



## Self-Sequential Attention Layer based DenseNet for Thoracic Diseases Detection

Roshan Shetty<sup>1\*</sup> Prasad Narasimha Sarappadi<sup>2</sup>

<sup>1</sup>*Department of Computer Science and Engineering, Karavali Institute of Technology, Mangaluru, India*

<sup>2</sup>*Department of Electronics & Communication Engineering, REVA University, Bengaluru, India*

\* Corresponding author's Email: roshan564956@gmail.com

---

**Abstract:** Thoracic diseases are the various collections of diseases that are associate with the cavity of the thorax such as lungs, heart, oesophagus, chest wall, and diaphragm. The Chest X-Ray is the widely used radiological examination in the diagnosis of such diseases. Many existing methods used entire Chest X-Ray images for the process of training however, it suffered from few limitations. The misalignment of potential or exist of unrelated objects in the Chest X-Ray images resulted in irrelevant noises which limited the network performance. The pre-processing steps of existing methods during training the neural network resulted in lower resolution images. Due to loss of information, it is very difficult to find the tiny lesion regions from the images. To solve such issues, the Self-Sequential Attention Layer based DenseNet (SAL-DN) model is proposed to enhance thoracic disease prediction on Chest X-Ray. The SAL uses the co-relation among the class labels and abnormal pathology by analysing the known feature maps from DenseNet-121. The SAL-DN has the advantages of providing the best representation of images and the capacity to handle large imbalanced datasets like Chest X-Ray14. The performance of the proposed SAL-DN model is compared with existing methods in-terms of the achieved AUC score. The experimental result shows that the proposed SAL-DN outperforms the Thorax-Net method by obtaining an Average AUC score of 0.8715, whereas Thorax-Net obtained an Average AUC score of 0.7876 in patient-wise official split.

**Keywords:** Attention layer, Chest X-ray, Chest X-ray14 dataset, DenseNet, Thoracic diseases.

---

### 1. Introduction

Thoracic diseases are a serious health problem in the world, pneumonia alone affected 450 million people approximately [1]. According to the report published by UNICEF in 2019, the medical issues associated with the cavities of the thorax are much common these days [2]. Chest radiography is a commonly utilized diagnosis in medical care because of its usage, low cost, easiness, and higher availability in the canters of outpatient [3]. The chest radiography images provide the abnormalities that occurred in the lung, rib cage, mediastinum, and heart that allows doctors to identify the cause of different illnesses and provide treatment to diseases [4]. By analysing the images of chest radiography, the physician will diagnose various conditions of diseases like atelectasis, cardiomegaly, pleurisy, effusion, pneumonia, nodule, infiltration, fractures,

cardiomegaly, bronchitis, pericarditis, pneumothorax, etc. [5]. The timely study is difficult when Chest X-Rays are not reviewed immediately and it occurs periodically when the queue of cases is waiting for evaluation result [6].

In Chest X-Ray images, the overlapping tissues and lower contrast resolution need to be separated from the lesion and neighbouring tissues, which may result in missed identifications and diagnosis [7]. Many research scholars have done experiments related to deep learning approaches using Convolutional Neural Network (CNN) for image recognition especially medical images such as Chest X-Ray. Additionally, the irrelevant objects and weak alignments in Chest X-Ray images lowers the performance of the neural network [8]. The Chest X-Ray images are resized to lower resolution images to reduce the computation problem. But, these processes resulted in the loss of details in Chest X-Ray images which are very much essential to diagnose

pathologies [9]. To solve such an issue, the SAL-DN model is proposed to enhance the thoracic disease prediction performance. The data pre-processing is done to resize the images into  $224 \times 224$  pixels, ImageDataGenerator class is used for augmentation and faster convergence of data during the training of the network. The resized images are applied as an input to the proposed SAL-DN to enhance the prediction of thoracic diseases. The proposed SAL-DN has the advantage of providing the best representation of images and has the capacity to handle a large dataset.

This research paper is organized as follows: a survey of the existing methods is given in Section 2; problems identified from the review of existing methods are explained in Section 3. The objective of the research is explained in Section 4 and the proposed SAL-DN model for disease prediction is given in Section 5. The experimental results are discussed in section 6 and the conclusion and future work of this research are given in section 7.

## 2. Literature review

Accurate interpretation of Chest radiograph by radiologist plays a main role in the diagnosis of thoracic diseases. Analysis of Chest X-Ray is challenging for radiologists and it consumes more time. Due to an inadequate number of radiologists, healthcare is delayed to patients especially in developing countries. Many existing methods were developed for the identification and classification of thoracic diseases from the Chest X-Ray images. Deep learning methods were applied to the Chest X-Ray images for efficient diagnosis of thorax diseases. The recent researches involved in developing the deep learning models for thorax diseases was reviewed in this section.

