Lung X-Ray Image Enhancement to Identify Pneumonia with CNN

Mr. Mohommad Faraaz Firoz Khan[1] Mr. Sakshi Kishor Jadhav[2] Mr. Rushikesh Sanjay Ohol[3] Mr. Mohommad Afraaz Firoz Khan[4] Prof. Uttam R. Patole[5]

[1,2,3,4]BE Student [5]Assistant Professor

[1,2,3,4,5]Department of Computer Engineering

[1,2,3,4,5]SVIT, Nashik, India

Abstract— Pneumonia is a life-threatening infectious disease affecting one or both lungs in humans commonly caused by bacteria called Streptococcus pneumonia. The COVID-19 can cause severe pneumonia and is estimated to have a high impact on the healthcare system. Early diagnosis is crucial for correct treatment in order to possibly reduce the stress in the healthcare system. Pneumonia has caused significant deaths worldwide, and it is a challenging task to detect many lung diseases such as like atelectasis, cardiomegaly, lung cancer, etc., often due to limited professional radiologists in hospital settings. The standard image diagnosis tests for pneumonia are chest X-ray (CXR) and computed tomography (CT) scan. Although CT scan is the gold standard, CXR are still useful because it is cheaper, faster and more widespread. Chest X-Rays which are used to diagnose pneumonia need expert radiotherapists for evaluation. Thus, developing an automatic system for detecting pneumonia would be beneficial and it can save lots of people's life and help stopping and curing and control for treating the disease without any delay particularly in remote areas. Due to the success of deep learning algorithms in analyzing medical images, Convolutional Neural Networks (CNNs) have gained much attention for disease classification. In addition, features learned by pre-trained CNN models on large-scale datasets are much useful in image classification tasks. In this work, we appraise the functionality of pre-trained CNN models utilized as feature-extractors followed by different classifiers for the classification of abnormal and normal chest X-Rays. We analytically determine the optimal CNN model for the purpose. Statistical results obtained demonstrates that retrained CNN models employed along with supervised classifier algorithms can be very beneficial in analyzing chest X-ray images, specifically to detect Pneumonia. This study aims to identify pneumonia caused from other types and also healthy lungs using only X-Ray images. The model's performance in pneumonia detection shows that the proposed model could effectively classify normal and abnormal X-rays in practice, hence reducing the burden of radiologists.

Keywords: X-Ray, CXR, COVID-19, Chest X-Ray Images, Pneumonia Detection; Convolutional Network (CNN), Image Enhance

I. INTRODUCTION

Pneumonia is an infectious and deadly illness in respiratory that is caused by bacteria, fungi, or a virus that infects the human lung with the load full of fluid or pus. Chest X-rays are the common method used to diagnose pneumonia and it needs a medical expert to evaluate the result of X-ray. The troublesome method of detecting the pneumonia cause a life loss due to improper diagnosis and treatment. And the diagnosis of this diseases can take time to do that and hospital must have good radiologist but in our country we can't afford it. So we must go to the automated system. With the emerging computer technology, development on an automatic system to detect pneumonia and treating the disease is now possible especially if the patient is in a distant area and medical services is limited. Pneumonia is a lung parenchyma inflammation often caused by pathogenic microorganisms, factors of physical and chemical, immunologic injury and other pharmaceuticals. There are several popular pneumonia classification methods: (1) pneumonia is classified as infectious and non-infectious based on different pathogeneses in which infectious pneumonia is then classified to bacteria, virus, mycoplasmas, chlamydial pneumonia, and others, while non-infectious pneumonia is classified as immune-associated pneumonia, aspiration pneumonia caused by physical and chemical factors, and radiation pneumonia. (2) Pneumonia is classified as CAP (community-acquired pneumonia), HAP (hospital-acquired pneumonia) and VAP (ventilator-associated pneumonia) based on different infections, among which CAP accounts for a larger part. Because of the different range of pathogens, HAP is easier to develop resistance to various antibiotics, making treatment more difficult. Pneumonia kills more than 800,000 children under five per year, with around 2200 deaths every day. There are more than 1400 children infected with pneumonia per 100,000 children. The Global Burden of Disease Study reported that lower respiratory tract infections, including pneumonia, were the second largest cause of death in 2013. In Europe, nearly 35% of patients in hospital are infected with pneumococcal disease and worldwide, the percentage is 27.3% .In India, the latest report of John Hopkins Bloomberg School of Public Health has said that India suffers the most pneumonia deaths and in 2015, there were about 2.97 lakh pneumonia and diarrhea deaths in children aged less than five years old. The COVID-19 can cause illness to the respiratory system, fever and cough and in some extreme cases can lead to severe pneumonia. Pneumonia is an infection that causes inflammation primarily in the lungs' air sacs responsible for the oxygen exchange. Pneumonia can be caused by other pathogens besides SARS-CoV-2, such as bacteria's, fungi and other viruses. Several characteristics can influence its severity: weak or impaired immune system, chronic diseases like asthma or bronchitis, elderly people and smoking. The treatment depends on the organism responsible for the infection, but usually requires antibiotics, cough medicine, fever reducer and pain reliever. Depending on the symptoms, the patient may need to be hospitalized; in severe cases the patient must be admitted into an intensive care unit (ICU) to use a mechanical ventilator to help breathing. By analyzing CXR images, we can observe that texture is one of the main visual attributes present in those images. So, we decided to extract features from CXR images by exploring some popular texture descriptors, and also a pre-trained CNN model, not to neglect the power of representation learning approaches. Thus, for the flat classification, using the extracted features,

