ISSN:0975-3583,0976-2833 VOL12,ISSUE05,2021

# Pneumonia Detection on Chest X-Ray Using Convolutional Neural Network and Transfer Learning

**Bhakgya M**
Student Department of Computer Applications PSG College of Technology Coimbatore, India
19mx152@psgtech.ac.in
**Manoranjitham A**
Assistant Professor (SG) Department of Computer Applications PSG College of Technology Coimbatore, India
**amr.mca@psgtech.ac.in**

**ABSTRACT**
Diagnosis of thorax diseases is most commonly done by examining the Chest X-rays (CXR). Computer-aided diagnosis (CAD) System assist radiologists in the interpretation of medical images. CADimproves the quality of diagnosis and leverages the productivity of radiologists. Several Deep Learning algorithms have been successfully implemented inorder to provide fully-automated, high precision Computer-aided diagnosis (CAD) Systems. In this paper, a dedicated X-ray network trained from scratch for Chest X-Ray image classification is explored. Further, transfer learning using powerful network architecture like MobileNet, InceptionV3, VGG19 and ResNet-50 are investigated indetail. Limited availability of annotated Chest X-ray images makes medical image classification challenging. Transfer learning approach with andwithout fine-tuning can help overcome this issue by transferring the knowledge gained from pre-trained networks to domain specific tasks.The experimental study is tested onthe dataset of Chest X-Ray images, consisting of 2 categories, Pneumonia and Normal.

**Key words:** Chest X-Ray, Computer-aided diagnosis, Deep Learning, Transfer Learning.

## 1. INTRODUCTION

Computer-aided diagnostics (CAD) provides accurate and efficient detection and classification of abnormalities in medical images. They are used as a “second opinion” complementary to that of a radiologist. CAD is comprised of two processes: Computer-aided Detection (CADe) and Computer-aided Diagnosis (CADx) [1].Chest X-Ray images (CXRAY), Computed Tomography (CT), Magnetic Resonance Imaging (MRI)and Ultrasound are some of the common medical images used in the diagnosis of pulmonary diseases. Among these, Chest X-ray is the most common type of radiology examination that plays an important role in the diagnosis of various thorax diseases like pneumonia, infiltration, pneumothorax etc.

Traditionally, several Machine Learning (ML) techniques like Support Vector Machines (SVM), K-nearest neighbour method (KNN), random forest etc have been used to provide CAD. There are two main disadvantages in using traditional ML models. These models require input features to be selected manually. And also, they have limited capacity of processing large scale image data. Therefore, researchers prefer Deep Learning (DL) to ML techniques as it facilitates automatic feature extraction with the help of Neural Networks.

With transfer learning, the knowledge gained from one problem domain can be transferred and applied to another task in a new domain resulting in more efficient results. In this paper, a dedicated deep neural network tailored to Chest X-RayImage classification is trained from scratch. Further, the performance of the various baseline CNN architectures like the VGG-16, ResNet, InceptionV3 and MobileNet with and without fine-tuning in the CXRAY classification problem is analysed and compared with the proposed architecture.The main aim is to minimize the computational and memory requirements without compromising the classification performance of the neural networks.

The organization of the paper is as follows. Section 2, describes some of the works that are previously done for the classification of X-Ray images are reviewed. Section 3 gives a detailed view about Transfer Learning and the different ways in which it can be achieved. In Section 4, the

ISSN:0975-3583,0976-2833 VOL12,ISSUE05,2021

architecture of the proposed system is presented. Section 5 discusses the experimental results obtained in terms of accuracy and other metrics. Section 6 concludes the work.

## 2. RELATED WORK

Various research works have been proposed so far in advancing the capability of Computer Vision (CV) tasksto automatically detect and classify disease from medical images. In the work of Zhang et al. classification of Magnetic Resonance Images is performed based on weighted Functional Fourier Transform and non-parallel support vector machines [2]. Vajda et al. used a complex machine learning pipeline that starts with an atlas-based lung segmentation algorithm, then extracts manually selected features such as shape and curvature descriptor histograms or the eigenvalues of the hessian matrix, and finally uses a classifier to diagnose the disease[3].