Wang [10] developed a Thorax-Net model based on a Deep Convolutional Neural Network for the diagnosis of 14 thorax diseases by utilizing chest radiography. Their model includes the classification and attention branch, the classification branch acts as a uniform feature extractor and a classifier architecture. The attention branch utilizes the correlation between class labels and the location of pathological anomalies by analysing the feature map obtained by the classification branch. The data set used in the Thorax-Net classification model was Chest X-Ray14. Thorax-Net model showed higher performance in detecting the diseases. The developed method showed performance degraded due to class imbalance nature of dataset.

Liu [11] developed Segmentation Based Deep Fusion Network (SDFN) for the classification of

thoracic diseases using the Chest X-Ray14 images. The SDFN controls the domain information and high resolution knowledge of the local lung region. The local regions of the lungs were detected and cropped using the lung region detector. Then, two CNN classification methods were used for the extraction of features to get specific features from the entire Chest X-Ray images and cropped lung regions. Finally, the features obtained were fused using feature fusion module for the classification of diseases. The SDFN method has higher performance compared to existing methods in terms of disease classification and reliable lesion detection. The computation complexity of the SDFN method was high in identifying the diseases.

Wang [12] developed an adaptive sampling method to detect thoracic diseases using the DenseNet model and adaptive data augmentation approach. The imbalance data were trained to bias the classes to a threshold level. The adaptive sampling method monitors the performance of method during the training phase and automatically increases the weights of weakly performed classes. The data set used in the adaptive sampling method was Chest X-Ray14. The adaptive sampling method was a cost-sensitive method and has the capacity to handle the larger dataset. The developed method has not identified the smaller lesion regions due to which the disease detection performance was reduced.

Allaouzi and Ahmed [13] developed a method that combines DenseNet for feature extraction from images and the supervised multiple label classifier to detect the thoracic disease from Chest X-Ray. The process of feature extraction was carried out using the DenseNet-121 architecture. The data set used in this CNN-based model are Chest X-Ray14 and Chexpert. This CNN-based model has lower efficiency on an imbalanced dataset.

Chen [14] developed a DualCheXNet model based on the dual asymmetric features learning for multiple labels thoracic disease classification from Chest X-Ray. The ResNet and DenseNet sub-networks were combined to extract the particular features with various abnormalities from the Chest X-Ray. The DualCheXNet method allows two feature fusion processes such as feature level fusion and decision level fusion to form the embedded complementary feature learning in the DualCheXNet. The complementary feature learning has been optimized and DualCheXNet model achieved higher performance in detecting diseases. The DualCheXNet model was unable to identify the various texture distributions in the lung portion.

Wang [15] created the Chest X-Ray8 database using NLP techniques and developed an imperfectly supervised classification and localization framework

of thoracic disease detection. The developed method utilized Stochastic Gradient Descent (SGD) for fine-tuning the CNN to shuffle the dataset into three subgroups. In this deep CNN only transition layer and prediction layer was trained from scratch. The global pooling and prediction layers were designed for the classification of pathologies and also to produce the likely-hood map of pathologies. The developed method showed better performance in the detection of the disease. However, the developed method used only image-level labels for the training which is difficult for computer-aided diagnosis. Later their Chest X-Ray8 database is extended to Chest X-Ray14 and made available to the public.

Gundel [16] developed Location-aware Dense Network (DNetLoc) for thoracic disease detection. The spatial information of pathologies in X-ray was incorporated during training and testing. The spatial location label was trained as independent and binary class using the function of cross entropy. The DNetLoc showed improved performance when the location information of pathologies was used explicitly. The DNetLoc method showed limitations in an official split of images while training and testing due to the larger inconsistency in the image labels.

### 3. Problem statement

The problem identified from the literature review of existing methods in detecting thoracic diseases are explained in this section:

The existing thoracic disease detection models have lower efficiency while training with a small and imbalanced dataset, so it is needed to increase the data instance in pathology classes to improve the overall performance. The DenseNet based model with adaptive data augmentation technique has lower efficiency in detecting tiny lesions and the cascade learning method needs to be included to increase the performance. The existing approaches have lower performance in analysing the various texture distributions in the lung portion and an effective parameter optimization method needs to be applied to increase the performance while detecting thoracic diseases.