we applied some well-known multi-class classification algorithms. Normal classification techniques are hampered by the fact that the data is typically both incomplete and heterogeneous. To address this two-fold obstacle, we propose a CNN variant (CNN) algorithm which accurately and efficiently classifies Pneumonia. The main two ideas behind the proposed algorithm are that for each instance to be classified it chooses the parameter K adaptively and calculates the distances to other instances in a novel way. The CNN was implemented and tested on a Pneumonia dataset from the Italian society of medical and intervention radiology society. It was also compared to three algorithms of its category. The test results show that the CNN can efficiently and accurately classify Pneumonia patients. The comparison results show that the algorithm greatly outperforms all its competitors in terms of four metrics: precision, recall, accuracy, and F-Score. This study intends to incorporate deep learning methods to alleviate the problem. Convolutional Neural Network is optimized to perform the complicated task of detecting diseases like pneumonia to assist medical experts in diagnosis and possible treatment of the disease. Medical personnel check the patient's radiograph of their chest to determine if they are infected with pneumonia or not. In addition, the usual method for finding pneumonia is through medical history and laboratory results of the patient. Radiograph of chest is penetrated through X-rays where the soft tissues produces a dark color and hard tissues like bones produces a bright color. Patients diagnosed with pneumonia shows the chest cavity signs of fluids filling the air sacs of lungs as for the radiograph picture appears brighter. The deep learning by means of the convolutional neural networks has ability in obtaining significant characteristics in image classification tasks and provides medical promising results in image analysis. CNN advantages is capable in assisting the identification of some features from an image and use this feature to generate probabilities in classifying specific input .The contribution of this study is developed an optimized deep learning models of CNN that can detect and classify pneumonia diseases efficiently . The work consists of an optimized CNN models and experimental analysis of each model towards the detection and classification of pneumonia diseases.

II. Literature Review

Many researchers have contributed to this field. Various combinations of existing technologies have been used.

