Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 218 (2023) 74–82

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering

An Attention-based Pneumothorax Classification using Modified Xception Model

Upasana C.[a1], *Anand Shanker Tewari*[b1], *Jyoti Prakash Singh*[c1]

[1]*National Institute of Technology Patna*, 800005, *India*

Abstract

Chest radiographs, among other medical imaging, are the most significant and effective diagnostic tools for detecting lung disorders. Numerous research is being done to develop reliable and automatic diagnostic systems for detecting diseases using chest radiographs. Pneumothorax is a potentially fatal condition that needs early diagnosis and treatment. Artificial Intelligence (AI) approaches have offered promising results in medical imaging. Different AI-based approaches for classifying pneumothorax using medical images have been proposed. However, there is limited medical imaging available for the identification of pneumothorax. This work aims to develop a model to detect pneumothorax in chest X-ray images by combining xception network with an attention module. The proposed model was experimented on 2,597 chest X-ray images and has achieved training accuracy of 99.18%, validation accuracy of 87.53% and average AUC (Area under the ROC Curve) of 90.00%.

Keywords: Pneumothorax; Xception; Deep Learning; Channel Attention; Spatial Attention; Transfer Learning

1. Introduction

Pneumothorax is a medical term for the condition that arises when there is free air present in the pleural space. It is a condition that could be fatal if not caught and treated right away. Patients with chest injuries have 30 to

E-mail address: [a]*upasanac.phd*19.*cs@nitp.ac.in*

E-mail address: [b]*anand@nitp.ac.in*

E-mail address: [c] *jps@nitp.ac.in*

1877-0509
10.1016/j.procs.2022.12.403

39 percent probability of developing a pneumothorax [1]. The quantitative analysis of chest imaging performed by medical professionals is helpful in the diagnosis the lung disorders. However, timely evaluation of such diseases can be complex due to queues of cases awaiting. So, an automatic and efficient method is urgently needed to increase the reliability and speed up the diagnosis of pneumothorax using X-ray images. The recent advancements in deep learning and big data technology benefit researchers engaged in medical image analysis and could aid radiologists in identifying and quantifying lung illnesses in interpretative situations [7]. Many researchers have noted AI, which has already proven its effectiveness in the diagnosis of other diseases such as tuberculosis [2], cancer [3, 4], leukemia [5], brain tumour [6] and so on. For instance, Kidney disease is one of the fatal diseases that leads to the death of people at a young age. Singh et al. [8] proposed an innovative deep learning method for the early detection and prediction of kidney disease. The authors developed a deep neural network and compared the results to other machine learning models already in existence. Similarly, leukaemia is another deadly disease whose mortality rate is 6.1 percent per 100,000 men and women per year [1]. Rastogi et al. [5] proposed a model "LeuFeatx" for precisely classifying leukocytes that cause leukaemia. LeuFeatx is a customized VGG16 model implemented on three additional datasets to check the robustness of the proposed model. It was found that LeuFeatx exhibited satisfactory results as compared to the state-of-the-art (SOTA) methods. In Our proposed model, we have used a deep learning model: xception and an attention mechanism to classify pneumothorax in chest x-ray images.

The main contributions of this paper are:

- The analysis is done on a publicly available dataset and we enriched the images using pre-processing.
- The work used the attention mechanism on the output features of xception net to increase the effectiveness of the model.
- The output of the proposed model is compared with the other SOTA deep learning models.

The paper is organised in the following manner: the second section covers the works relevant to this study. The proposed method is explained on a deeper level in the third part. The fourth section provides an illustration of the execution and evaluation of the recommended technique, and the fifth section consists of the conclusion as well as the future work objectives.

