# Deep Learning Approach for Automatic Classification of X-Ray Images using Convolutional Neural Network

Sushavan Mondal,
*B.Tech Student, School of Computer Science & Engineering, Lovely Professional University, India.*
Email: sushavanmessi@gmail.com

Krishna Agarwal,
*B.Tech Student, School of Computer Science & Engineering, Lovely Professional University, India.*
Email: krishnaagarwalyh@gmail.com

Mamoon Rashid,
*Assistant Professor, School of Computer Science & Engineering, Lovely Professional University, India.*
Email: mamoon873@gmail.com

***Abstract*—Medical Image diagnosis has increased drastically in recent years. It may be a relatively simple task to read and diagnose chest X-ray images but the complexity of the images sometimes lead to improper diagnosis. After the success of deep learning in many industries, it also has proven results in the form of better accuracies for medical imaging. In this paper, the authors used Xception model which is a pre-trained model of the state-of-the-art deep learning image classifiers. Our deep learning approach has helped to build highly accurate prediction model for the identification of different diseases from chest X-rays. The classification accuracy will benefit the patients who do not have access to radiologists to read their chest x-rays. The experimentation has been done using the Keras library built under Tensorflow backend.**

***Keywords—Medical Image, Tensorflow, Deep Learning, Xception, Classification, Dropout.***

## I. INTRODUCTION

Chest X-rays is considered to be one of the most cost-effective medical image examinations available. A lot of advanced radiological techniques is being introduced for examining the chest condition reports e.g., Magnetic Resonance Imaging (MRI), Computed Topography (CT), Positron Emission Topography (PET), etc. [1]. Automatic medical image analysis is an important experimentation which is being done by all researchers worldwide to come out with best classification results. It has helped us to face the challenges of clinical diagnosis which seems to be more difficult than the chest CT imaging. Ample of X-rays are performed all over the world each year which are constantly increasing over the past decade. These X-rays images are used to identify plethora of diseases such as Pneumonia, Hernia, Cardiomegaly etc. In this study, the authors aim to use deep learning approach to predict 14 categories of diseases. Our work is a little effort to harness the advances of artificial intelligence and machine learning in the field of medical science. Our experimentation has been performed using the Keras, which is an open source neural network library written in Python. It can run on top of Tensorflow, Theano, Microsoft Cognitive Toolkit, or PlaidML, whereas our work is built on top of tensorflow backend. These open source libraries have a great contribution in producing best results in form of accuracies in all type of deep learning problems.

The outline of the paper is structured as follows: Section II gives the background of X-ray image classification work in Deep Learning. Section III discusses about the dataset used in proposed model. In Section IV, the features of importance are selected and pre-processing is done. Implementation of the proposed model using Convolutional Neural Networks is done is Section V and result analysis is given in Section VI. Conclusion of implemented model is given in section VII.

## II. BACKGROUND

Deep learning has been extensively used in the field of Computer-Aided-Diagnosis (CAD). Abnormality detection in Chest X-rays were performed by the implementation of pre-trained models where GoogLeNet outperforms as the best than other models of InceptionV3 and Residual network architectures [2]. CheXNet is a 121-layer convolutional neural network which is trained on ChestX-ray14 dataset which exceeds the average radiologist performance on the F1 metric to detect all 14 diseases [3]. The previous generation models faced several difficulties to come out with outstanding results for high-dimensional data. Convolutional Neural Network is considered to be achieving the most success in the field of medical image diagnosis which is also known as ConvNets [4]. Breaking down multi-label classification problems into independent binary classification problems, one for each label was proposed in [5]. A pattern based approach was used to classify lung nodules was used for representing essential features. It uses support vector machine which is a linear classifier to come out with excellent classification results [6]. Drug prediction model is proposed with storage of drugs in hadoop and processing of stored data with the help of hybrid machine learning approach [7]. Deep learning approach is used for assessing skeletion bones in X-ray images [8]. This model claims for providing answers for questions related to medical images using deep learning. Deep Convolutional Neural Networks were used classification of X-ray images for screening of baggage security [9]. This research claims an accuracy of 98.92% for detection outperforming other techniques of machine learning.