1) In 2016, Redmon et al. proposed YOLO, which does not require a separate region proposal network, so its detection speed is extremely fast and can reach 45 FPS. In the same year, Liu et al. [11] proposed the SSD algorithm. Both SSD and YOLO win in detection speed, but SSD uses a multiscale feature map to detect independently, the spatial resolution of images in deep networks has been significantly reduced, and it may not be possible to locate small targets that are difficult to detect in low resolution, reducing the accuracy of detection. YOLO does not use multiscale feature maps for independent detection. It smoothes the feature map and splices it with another lower-resolution feature map, but it treats the detection only as a regression problem and the detection accuracy is low. In 2014, Girshick et al. proposed R-CNN, which greatly improved the speed of training. On the PASCAL VOC 2010 dataset, the mAP improved from 35.1% to 53.7%.[1]
2) In 2018, Lee et al. Proposed DetNet, which was designed specifically for target detection and achieved better detection results with fewer layers. To avoid the large computational complexity and memory consumption caused by the high-resolution feature map, the network adopts a low-complexity dilated bottleneck structure; a higher resolution of the feature map is ensured while obtaining a higher subtractive field. This paper draws on the idea of DetNet and the framework of Faster R-CNN to study the detection of pneumonia. [2]
3) In recent years, many scholars have made efforts to detect pneumonia. Abiyev and Ma'aitah apply a convolutional neural network (CNN) for the diagnosis of chest X-ray diseases. Compared to BPNN and RNN, CNN gets higher precision but longer training time. Vijendran and Dubey combine multilayer extreme learning machine (MLELM) and online sequential extreme learning machines (OSELM) to detect pneumonia on the chest X-ray image. Abiyev and Ma'aitah explore the features extracted from layers of the CNN along with a set of classical features, including GIST and bag of words on a dataset of more than 600 radiographs [4]
4) Chowdhury et al. worked with chest X-ray images to develop a novel framework named PDCOVIDNet based on parallel-dilated CNN. In the proposed method, the authors used a dilated convolution in the parallel stack that could capture and stretch necessary features for obtaining a detection accuracy of 96.58%. [5]
5) Abbas et al. proposed and validated a deep convolutional neural network called decompose, transfer, and compose (DeTraC) to detect COVID-19 patients from their chest X-ray images. They proposed a decomposition mechanism to check irregularities from the dataset by investigating class boundaries for obtaining a high accuracy (93.1%) and sensitivity (100%).[6]

III. Problem Statement

To make an efficient use of Machine Learning techniques. Provide solution with least hardware requirement. To develop an application that is cost efficient. Minimize the use of Treatment as Normal People can't afford costly equipment. Easy to use and accurate so that medical can adopt the application quickly. To implement application to Detect Pneumonia in early stage by using Deep Learning Classification algorithm are executed using Convocational Neural Network (CNN).

IV. Objectives

The objectives are as follows:

1) Different models of deep learning and transfer learning are analyzed in this work for the image classification application. Using Deep Learning features are extracted from the images and are used for classification of x-rays having pneumonia.

2) A convolutional neural network architecture is made and it is trained with the images of x-rays with pneumonia and normal x-rays which will be used for further classification.
3) CNN models have been created from scratch and trained on Chest X-Ray Images (Pneumonia) dataset on Kaggle.
4) To discover this disease as early as possible.
5) If we discover this disease earlier, then the treatments are more likely to improve the quality life of the patients and their families.
6) Develop predictive models to differentiate between healthy people and people with Pneumonia Disease.
7) Study and analyses different learning models, including CNN.