Several Machine learning approaches like Support Vector Machine (SVM), KNN, random forest etc., have been previously implemented in automatic classification of digitised chest images. For example, in [5]discrimination of malignant and benign nodules is done by calculating three statistical features from lung texture using a Support Vector Machine *SVM* classifier. Agrawal et al. used artificial bee colony algorithm combining with k-nearest neighbour algorithm and support vector machine to classify 271 computed tomography (CT) images of cervical cancer [4].

Delen et al. proposed the application of data mining techniques: artificial neural networks, decision trees and logistic regression in developing models to predict breast cancer survivability. Islam et al. explored ensemble models for abnormality detection in frontal CXR. They found that combining DCNN models with rule-based models reduces the accuracy and they concluded that while using DCNN models alone, accuracy can besignificantly improved by using ensemble models when compared to a single model.

With the advancement of technology, the large-scale annotated images are easily available for any domain these days. Traditional machine learning algorithms are being replaced by deep learning approaches because of their ability to produce more accurate results with less effort. The ImageNet Large Scale Visual Recognition Challenge (ILSVRC) is one of the most important challenges in history. It led to the development of baseline models like VGG-16, Residual Networks (ResNet), Densely connected networks (DenseNet) etc. These baseline models can be used to implement new systems for CAD.

## 3. TRANSFER LEARNING

Deep learning has proved to be the most promising ML technique in medical image analysis tasks such as image classification and segmentation. Neural networks are the baseline of deep learning methods. A neural network consists of multiple layers comprised of neurons. The layers include input layer, output layer and the hidden layers: the layers in between input and output layers. Convolutional Neural Network (CNN) is the widely used deep learning architecture because of its ability to automatically learn and preserve feature relationships and its highly parallelizable nature making it suitable for dealing with large training datasets.

Application of CNN for Chest X-Ray diagnosis popularized with the release of Chest-Xray14 dataset (CX14) consisting of 108,948 posterior-anterior (PA) CXRAYs by Wang et al. in 2017. Since then, an enormous number of CNN approaches for CXRAY classification and segmentation have been proposed. In 2019, Irvin et al. a new large CXRAY dataset named CheXpert dataset comprising of 224,316 CXRAYs and also provided a baseline result for classifying 14 different classes.

A CNN can be trained in two primary ways. It can be an end-to-end network that requires huge annotated images dataset to train the network or transfer learning can be used to overcome the need for large scale annotated images. With transfer learning, knowledge from pre-trained CNNs can be transferred to themedical image classification task providing an effective and robust solution with the limited number of annotated medical images. Further, transfer learning can be achieved through 3 techniques: shallow-tuning, fine-tuning and deep-tuning. In Shallow tuning only the last layer is modified to carry on the customized task while freezing the parameters of the previous layers.In Fine tuning approach, the layers are gradually trained by fine tuning the learning parameters until a significant is achieved. Deep tuning is retraining all the layers of the pre-trained network.

ISSN:0975-3583,0976-2833 VOL12,ISSUE05,2021

The annual ImageNet Large Scale Visual Recognition Challenge (ILSVRC) has led to the development of a significant number of high-performing models for image classification.VGG19, ResNet50, MobileNet and InceptionV3 are some of the ImageNet pre-trained CNN networks. These models have been trained on more than 1.4M images belonging to 1000 categories and learnt to detect generic features. Image classification can be done by embedding a pre-trained model into an application. On the other hand, they can also be used as feature extractors and a new totally different classifier can be implemented. The entire pre-trained model or a portion of it can be integrated with the customized neural network. While using this approach, the weights of the pre-trained model can either be frozen or the pre-trained model can be used as a weight initializer while the training the new Convolutional Neural Network with the application specific dataset.

## 4. IMPLEMENTATION

This section explains in detail the proposed methods for classifying pneumonia and normal cases from Chest X-Ray images. First the overview of the Neural Network architecture is given, following whichthe description of the different components of the method is discussed and finally the section describes the workflow and formalises the method.

