International Journal of Engineering Research and Applications *www.ijera.com*
ISSN: 2248-9622, pp. 01-04

RESEARCH ARTICLE OPEN ACCESS

Ubiquitous Platform for classifying Chest diseases using Chest X-Ray

Karan Idnani, Mohit Khiani, Sourav Mantri, Sahil Rajpal, Mrs. Lifna CS
Karan Idnani, Student, VES Institute of Technology Chembur Mumbai 400074
Mohit Khiani, Student, VES Institute of Technology Chembur Mumbai 400074
Sourav Mantri, Student, VES Institute of Technology Chembur Mumbai 400074
Sahil Rajpal, Student, VES Institute of Technology Chembur Mumbai 400074
Lifna C S, Assistant Professor, VES Institute of Technology Chembur Mumbai 400074

ABSTRACT: *Around 10 million people are affected by Pneumonia and 2.7 million are affected by Tuberculosis in India every year. Also, COVID-19 affects people on a large scale and has made it difficult for doctors to distinguish between these diseases as they mainly affect the lungs. The model proposed in this paper helps to classify a given chest X-Ray as to whether it is a case of Pneumonia or Tuberculosis. Pneumonia and Tuberculosis are infectious diseases and affect a large number of individuals throughout the year. The treatment is available for both diseases, however, the treatment period may vary. The model uses frontal chest X-Ray images to characterize the sickness with the help of Transfer Learning and Convolution Neural Networks (CNN). The model has twelve layers of Convolution Networks with a 'ReLU' activation function to extend non-linearity in the system. The paper compares three different models and uses the one with the most effective accuracy in developing the web application which can be used by Pulmonologists.*

KEYWORDS: Pneumonia, Tuberculosis, Deep Learning, Chest X-Ray, Transfer Learning, CNN

I. INTRODUCTION

In the current scenario of COVID, there's a lot of stress on health workers (especially, doctors and nurses) as it isn't feasible to admit and prioritize patients suffering from diseases other than coronavirus. The need to develop a platform for doctors to ease their work and reduce the time required to assess piles of reports is of utmost importance. Moreover, the platform can be used remotely and necessary decisions can be taken to evaluate the condition of the patient (whether serious or trivial).

Pneumonia and Tuberculosis are curable with the appropriate treatments and diagnosis requires a study of the chest X-Ray of the person. Detection of Pneumonia and other lung diseases can be done by CT Scan and Chest X-Ray which can be costly. To reduce the cost, the paper proposes a prototype that can detect Pneumonia and Tuberculosis, using a scanned image of Chest X-Ray. The other cause behind developing such a model is to reduce the time required by the pathological labs to conclude the results, especially in the times of pandemics such as COVID.

There exist different models that categorize specific diseases by using chest X-Rays. Some of the examples include detecting pneumonia using deep learning[4] and detecting tuberculosis using deep learning, segmentation, and visualization[7]. These models are built specifically for these individual diseases such as tuberculosis, pneumonia. However, there is no such system that is integrated for two or more than these types of diseases and even if there is an attempt for such a system, it lacks certain requirements. The drawbacks in the existing systems are that there is no such system that classifies two or more chest diseases and the systems that classify these diseases have low accuracy. The additional time that the existing system takes can be harmful in the time of COVID whose symptoms are similar to the mentioned chest diseases.

Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm that can take in an info picture, assign significance (learnable loads) to different viewpoints/objects in the picture, and have the option to separate one from the other. The pre-processing needed in a ConvNet is a lot lower when contrasted with other order calculations. While in crude strategies channels are hand-designed, with enough preparation and training, ConvNets can gain proficiency with these filters.[9] Transfer learning is a popular approach where previously trained models are used as a starting point for computer vision and natural language processing activities given large-scale computational resources and the time required to develop neural network models in these problems.[8] After comparing the three models, which are based on the above-mentioned techniques, the model with the best accuracy was used to develop a web application for detecting

whether the input chest X-Ray has pneumonia, tuberculosis, or is a normal chest X-Ray.

II. LITERATURE SURVEY

The papers related to this topic mention different techniques to detect and classify abnormalities in the lungs. In one research paper[3], to detect tuberculosis - a three-level hypothesis is being used where in those levels geometrical figures like shape, size are extracted from the lungs, and then the algorithm looks for any deformation present. To improve the detection performance, a union of shape-based geometrical figures and statistical texture features is used.

In paper[7], the author tries to reduce the wait time of the patients who are being diagnosed with X-Ray diseases with the help of machine learning models. The technologies used here are Deep Learning, Deep CNN, and Region-based CNN. The paper uses datasets that are manually collected and by ImageCLEF collection. The authors lay out an X-Ray database that is devoted to Tuberculosis detection and they propose a helpful and structured computational model for Tuberculosis categorization.

