**Developing a Pneumonias Disease Prediction System using CNN-based Deep Neural Network**

**Fatima Zulfiqar #1, Rehan Raza \*2**

***# Department of Computer Science, COMSATS University Islamabad***

***Lahore, Pakistan***

**1 *SP20-RCS-005@cuilahore.edu.pk***

**\* *Department of Computer Science, COMSATS University Islamabad***

***Lahore, Pakistan***

**2 SP20-RCS-018@cuilahore.edu.pk**

**Abstract**

Pneumonia is one of the life threatening very common disease and needs proper diagnosis at an early stage for proper treatment of recovery. Chest X-ray is used as an imagining modality to identify the disease by a professional radiologist. This report proposed a deep learning based Convolutional Neural Network framework for automatic detection of pneumonia from X-ray images to assist the medical practitioners. Evacuation was carried out on Pneumonia Disease Dataset using Accuracy, Precision, Recall and F1 measures. Extensive simulation results show that the proposed approach enables detection of pneumonia with 97% test accuracy.

**1. Introduction**

According to the World Health Organization (WHO), pneumonia kills about 2 million children under 5 years old every year and is consistently estimated as the single leading cause of childhood mortality killing more children than HIV/AIDS, malaria, and measles combined. . Pneumonia is a lung infection which can be caused by either bacteria or viral pathogens but require very different forms of management. Pneumonia requires urgent referral for immediate antibiotic treatment with supportive care. Therefore, accurate and timely diagnosis is imperative. One key element of diagnosis is radiographic data, since chest X-rays are routinely obtained as standard of care and can help to classify the pneumonia. However, rapid radiologic interpretation of images is not always available, particularly in the low-resource settings where childhood pneumonia has the highest incidence and highest rates of mortality. To this end, we developed an efficient model that predict the pneumonia disease from chest X-ray images.

**2. Approaches**

**2.1. Baseline Approach**

As baseline approach, we used Majority Class Categorization (MCC) approach. MCC is calculated by assigning the label of Majority Class to all the Test Instances.

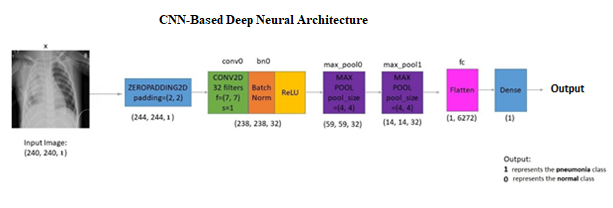
* 1. **Proposed Approach**

Deep neural network models have conventionally been designed, and experiments were performed upon them by human experts in a continuing trial-and-error method. Such process requires enormous time, know-how and resources. To overcome this problem, a novel but simple model is introduced that automatically perform optimal classification tasks with deep neural network architecture. This neural network architecture is specifically designed for pneumonia image classification tasks. The proposed technique is based on the convolutional neural network (CNN) algorithm, utilizing a set of neurons to convolve on a given image and extract relevant features from them.

|  |  |
| --- | --- |
| Main Parameters | |
| Number of Epoch | 15 |
| Optimizer | Adam |
| Learning Rate | 0.0001 |
| Loss Function | Binary\_crossentropy |

**Table 1:** Main Parameters of Model

CNN model consists of two major parts: the feature extractors and a classifier (ReLU activation function). Each layer in the feature extraction layer takes its immediate preceding layer's output as input, and its output is passed as an input to the succeeding layers. The convolution layer and pooling layers are responsible for feature extraction whereas dense layers which are also called fully connected layers are consists of multiples neuron units and last an output layer is responsible for classification. The proposed architecture in Figure 1 consists of the convolution, max-pooling, and classification layers combined together.

**Figure 1:** Architecture of CNN

The feature extractors comprise conv3 × 3, 32; conv3 × 3, 64; conv3 × 3, 64; conv3 × 3, 128; conv3 × 3, 256, max-pooling layer of size 2 × 2, and a RELU activator between them. The output of the convolution and max-pooling operations are assembled into 2D planes called feature maps.

It is worthy to note that each plane of a layer in the network was obtained by combining one or more planes of previous layers. The classifier is placed at the far end of the proposed convolutional neural network (CNN) model. It is simply an artificial neural network (ANN) often referred to as a dense layer. This classifier requires individual features (vectors) to perform computations like any other classifier. Therefore, the output of the feature extractor (CNN part) is converted into a 1D feature vector for the classifiers. This process is known as flattening where the output of the convolution operation is flattened to generate one lengthy feature vector for the dense layer to utilize in its final classification process. The classification layer contains a flattened layer, one dense layers of size 512 and 1, respectively, a RELU between the two dense layers and a sigmoid activation function that performs the classification tasks.