V. METHODOLOGY

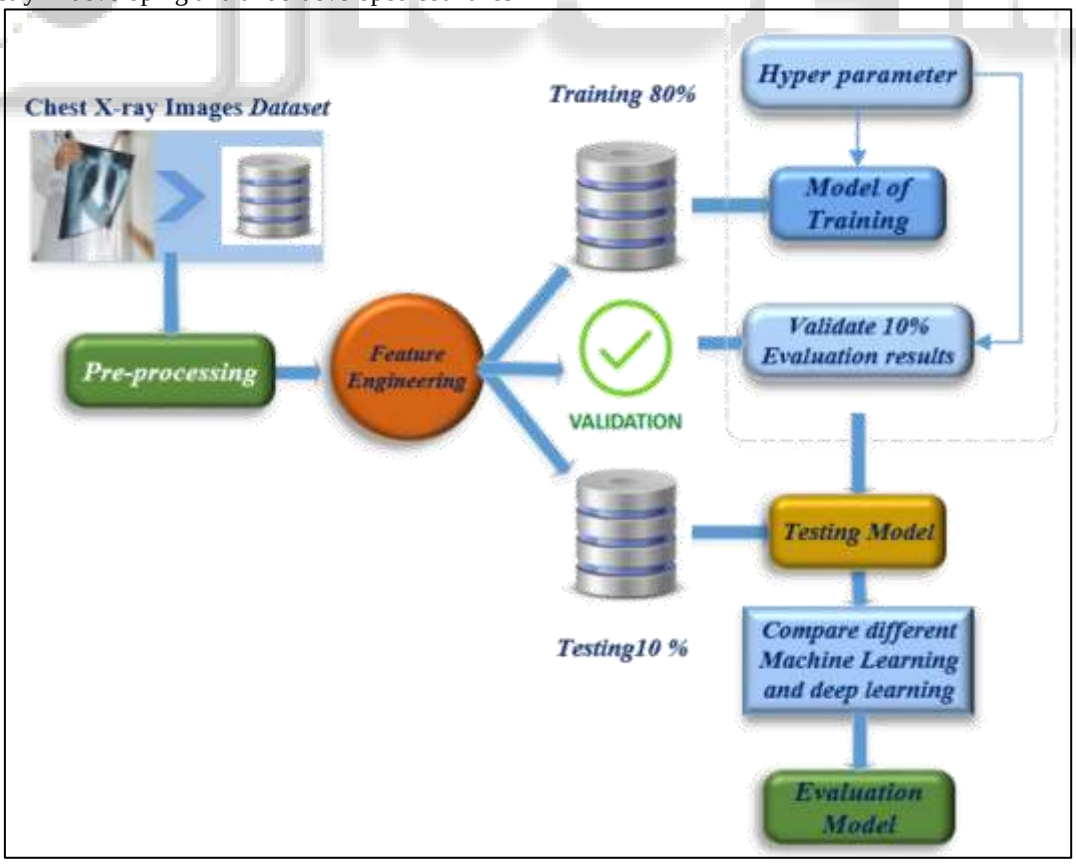
CNN models have been created from scratch and trained on Chest X-Ray Images (Pneumonia) dataset on Kaggle. Keras neural network library with Tensor Flow backend has been used to implement the models. Dataset consists of 5216 training images, 624 testing images and 16 validation images. Data augmentation has been applied to achieve better results from the dataset. The four models have been trained on the training dataset, each with different number of convolutional layers. Each model was trained for 20 epochs, with training and testing batch sizes of 32 and 1, respectively. The following sub-headings further explain the above stages in depth

VI. APPROACH

Pneumonia affects a large number of individuals, especially children, mostly in developing and underdeveloped countries characterized by risk factors such as overcrowding, poor hygienic conditions, and malnutrition, coupled with the unavailability of appropriate medical facilities. Early diagnosis of pneumonia is crucial to cure the disease completely. Examination of X-ray scans is the most common means of diagnosis, but it depends on the interpretative ability of the radiologist and frequently is not agreed upon by the radiologists. Thus, an automatic CAD system with generalizing capability is required to diagnose the disease. To the best of our knowledge, most previous methods in the literature focused on developing a single CNN model for the classification of pneumonia cases, and the use of the ensemble learning paradigm in this classification task has not been explored. However, the ensemble learning model incorporates the discriminative information from all the constituent base learners, allowing it to make superior predictions, and thus was implemented in this study. To handle the low amount of available biomedical data, transfer learning models were used as base learners, the decision scores of which were ensemble

