

Enhanced Lung Segmentation from Chest X-Ray Images using Attention Based FCNN

Pradeep Kumar¹, Pramod Kumar Soni^{*2}, Linesh Raja³

Submitted: 27/01/2024 Revised: 05/03/2024 Accepted: 13/03/2024

Abstract: Chest Radiographs are extensively used imaging tool for retrieving visual features of the affected area. Detection of abnormalities visually from chest radiographs is a very challenging task for medical practitioners as the thoracic cavity comprises many sensitive organs like the lungs, heart, sternum, etc. The technological advancements in computational technology have facilitated medical experts to improve diagnosis accuracy. Recently deep learning (DL)based architectures have gained popularity among radiologists for better diagnosis. In this research article, an attention based FCNN model is presented for segmenting lungs from Chest radiographs. The proposed model eliminates the computation overhead by eliminating the irrelevant features generated during feature extraction by including the attention mechanism in the decoder architecture of the proposed FCNN. This further enhance the model's performance and computational complexity. The performance of the proposed model is evaluated on chest radiographs obtained from JSRT dataset and measured on different evaluation metrics likewise precision, Recall, F1-score, accuracy, and the Jaccard Similarity coefficient (JSC). The proposed model has obtained an 98% accuracy during the training and ~97% accuracy during the testing stages. Furthermore, the comparison of proposed model with baseline U-Net is performed.

Keywords: X-ray, FCNN, attention U-Net, image segmentation

1. Introduction

Lung Diseases nowadays are very common because of various environmental factors such as air pollution etc. emitted by industries. They are the major causes for mortality and hospitalization throughout the world. So, it seems essential to diagnose different lung diseases like TB, pneumonia, asthma, etc. Information about the conditions or severity of the lungs is obtained through various medical imaging sensory system like X-ray, CT etc. The X-ray images were discovered much earlier than CT images. Although CT images are high in quality and have abundant information as compared to X-rays, the usage of CT images is limited by the availability of CT machines and their cost. On the contrary X-rays images requires less computing facility and cost. Due to the above-mentioned reasons, X-ray images are widely used. Chest X-rays (CXR) are a very popular modality usually utilized by medical practitioners in imaging thoracic illnesses. CXR is a less sensitive form of imaging in the identification of lung infection as compared to other modalities. Frequently reported Chest X-ray includes lung consolidation of COVID-19 disease. medical domain relies upon CXR due to its full accessibility and less infection control. Infectious disease like pneumonia has some specific features like peripheral lung involvement

which mirrors another incendiary process that can be instantly detected in CXR.

CXR makes available general positioning of infection or abnormality for some complications pertaining to chronic lung conditions like cystic and pulmonary fibrosis, bronchial carcinoma[1]. Both fixed and portable X-ray equipment is Currently in use which is entirely digital and replaced the classical film-foil combinations. The CXR images of lungs obtained from the JSRT database are exhibited in Fig. 1

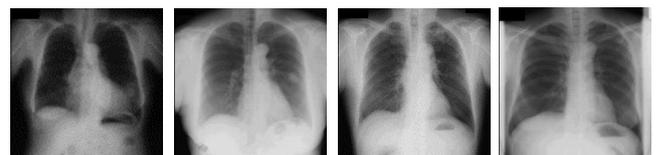


Fig. 1. CXR image obtained from JSRT dataset.

The CXR images can help the medical practitioners in diagnosing various diseases. For this task, the information must be extracted from unprocessed CXR images. The general flow of information extraction from medical images is shown in Fig. 2 which includes steps such as preprocessing, segmentation, postprocessing and evaluation of results.

¹ Manipal University Jaipur, Jaipur – 303007, INDIA
ORCID ID: 0000-0002-1320-3910

² Manipal University Jaipur, Jaipur – 303007, INDIA
ORCID ID: 0000-0003-0632-5044

³ Manipal University Jaipur, Jaipur – 303007, INDIA
ORCID ID: 0000-0002-3663-1184

* Corresponding Author Email: pramod.soni@jaipur.manipal.edu

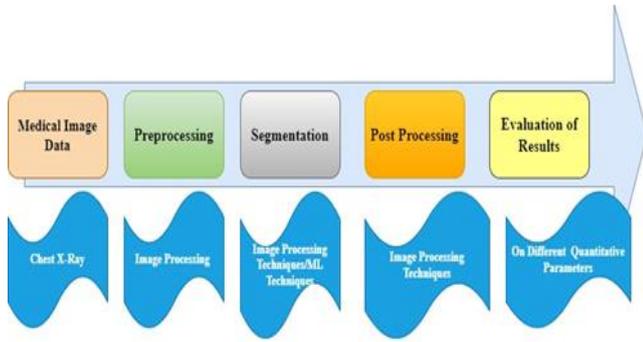


Fig 2: Steps used in medical image segmentation.