### 4. Objective

The self-sequential attention layer is added in the DenseNet method to train the classes to the threshold level to handle the imbalanced dataset. The cascade learning method is applied to the proposed SAL-DN method to detect the small lesion in the image. The proposed SAL-DN method is established to increase the thoracic disease detection performance on the Chest X-Ray14 dataset.

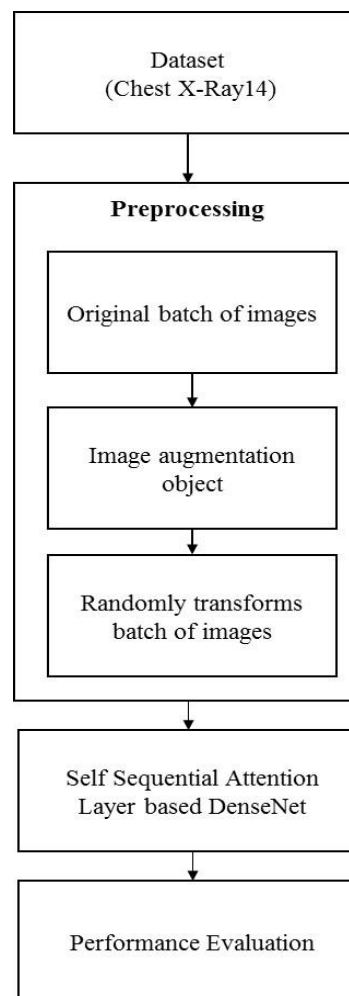


Figure. 1 Block diagram of proposed SAL-DN for the detection of thoracic diseases

### 5. Proposed method

This section explains the framework of the proposed SAL-DN, the global structure, and a detailed description of the framework is also outlined. The tasks performed by the proposed system are given as follows:

Deep learning approaches are widely used to get an accessible and accurate diagnoses of thorax diseases on chest radiographs. In this research, the SAL-DN model is proposed to enhance the performance of thorax disease detection. The Chest X-Ray14 dataset is collected which consists of 14 types of thoracic disease labels obtained from 30,805 patients. Then, the data pre-processing is done to resize the images into  $224 \times 224$  pixels instead of the whole image. ImageDataGenerator is used for dataset augmentation to make faster convergence while training the network. These resized images are applied as input to the SAL-DN for the thoracic diseases detection. The block diagram of the proposed SAL-DN method is shown in Fig. 1.

## 5.1 Data collection

The Chest X-Ray14 dataset includes 112,120 frontal view X-Ray images of 30,805 different patients aged from 1 to 95 years. Among these 60,412 images are normal cases and do not include any disease labels and the remaining 51,708 images consist of more than one disease label which are mined from the particular radiology text report of the patient by using Natural Language Processing (NLP) approaches. The 14 thoracic disease labels include nodule, atelectasis, edema, effusion, fibrosis, pleural thickening, consolidation, pneumonia, emphysema, mass, infiltration, cardiomegaly, pneumothorax, and hernia. These Chest X-Ray14 images were originally saved in Portable-Network-Graphics (PNG) format and rescaled into  $1024 \times 1024$  size. This dataset is split into the training subset of 80% and the testing subset of 20% patient-wise official split. The 36024 images are having more than one pathology label and 50500 images are normal case in the training dataset. The 15684 images with more than one pathology label and 9912 normal case images in testing dataset.

## 5.2 Pre-processing

Data pre-processing is carried out to format the ChestX-Ray14 images, where the images are resized to  $224 \times 224$  pixels. In a pre-processing step, the ImageDataGenerator tool is used for data augmentation to make faster convergence while training the network. The ImageDataGenerator tool is used to generate the maximum training data from the original images to overcome the problem of overfitting. The data augmentation is used to increase the size of the Chest X-Ray14 dataset that solves the overfitting problem and generalise the model better.

### 5.2.1. Data augmentation

Data augmentation is the method used to expand the size of the training dataset by creating a modified version of the original image and which includes position augmentation and colour augmentation.

- Position augmentation changes the pixel's positions in the images such as resizing, cropping, translation, padding, and rotation.
- Colour augmentation alters the properties of colour in the images by altering the pixel values like brightness, contrast, and saturation.

In resizing, the image is resized to the required size by reducing pixels of Chest X-Ray14 images. In cropping, a particular portion of the image is selected to identify thoracic diseases. In flipping and rotation, the image is rotated vertically or horizontally as

required and it allows to display of the selected images properly. In padding, the image is padded with the given values on all the sides and it is added to the frame of the Chest X-Ray14 images to allow much space for the kernel to cover the images. In translation, the image is moved to the x-axis or y-axis and the brightness makes the resultant image lighter or darker as compared to the original image. Brightness is calculated by using Eq. (1).

$$Y' = 0.299R + 0.587G + 0.114B \quad (1)$$

where  $R, G, B$  are the standard Red Green Blue (sRGB) coordinates according to International Telecommunication Union Radio communication (ITU-R).