2. Related Work

Different classification models that incorporate machine learning have been developed in past years. Deep learning techniques are most widely used among them. In this section, a quick summary of detection of lung diseases is discussed:

Wang et al. [9] have made available a ChestX-ray8 dataset with eight thoracic disorders that contains over 100,000 chest radiographic images. They evaluated the performance of pre-trained CNN models: GoogLeNet, AlexNet, and VGG16, to determine the classification performance. However, the AUC for detecting the pneumothorax was only 80% because of the lack of advanced optimization. Cicero et al. [10] have done six class classifications, including five pathologies (Cardiomegaly, Consolidation, Edema, Effusion and Pneumothorax) and a normal class, using a deep learning technique called GoogleNet. As a preprocessing, they converted the 16-bit DICOM files to 8-bit PNG (Portable Network Graphics) format and down-sampled them to 256×256 to fulfil the required criteria. An explicit limitation was that radiographs with "minimal" findings other than the five anomalies were regarded as normal. Taylor et al. [11] examined the performance of four deep learning models (VGG16, VGG19, Xception, and Inception) for moderate and large pneumothorax detection on a customized dataset. In order to avoid lethal delays in radiologists' reviews during emergencies. The pneumothorax segmentation framework proposed by Ouyang et al. [12] included images with annotations at the pixel level and on a weak image level. The authors trained ResNet-101, an image classification model, with weakly labelled data to generate attention maps. The above attention maps were combined with pixel-level annotations to train the Tiramisu segmentation model. To categorize lung diseases based on the data

[1] https://seer.cancer.gov/statfacts/html/leuks.html

provided by ChestX-ray8, a two-stage neural network model was proposed by Yao et al. [13]. This model combined a densely linked image encoder with an RNN (Recurrent Neural Network) decoder. In comparison to Wang et al. [9] study, Yao et al. result showed enhanced classification performance, and AUC for identifying pneumothorax was 84.1%. Jun et al. [14] used the ensemble technique to classify the chest radiography images. The authors downsized the original dataset into three sizes: 512×512, 384×384, and 256×256, then applied the ResNet50 model to each dataset and averaged the results to get the final classification result. Jakhar K, Kaur A, and Gupta D [15] developed a segmentation approach to precisely predict the lung region damaged by pneumothorax. Their model generates a binary mask to help radiologists to identify the position and magnitude of pneumothorax using the U-Net architecture. Park et al. [16] employed a 26-layer CNN model to diagnose pneumothorax on chest X-ray images taken after the needle biopsy to improve clinical workflow especially for patients who need additional treatment.

3. Proposed Method

The proposed approach has combined Xception model with Attention module. The Xception model fetches the pre-processed data for feature extraction. Weights of the Xception model are pre-trained on the Imagenet dataset [17]. In order to refine the features, we have used the attention module, which is consist of two parts: *Channel Attention* and *Spatial Attention*. We have discussed the structure of both Xception model and Attention module in subsections 3.1, 3.2 and 3.3, respectively.

3.1. Xception model

The pre-processed images are passed to the Xception model. It is a 71-layer [2] deep CNN (Convolution Neural Network) inspired by google's Inception model and based on a radical interpretation of the Inception model. It employs a depth-wise separable convolutional layer [18]. It starts with a 1x1 point-wise convolution and then moves to a 3x3 depth wise convolution followed by a logistic regression. Figure 1 depicts the separable convolution in Xception model.

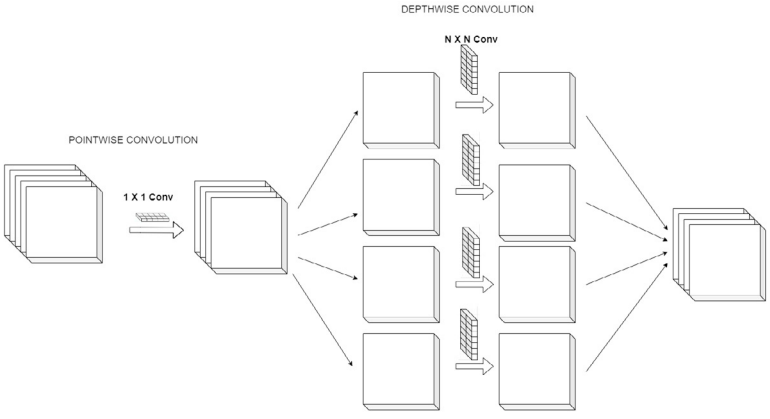

Fig. 1. Separable convolution in Xception model.