Paper[2] centers around choosing the treatment of tuberculosis. The information utilized in this paper has been driven from assembled records by wellbeing specialists, medical attendants, and doctors that control focuses all through Iran. The innovations utilized in this paper are problem modeling, decision tree, Bayesian networks, logistic regression, artificial neural networks, radial basis functions, and support vector machines. The paper suggests utilizing a Decision Tree with the flowchart-type structure that is the more probable strategy to be perceived by broad users with a low degree of particular information about Tuberculosis..

III. METHODOLOGY AND PROPOSED SYSTEM

The user would upload their chest X-Ray and also the ML model would classify diseases like Pneumonia, Tuberculosis. The web app would be made using Django Web-Framework. For Tuberculosis, a technique is put forward to detect images that are abnormal i.e. images that have pathological features such as abscess formation in the alveoli sac of the lungs, lung inflammation. The technique is prognosticated on feature extraction using the VGG 16 model.

In this paper, Deep Learning is used which's a subset of machine learning in AI. It's based upon artificial neural networks and representation learning because it is capable of implementing a function that's accustomed to mimic the functionality of the brain by creating patterns and processing data. A Convolutional Neural Network is a Deep Learning algorithm that can take in an info picture, assign significance (learnable loads) to different viewpoints/objects in the picture, and have the option to separate one from the other. The pre-processing needed in a ConvNet is a lot lower when contrasted with other order calculations. It reduces overfitting which is the main problem of Multilayer Perceptrons (MLP's), but that does not seem to be true in the case of small datasets with large numbers of features. CNN can handle variance in images like translation, rotation, illumination, size, etc. Moreover, it is used for automatic feature extraction which is an added advantage. For the primary model, this paper uses four layers of convolution networks with the activation function of 'relu' to extend non-linearity in our system. The algorithm uses a Max pooling of two to scale back the number of parameters with a stride of two. This paper uses data augmentation to extend the range of our training set by applying random (but realistic) transformations like image flipping. With the assistance of this model, an accuracy of 82% was achieved for detecting both diseases. One important point to note is that CNN does not work well with small datasets which tends to have an adverse impact on the accuracy of the model because it fails to learn veracious data distribution.

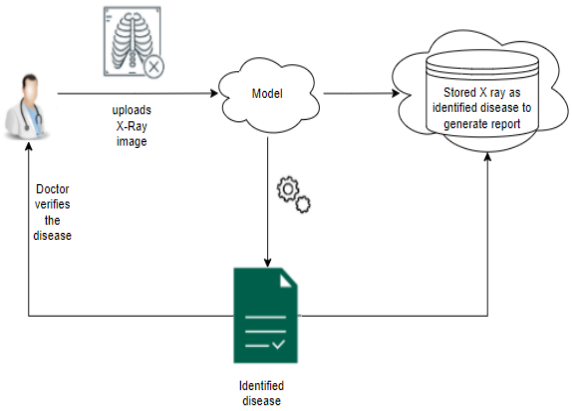

Figure 1 - Modular diagram

The dataset should contain valid facts, information, and pictures about chest X-Rays for identification.

For the second model, the paper optimizes our algorithm of CNN using Keras Tuning which increased the accuracy to 87% for detecting both of the diseases.

For the third model, the paper uses transfer learning, which is a popular approach where bottleneck features of previously trained models are used as a starting point. By using transfer learning, the accuracy of 97.59 & 99.4 has been achieved for Pneumonia and Tuberculosis.

Thus it had been evident that the model

with transfer learning provided the simplest results and so it had been included within the web app developed which processed the image and provided us with the identified lung disease just in case of any lung disease and normal in the case where the input was a standard chest X-Ray. Identified disease together with the input image is stored in a very database for generation of report and record keeping.

IV. EXPERIMENTAL SETUP AND RESULTS

Data Acquiring:

The dataset consists of 5,800 images and two categories namely, Pneumonia & Normal. The anterior-posterior image is selected of pediatric patients from Guangzhou. The chest X-Ray imaging of patients was performed as a component of routine clinical care.[11]

A database of 3,500 TB chest X-Rays and general chest X-Rays were compiled by a team of researchers from Qatar University and the University of Dhaka, Bangladesh, and colleagues from Malaysia and medical doctors from Hamad Medical Corporation.[12]

These two datasets were then used to test and train the three models, the first model uses four layers of convolution networks with the activation function of 'relu' to extend non-linearity in the system. The second model is an improvement on the first model by using Keras Tuning to increase the accuracy. The third model uses Transfer Learning where a pre-developed model for a task is reused.

Before using the model, the chest X-Ray images are scanned by a radiologist for vetting by removing unreadable scans. The diagnoses for the pictures were then approved by an expert physician before being cleared for training the model.