The contrast is known as the degree of separation among the darker and brighter areas of an image. The contrast of the image can also be changed and it is carried out to make the features of images more clear. The contrast is calculated using Eq. (2).

$$\text{Contrast} = \frac{\text{Change in luminance}}{\text{Average luminance}} \quad (2)$$

Saturation is the separation between the colours in an image, it is carried out to make it appear more vivid. Saturation is calculated using Eq. (3).

$$\text{Saturation} = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \quad (3)$$

The augmented images are applied as an input to the proposed SAL-DN for the detection of diseases.

## 5.3 Architecture of SAL-DN

The architecture of Self-sequential Attention layer based Dense Convolutional Network (SAL-DN) is shown in Fig. 2. Densenet-121 is used as a backbone network, where every layer is connected to each other hence it is also called densely connected CNN or DenseNet. The idea of DenseNet is to allow maximum details to flow among all the layers in the network by connecting with each other. For  $l$  layers in DenseNet it includes  $l(l+1)/2$  direct connections. In every layer, the feature maps of the preceding layers will be used as input, and the feature maps of its own layer will be utilized as input in every subsequent layer. The feature maps are used to identify the thoracic diseases and exploited to create the class discriminative attention maps using SAL.

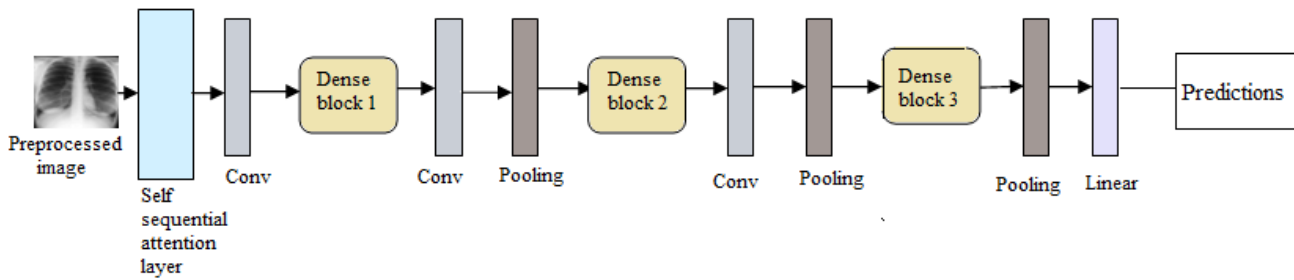


Figure. 2 Architecture of self-sequential attention layer based DenseNet-121

Lesser resolution feature maps in deep CNNs includes a greater level of visual characterization and higher semantic details.

Further, the higher resolution feature maps from the shallow layer include stronger information about abnormalities, so the resolutions of feature maps are necessary to create the class discriminative attention maps using SAL.

In the proposed SAL-DN method self-sequential attention layer, Convolution, Pooling, three dense blocks, and linear layer are used to detect thoracic diseases. The SAL uses the co-relation among the class labels and abnormal pathology by analysing the known feature maps from DenseNet-121. These feature maps are created by DenseNet-121 and are given as input to the convolution layer. The layers between the dense blocks are transition layers which do convolution and pooling in the proposed SAL-DN model. The convolution layer produces many feature maps which perform the convolution process on pre-processed images using a linear filter and then passes the feature maps to pooling layer. i.e., every neuron in feature maps get input from  $N \times N$  regions which are a subset of all the feature maps. The pooling layer follows the convolution layer activated by a non-linear activation function which is responsible for reducing the dimensionality of feature maps produced by the convolutional layer. In dense blocks, every layer will be connected to each other and the size of feature maps inside the block will remain the same. The feature maps from the pooling layer will be forwarded to the linear activation layer it performs biasing and predicts thoracic diseases. The SAL-DN is convolutional feed-forward network connects the outcome of the  $l^{th}$  layer as input to the  $(l + 1)^{th}$  layer as shown in Eq. (4).

$$X_l = H_l(X_{l-1}) + X_{l-1} \quad (4)$$

where  $X_l$  is the output of  $l^{th}$  layer,  $H_l$  is function produced by feature maps, the  $X_{l-1}$  is the feature map at layer  $l - 1$ . To improve the information flow among the layers, different connection patterns are introduced which directs the information's from any layer to every subsequent layer. The  $l^{th}$  layer gets the feature maps from all preceding layers,  $X_0, \dots, X_{l-1}$  and it is explained in Eq. (5).

$$X_l = H_l([X_0, X_1, \dots, X_{l-1}]) \quad (5)$$

where  $X_l$  is the feature maps at  $l^{th}$  layer,  $[X_0, X_1, \dots, X_{l-1}]$  is the concatenation of feature maps produced in layers from  $0, \dots, l - 1$ .