Before performing segmentation or information task the CXR images needs to be pre-processed in which the various features like colour, shape, texture, moment, etc are extracted. These features help the in investigation and analysis of certain applications for diagnosis in advance. Various image processing techniques are employed to extract features basis on textures and shapes from the masked image [2]. The Gaussian blur filter method is used for the removal of noise and duplicate data for X-ray image preprocessing.

After feature extraction, these features are utilized in a higher-order tasks such as segmentation which is a process of demarcating important areas of CXR images. Lung images are split into some regions or areas in the process of segmentation. The right Lungs can be segmented into three upper lobes, two middle lobes and five lower lobes. Similarly, the left lung is segmented into four upper and four lower lobes. CXR segmentation has several challenges such as noise, partial aliasing and sampling artifacts that cause disconnect or indistinct boundaries [3]. The segmentation is followed by the post-processing step, in which the false segmented areas or pixels are eliminated so that the overall accuracy of the segmentation or information task increases. The post-processing is generally based on features like shape which are performed by using techniques such as morphological filtering. Finally, the performance is evaluated on various evaluation indicators likewise Accuracy, Jaccard index etc.

CXR image segmentation techniques include thresholding-based [4], region-based [5], graph-based, and cluster-based [6], ML and DL-based segmentation [7]for accurate and reliable lung segmentation of critical patients. These techniques are used for handling diverse features of CXRs to enable more generalized lung segmentation. Conventional techniques used for chest radiograph segmentation based on heuristics are quick and simple, but they require fine-tuning for supporting certain manual use cases and are generally low in accuracy for complex images. current segmentation techniques use ML to improve accuracy and making more flexible .

DL techniques are more effective in lung radiograph segmentation. DL-based techniques are used to automate

numerous time-consuming tasks executed by radiology experts like segmentation, predicting the response of treatment, lesion or patch detection, and classification. Analysis of CXR images includes the number of images obtained, Interpretation of their complexities and clinical values in practice. Convolution Neural Networks (CNN) are the most popular DL-based architecture in the medical imaging domain [8]. DCNN or Deep convolutional neural network is a variant of the frequently used CNN. The architecture of DCNN is utilized for automatic multiclass segmentation of CXRs of clavicles, lungs, and heart. An FCNN has two main parts. (i) First it has a contraction path, which attains the information of the image and (ii) the expansion path used to encode information for creating the segmentation output .

A lightweight Deep convolution neural network(DCNN) model is investigated using the residual connection for automatic lung region segmentation from CXR images obtained from multiple data sources [9]. A U-Net fusion framework incorporated with dialled convolution is proposed for lung segmentation from CT images.DeepLabv3 employed Atrous Spatial Pyramid Pooling (ASPP) architecture consisting of dilated convolution is utilized for lung segmentation from CXR images collected from various benchmark datasets for early detection of severe lung abnormalities. A dual attention U-Net model is proposed for automatic lesion segmentation of lung infection in grayscale CT images collected from different independent datasets. The model is based on gate attention and decoder attention with a hybrid dilated Convolution approach[10]. A hybrid working mechanism based upon DL, weighted sum and genetic algorithm is employed for extracting lungs to automatically and fast diagnose disease. The designed model is implemented on a unified CXR image dataset. Dimensional Interactive U-Net(DI-U-Net) is developed which is a self-attention-based modified U-Net model implemented over the ISIC-2018 challenge dataset of CXR [11]. A modified model was proposed which was based upon multi-dropout in all DeepConv layers and trained on the JSRT image dataset of CXR using modified architecture of U-Net [12]for Lung segmentation in CXR images. Further, a lightweight CNN named LigtEfficientNetV2 is designed for Covid detection in CXR and CT images. Various modified U-Net architecture such as Dense U-Net, Adversarial U-Net [13] were designed and implemented for Lung extraction from various medical imaging modalities.

Various DL based models have been used for segmenting lung from CT and CXR images. Most of the proposed models from the above-mentioned studies aim to segment lungs from various image modalities with higher accuracy but with less emphasis on computational overhead on time and cost parameters. To resolve these flows in this work, an attention based FCNN is proposed. The prime motive of this

work is to segment lung from CXR images by computationally efficient DL-based FCNN method. The proposed method performs lung segments using FCNN by making use of the small number of images used for training and testing in comparison with other models which require abundant number of resources and training images.

Remaining part of this work is structured as follows: proposed model and dataset used in our work are illustrated in Section 2. Experimental findings are displayed and conferred in Section 3. The future scope and conclusion are discussed in Section 4.

2. Research Methodology

In this section, a lung segmentation method is proposed to segment lung region from CXR using a DL-based technique and the dataset used is described. The flow diagram of proposed model is depicted in Fig 3. The proposed model is consisting of phases namely pre-processing, segmentation and evaluation.