VII. EXISTING SYSTEM

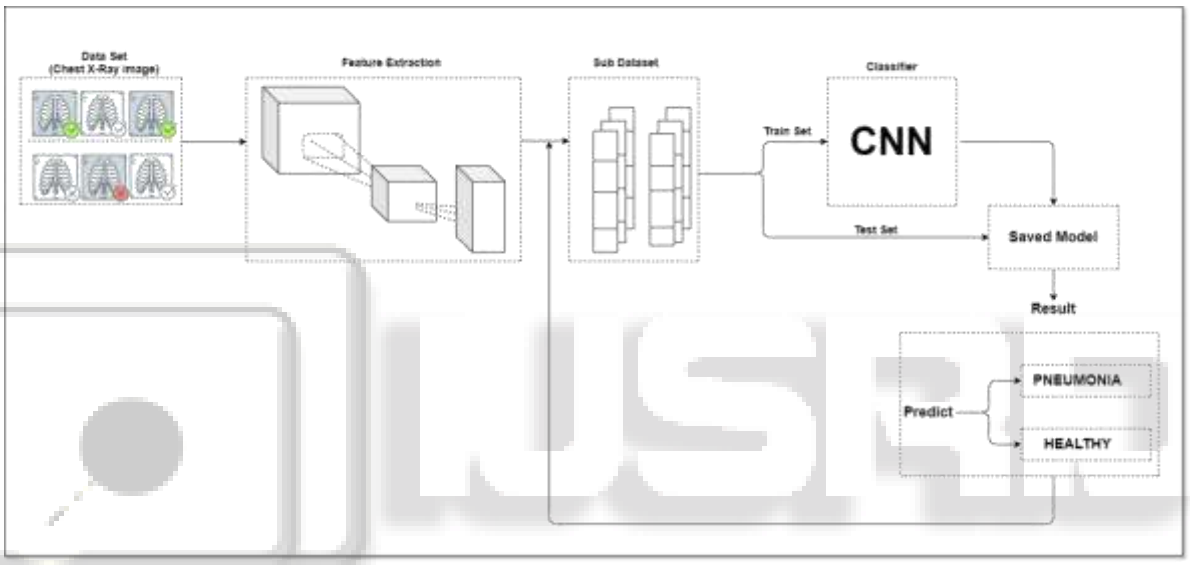
The main diagram of our existing full system. As seen in this figure, the system is composed of several steps including the Data collection and preprocessing description, building the classification models, and extracting the required features. These steps can be divided into four phases: Data Set, data preprocessing, building and validating classification models, and feature extraction.