In this approach, a CNN architecture is built from scratch. The network takes raw X-ray images as inputand classifies them as Pneumonia case or a normal case. In order to so, the CNN assigns class scores or probabilities to the images that determines to which class the image data belongs. The image data goes into the Input layer as input to the network and the output layer gives the classification result. All the computation works are done in the hidden layers comprised of the convolution layers, pooling layers and the rectified linear unit (ReLU) layer.

Convolutional layers extract features from previous layers and pass them as input to the next layer. Number of filters for detecting patterns at each convolutional layer is specified. Pooling layers reduce computational complexity, down-samples an input representation decreasing its dimensionality by retaining the activated features and the dense layer at the end outputs the classification result.

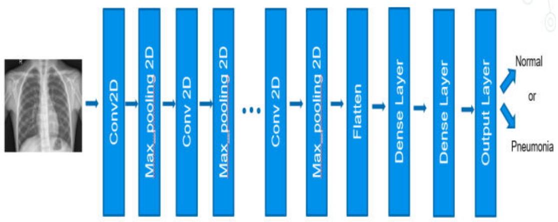



Fig 4.1 Layers of the ConvNet Model

Figure 4.1 shows the stacked layers of the Neural network architecture built. The model is constructed with a layer of Conv2D followed by a MaxPooling layer. The flattening layer is added to the end these set of Conv2D and MaxPooling layers. It creates a 1D vector of pixels which is then given to a sigmoid function that gives the classification output in a range of 0 to 1. All values in the range 0.5 to 1 are categorised as 0 (normal) and those that are less than 0.5 as 1 (pneumonia).

There are a dozen top-performing models trained on the ImageNet dataset (a large dataset of 1.4M images and 1000 classes)for image classification. Four of the popular models namely,

1. VGG19,
2. GoogLeNet (InceptionV2),
3. MobileNet and
4. Residual Network (ResNet50)

are implemented in this work.

The VGG model,a successor of AlexNetwas built by a group called Visual Graphics Group (VGG) at Oxford and hence the name VGG. VGG19 is a variation of VGG consisting of 19 layers. It was published in the paper titled “Very Deep Convolutional Networks for Large-Scale Image Recognition” in 2014. The model expects the images to be rescaled to the size of 224 x 224 by default.

The InceptionV3 model, one of the CNN from Inception family was first presented in a paper titled “Rethinking the Inception Architecture for Computer Vision” by researchers at Google. It was

ISSN:0975-3583,0976-2833 VOL12,ISSUE05,2021

developed for the GoogLeNet Model. It is the third iteration of the inception architecture. The model requires the images to be of shape 299 x 299.

The Residual Network or ResNetinshort was developed by Microsoft and was described in “Deep Residual Learning for Image Recognition” in 2015. The model expects images have a shape of 224 x 224.

MobileNet was also developed at Google and described in “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications”, Howard et al, 2017.It is a lightweight architecture that uses depth wise separable convolutions.

The features of the pre-trained CNN models are used and Shallow tuning is performed byonly training the final classification layer with the Chest X-ray images.This very last layer before the flatten operation is used for feature extraction.

## 5. EXPERIMENTAL RESULTS

This section discusses the dataset used for the training the models and the results of the experiment.

The dataset used is Chest X-Ray images consisting of 2 categories, Pneumonia and Normal [6]. The dataset is organized into 3 folders: train, test and val. Each folder is sub divided into opacity (viz pneumonia) and normal. The total number of observations(images) is 5,856 with 4,192 training, 1,014 validation and 624 test observations. Figure 5.1 shows the count plot of the number of pneumonia and normal images in train, test and validation folders.

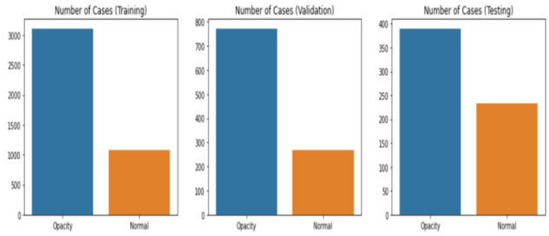
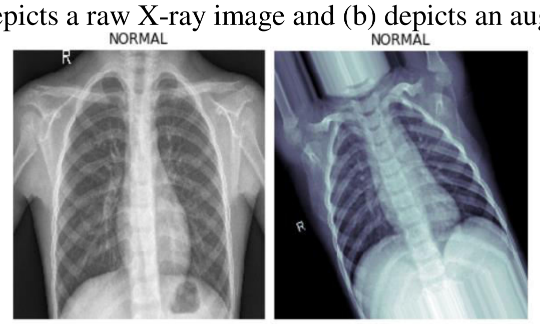