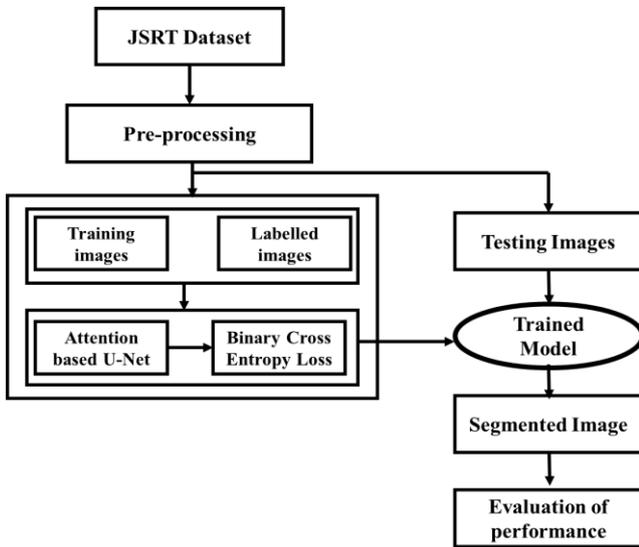


Fig 3: Flow chart of proposed attention based FCNN for Lung Segmentation

2.1. Dataset

The images used in the current work is acquired from the Japanese Society of Radiological Technology (JSRT) [14] dataset. The JSRT dataset contains chest radiographs in different formats such as grey-scale, DICOM and RGB of varying dimensions. The JSRT society recently introduced images that are suitable for performing experiments using DL-based archetype. This work uses chest radiographs of dimension 256 x 256 and corresponding labelled image in RGB formats which is also exhibited in Fig. 1.

2.2. Pre-processing

The segmentation of lungs from chest radiographs is a binary classification process which means that the lung portion must be extracted, and the remaining portion is not

considered. Segmenting lungs from chest radiographs using DL architecture requires the labelled image having information of lungs and the remaining part only (binary information). The chest X-Ray image and corresponding label image in the dataset are in RGB formats. So, the labelled mask is persuaded to a grayscale image having binary information only, and the pixels containing lungs regions are set to 255 and remaining region is set to 0. This transformation is performed as per Eq. (1).

$$G_I(p, q) = 0.3 * I_R(p, q) + 0.59 * I_G(p, q) + 0.11 * I_B(p, q) \quad (1)$$

2.3. Network Architecture

In proposed work, a FCNN model for segmentation lungs from CXR image is designed. The proposed model uses modified U-Net [15] as a backbone (encoder and decoder architecture) and attention gate mechanism. U-Net is massively used for image segmentation in different application domains like remote sensing images, medical images, and engineering applications. The proposed FCNN architecture is designed by incorporating an attention unit in baseline U-Net architecture. The initial input of architecture is 256 x 256 x 3 chest X-Ray image. The encoder is a CNN comprising repeated CNN blocks having two convolutional layers of 3 x 3 preceded by an activation function given as in Eq (2) and batch normalization (BN) layer. Additionally, each block has a max-pooling layer of 2 x 2 and a dropout layer.

$$\sigma 1(x^L_{i,c}) = \max(0, x^L_{i,c}) \quad (2)$$

Where $x^L_{i,c}$ is the feature map at layer L having i-pixel location, C channel dimensions, and the feature activations can be represented as given in Eq. (3)

$$x^L_c = \sigma 1(\sum_{c' \in F_L} x^{L-1}_{c'} * k_{c,c}) \quad (3)$$

Where * is the convolution operation and F_L corresponds to the feature map at layer L. In the decoder, each block corresponding to the encoder consists of 2 x 2 deconvolution, concatenation of attention-based feature map from the encoder part, and 3 x 3 convolutional followed by activation BN layers. In the final layer, a convolution of 1 x 1 based on the non-linear activation function as given in Eq. (4) is employed to generate the required segmented image. The decoder part performs the interpolation-based up-sampling and concatenation. The proposed architecture is shown in Fig.4 and Fig.5