MODEL	ACCURACY
CNN	82%
CNN + Keras Tuning	87%
Transfer Learning	97.59 & 99.4

Figure 2 - Model Results

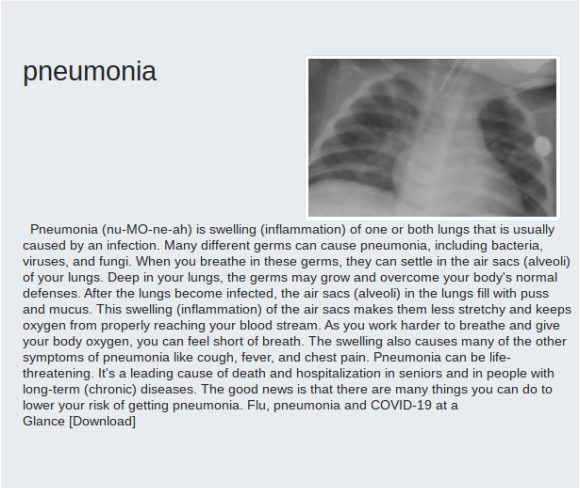

Figure 3 - Prediction of Pneumonia

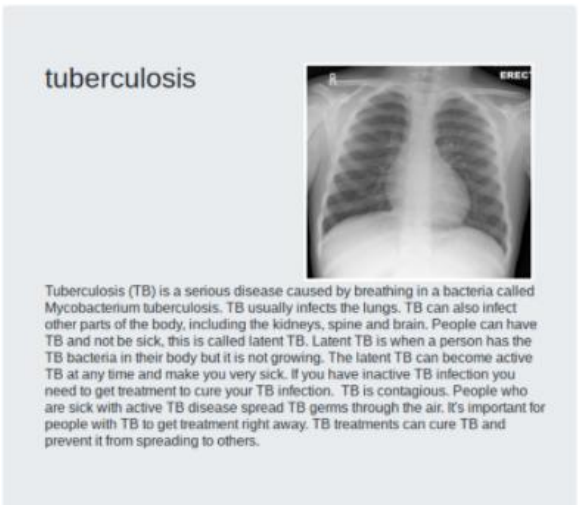

Figure 4 - Prediction of Tuberculosis

V. CONCLUSION & FUTURE SCOPE

Out of the three models which are tested the best accuracy was obtained by Transfer Learning. By using Transfer Learning accuracy of 97.59% & 99.4% has been achieved for Pneumonia

and Tuberculosis. The model used to overfit a lot, so used data augmentation, batch norm, and dropouts to reduce variance. A web app has been developed that currently detects whether a person has Pneumonia or Tuberculosis with the help of an X-Ray image.

As an extension, the proposed system can be revitalized to address more chest diseases. Also, the accuracy of the existing models can be improvised. Incorporating the above-mentioned changes, the system then can be delivered to Pulmonologists for evaluation in terms of user interface & features required to detect diseases and suggest areas of improvement to benefit the general public. Once this evaluation is done, the authors are planning to approach the medical bodies such as ICMR, NHS, CDC, etc to widen the scope of the project.

ACKNOWLEDGEMENTS

The authors are thankful for all the help that Dr. Haresh Kukreja (M.B.B.S, Maharashtra Medical Council) and Dr. Meenakshi Thakur (M.D., Radiology, Tata Memorial Hospital) for their support extended to the authors in understanding the technical features.

REFERENCES

[1]. Ke, Qiao, et al. "A neuro-heuristic approach for recognition of lung diseases from X-Ray images." Expert Systems with Applications 126 (2019): 218-232.

[2]. Kalhori, S.R.N. and Zeng, X.J., 2013. Evaluation and comparison of different machine learning methods to predict the outcome of the tuberculosis treatment course.

[3]. Chandra, Tej Bahadur, et al. "Automatic Detection of Tuberculosis Related Abnormalities in Chest X-Ray Images using Hierarchical Feature Extraction Scheme." Expert Systems with Applications (2020): 113514.

[4]. Rajpurkar, Pranav, et al. "Chexnet: Radiologist-level pneumonia detection on chest X-Rays with deep learning." arXiv preprint arXiv:1711.05225 (2017).

[5]. Siamak, Aram, et al. "Diagnosing Heart Disease Types from Chest X-Rays Using a Deep Learning Approach." 2019 International Conference on Computational Science and Computational Intelligence (CSCI). IEEE, 2019

[6]. Solti, Imre, et al. "Automated classification of radiology reports for acute lung injury: comparison of keyword and machine learning-based natural language processing approaches." 2009 IEEE International Conference on Bioinformatics and Biomedicine Workshop. IEEE, 2009.

[7]. Cao, Yu, et al. "Improving tuberculosis diagnostics using deep learning and mobile health technologies among resource-poor and marginalized communities." 2016 IEEE first international conference on connected health: applications, systems and engineering technologies (CHASE). IEEE, 2016.

[8]. WHO - World Health Organization [Online]. Available : www.who.int

[9]. Shreyak.Transfer Learning.Medium [Online] Available : https://bit.ly/3db0A9V

[10]. Sumit Saha.A Comprehensive Guide to Convolutional Neural Networks- the ELI5 way. [Online] Available: https://bit.ly/39hndbh

[11]. Kaggle [Online] Available : https://cutt.ly/Gbm1bRT

[12]. Kaggle [Online] Available : https://cutt.ly/ibm1Ubr