3.2. Channel Attention

The proposed model has used the channel attention to the output of the xception to improve the capabilities of a deep convolution neural network. channel attention concentrates on "what" is relevant given an input image, as

[2] https://www.mathworks.com/help/deeplearning/ref/xception.html#:~:text=Xception%20is%20a%20convolutional%20neural,%2C%20pencil%2C%20and%20many%20animals.

each channel of a feature map is is regarded as a feature detector. We minimize the spatial dimension of the input feature map to compute channel attention with more efficiency. So, channel attention is primarily concentrated on the essential channels over the other channels. The simplest approach to do this is to increase the value of the more relevant channels. Figure 2 describes the working of channel attention.

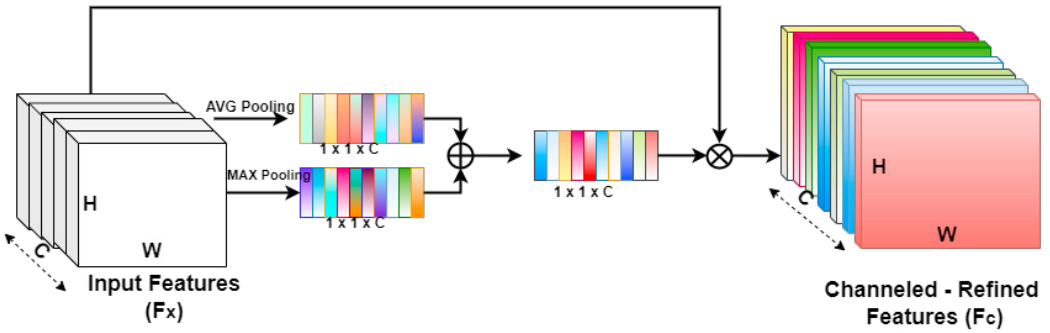

Fig. 2. Channel attention.

In channel attention, we use both global max pooling (GMP) and global average pooling (GAP) and the addition of both the poolings is multiplied with the input features (F_x) that we obtained from the xception model. The channelled-refined features (F_c) are then passed through the spatial attention module which is discussed in the next subsection.

3.3. Spatial Attention

The channelled-refined features (F_c) are then passed through the spatial attention module. It complements the channel attention, which primarily concentrates on evaluating "where" is an informative portion. Spatial attention is the concatenation of max-pooling and average-pooling along the channel axis, and in order to generate the spatial attention map, the Conv2D layer is applied. The computation for spatial attention is shown in equations 1, and 2 [19]

$$M_s(F) = \sigma(f^{3\times3}([Avg\_Pool(F); Max\_Pool(F)])) \tag{1}$$

$$M_s(F) = \sigma(f^{3\times3}([F^s_{avg}; F^s_{max}])) \tag{2}$$

Here, σ denotes the sigmoid function, $f^{3\times3}$ is a convolution operation with a filter size 3×3, and M_s(F) represents the spatial attention map features. The operation of the spatial attention module is illustrated in figure 3.

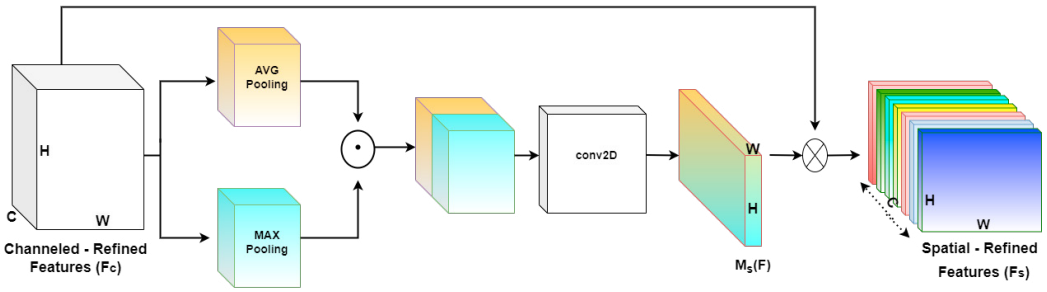

Fig. 3. Spatial attention.