$$\sigma 2(x_{i,c}) = \left\{ \frac{1}{\exp(-x_{i,c})} \right\} \quad (4)$$

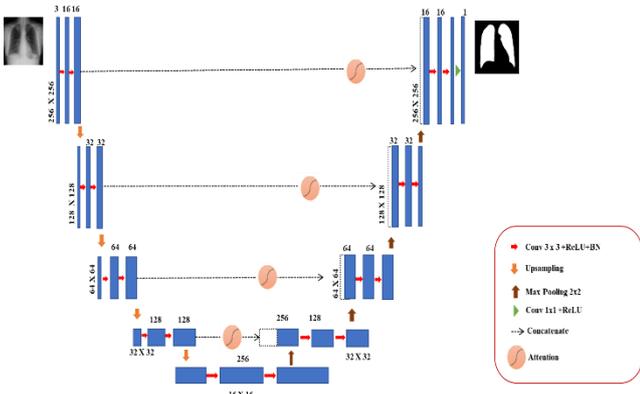


Fig. 4 : Architecture of attention based FCNN

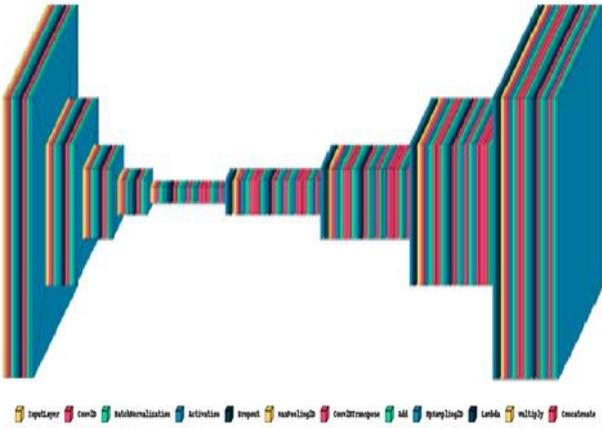


Fig. 5 Visualization of the proposed architecture

Traditional FCNN models consist of two components namely encoder and decoder, the major function of first component is to extract prominent features which capture the semantic contextual information for this purpose feature map is down sampled. During this process, most of the false positive information remains there which further reduces the performance of FCNN model and increases computational complexity. To avoid irrelevant features the attention gates [16] are incorporated into the network. The attention gate structure is shown in Fig.6. depicts the building block of the attention gate is presented in It takes two inputs, a feature map from the dense layer (x) another is the attention signal (g). The input features are scaled as per the attention coefficients (α) of the resampled grids obtained by interpolation. The attention factor (α) $\in \{0,1\}$ used to extract the focus areas. The output of attention gate is a convolution of the input feature map ($x_{i,c}^l$) and α and can be represented as,

$$\hat{x}_i^l = x_{i,c}^l \cdot \alpha_i^l \quad (5)$$

The gating vector $g_i \in \mathbb{R}^{F_g}$ as shown in Fig. 6 is used for each pixel i to determine focus regions and eliminates the

lower-level feature responses.

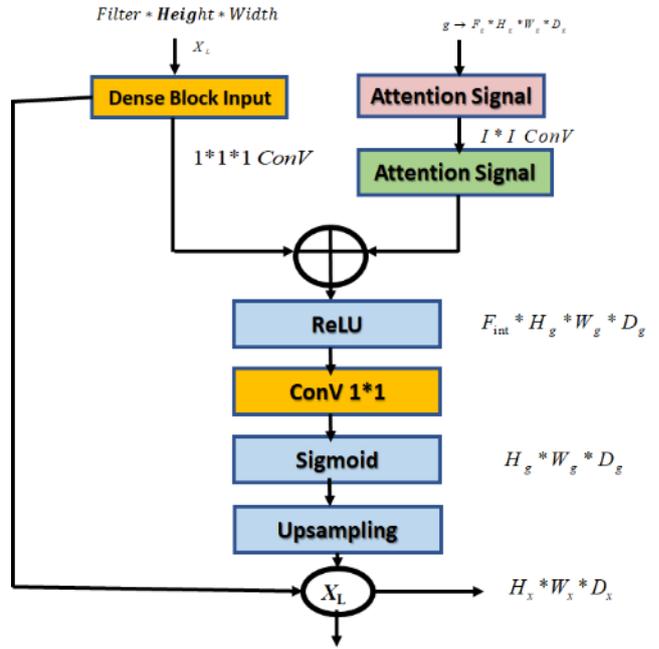


Fig 6. structure of attention gate

The gating coefficient can be calculated as, The gating vector $g_i \in \mathbb{R}^{F_g}$ as shown in Fig. 6 is used for each pixel i to determine focus regions and eliminates the lower-level feature responses. The gating coefficient can be calculated as,

$$Q_{at}^l = \Phi^T(\sigma_1(W_x^T x_i^l + W_g^T g_i)) + b_\Phi \quad (6)$$

$$\alpha_i^l = \sigma_2(Q_{at}^l(x_i^l, g_i; \theta_{at})) \quad (7)$$

Where $\sigma_2(x_{i,c})$ is non-linear activation function as discussed in Eq (4), θ_{at} is a linear transformation of parameters $W_x \in \mathbb{R}^{F_L \times F_{int}}$, $W_g \in \mathbb{R}^{F_g \times F_{int}}$ and bias term b .

The integration of the attention gate eliminates irrelevant information and enhances the local details and ensures the maximum transmission of important features between all the layers of the network. The specification of the proposed model is shown in Table 1.