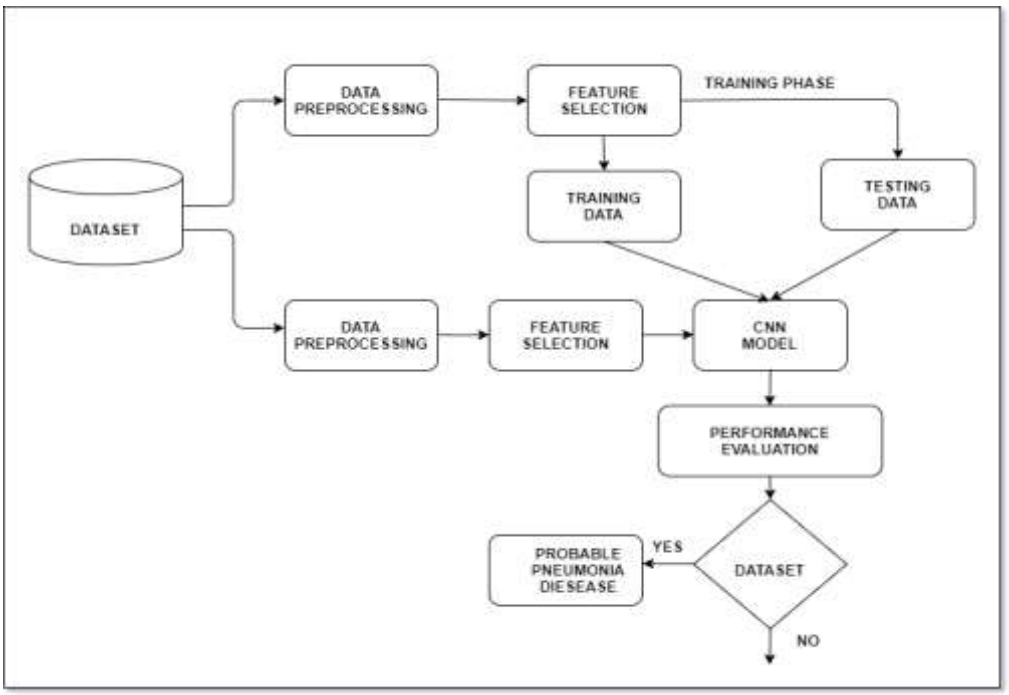
VIII. PROPOSED SYSTEM

The standard image diagnosis tests for pneumonia are chest X-ray (CXR) and computed tomography (CT) scan. Although CT scan is the gold standard, CXR are still useful because it is cheaper, faster and more widespread. Chest X-Rays which are used to diagnose pneumonia need expert radiotherapists for evaluation. Thus, developing an automatic system for detecting pneumonia would be beneficial and it can save lots of people's life and help stopping and curing and control for treating the disease without any delay particularly in remote areas. Due to the success of deep learning algorithms in analyzing medical images, Convolutional Neural Networks (CNNs) have gained much attention for disease classification. In addition, features learned by pre-trained CNN models on large-scale datasets are much useful in image classification tasks. In this work, we appraise the functionality of pre-trained CNN models utilized as feature-extractors followed by different classifiers for the classification of abnormal and normal chest X-Rays. We analytically determine the optimal CNN model for the purpose. Statistical results obtained demonstrates that retrained CNN models employed along with supervised classifier algorithms can be very beneficial in analyzing chest X-ray images, specifically to detect Pneumonia. This study aims to identify pneumonia caused from other types and also healthy lungs using only X-Ray images. The model's performance in pneumonia detection shows that the proposed model could effectively classify normal and abnormal X-rays in practice, hence reducing the burden of radiologists.

A. Block Diagram:



B. System Architecture:



C. Algorithm-

1) Convolutional Neural Network(CNN Model):

We introduced the basic scheme of a neural network showing the feedforward neural network structure where each node is connected to all the previous layer nodes. But this structure is not viable for working with images. The image is a multidimensional input where the pixels that compose it maintain a relation of local correlation with the neighboring pixels, if we used feedforward neural network the number of necessary parameters would make its training unfeasible. Therefore, to work with images arise convolutional neural networks [46]. Convolutional neural networks are a type of neural network specialized for working with images. These networks are inspired by the processing of Visual Cortex when the human brain processes an image. Each node covers a region of the image and together with the rest of the nodes the whole image is covered. In a convolutional neural network 3 types of main layers are combined:

a) Convolutional layer: Acts as a feature extractor for the input image. It applies a series of filters or kernels, which contain the trainable weights, and which run through the entire image executing convolution operations and generating a feature map. Each of the filters ends up specializing in detecting certain characteristics. In the first layers they detect simple patterns such as different lines or curves and in the later layers they detect more complex characteristics or patterns such as objects or shapes. After a convolutional layer an activation function is applied, most commonly using the rectified linear unit (ReLU).

b) Pooling layer: This layer has the function of reducing the dimensionality of features maps. The most common operations are on the one hand to detect the maximum value of the subsample region discarding the rest to this operation or to calculate the average of the elements of each region.

c) Fully-connected layer: This is one or more fully connected layers and is located close to the output. These layers learns the relationships between features maps and obtains output predictions that minimize the cost function.

In Figure the different layers are combined following a logical sequence. The convolutional layer and pooling layer take care of extracting the features maps from the input image. And in the last layers are located the fully-connected layers that carry out the classification function, reaching a prediction from the feature map extracted from the image.