The spatial attention map features are multiplied with the (F_c) to generate spatial-refined features (F_s). F_s is then passed through a GAP layer. Afterwards, the classification layer is used with a sigmoid activation function for the classification. Figure 4 depicts the architecture of our proposed model.

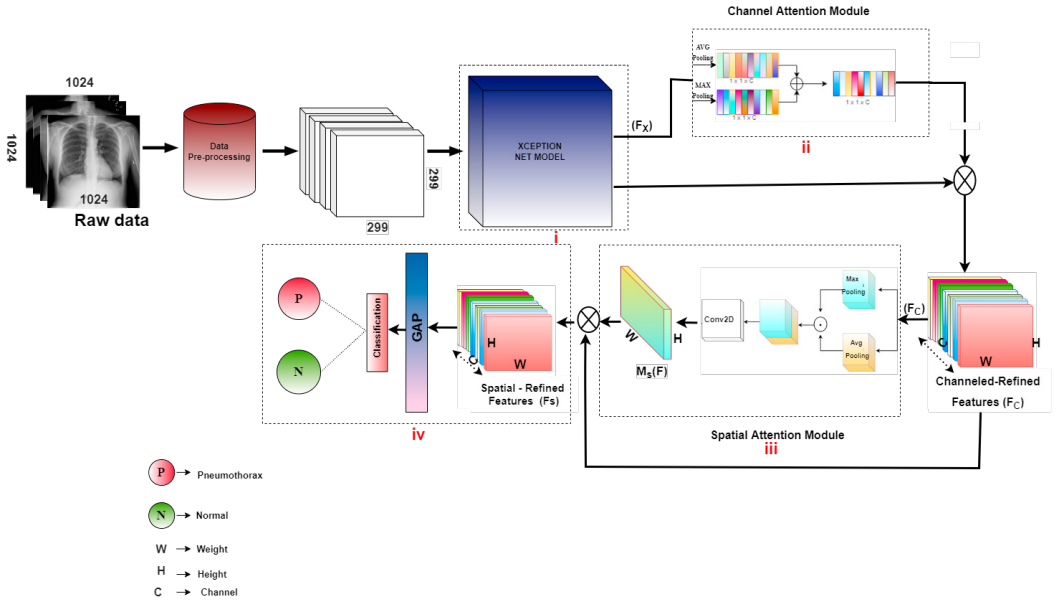

Fig. 4. Proposed architectural model.

Table 1. Hyperparameters used in our proposed model.

Sl No.	Properties	Values
1	Epochs	50
2	Learning rate	0.001
3	Optimizer	Adam
4	Activation	Sigmoid
5	Loss function	Binary_crossentropy
6	Batch size	32

4. Implementation and Evaluation

The proposed model is implemented in python, with *keras* deep learning libraries [20]. The experiments are conducted on core i7 11^{th} generation processor with 16 GB DDR-4 RAM. In the proposed model *binary cross-entropy* is used to calculate the loss and other hyper-parameters used are discussed in table 1.

4.1. Dataset and Pre-processing

Wang et al. [9] released a publicly available dataset containing 14 different lung diseases. There are 112,120 frontal X-ray images in the dataset, and all the images are in PNG format of size 1024 × 1024. There are 30,806 different patients included in the dataset. We have taken 2,597 X-ray images, out of which 1,313 images are labelled with "pneumothorax", and 1,284 are "normal" images. Figure 5 shows the sample of diseased and normal X-ray images. The training, validation and testing data are divided into the ratio of 6:2:2. The following are the pre-processing steps employed in this work:

- In the pre-processing phase, we have downscaled the image size to 299×299 to make the model computationally efficient.
- We considered 1,661 images as training data, 416 images as validation data, and 520 images as test data for the evaluation of the model.