Table 1 Parameter specification of the proposed model

Layer Number	Input Dimension	Output Dimension	Number of Parameters
Layer-1	[256 x 256 x 3]	[128 x 128 x 16]	2896
Layer-2	[128 x 128 x 16]	[64 x 64 x 32]	14144
Layer-3	[64 x 64 x 32]	[32 x 32 x 64]	55936
Layer-4	[32 x 32 x 64]	[16 x 16 x 128]	222464
Layer-5	[16 x 16 x 128]	[16 x 16 x 256]	296152
Layer 6	[16 x 16 x 256]	[16 x 16 x 128]	1084800
Layer 7	[16 x 16 x 128]	[32 x 32 x 64]	723521
Layer 8	[32 x 32 x 64]	[64 x 64 x 32]	181537
Layer 9	[64 x 64 x 32]	[128 x 128 x 16]	45585
Layer 10	[128 x 128 x 16]	[256 x 256 x 1]	10790

2.4. Loss function

In this work, the set of training images are represented as I and corresponding lungs area labels are represented as G . The objective function of lung segmentation is defined as follows:

$$Min E(I, G, P) = \sum_{i=1}^N \| I_i * P - G_i \|^2 \tag{8}$$

Where N is the number of training images and P is the parameter of network. Binary cross entropy is utilized as loss function during training which is represented as follows:

$$L_{FCN_w}(I) = - \sum_{i=1}^N \sum_{j=1}^r \sum_{k=1}^c G_i(i, j) . \log FCN_w(I_i(J, K)) + (1 - G_i(J, K) . \log(1 - FCN_w(I_i(J, K)))) \tag{9}$$

Where (r, c) is the shape size of I and FCN_w represents the model at (j, k) position. ADAM(adaptive moment estimation) optimizer is used during training.

2.5. Evaluation Metrics

The performance of proposed method for lungs segmentation from CXR, various evaluation metrics such as precision ($L_{Precision}$), recall (L_{Recall}), $F1$ -score, Accuracy (L_{Acc}) and Jaccard Similarity coefficient (JSC) are used which can be obtained as follows,

$$L_{Precision} = \frac{NP_{TP}}{NP_{TP} + NP_{FP}} \tag{10}$$

$$L_{Recall} = \frac{NP_{TP}}{NP_{TP} + NP_{FN}} \tag{11}$$

$$L_{Acc} = \frac{NP_{TP} + NTN}{N} \tag{12}$$

$$F1 - Score = \frac{2 * L_{Precision} * L_{Recall}}{L_{Precision} + L_{Recall}} \tag{13}$$

$$JSC = \frac{NP_{TP}}{NP_{TP} + NP_{FP} + NP_{FN}} \tag{14}$$

Where NP_{TP} denotes the number of pixels which are present in reference and segmented image both, NP_{FP} is count of pixels which are not present in reference image but these pixels are present in segmented image, NP_{FN} is the count of pixels which are present in reference image but these pixels are absent in the segmented image and N is total count of pixels present in image. The evaluation method is shown in Fig. 7.

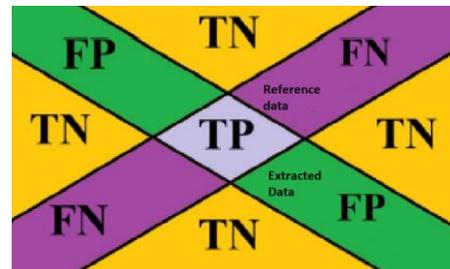


Fig 7: Confusion Matrix

The L_{Acc} signifies the count of pixels which are correctly classified, whereas $F1$ -score is comparatively better metric than L_{Acc} because it is harmonical average of L_{Acc} and L_{Recall} and combines the characteristics of both markers. JSC is the extent of similarity and diverseness of reference images and segmented images. It indicates the perceptual quality of segmentation.

3. Results and Discussion

For performing segmentation of lungs from chest radiographs, experiments are performed on a dedicated GPU environment. The 256 x 256 color images are provided as input into the attention based FCNN model as presented in Fig. 4 and 5. BN is applied to each convolutional layer

followed by activation by ReLU. The output of proposed model is an image which is in binary format containing the lung region and the region not containing lungs. In order to achieve this, a sigmoid function with class count 2 (0 or 1) is used. The dataset is bifurcated into testing and training in 10:90 ratio with respective masks.

The proposed model is compiled with Adam optimizer with learning rate 0.001. The training parameters are optimized, and the designed model is trained for 50 epochs of a batches of 4. Results are presented in Fig. 8. The RGB X-ray images are exhibited in Fig. 8 (a1) - (a6) and the corresponding ground truth or reference images are shown in Fig. 8 (b1) - (b6) . The experimental results of lungs segmentation from chest radiographs by the proposed attention-based FCNN are presented in Fig.8 (c1) - (c6).

