yao Q, Itoh H. Lung segmentation in chest radiographs by means of Gaussian kernel-based FCM with spatial constraints. In: Sixth International Conference on Fuzzy Systems and Knowledge Discovery; 14 August 2009. pp. 428-432.

[16] Kalinovsky A, Kovalev, V. Lung image segmentation using deep learning methods and convolutional neural networks. In: 13th International Conference on Pattern Recognition and Information Processing; 2016; Minsk, Belarus.

[17] Arbabshirani MR, Dallal AH, Agarwal C, Patel A, Moore G. Accurate segmentation of lung fields on chest radiographs using deep convolutional networks. In: Proceedings of SPIE Medical Imaging, International Society for Optics and Photonics; 24 February 2017. p. 10133051-6.

[18] Novikov AA, Lenis D, Major D, Hladůvka J, Wimmer M, Bühler K. Fully convolutional architectures for multi-class segmentation in chest radiographs. IEEE T Med Imaging (in press).

[19] Van Ginneken B, Frangi AF, Staal JJ, ter Haar Romeny BM, Viergever MA. Active shape model segmentation with optimal features. IEEE T Med Imaging 2002; 21: 924-33.

[20] Shi Y, Qi F, Xue Z, Chen L, Ito K, Matsuo H, Shen D. Segmenting lung fields in serial chest radiographs using both population-based and patient-specific shape statistics. IEEE T Med Imaging 2008; 27: 481-94.

[21] Annangi P, Thiruvenkadam S, Raja A, Xu H, Sun X, Mao L. A region based active contour method for x-ray lung segmentation using prior shape and low level features. In: International Symposium on Biomedical Imaging: From Nano to Macro; 14 April 2010. pp. 892-895.

[22] Candemir S, Jaeger S, Palaniappan K, Musco JP, Singh RK, Xue Z, Karargyris A, Antani S, Thoma G, McDonald CJ. Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration. IEEE T Med Imaging 2014; 33: 577-590.

[23] Springenberg JT, Dosovitskiy A, Brox T, Riedmiller M. Striving for simplicity: the all convolutional net. arXiv preprint, arXiv:1412.6806, 2014.

[24] Shiraishi J, Katsuragawa S, Ikezoe J, Matsumoto T, Kobayashi T, Komatsu KI, Matsui M, Fujita H, Kodera Y, Doi K. Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. Am J Roentgenol 2000; 174: 71-74.

[25] Jaeger S, Candemir S, Antani S, Wáng YX, Lu PX, Thoma G. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. Quant Imag Med Surg 2014; 4: 475-477.