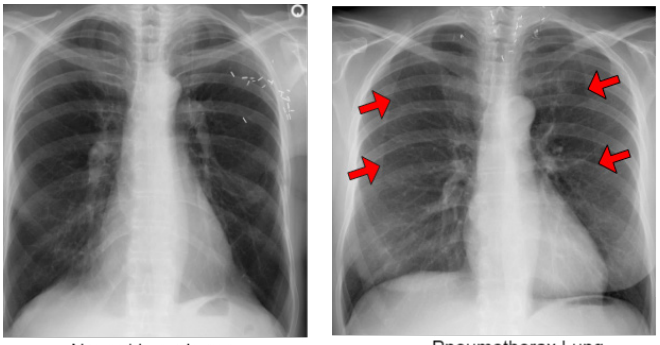

Fig. 5. Normal Vs Pneumothorax lung images.

- Data augmentations: Flipping by 7^o and zooming by 20% is applied to the images.

4.2. Evaluation metrics

Five evaluation metrics are used to assess the proposed model's effectiveness: accuracy, recall, precision, F1 score and AUC. The patients accurately classified as positive are true positives (TP), while incorrect positives are called false positives (FP). Those appropriately identified as negative are true negatives (TN). Finally, patients incorrectly diagnosed as negative were given the term false negative (FN). Equations 3 - 7 represent the formulation of each metric.

$$Accuracy = \frac{TN + TP}{TN + TP + FP + FN} \tag{3}$$

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{6}$$

$$Area\ under\ curve = \frac{1}{2}\left(\frac{TP}{P} + \frac{TN}{N}\right) \tag{7}$$

4.3. Result

We have implemented seven deep learning models, namely, VGG16, VGG19, ResNet_50, InceptionV3, EfficientNet_B0, Densnet121, Xception and compared their performance with the proposed model based on the evaluation metrics. As the table 2 shows, xception model gives the highest precision of 77%, but the DenseNet121 model shows the highest recall of 77%. Despite this, the F1-score for both the xception model and the DenseNet121 model is same i.e. 76%, and the xception model shows the highest AUC of 83%. Compared to the other deep learning models, the xception produces better outcomes. In order to boost the effectiveness of the xception model, we have added channel attention followed by spatial attention to it, which results in the improvement of the precision, recall and F1-score values by 86%, 85%, and 86%, respectively.

The validation accuracy of all the models is illustrated in figure 6, and the proposed method's validation accuracy, validation loss, confusion matrix, and AUC curve are represented in figure 7.

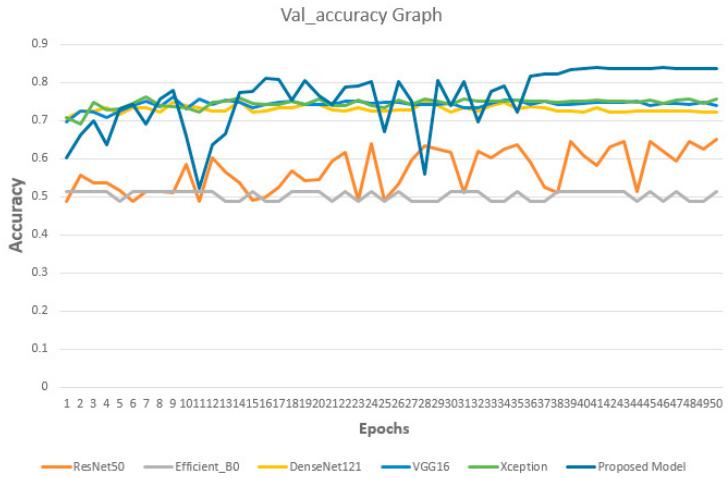

Fig. 6. Accuracy graph.

Table 2. comparison with respect to the SOTA models.