To facilitate down-sampling in SAL-DN architecture, the network is split into various densely connected dense blocks. The convolution and pooling process is carried out among the blocks of transition layers. The transition layer utilized in the proposed method includes a  $1 \times 1$  convolutional layer,  $3 \times 3$  convolution layer, and  $2 \times 2$  average pooling layer. Every  $H_l$  function produces the  $K$  feature maps and it follows the  $l^{th}$  layer as shown in Eq. (6).

$$K_l = K_0 + K \times (l - 1) \quad (6)$$

where  $K_0$  is the overall number of channels in the input layer and  $K$  hyperparameter is the growth rate of the network and every layer adds  $K$  feature maps to its state.

The growth rate tells the amount of new information each layer generates to the global state. The global state can be accessed anywhere in the network and it is not true in the traditional network architectures. It is not necessary to replicate from layer to layer, the  $1 \times 1$  convolution is introduced as bottleneck layer before the  $3 \times 3$  convolution layer. This improves a number of input feature maps and also improves computational efficiency.

## 6. Experimental setup and result analysis

Thoracic diseases infect millions of people every year. The first step in diagnosing these diseases is a procedure called Chest X-Ray. In this research, the SAL-DN method is proposed for the thoracic disease detection and classification from the Chest X-Ray14 dataset. This section gives the details of the dataset, experimental setup, performance, and comparison analysis.

### 6.1 Dataset

The Chest X-Ray14 dataset is utilized to evaluate the proposed SAL-DN method performance. The Chest X-Ray14 dataset includes 112,120 front views of chest radiographs that are obtained from 30,805 patients with 14 types of disease image labels. The dataset is mined from associated reports of radiology by utilizing natural language processing. The release of such a large scale new Chest X-ray14 database to public triggered a lot of interest and progress in the field of chest radiology.

### 6.2 Experimental setup

The proposed SAL-DN method in thoracic disease detection is evaluated by utilizing the Intel i7 processor system configurations, 500GB of hard disk and 8GB of RAM, GPU configuration is 22GB NVIDIA RTX 2080Ti. The proposed SAL-DN method is tested and developed by python 3.

### 6.3 Metrics

The performance of the proposed SAL-DN method is evaluated in a Chest X-Ray14 dataset using the following metrics:

- **False positive rate (FPR)**

False Positive Rate is an accuracy metric that are measured in the subsets of machine learning. The equation for False Positive Rate is shown in Eq. (7).

$$FPR = \frac{FP}{TN + FP} \quad (7)$$

- **True positive rate (TPR)**

True positive rate is the ratio of true instances, which means that the feature vectors of a malicious applications were classified correctly. The true positive rate is calculated using the Eq. (8).

$$TPR = \frac{TP}{TP + FN} \quad (8)$$

- **Area under curve (AUC)**

The area under curve calculates entire two-dimensional areas under the ROC curve. The area under the curve is calculated using the Eq. (9).

$$AUC = \int_a^b f(x)dx \quad (9)$$

Where, to determine the area under the curve  $y = f(x)$  among  $x = a$  and  $x = b$ , with the integration of  $y = f(x)$  among the limits of  $a$  and  $b$ .

- **Receiver operating characteristic curve (ROC)**

The ROC curve is plotted by creating TPR against FPR at various threshold values. The mathematical expressions for FPR, and TPR are shown in Eq. (7) and (8). Where,  $TP$  = True Positive,  $TN$  = True Negative,  $FP$  = False Positive and  $FN$  = False Negative.

### 6.4 Results and comparison

Table 1 shows result and performance comparison in terms of AUC scores of the proposed SAL-DN method with existing methods such as Thorax-Net [10], Wang [15], and Gundel [16] by using a Patient-wise official split of dataset. The Patient-wise official split divides the dataset among training set of 86,524 images and testing set of 25,596 images. The training set includes 50,500 normal and 36,024 diseased images, similarly testing set includes 9912 normal and 15,684 diseased images. This split includes 80% as a training set, 20% as a testing set, and 10% of the training set is utilized for validation dataset. The uniqueness of this split is images of the same patient must appear either in training or in the test set not in both. The method utilized in [16] was utilized more than 180,000 images for training from PLCO dataset along with Chest X-Ray14 data as extra training data. The proposed method outperformed the other three models without being trained by external data.