## III. DATASET

NIH clinical center provides one of the largest publicly available datasets of X-rays for research purpose [10]. The dataset is comprised of 112,120 X-ray images with disease labels from 30,805 unique patients. Our experiment is performed on the random 5% of the full dataset which is publicly available on Kaggle. It contains 5606 images with size 1024*1024 pixels, each of which originally has respective information such as patient age, gender, ID, and number of follow up visits to the hospital. There are 15 classes of common thorax diseases in the dataset which include Atelectasis, Cardiomegaly, Consolidation, Edema, Emphysema, Effusion, Fibrosis, Hernia, Infiltration, Mass, Nodule, Pleural Thickening, Pneumonia and Pneumothorax, No Finding. The image labels are NLP extracted so there could be some erroneous labels but the NLP labelling accuracy is estimated to be >90%. We have performed our experiment considering 75% as the training value of our system and 25% as our testing value.

## IV. DATA PREPROCESSING

In this experiment, The authors have loaded a dataset which is having 5,606 images and is having a number of attributes i.e. Finding labels, Follow-up Ids, Patient Id, Patient Age, Patient Gender, View Position, OriginalImageWidth, OriginalImageHeight, OriginalImagePixelSpacing_x, OriginalPixelSpacing_y, path. There are 15 classes of common thorax diseases in the dataset which include Atelectasis, Cardiomegaly, Consolidation, Edema, Emphysema, Effusion, Fibrosis, Hernia, Infiltration, Mass, Nodule, Pleural Thickening, Pneumonia, Pneumothorax, and No Finding. We have removed the No Finding label because it was superfluous to our task. We have performed a multi-label classification to predict what image category was presented to a subject based on chest X-rays information. Many image inputs were categorized to more than one label so we did split the column consisting the labels and performed a one-hot encoding by which categorical variables can be read into binary values (0, 1) so that we can feed them into the network for better performance. We have made individual list of one-hot encoding for each input image. Python's Scikit-learn library plays an important role in image pre-processing [11]. An example of X-ray images taken as input is shown in Fig. 1

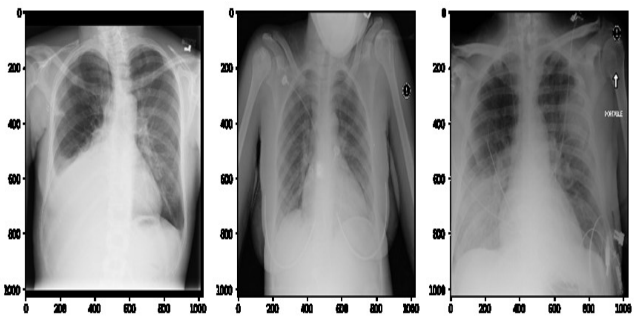

Fig 1. Example of X-ray Images as Input

We have implemented various data augmentation techniques on this large dataset to prevent overfitting and to make use of the whole data available [12] .Each images was flipped in 0, 90, 180, 270 degrees and left or right additionally.

We have normalized the pixel intensities of the input images between 0 and 1 as shown in Fig. 2. It helps us to deter the influence of high frequency and low noise. It is better to normalize the pixel intensities of input data instead of using raw pixel values. It helps to avoid gradient values that could make training difficult [13]. All input images are of grayscale in colour which is represented as 1. The input images are reshaped from 1024x1024 to 128x128 to feed into the network for faster processing.

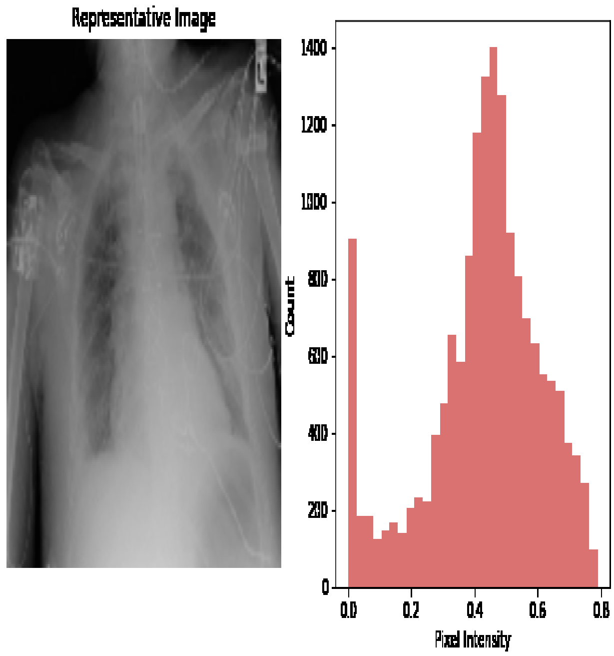


Fig. 2. Normalization of Pixel Intensities between 0 and 1.