1. **Experimental Setup**

This section present the experimental setup including Dataset, Technique, Evaluation Methodology and Evaluation Measure.

**3.1. Data Set**

The original dataset consists of three main folders (i.e., training, testing, and validation folders) and two subfolders containing pneumonia (P) and normal (N) chest X-ray images, respectively taken form Kaggle (<https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>). A total of 5,856 X-ray images of anterior-posterior chests were carefully chosen from retrospective pediatric patients between 1 and 5 years old shown in Table.

|  |  |
| --- | --- |
| Class | Number of Instances |
| Normal | **1,583 (27%)** |
| Pneumonia | **4,273 (73%)** |
| Total | **5,856** |

**Table:** Dataset

The entire chest X-ray imaging was conducted as part of patients’ routine medical care. To balance the proportion of data assigned to the training, testing and validation set, the original data category was modified. We rearranged the entire data into training, testing and validation set only. The total of 80% images are used for training set, 10% for testing set and 10% for validation set. A total of 4,684 images were allocated to the training set, 587 images were assigned to the testing set and 585 images were assigned to the validation set to improve validation accuracy. A complete distribution of dataset is given in Table 2.

|  |  |  |  |
| --- | --- | --- | --- |
| Classes | Training Set | Testing Set | Validation Set |
| Normal | 1266 | 159 | 158 |
| Pneumonia | 3418 | 428 | 427 |
| Total | **4684** | **587** | **585** |

**Table 2:** Distribution of Dataset

**3.2. Technique**

|  |  |
| --- | --- |
| System Settings | |
| Developer Name | Ms. Fatima Zulfiqar & Mr. Rehan Raza |
| Programming Language | Python 3.7.4 |
| IDE | Google Colab Notebook |
| Deep Learning Toolkit | Tensorflow 2.0  Keras 2.2 |
| Code Version | 1.0 |
| Date | 01 – December – 2020 |

We present the detailed experiments and evaluation steps undertaken to test the effectiveness of the proposed model. Our experiments were based on a chest X-ray image dataset. We deployed Keras open-source deep learning framework with tensorflow backend to build and train the convolutional neural network model. All experiments were run on a standard PC. We perform pre-processing on training, testing and validation data.

Several preprocessing steps were performed on X-ray images before fitting them into the model. Note that preprocessing steps might vary from problem to problem. In this project X-ray images were resized into a fix dimension of 224×224 pixels and then resized RGB X-ray images were converted into grayscale images. These preprocessing steps were performed on training, validation and testing set respectively. Instances in each set were converted into a numpy array and then splitted into input feature vector and output labels. After that input feature vectors were then normalized to scale down the range of intensity values of pixels from 0-255 to 0-1 and a datatype of float 32. These normalized feature vectors were then flattened to fit into the model.

**3.3. Evaluation Methodology**

Pneumonia disease prediction problem is treated as the supervised machine learning problem. We deal the problem as a binary classification problem because our goal is to distinguish between two classes: Pneumonia and Normal. Furthermore, we have used CNN based approach to train the model that is based on automated feature extraction method.

**3.4. Evaluation Measure**

The models is tested on the test dataset after the completion of the training phase. The performance is evaluated using the accuracy, recall, precision, F1. The dataset that we used for the training of model is highly unbalanced therefor F1 and precision is the better evaluation measure for the deeply analysis of the prediction results. As our problem is on medical images we also used recall and sensitivity and also used confusion matrix as well.

* Accuracy is defined as the proportion of correctly classified Test instances
* Precision (P) is defined as the proportion of the predicted Positive cases that were correct
* Recall is defined as the proportion of Positive cases that were correctly classified
* F1 – measure is defined as Harmonic mean of Precision and Recall, When we assign same weights to Precision and Recall the F – measure become F1 – measure