Fig. 3 shows the comparison graph of the average AUC score of the proposed model with the existing three models under Patient-wise official split. Fig. 5 shows the ROC curve of diagnosis obtained by SAL-DN under patient-wise official split of the dataset. Proposed SAL-DN showed higher average AUC

Table 1. Performance comparison of SAL-DN with existing methods in terms of AUC Scores under patient-wise official split of chest X-Ray14 dataset.

Thoracic diseases	Wang [15]	Gundel [16]	Thorax-Net[10]	SAL-DN
Atelectasis	0.7003	0.767	0.7505	0.85
Cardiomegaly	0.8100	0.883	0.8710	0.95
Effusion	0.7585	0.828	0.8181	0.91
Mass	0.6933	0.821	0.7994	0.91
Pneumonia	0.6580	0.731	0.6938	0.80
Consolidation	0.7032	0.745	0.7147	0.85
Edema	0.8052	0.835	0.8354	0.93
Emphysema	0.8330	0.895	0.8428	0.94
Fibrosis	0.7859	0.818	0.8040	0.86
Infiltration	0.6614	0.709	0.6815	0.75
Nodule	0.6687	0.758	0.7147	0.81
Pleural Thickening	0.6835	0.761	0.7463	0.84
Pneumothorax	0.7993	0.846	0.8254	0.93
Hernia	0.8717	0.896	0.9022	0.98
Average	0.7451	0.807	0.7876	0.8715

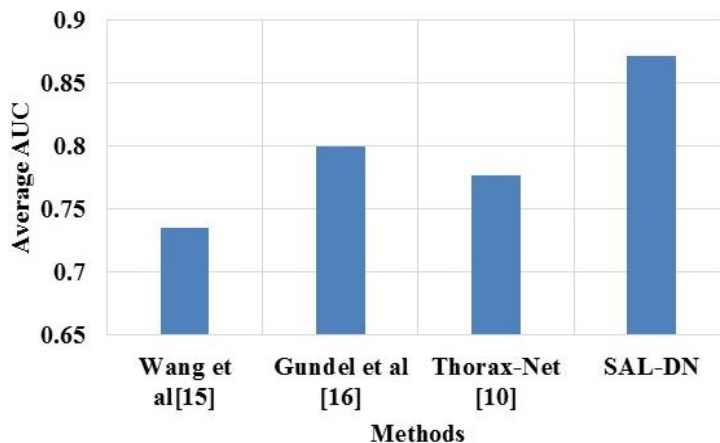


Figure. 3 Comparison graph of proposed SAL-DN with existing methods in terms of average AUC scores by using patient-wise official split

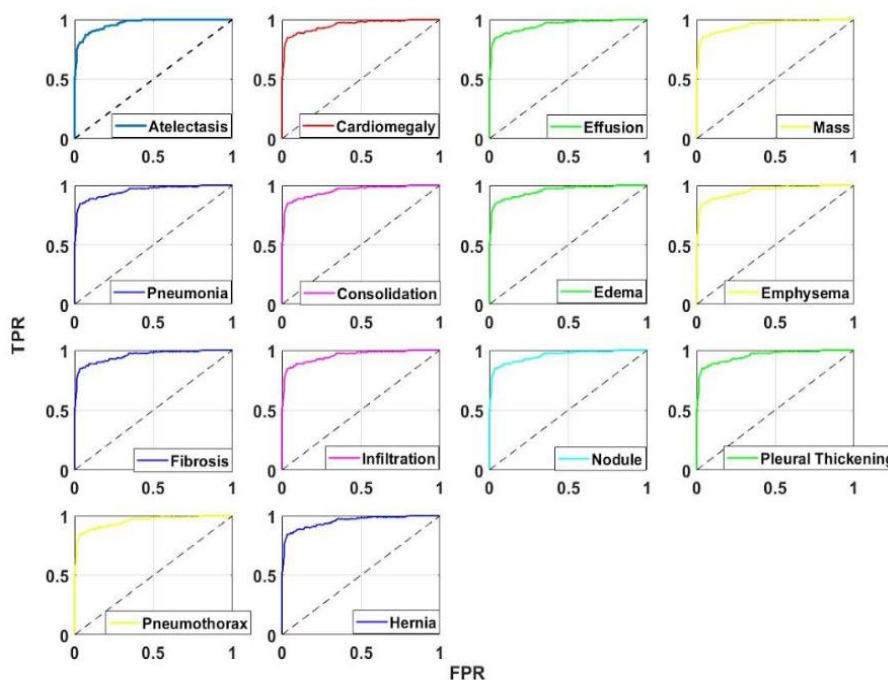


Figure. 4 The ROC curves of diagnosis obtained by SAL-DN on patient-wise official split of chest X-Ray14 dataset