## V. IMPLEMENTATION

We have run our CNN model on Tesla K80 GPU. The data was split into 75:25 training and testing data respectively using the Scikit-learn library. To build our architecture, we have used a pre-trained model called Xception which is based on depthwise seperable convolutional layers [14]. It is a deep convolutional neural network which is trained on a popularly large dataset known as ImageNet which enables the network to classify images with good accuracies. The default input image size for this Xception model is 299x299.We didn't used weights as ImageNet for the network. The model architecture of implemented work is shown in Fig. 3.

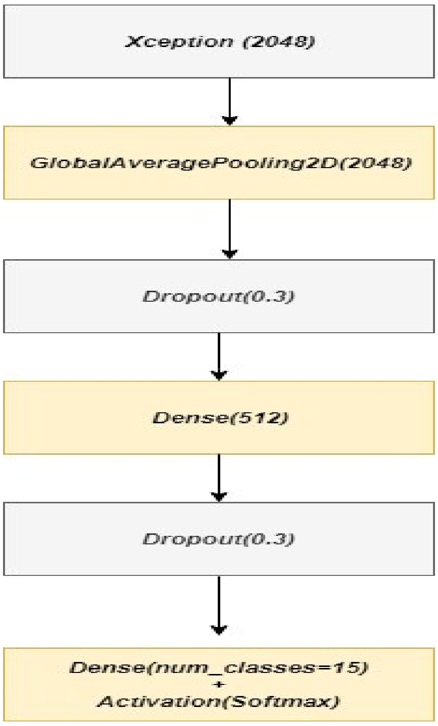


Fig 3. Model Architecture for Proposed Work

To simplify the architecture given in Fig. 3, the input chest X-ray image [128x128x1] is flattened into [2048]. The classification model is implemented using the sequential API model of Keras. The model is built with Xception as its base along-with 2-layer MLP. In Keras, the MLP layer's used in the neural network architecture is known as Dense [15]. Global average pooling is used where we take the average of each map, and the resulting vector is directly fed into the softmax layer which is considered to more beneficial instead of adding fully connected layers on the top of the feature maps [16]. Dropout of 0.3 is added after each layer to prevent overfitting and provides a way of approximately combining exponentially many different neural network architectures efficiently. It basically removes the specified fraction of units to participate in the next layer. The last layer or the output layer is followed by a softmax activation function which provides the probabilities of the categories present in the output layer. In order to build our model we need to estimate the Θ1, Θ2, Θ3… Θ15 € R15 parameters. In a softmax regression the probability, given an image x to be classified as y, is given in equation 1:

$$p\left(y^{(i)} = j \mid x^{(i)};\theta\right) = \frac{\exp\left(\theta_j^T x^{(i)}\right)}{\Sigma_{1\leq i\leq k} \exp\left(\theta_i^T x^{(i)}\right)} \tag{1}$$

If M>2 for multiclass classification, we calculate a separate loss for each class label per observation and sum the result [17]. The loss is calculated by using binary_crossentropy and is given in equation 2:

$$-\sum_{C=1}^{M} y_{o,c} \log(p_{o,c}) \tag{2}$$

where, M=number of classes.
log = the natural log.
y =binary indicator (0 or 1) if class label c is the correct classification for observation o.
p=predicted probability observation o is of class c.

We use optimizers in our neural network architecture to minimize the loss and to maximize the weights. Our experiment consists of Adam optimizer with an initial learning rate of 0.001 which is being used instead of stochastic gradient descent which is used to update weights iteratively on the training data.