From the shown parameters it comes in observation that proposed model obtained the maximum L_{Acc} which is depicted in Fig. 8 (a2) and highest JSC in Fig. 8(a2). The lowest L_{Acc} and JSC are obtained in Fig .8 (a6). This decline in the L_{Acc} and JSC clearly reflected by visualizing Fig. 8 and 9.

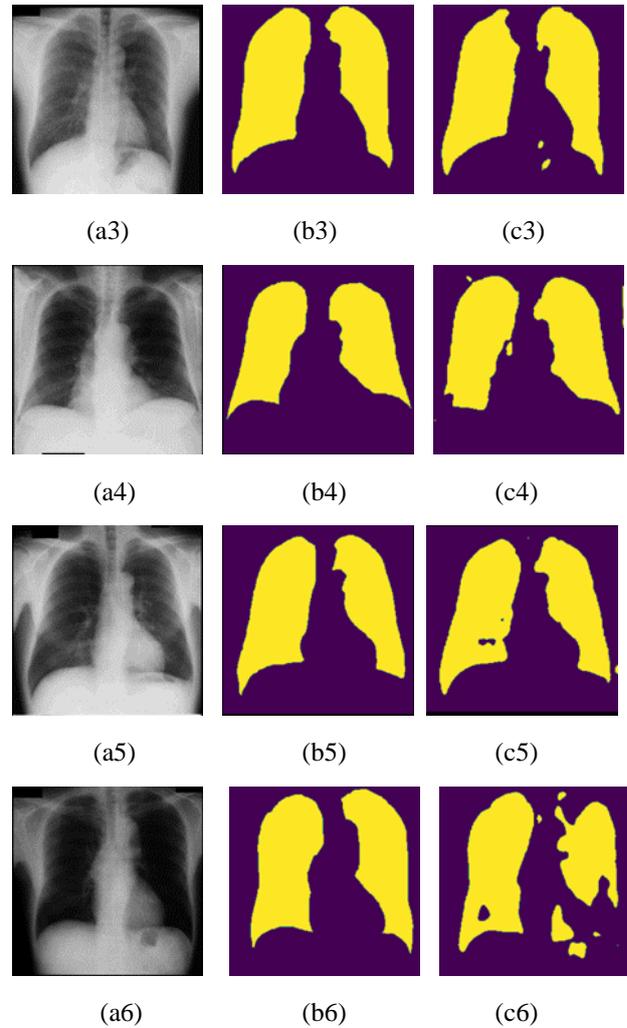
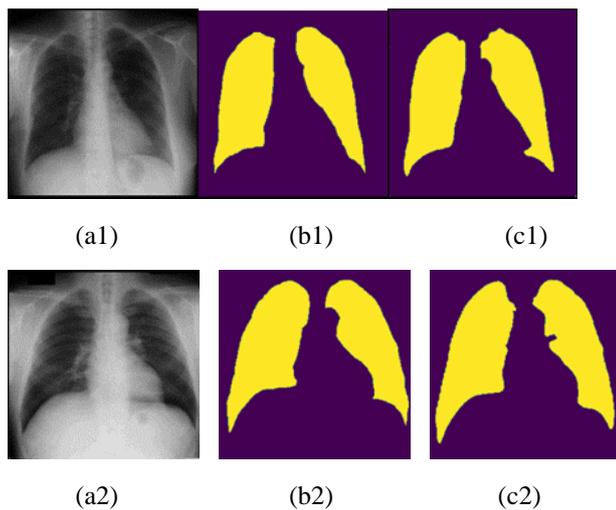


Fig 8. Experimental evidence of the proposed attention based FCNN. RGB CXR image (a1)-(a4), corresponding ground truth (b1)-(b4), segmented image by proposed attention based FCNN (c1) –(c4)

By visualizing the quality of image segmentation looks promising for practical use, furthermore, the performance of proposed model is evaluated on quantitative parameters such as $L_{Precision}$, L_{Recall} , $F1-score$, L_{Acc} and JSC which are shown in Table 2.

Table 2 performance of proposed attention based FCNN on different evaluation indicators

	$L_{Precision}$	L_{Recall}	$F1-score$	L_{Acc}	JSC
Fig. 8 (a1)	0.88	0.81	0.84	0.8988	0.93
Fig. 8 (a2)	0.97	0.96	0.96	0.976	0.94
Fig. 8 (a3)	0.94	0.95	0.94	0.9636	0.9
Fig. 8(a4)	0.95	0.95	0.96	0.967	0.89
Fig. 8 (a5)	0.93	0.98	0.97	0.9718	0.92
Fig. 8 (a6)	0.74	0.86	0.92	0.8942	0.73

The Fig. 9 depicts the accuracy of proposed attention based FCNN with increasing count of epochs, the utmost training

accuracy is 98.76% and the validation accuracy is 96.74%. In Fig. 10, the validation and training loss that come off

during training time and testing time with an increasing number of epochs is shown.