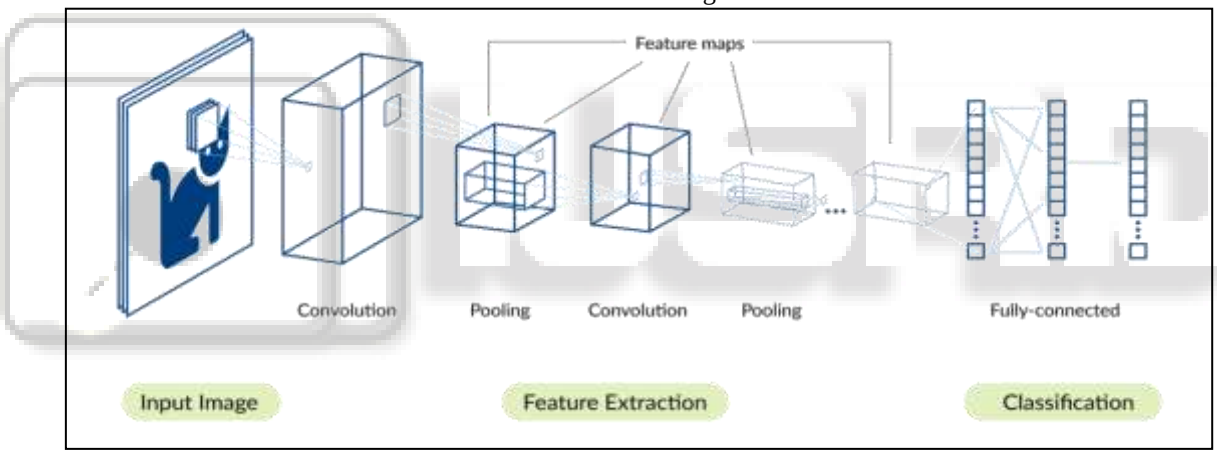
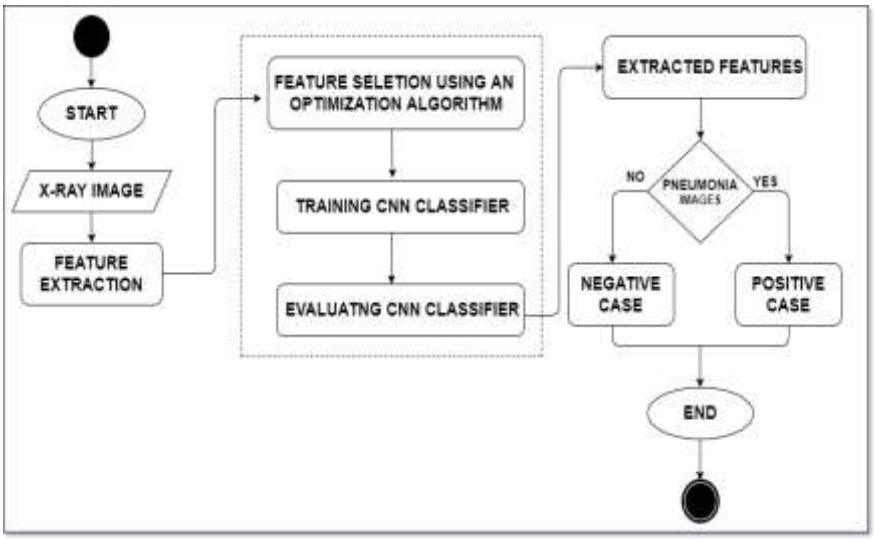


Fig. Basic structure of a neural convolutional network.

D. Flow Chart of Algorithms:



CNN models have been created from scratch and trained on Chest X-Ray Images (Pneumonia) dataset on Kaggle. Keras neural network library with Tensor Flow backend has been used to implement the models. Dataset consists of 5216 training images, 624 testing images and 16 validation images. Data augmentation has been applied to achieve better results from the dataset. The four models have been trained on the training dataset, each with different number of convolutional layers. Each model was trained for 20 epochs, with training and testing batch sizes of 32 and 1, respectively. The following sub-headings further explain the above stages in depth.

IX. MATHEMATICAL MODEL

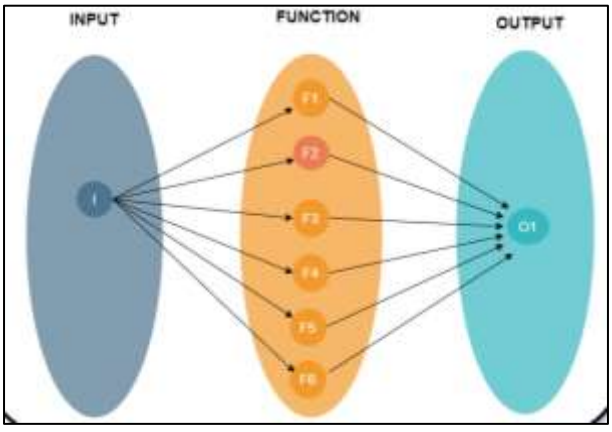
A. System Description:

S= I, O, F, DD, NDD, Failure, Success
Where,
S=System
I= Input
O=Output
F=Failure
S=Success
I is Input of system
Input I = set of Inputs
Where,
I= {Image File}
F is Function of system
F = set of Function
Where,
F1= {Browse Image or X-Ray}
F2= {Feature Extraction}
F3= {Train Dataset}
F4= {Feature Matching}
F5= {Test Dataset}
F6= {CNN}

O is Output of system
Output O = {Pneumonia Detection}

- Success Conditions: Pneumonia detected successfully.
- Failure Conditions: if internet connection Unavailable.