Table 2. Performance comparison of SAL-DN with existing methods in terms of AUC scores under Image-wise random split of chest X-Ray14 dataset

Thoracic diseases	Wang [15]	Thorax-Net[10]	SAL-DN
Atelectasis	0.716	0.856	0.871
Cardiomegaly	0.807	0.957	0.961
Effusion	0.784	0.919	0.933
Mass	0.706	0.905	0.919
Pneumonia	0.633	0.869	0.875
Consolidation	0.708	0.870	0.873
Edema	0.835	0.943	0.949
Emphysema	0.815	0.959	0.961
Fibrosis	0.769	0.889	0.896
Infiltration	0.609	0.776	0.782
Nodule	0.671	0.832	0.839
Pleural Thickening	0.708	0.883	0.892
Pneumothorax	0.806	0.941	0.943
Hernia	0.767	0.951	0.968
Average	0.738	0.896	0.904

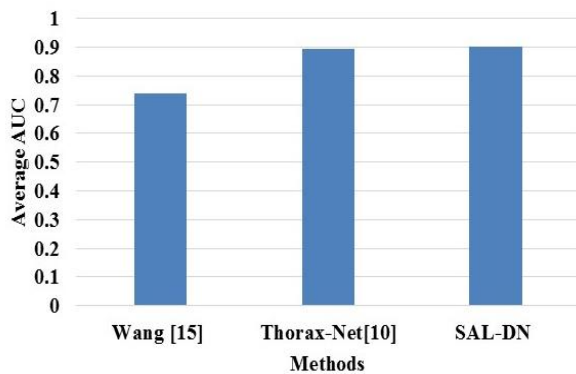


Figure. 5 Comparison graph of proposed SAL-DN with existing methods in terms of average AUC scores under image-wise random split

scores of 0.8715 for the detection of 14 thoracic diseases compared to other three methods.

Table 2 shows the AUC score comparison of SAL-DN with existing Wang [15] and Thorax-Net [10] method under Image-wise random split of Chest X-Ray14 dataset. The complete dataset is randomly split into 70% for training process, 10% for validation process, and 20% for testing process. When Image-wise random split of data is utilized, the radiography from the similar subject appears simultaneously in training and testing sets. The proposed SAL-DN showed effective performance using Image-wise random split than the existing methods by achieving an average AUC of 0.904.

In both split, SAL-DN shows higher average AUC values compared to other competitors. The major reason is the incorporation of the self-sequential attention layer with the deep learning model such as DenseNet-121. The self-sequential attention layer guide DenseNet to focus more on abnormal regions rather than focusing equally on the

whole image. Also learned features from DenseNet with the help of the self-sequential attention layer had enough discriminatory powers to do the effective classification of pathologies.

## 7. Conclusion and future work

In this research, the SAL-DN method is proposed to enhance the thorax disease prediction. Data pre-processing is employed to resize the images into  $224 \times 224$  pixels. The ImageDataGenerator is utilized for dataset augmentation to make faster convergence while training the network. The data augmentation such as position augmentation and colour augmentation are applied to the images. The augmented images are applied as input to the proposed SAL-DN to predict thoracic diseases. The proposed method showed advantage of providing the best representation of images and the capacity to handle large dataset like Chest X-Ray14. The self-sequential attention layer can identify abnormal regions in X-ray and guide DenseNet-121 sequentially to improve classification performance. The proposed SAL-DN method was compared to three state-of-the-art deep learning methods using the Patient-wise official split of the dataset and also against the other two deep learning methods using the Image-wise random split of data. The proposed model achieved a higher average AUC per-class of 0.8715 and 0.904 in both the split, respectively. In future work, the thoracic disease detection performance can be enhanced with a lesser number of training data by adding a softmax function on the last layer and Rectified Linear Unit (ReLU) on the rest of the layers. Also need to address the class imbalance issue of this dataset which is not considered here.



## Conflicts of Interest

The authors declare no conflict of interest.

## Author Contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, have been done by 1<sup>st</sup> author. The supervision, and project administration, have been done by 2<sup>nd</sup> author.