Fig 5.1 Count plot of dataset

Data augmentation technique is used to artificially increase the size and variation of the dataset where a modified version of the existing dataset is created to improve the model ability to predict new images. Different data augmentation techniques like rescaling, flipping, zooming and rotating are applied on the training data and the validation data. The test data is not manipulated beforehand.Figure 5.2 (a) depicts a raw X-ray image and (b) depicts an augmented X-ray image.




(a) (b)

Fig 5.2 (a) X-ray Image without augmentation (b) Augmented X-ray image

Overfitting is one of the major issues faced while training a Deep Neural Network. This is overcome with the help of Early Stopping technique. It stops the training epochs based on some conditions given. Here in this experiment, ‘val_loss’ is taken as a metric for early stopping. Also, learning rate is reduced when the model has stopped learning or improving. Class weights are assigned so that the weight of the minority class is emphasised to aid the model in learning from all the classes equally.

The ConvNet model developed from scratch is compared with the state-of-the-art baseline models VGG, ResNet, InceptionNet and MobileNet in terms of precision, recall and F1-score. The Performance measures of classification results for all models are shown in Table 5.1

| Model | Precision | Recall | F1Score |
|---|---|---|---|

ISSN:0975-3583,0976-2833 VOL12,ISSUE05,2021

| | | | |
|---|---|---|---|
| CNN (From scratch) | 0.926 | 0.964 | 0.923 |
| VGG19 | 0.894 | 0.982 | 0.907 |
| ResNet50 | 0.886 | 0.897 | 0.854 |
| InceptionNet | 0.886 | 0.943 | 0.879 |
| MobileNet | 0.868 | 0.964 | 0.873 |

Table 5.1 Performance measures of different models on the test set

A confusion matrix describes the performance of a classification model on the test data for which the actual labels are known.In case of binary classification, there are four possible outcomes.
True Positive (TP) - Predicted as Normal and actual label is also Normal
True Negative (TN) -Predicted as Pneumonia and the actual label is also Pneumonia
False Positive (FP) -Predicted as Normal but the actual case was Pneumonia
False Negative (FN) -Predicted as Pneumonia but the actual case was Normal
Figure 5.3 (a-e) depicts the confusion matrix of the different Neural Network Models implemented on the Chest X-Ray classification dataset.