Xception is a deep convolutional neural network architecture which is built in reference to inception modules, where inception modules are replaced by depthwise separable convolutions [14]. The architecture of Xception is given in Fig. 4. Xception outperforms the Inception V3 when experimented on the ImageNet dataset which is a large dataset comprising 350 million images and 70,000 classes. Xception V3 helps to make the most efficient use of hyperparameters. Depthwise separable convolutions are implemented with linearity and residual connections. The architecture consists of 36 convolutional layers which forms the feature extraction base of the network. The architecture makes us easy to modify and implement using 30 to 40 lines of code by the help of deep learning libraries such as Keras or Tensorflow. An open source implementation of Xception using Keras or Tensorflow is readily available as part of keras application module, under the MIT license [18].

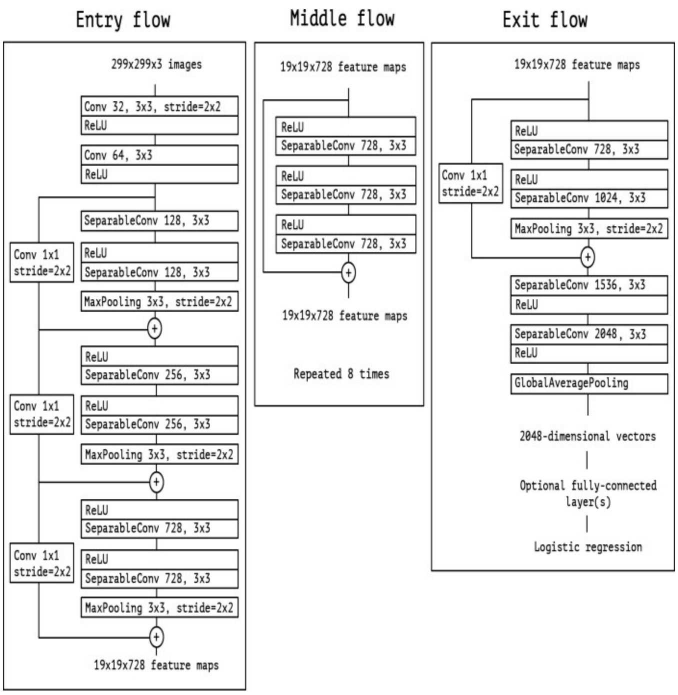


Fig. 4. Deep Convolutional Neural Network Xception Architecture [18]

Simplifying the Xception architecture, the data is passed through the entry flow followed by the middle flow which is iterated eight times and then passes through the exit flow.

## VI. RESULT ANALYSIS

We have used AUC scores, which is in reference to the area under the ROC curve is used, to evaluate our model's performance in performing multi-label classification of predicting the diseases. ROC tells us the probability of the occurrence of diseases whereas AUC helps us to measure the degree of separability. According to AUC, we can say that higher the scores better the model in differentiating between classes. The curve is plotted on True Positive Rate v/s False Positive Rate. These are calculated as given in equations 3, 4 and 5:

True Positive Rate= True Positive / (True Positive+ False Negative) (3)

Specificity= True Negative/ (True Negative+ False Positive) (4)

False Positive Rate= 1- Specificity (5)

Our proposed model has achieved better AUC scores for most diseases and Hernia's score is considered to be the best. Our model have been evaluated on 75% images for training and 25% images for testing with an input shape of 128x128.The achieved AUC scores of the predicted labels are given in TABLE I.

TABLE I. AUC SCORES OF PREDICTED VALUES

| S.No | Diseases | AUC scores |
|---|---|---|
| 1. | Atelectasis | 0.57 |
| 2. | Cardiomegaly | 0.38 |
| 3. | Consolidation | 0.60 |
| 4. | Edema | 0.72 |
| 5. | Effusion | 0.61 |
| 6. | Emphysema | 0.52 |
| 7. | Fibrosis | 0.69 |
| 8. | Hernia | 0.83 |
| 9. | Infiltration | 0.54 |
| 10. | Mass | 0.56 |
| 11. | Nodule | 0.57 |
| 12. | Pleural_Thickening | 0.56 |
| 13. | Pneumonia | 0.58 |
| 14. | Pneumothorax | 0.54 |

The AUC-ROC curve of these predicted labels is shown in Fig. 5.