Model Name	Precision	Recall	F1-score	AUC
VGG16	74	74	73	74
VGG19	75	75	74	82
ResNet_50	65	61	56	76
InceptionV3	75	75	75	82
EfficientNet_B0	30	55	39	64
DenseNet121	76	77	76	81
Xception	77	76	76	83
Proposed Method	86	85	86	90

5. Conclusion and Future Work

Chest radiography is one of the convenient and cost-effective procedures for diagnosing lung disorders. Researchers have been inspired to automate the diagnosis of chest radiographs by the effective implementation of deep learning techniques in a variety of healthcare sectors. In this work, pneumothorax was diagnosed using a deep learning-based method. The suggested approach efficiently classifies X-ray images of pneumothorax utilising the xception and attention-based mechanism. Before deploying the models, we enriched images using pre-processing. The evaluation was conducted on 2,597 X-ray images. The proposed method beats various deep learning models, including VGG16, VGG19, ResNet 50, InceptionV3, EfficientNet B0, Densnet121, and Xception, with an average AUC of 90%, training accuracy of 99.13%, and validation accuracy of 87.53%. Future research will include (1) region-based categorization and (2) examining the potential associations between pneumothorax and other thoracic conditions.

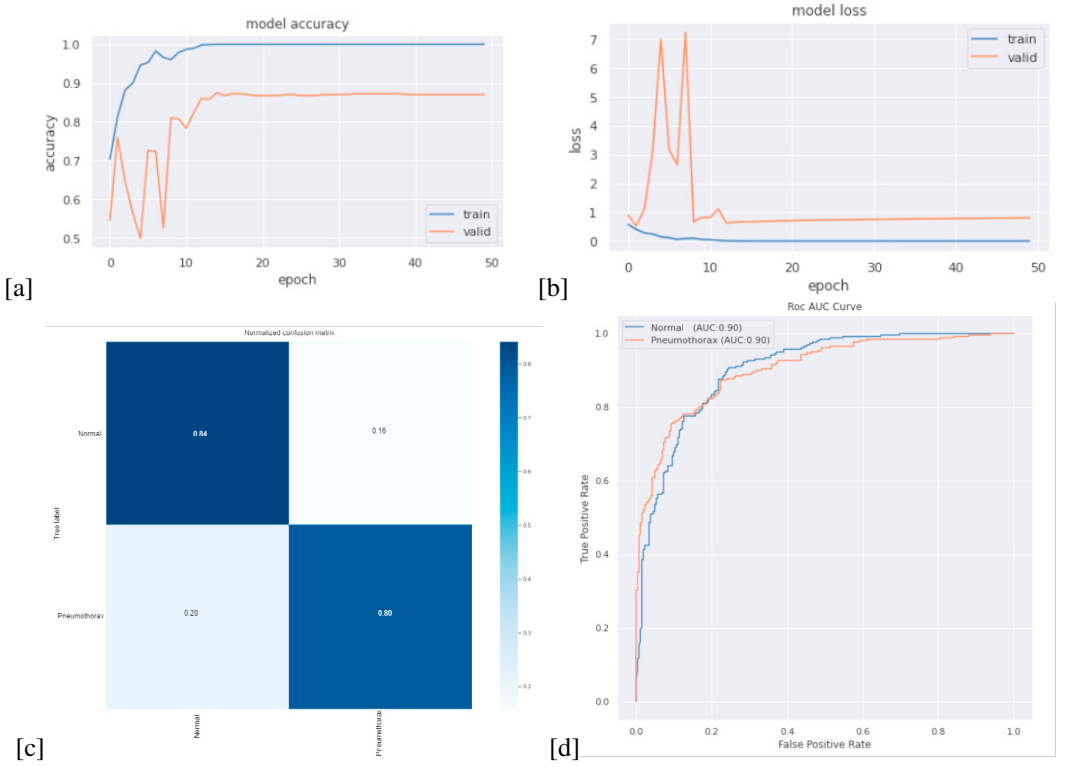

Fig. 7. The performance of pneumothorax classification from chest X-ray images; (a) The correlation between validation accuracy and training accuracy; (b)The relationship between the validation loss and training loss; (c) The confusion matrix of our method. (d) The AUC of our method.

References

[1] Shorr RM, Crittenden MI, Indeck M, Hartunian SL, Rodriguez AU. Blunt thoracic trauma. Analysis of 515 patients. Annals of surgery. 1987 Aug;206(2):200.