1. **Results and Analysis**

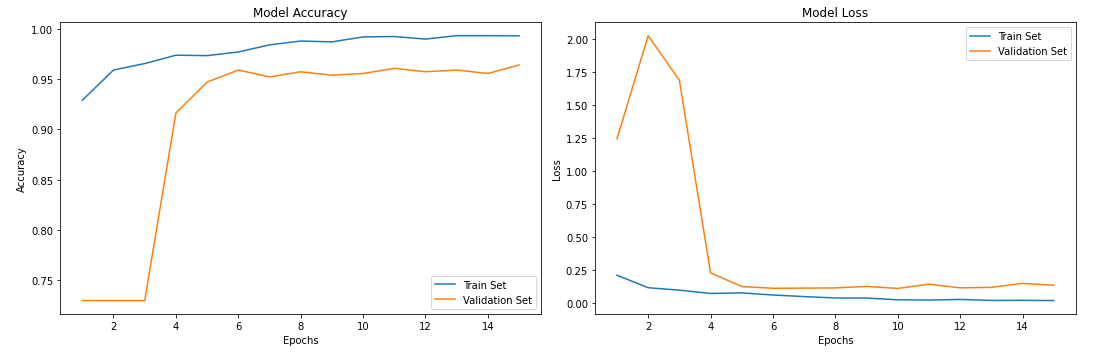
To evaluate and validate the effectiveness of the proposed approach, we conducted the experiments 10 times each for one hours, respectively. Parameter and hyper parameters were empirically turned to increase the performance of the model. Different results were obtained, with baseline approach applying on test data, we got 0.73

But in proposed approach, the final results obtained are training loss = 0.029, training accuracy = 0.9710, validation loss = 0.1337, and validation accuracy of 0.9641.

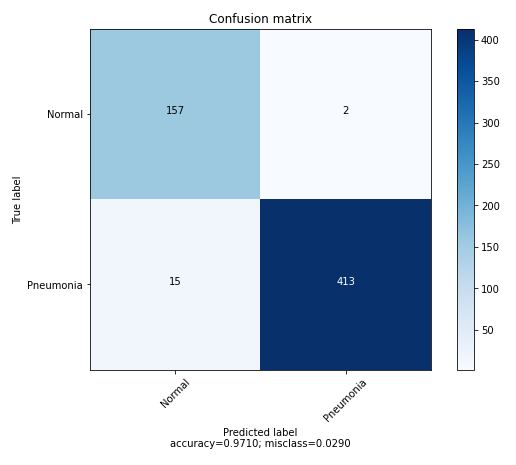
In the comparison of chest X-rays presenting as pneumonia versus normal, we achieved an accuracy of 97%, with a F1 of 98% and recall of 96%. The training and validation accuracy and loss curve of the purposed CNN model is shown in Figure 2.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Technique | Class | Accuracy | Precision | Recall | F1 – Score |
| MCC | **Normal** | - | - | - | - |
|  | **Pneumonia** | - | - | - | - |
|  | **Macro Average** | 0.73 | - | - | - |
| CNN | **Normal** | 0.99 | 0.91 | 0.99 | 0.95 |
|  | **Pneumonia** | 0.96 | 1.00 | 0.96 | 0.98 |
|  | **Macro Average** | 0.98 | 0.95 | 0.98 | 0.96 |

**Table 3:** Classification Report



**Figure 2:** Training and validation accuracy and loss curve of the proposed CNN model



**Figure 3:** Confusion Matrix of CNN Model

First of all for better results and accuracy depth CNN model with hidden layer play a vital role. By increasing the convolutional and pooling layer more complex and higher level features can be extracted. As bottom layers are responsible for low level feature extraction and top layers are responsible for high level feature extraction. Secondly, we added dropout layer to simplify the CNN model and to avoid the model from overfitting. Thirdly, we keep the learning rate very small for better convergence.

We developed a model to predict pneumonia from chest X-ray images taken from frontal views at high validation accuracy. The algorithm begins by transforming chest X-ray images into sizes smaller than the original. The next step involves the identification and classification of images by the convolutional neural network framework, which extracts features from the images and classifies them. Due to the effectiveness of the trained CNN model for identifying pneumonia from chest X-ray images, the validation accuracy of our model is 97%. To affirm the performance of the model, we repeated the training process of the model several times, each time obtaining the same results.

1. **Findings:**
2. Proposed Technique (CNN Model) outperforms the Baseline Technique (MCC)
3. Accuracy of CNN Model is 0.97 compared to 0.73 obtained with MCC. This indicates a huge improvement in performance
4. The most problematic Class is Pneumonia
5. Google Colab is a good platform to run experiments (on large Data) particularly using CNN-based Deep Neural Networks
6. **Conclusion and Future Work**

We have demonstrated how to classify positive and negative pneumonia data from a collection of X-ray images. We build our model from scratch, which separates it from other methods that rely heavily on transfer learning approach. In the future, this work will be extended to detect and classify X-ray images consisting of lung cancer and pneumonia. Distinguishing X-ray images that contain lung cancer and pneumonia has been a big issue in recent times, and our next approach will tackle this problem. More ever we can increase the performance of our model by performing different data augmentation technique to increase the training data. And we can conduct several experiments using transfer learning to improve the accuracy of the model.