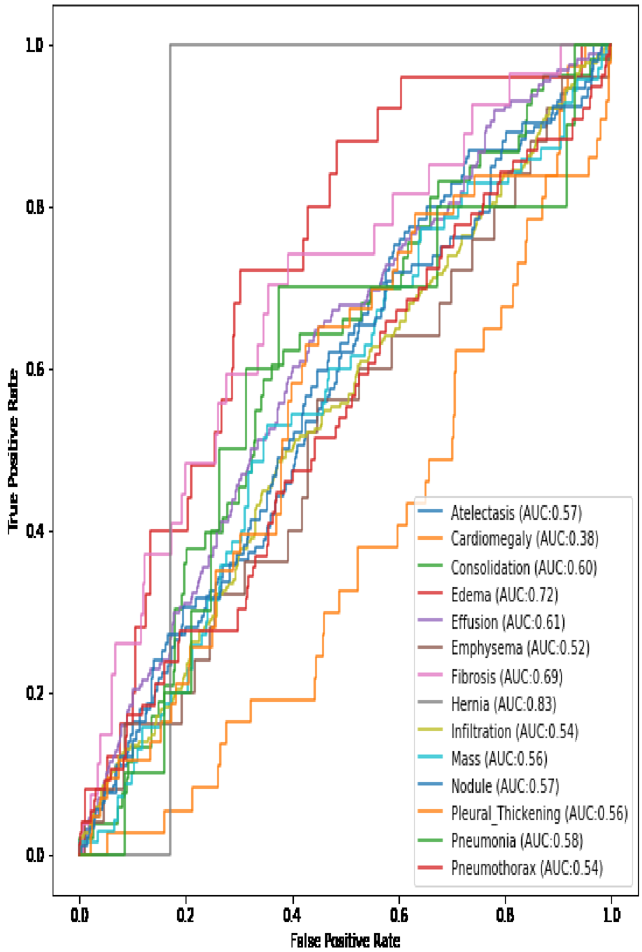


Fig. 5. AUC-ROC curve of the Predicted Labels

Each random image is having more than one classification categories. Our job in this experiment is to predict the probabilities of each chest X-ray images classification scores in terms of percentage for their respective categories. The random samples showing predicted labels along with probabilities are shown in Fig. 6. We have trained our model for 28 epochs and from the graphs in Fig. 7, we can clearly see that the test accuracy has dramatically changed and lastly attained an accuracy of 88.76. The testing loss also got lessened to provide a better accuracy in terms of results as shown in Fig. 8.