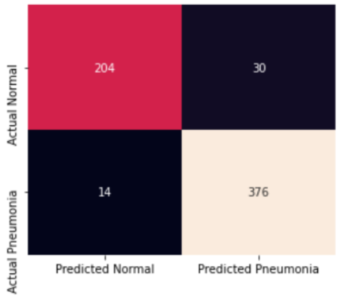



Fig 5.3 (a) Confusion Matrix- CNN

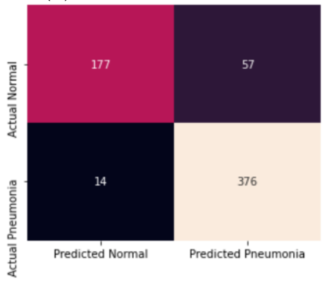



Fig 5.3 (b) Confusion Matrix - MobileNet

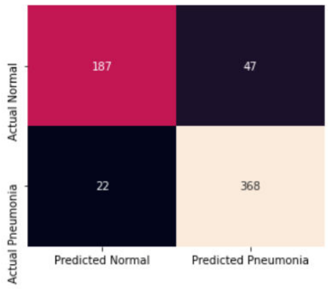



Fig 5.3 (c) Confusion Matrix – InceptionV3

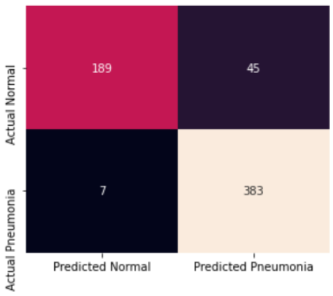


Fig 5.3 (d) Confusion Matrix – VGG19

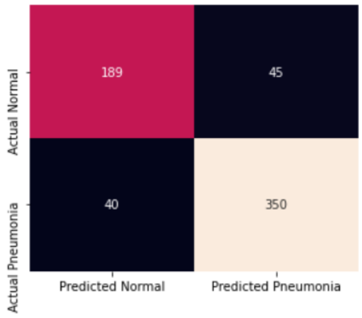



Fig 5.3 (e) Confusion Matrix – ResNet50

| Model | Test Accuracy (%) |
|---|---|
| CNN (From scratch) | 92.9 |
| VGG19 | 91.6 |
| ResNet50 | 86.3 |
| InceptionNet | 88.9 |
| MobileNet | 88.6 |

Table 5.2 Test Accuracy of classification models

Table 5.2 shows the test accuracy of the classification models. The CNN built and trained from the scratch with Chest X-Ray images for pneumonia detection task worked better when compared to the baseline models. This is because these models are trained and optimized on millions of images and this makes them prone to overfitting and they are less likely to generalize when applied to medical tasks with limited amount of data.

## 6. CONCLUSION AND FUTURE WORK

In this paper, a CNN trained particularly with Chest X-Ray dataset and a few transfer learning models with shallow tuning for automatic classification of chest x-rays as pneumonia cases or not is presented. Further, the performance of the models can be enhanced by deep tuning the pre-trained models.

## ACKNOWLEDGMENT

I would like to thank PSG College ofTechnology, Coimbatore, for permitting me in submitting this paper. I take this opportunity to thank Dr. A. Chitra, Professor and Head of the Department of Computer Applications forallowing me to do this work. Also, I would liketo express my gratitude to, my guide Mrs. Manoranjitham A, Assistant Professor (SG), Department of Computer Applications for her support to complete the worksuccessfully.

## REFERENCES

1. Doi, K. Computer-Aided Diagnosis in Medical Imaging: Historical review, current status and future potential. Comput. Med. Imaging Graph. 2007, 31, 198–211.
2. Zhang, Y.D.; Chen, S.; Wang, S.H.; Yang, J.F.; Phillips, P. Magnetic resonance brain image classification based on weighted-type fractional Fourier transform and nonparallel support vector machine. Int. J. Imaging Syst. Technol. 2015, 25, 317–327.
3. Vajda, S. et al. Feature Selection for Automatic Tuberculosis Screening in Frontal Chest Radiographs. J. Med. Syst. 42 (2018).
4. Agrawal, V.; Chandra, S. Feature selection using Artificial Bee Colony algorithm for medical imageclassification. In Proceedings of the 2015 Eighth International Conference on Contemporary Computing (IC3), Noida, India, 20–22 August 2015; pp. 171–176.
5. JAIN, M.; K, V. Paris Attack in Wireless Ad Hoc Network. IARS' International Research Journal, Vic. Australia, v. 1, n. 1, 2011. DOI: 10.51611/iars.irj.v1i1.2011.2.

6. Kuruvilla J, Gunavathi K (2014) Lung cancer classification using neural networks for ct images. Comput Meth Prog Biomed 113(1):202–209
7. Kermany, Daniel; Zhang, Kang; Goldbaum, Michael (2018), "Labelled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification", Mendeley Data, v2, doi: 10.17632/rscbjbr9sj.2
8. 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. IEEE CVPR 2017.
9. KAKKAR, A.; TAYAL, P.; PUNHANI, R. Future of healthcare vis-a-vis building trust in major stakeholders through Information Security Management. IARS' International Research Journal, Vic. Australia, v. 3, n. 2, 2013. DOI: 10.51611/iars.irj.v3i2.2013.32.
10. ALYAS, T.; TABASSUM, N.; NASEEM, S.; AHMED, F.; EIN, Q. T. Learning-Based Routing in Cognitive Networks. IARS' International Research Journal, Vic. Australia, v. 4, n. 2, 2014. DOI: 10.51611/iars.irj.v4i2.2014.40