B. Venn diagram:



Where,
I= {Image File}
F1= {Browse Image or X-Ray}
F2= {Feature Extraction}
F3= {Train Dataset}
F4= {Feature Matching}
F5= {Test Dataset}
F6= {CNN}
O = {Pneumonia Detection}

C. Software/Hardware Required

1) Hardware Requirements-

- System: Pentium Dual Core
- Hard Disk: 120 GB.
- Monitor: 15" LED
- Input Devices: Keyboard, Mouse
- RAM: 1GB.

2) Software Requirements-

- Operating System: Windows XP and later versions
- Front End: HTML,CSS
- Programming Language: Python
- Database: MySQL
- Algorithm: CNN

D. Development Environment:-

We are developing application in .net framework. Web cam will be used for video capturing frames from video taken for face detection and recognition, if missing face matched then we will generate notification to police.

E. Testing Environment:- OS: Windows 7,8 RAM: 2 GB

F. Advantages & Disadvantages:

1) Advantages:

1) User Friendly.
2) Reduce the increasing count of Pneumonia.
3) Detect Pneumonia
4) Increased accuracy for effective Pneumonia disease diagnosis.
5) Handles roughest(enormous) amount of data using random forest and CNN algorithm and feature selection
6) Reduce the time complexity of doctors.
7) Cost effective for patients.

2) Disadvantages:

1) Required Internet Connection

3) Application Areas:

1) Used to detect Pneumonia at early stage.
2) Medical System

X. CONCLUSION

We have successfully describes a CNN-based model aiming to diagnose pneumonia on a chest X-ray image set. This will help our early detection of pneumonia quickly. Our project successfully provides with a CNN based approach for detection of pneumonia automatically.

REFERENCE

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779–788, Las Vegas, America, 2016.

[2] Z. Li, C. Peng, G. Yu, X. Zhang, Y. Deng, and J. Sun, "Detnet: design backbone for object detection," in Proceedings of the European Conference on Computer Vision (ECCV), pp. 334–350, Munich, Germany, 2018.
[3] S. Vijendran and R. Dubey, "Deep online sequential extreme learning machines and its application in pneumonia detection," in 2019 8th International Conference on Industrial Technology and Management (ICITM), pp. 311–316, Cambridge, England, 2019.
[4] R. H. Abiyev and M. K. Ma'aitah, "Deep convolutional neural networks for chest diseases detection," Journal of healthcare engineering, vol. 2018, 12 pages, 2018.
[5] Chowdhury, N.K.; Rahman, M.M.; Kabir, M.A. PDCOVIDNet: A parallel-dilated convolutional neural network architecture for detecting COVID-19 from chest X-ray images. Health Inf. Sci. Syst. 2020, 8, 1–14.
[6] Abbas, A.; Abdelsamea, M.M.; Gaber, M.M. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. Appl. Intell. 2021, 51, 854–864.
[7] Anthimopoulos, M., Christodoulidis, S., Ebner, L., Christe, A., Mougiakakou, S.: Lung pattern Classification for interstitial lung diseases using a deep convolutional neural network. IEEE Trans. Med. Imaging 35(5), 1207–1216 (2016)
[8] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
[9] Glozman, T., Liba, O.: Hidden Cues: Deep Learning for Alzheimer's disease Classification CS331B project final report (2016)
[10] Hemanth, D.J., Vijila, C.K.S., Selvakumar, A.I., Anitha, J.: Performance improved iterationfree artificial neural networks for abnormal magnetic resonance brain image classification. Neurocomputing 130, 98–107 (2014)