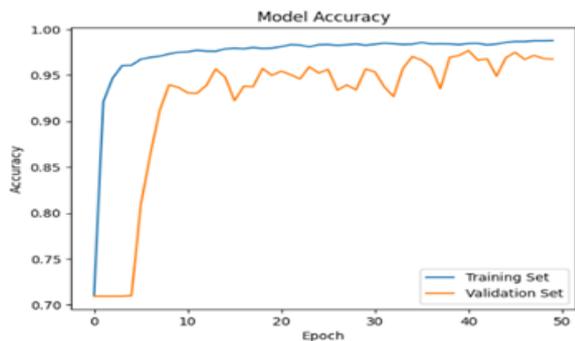


Fig 9: Segmentation accuracy (output of proposed FCNN)

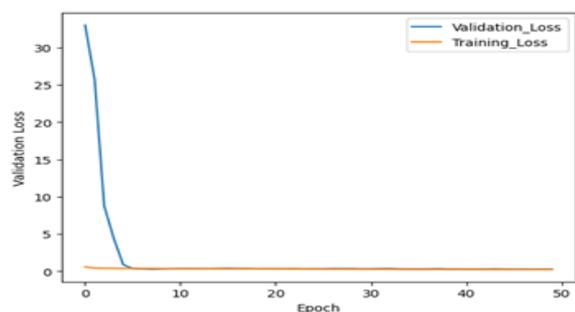


Fig 10: Loss vs Epochs

Furthermore, to show the improvements in experimental results, the proposed attention based FCNN is matched with baseline U-Net Model with identical parameters likewise epochs, learning rate, batch-size, optimizer etc. The visual comparison of the proposed method and baseline method on

all images is shown in Fig. 11.

Which reflects the significant improvement in segmentation results.

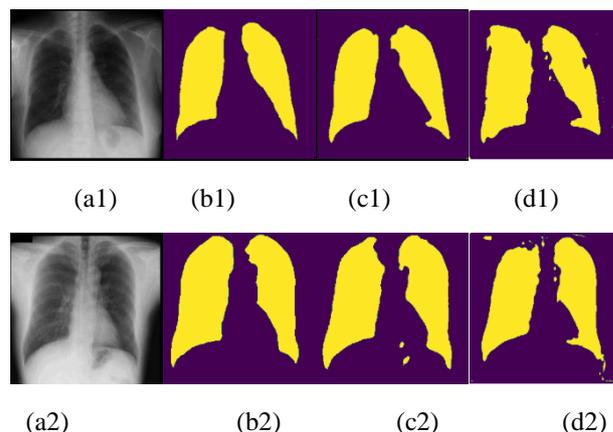


Fig 11: Visual comparison of proposed method; Lung image (a1)-(a2), reference image (b1)- (b2), segmented by proposed method (c1)-(c2) and segmented by baseline U-Net (d1)-(d2).

The experimental results are matched on L_{Acc} and JSC parameters on different images used to test the proposed model’s performance. In all images except the image exhibited in Fig. 9(a2), the L_{Acc} of the proposed model is better by a significant margin as compared to baseline model. While in JSC metric the proposed model’s is superior in comparison to counterpart as shown in Table 3.

Table 3 Comparative study of state of the art method

	Baseline UNet		Proposed Attention-based FCNN	
	L_{Acc}	JSC	L_{Acc}	JSC
Fig. 8 (a1)	0.8258	0.77	0.8988	0.93
Fig. 8 (a2)	0.9679	0.90	0.9760	0.94
Fig. 8 (a3)	0.9573	0.84	0.9636	0.90
Fig. 8 (a4)	0.9417	0.87	0.9670	0.89
Fig. 9 (a1)	0.9619	0.90	0.9718	0.92
Fig. 9 (a2)	0.9212	0.64	0.8942	0.73

4. Conclusion

In this paper, an attention based FCNN model is proposed for segmenting lung from CXR images. The proposed model encompasses preprocessing, segmentation and evaluation. The Chest X-Ray images are obtained from JSRT database that are transformed into grayscale images having binary information in the preprocessing stage followed by segmentation process using attention based FCNN. The proposed model is computationally very efficient since it removes non-relevant features through attention module. The proposed model is trained with small

number of CXR images instead of a huge dataset of images used in other DL models. The experimental results show the ability of the proposed model in comparison to traditional FCNN models such as U-Net both visually and quantitatively shown in Fig.11 and in Table 3. The proposed model has obtained L_{Acc} of ~97%. The proposed model uses the training weights on small dataset; however the performance can be enhanced by using transfer learning with pretrained weights of existing trained models.