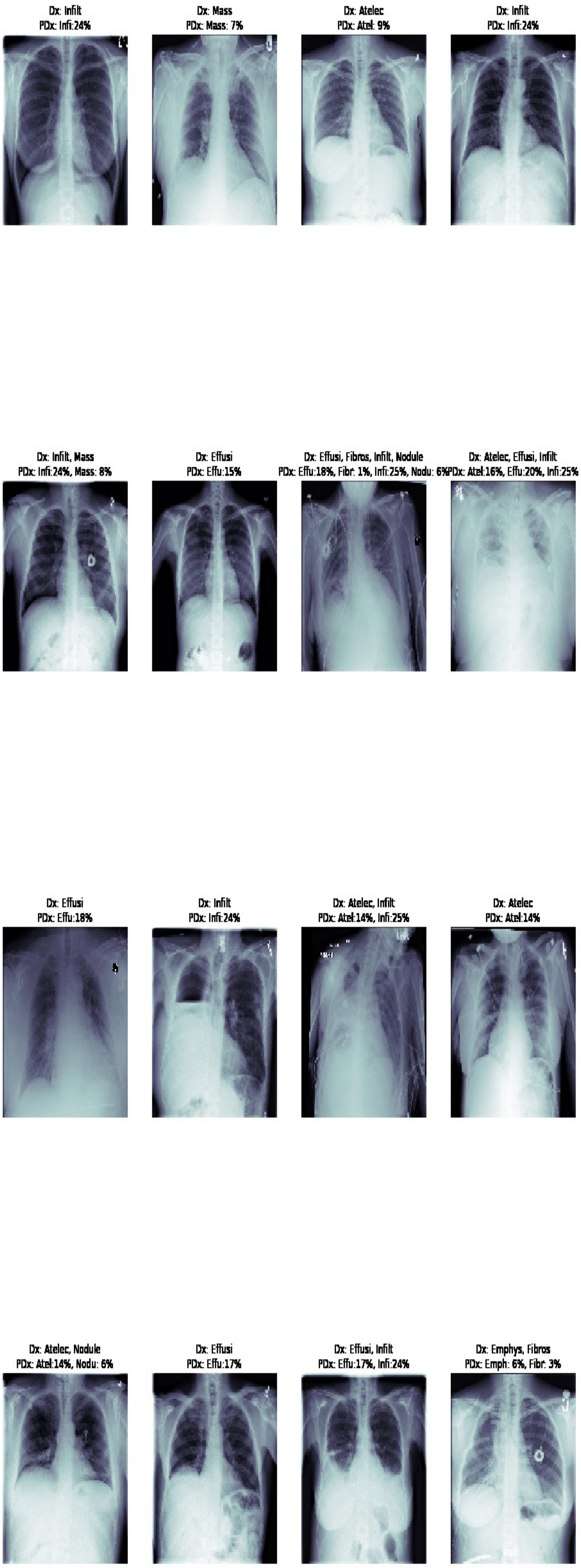


Fig. 6. Random Samples showing Predicted Labels along with the Probabilities.

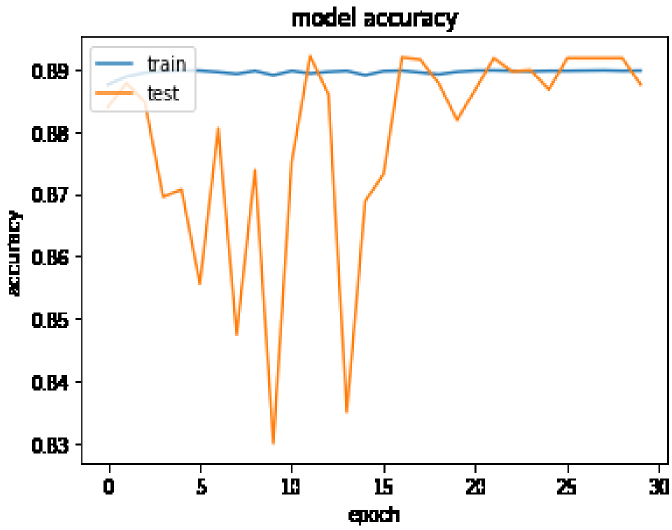


**Fig 7. Model Accuracy with respect to Epochs.**

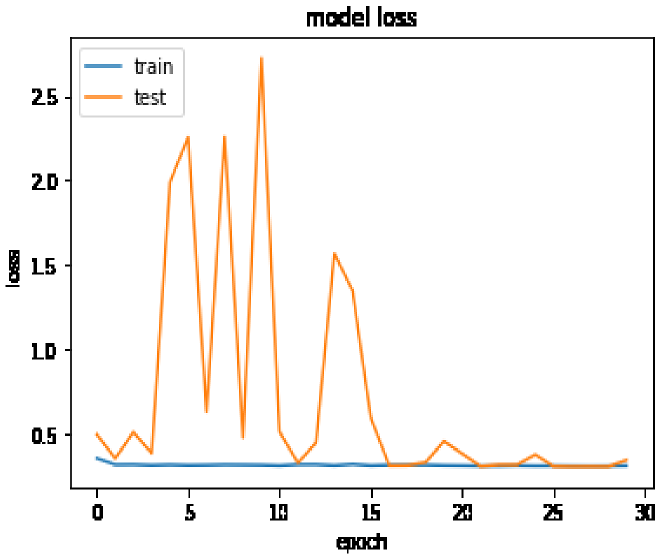


**Fig 8: Model Loss with respect to Epochs.**

The accuracy plot depicts that the model could be trained a little more in accordance to the trend of accuracy. In comparison to the both datasets, we can see that the accuracy is having a hike at the last few epochs as shown in Fig. 7. Whereas, from the plot of loss we can see a comparable performance on both the train and validation datasets which is labelled as test. If the loss plot cease consistently, it might be a good sign to stop training at an earlier epoch for attaining a higher accuracy of our model as shown in Fig. 8.