## References

- [1] H. Wang, S. Wang, Z. Qin, Y. Zhang, R. Li, and Xia, “Triple attention learning for classification of 14 thoracic diseases using chest radiography”, *Medical Image Analysis*, Vol. 67, pp. 101846, 2020.
- [2] H. Mane, P. Ghorpade, and V. Bahel, “Computational Intelligence Based Model Detection of Disease using Chest Radiographs”, In: *Proc. of the 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*, pp. 1-5, 2020.
- [3] C. H. Liang, Y. C. Liu, M. T. Wu, F. Garcia-Castro, A. Alberich-Bayarri, and F. Z. Wu, “Identifying pulmonary nodules or masses on chest radiography using deep learning: external validation and strategies to improve clinical practice”, *Clinical Radiology*, Vol. 75, No. 1, pp. 38-45, 2020.
- [4] I. Chamveha, T. Tongdee, P. Saiviroonporn, and W. Chaisangmongkon, “Local Adaptation Improves Accuracy of Deep Learning Model for Automated X-Ray Thoracic Disease Detection: Thai Study”, *arXiv preprint arXiv:2004.10975*, 2020.
- [5] R. H. Abiyev and M. K. S. Ma’aitah, “Deep convolutional neural networks for chest diseases detection”, *Journal of healthcare engineering*, 2018.
- [6] X. Li, J. H. Thrall, S. R. Digumarthy, M. K. Kalra, P. V. Pandharipande, B. Zhang, C. Nitiwarangkul, R. Singh, R. D. Khera, and Q. Li, “Deep learning-enabled system for rapid pneumothorax screening on chest CT”, *European journal of radiology*, Vol. 120, pp. 108692, 2019.
- [7] T. K. K. Ho and J. Gwak, “Utilizing Knowledge Distillation in Deep Learning for Classification of Chest X-Ray Abnormalities”, *IEEE Access*, Vol. 8, pp. 160749-160761, 2020.
- [8] S. Bharati, P. Podder, and M. R. H. Mondal, “Hybrid deep learning for detecting lung diseases from X-ray images”, *Informatics in Medicine Unlocked*, Vol. 20, pp. 100391, 2020.
- [9] K. Wang, X. Zhang, S. Huang, F. Chen, X. Zhang, and L. Huangfu, “Learning to Recognize Thoracic Disease in Chest X-Rays with Knowledge-Guided Deep Zoom Neural Networks”, *IEEE Access*, Vol. 8, pp. 159790-159805, 2020.
- [10] H. Wang, H. Jia, L. Lu, and Y. Xia, “Thorax-Net: An Attention Regularized Deep Neural Network for Classification of Thoracic Diseases on Chest Radiography”, *IEEE journal of biomedical and health informatics*, Vol. 24, No. 2, pp. 475-485, 2019.
- [11] H. Liu, L. Wang, Y. Nan, F. Jin, Q. Wang, and J. Pu, “SDFN: Segmentation-based deep fusion network for thoracic disease classification in chest X-ray images”, *Computerized Medical Imaging and Graphics*, Vol. 75, pp. 66-73, 2019.
- [12] H. Wang, Y. Y. Yang, Y. Pan, P. Han, Z. X. Li, H. G. Huang, and S. Z. Zhu, “Detecting thoracic diseases via representation learning with adaptive sampling”, *Neurocomputing*, Vol. 406, pp. 354-360, 2020.
- [13] I. Allaouzi and M. B. Ahmed, “A novel approach for multi-label chest X-ray classification of common thorax diseases”, *IEEE Access*, Vol. 7, pp. 64279-64288, 2019.
- [14] B. Chen, J. Li, X. Guo, and G. Lu, “DualCheXNet: dual asymmetric feature learning for thoracic disease classification in chest X-rays”, *Biomedical Signal Processing and Control*, Vol. 53, pp. 101554, 2019.
- [15] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, “Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases”, In: *Proc. of the IEEE conference on computer vision and pattern recognition*, pp. 2097-2106, 2017.
- [16] S. Guendel, S. Grbic, B. Georgescu, S. Liu, A. Maier, and D. Comaniciu, “Learning to recognize abnormalities in chest x-rays with location-aware dense networks”, In: *Proc. of Iberoamerican Congress on Pattern Recognition*, pp. 757-765, 2018.
- [17] O. Ruuskanen, E. Lahti, L. C. Jennings, and D. R. Murdoch, “Viral pneumonia”, *The Lancet*, Vol. 377, No. 9773, pp. 1264-1275, 2019.