Acknowledgements

funding was not obtained for this research-based study.

Author contributions

Pradeep kumar has performed all the experiments under the supervision of Linesh Raja and Pramod Kumar Soni. The idea is suggested by Linesh Raja and Pramod Kumar Soni. Each author has read and reviewed the manuscript.

Conflicts of interest

The authors hereby declare that they do not have any known competing financial interests.

References

- [1] M. O. Wielpütz, C. P. Heußel, F. J. F. Herth, and H.-U. Kauczor, "Radiological Diagnosis in Lung Disease," *Dtsch Arztebl Int*, Mar. 2014, doi: 10.3238/arztebl.2014.0181.
- [2] R. Ali, A. Hussain, and M. Man, "Feature Extraction and Classification for Multiple Species of Gyrodactylus Ectoparasite," *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol. 13, no. 3, pp. 503–511, Mar. 2015.
- [3] Qureshi et al., "Medical image segmentation using deep semantic-based methods: A review of techniques, applications and emerging trends," *Information Fusion*, vol. 90, pp. 316–352, Feb. 2023, doi: 10.1016/j.inffus.2022.09.031.
- [4] J. John and M. G. Mini, "Multilevel Thresholding Based Segmentation and Feature Extraction for Pulmonary Nodule Detection," *Procedia Technology*, vol. 24, pp. 957–963, 2016, doi: <https://doi.org/10.1016/j.protecy.2016.05.209>.
- [5] D. Barbosa, T. Dietenbeck, J. Schaerer, J. D'hooge, D. Friboulet, and O. Bernard, "B-Spline Explicit Active Surfaces: An Efficient Framework for Real-Time 3-D Region-Based Segmentation," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 241–251, 2012, doi: 10.1109/TIP.2011.2161484.
- [6] B. N. Li, C. K. Chui, S. Chang, and S. H. Ong, "Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation," *Comput Biol Med*, vol. 41, no. 1, pp. 1–10, 2011, doi: <https://doi.org/10.1016/j.compbimed.2010.10.007>.
- [7] Md. Z. Islam, Md. M. Islam, and A. Asraf, "A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images," *Inform Med Unlocked*, vol. 20, p. 100412, 2020, doi: <https://doi.org/10.1016/j.imu.2020.100412>.
- [8] B. Sahiner et al., "Deep learning in medical imaging and radiation therapy," *Med Phys*, vol. 46, no. 1, pp. e1–e36, Jan. 2019, doi: 10.1002/mp.13264.
- A. Maity, T. R. Nair, S. Mehta, and P. Prakasam, "Automatic lung parenchyma segmentation using a deep convolutional neural network from chest X-rays," *Biomed Signal Process Control*, vol. 73, p. 103398, 2022, doi: <https://doi.org/10.1016/j.bspc.2021.103398>.
- [9] X. Zhao et al., "D2A U-Net: Automatic segmentation of COVID-19 CT slices based on dual attention and hybrid dilated convolution," *Comput Biol Med*, vol. 135, p. 104526, 2021, doi: <https://doi.org/10.1016/j.compbimed.2021.104526>.
- [10] Y. Wu, G. Wang, Z. Wang, H. Wang, and Y. Li, "DI-UNet: Dimensional interaction self-attention for medical image segmentation," *Biomed Signal Process Control*, vol. 78, p. 103896, 2022, doi: <https://doi.org/10.1016/j.bspc.2022.103896>.
- [11] S. Arvind, J. V. Tembhurne, T. Diwan, and P. Sahare, "Improvised light weight deep CNN based U-Net for the semantic segmentation of lungs from chest X-rays," *Results in Engineering*, vol. 17, p. 100929, 2023, doi: <https://doi.org/10.1016/j.rineng.2023.100929>.
- [12] S. Tyagi and S. N. Talbar, "CSE-GAN: A 3D conditional generative adversarial network with concurrent squeeze-and-excitation blocks for lung nodule segmentation," *Comput Biol Med*, vol. 147, p. 105781, 2022, doi: <https://doi.org/10.1016/j.compbimed.2022.105781>.
- [13] J. Shiraishi et al., "Development of a Digital Image Database for Chest Radiographs With and Without a Lung Nodule," *American Journal of Roentgenology*, vol. 174, no. 1, pp. 71–74, Jan. 2000, doi: 10.2214/ajr.174.1.1740071.
- [14] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds., Cham: Springer International Publishing, 2015, pp. 234–241.
- [15] O. Oktay et al., "Attention U-Net: Learning Where to Look for the Pancreas," Apr. 2018, Accessed: Apr. 05, 2023. [Online]. Available: <http://arxiv.org/abs/1804.03999>