## VII. Conclusion and Future Work

In this paper, the authors have developed novel deep learning model for the prediction of chest diseases on the basis of X-ray images using Xception deep convolutional neural network architecture. Our model achieved an accuracy of 88.76% and model loss also attained level which is quite acceptable in terms of performance. Moreover our model achieved better AUC Scores for most of the diseases with 0.83 for Hernia as highest AUC Score. In future, the authors are planning to use this deep learning architecture for other medical techniques of Functional Magnetic Resonance Imaging for making classification of brain images.

## REFERENCES


[1] D. S. Gupta, "Assessment of spectrum of Chest Radiographic Manifestations of Legionnaires' Disease: A retrospective study," *Journal of Medical Science And clinical Research*, vol. 6, no. 5, 2018.

[2] X. Zhang, S. Huang, X. Zhang, W. Wang, Q. Wang, and D. Yang, "Residual Inception: A New Module Combining Modified Residual with Inception to Improve Network Performance," *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018.

[3] S. Rajaraman, G. Thoma, S. Antani, and S. Candemir, "Visualizing and explaining deep learning predictions for pneumonia detection in pediatric chest radiographs," *Medical Imaging 2019: Computer-Aided Diagnosis*, 2019.

[4] C. D. Medio, F. Gasparetti, C. Limongelli, F. Sciarrone, and M. Temperini, "A Machine Learning Approach to Identify Dependencies Among Learning Objects," *Proceedings of the 8th International Conference on Computer Supported Education*, 2016.

[5] Mao, K. Mao and Z. Deng, "Lung Nodule Image Classification Based on Local Difference Pattern and Combined Classifier," Computational and Mathematical Methods in Medicine, vol. 2016, pp. 1–7, 2016.

[6] Rashid, Mamoon, Aamir Hamid, and Shabir A. Parah. "Analysis of Streaming Data Using Big Data and Hybrid Machine Learning Approach." Handbook of Multimedia Information Security: Techniques and Applications. Springer, Cham, 2019. pp. 629-643.

[7] M. Rashid, A. Hamid, and S. A. Parah, "Analysis of Streaming Data Using Big Data and Hybrid Machine Learning Approach," Handbook of Multimedia Information Security: Techniques and Applications, pp. 629–643, 2019.

[8] Sp C. Spampinato, S. Palazzo, D. Giordano, M. Aldinucci, and R. Leonardi, "Deep learning for automated skeletal bone age assessment in X-ray images," Medical Image Analysis, vol. 36, pp. 41–51, 2017.

[9] S. Akcay, M. E. Kundegorski, M. Devereux, and T. P. Breckon, "Transfer learning using convolutional neural networks for object classification within X-ray baggage security imagery," 2016 IEEE International Conference on Image Processing (ICIP), 2016.

[10] 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," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[11] Á. L. García, "DEEPaaS API: a REST API for Machine Learning and Deep Learning models," Journal of Open Source Software, vol. 4, no. 42, p. 1517, 2019.

[12] M. Moradi, K. C. L. Wong, T. Syeda-Mahmood, and J. T. Wu, "Identifying disease-free chest x-ray images with deep transfer learning," Medical Imaging 2019: Computer-Aided Diagnosis, 2019.

[13] J. Moolayil, "An Introduction to Deep Learning and Keras," Learn Keras for Deep Neural Networks, pp. 1–16, Jul. 2018.

[14] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[15] S. Alla and S. K. Adari, "Introduction to Deep Learning," *Beginning Anomaly Detection Using Python-Based Deep Learning*, pp. 73–122, 2019.

[16] Lin, Min, Qiang Chen, and Shuicheng Yan. "Network in network, 2013.

[17] J. Moolayil, "An Introduction to Deep Learning and Keras," *Learn Keras for Deep Neural Networks*, pp. 1–16, Jul. 2018.

[18] N. K. Manaswi, "Understanding and Working with Keras," Deep Learning with Applications Using Python, pp. 31–43, 2018.