[2] Rahman T, Khandakar A, Kadir MA, Islam KR, Islam KF, Mazhar R, Hamid T, Islam MT, Kashem S, Mahbub ZB, Ayari MA. Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization. IEEE Access. 2020 Oct 15;8:191586-601

[3] Han G, Liu X, Zhang H, Zheng G, Soomro NQ, Wang M, Liu W. Hybrid resampling and multi-feature fusion for automatic recognition of cavity imaging sign in lung CT. Future Generation Computer Systems. 2019 Oct 1;99:558-70.

[4] Rastogi P, Khanna K, Singh V. Gland segmentation in colorectal cancer histopathological images using U-net inspired convolutional network. Neural Computing and Applications. 2022 Apr;34(7):5383-95.

[5] Rastogi P, Khanna K, Singh V. LeuFeatx: Deep learning–based feature extractor for the diagnosis of acute leukemia from microscopic images of peripheral blood smear. Computers in Biology and Medicine. 2022 Mar 1;142:105236.

[6] Michael Mahesh K, Arokia Renjit J. Evolutionary intelligence for brain tumor recognition from MRI images: a critical study and review. Evolutionary Intelligence. 2018 Oct;11(1):19-30.

[7] Rastogi P, Singh V, Yadav M. Deep learning and big datatechnologies in medical image analysis. In2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC) 2018 Dec 20 (pp. 60-63). IEEE.

[8] Singh V, Asari VK, Rajasekaran R. A Deep Neural Network for Early Detection and Prediction of Chronic Kidney Disease. Diagnostics. 2022 Jan 5;12(1):116.

[9] Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. InProceedings of the IEEE conference on computer vision and pattern recognition 2017 (pp. 2097-2106).

[10] Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. Imagenet: A large-scale hierarchical image database. In2009 IEEE conference on computer vision and pattern recognition 2009 Jun 20 (pp. 248-255). IEEE.

[11] Taylor AG, Mielke C, Mongan J. Automated detection of moderate and large pneumothorax on frontal chest X-rays using deep convolutional neural networks: A retrospective study. PLoS medicine. 2018 Nov 20;15(11):e1002697.

[12] Ouyang X, Xue Z, Zhan Y, Zhou XS, Wang Q, Zhou Y, Wang Q, Cheng JZ. Weakly supervised segmentation framework with uncertainty: A study on pneumothorax segmentation in chest x-ray. InInternational Conference on Medical Image Computing and Computer-Assisted Intervention 2019 Oct 13 (pp. 613-621). Springer, Cham.

[13] Yao L, Poblenz E, Dagunts D, Covington B, Bernard D, Lyman K. Learning to diagnose from scratch by exploiting dependencies among labels. arXiv preprint arXiv:1710.10501. 2017 Oct 28.

[14] Jun TJ, Kim D, Kim D. Automated diagnosis of pneumothorax using an ensemble of convolutional neural networks with multi-sized chest radiography images. arXiv preprint arXiv:1804.06821. 2018 Apr 18.

[15] Jakhar K, Kaur A, Gupta D. Pneumothorax segmentation: deep learning image segmentation to predict pneumothorax. arXiv preprint arXiv:1912.07329. 2019 Dec 16.

[16] Park S, Lee SM, Kim N, Choe J, Cho Y, Do KH, Seo JB. Application of deep learning–based computer-aided detection system: detecting pneumothorax on chest radiograph after biopsy. European radiology. 2019 Oct;29(10):5341-8.

[17] Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. Imagenet: A large-scale hierarchical image database. In2009 IEEE conference on computer vision and pattern recognition 2009 Jun 20 (pp. 248-255). Ieee.

[18] Chollet F. Xception: Deep learning with depthwise separable convolutions. InProceedings of the IEEE conference on computer vision and pattern recognition 2017 (pp. 1251-1258).

[19] Woo S, Park J, Lee JY, Kweon IS. Cbam: Convolutional block attention module. InProceedings of the European conference on computer vision (ECCV) 2018 (pp. 3-19).

[20] Erickson BJ, Korfiatis P, Akkus Z, Kline T, Philbrick K. Toolkits and libraries for deep learning. Journal of digital imaging. 2017 Aug;30(4):400-5.