**A Neural Network Architecture using Separable Neural Networks for the identification of “Pneumonia” in digital chest radiographs**

# N. Sarada

*Research Scholar Computer Science & Engineering*

K L E F, Guntur

[saradab](mailto:saradabusi@gmail.com)[usi@gmail.com](mailto:usi@gmail.com)

Dr. K. Thirupathi Rao

*Professor & BOS Chairman Computer Science & Engineering*

Deputy Dean-Academics, K L E F, Guntur

***Abstract***—**In recent years, convolutional neural networks created a huge impact in the fields of medical image processing. Image Semantic Segmentation and Image Classification have been the main challenges in these fields. Using these two techniques, we have been seeing quite a lot of improvement in medical surgeries which are being carried out by robots and autonomous machines. In this work, we will be focusing on a convolutional model to detect pneumonia in the given chest X-Ray scan. In addition to the convolution model, the proposed model consists of deep separable convolution kernels that replace a few convolutional layers. One primary advantage is, these take in a small number of parameters and filters. Our described model will be efficient, robust and fine-tuned than the previous models developed using Convolutional Neural Networks. We also benchmarked the present model with the CheXnet model that predicts over 16 abnormalities in the given chest-x-rays.**

***Index Terms—Deep Convolutional Networks, Separable CNN’s, Artificial Neural Network, Optimization, Hyperparameters.***

1. Introduction

Pneumonia is one of the dreadful and viral diseases which is mainly caused because of micro bacteria, virus, and Mycoplasma. It grows dangerous fluidic material inside the air sacs making breathing more difficult. Pneumonia can easily survive in a healthy throat and can multiply, and work their way into the lungs. Pneumonia also has few symptoms including a mild headache, shaking chills, shortness of breathing, loss of appetite, and in older people, confusion is also observed as one of the symptoms for pneumonia. There is a high chance of an increase in the pneumonia severity if the patient smokes tobacco or if he’s a frequent swimmer. People living in the hospital environment and nursing homes are vulnerable to pneumonia. Due to these reasons, millions of people are being hospitalized across the world. To overcome such dreadful scenarios, a fast and accurate diagnosis is the need of the hour.

The traditional methods currently being used mainly revolve around radiographic imaging as the primary tool. Chest X-rays seem to be the only reliable source in detecting pneumonia. Some thresholds are set to identify the affected regions. However, to get accurate results from that, expert radiologists are required. This might not always be the case in remote, rural areas. Also, a doctor manually examining chest x-ray might not always be correct if at all the patterns are blurred. This could lead to an incorrect diagnosis. Hence, these aren’t the best and reliable techniques that can yield dependable results.

In this work, we will be using these symptoms as the parameters for the network and build a highly efficient classifier. We shall be using neural networks, which are a class of machine learning algorithms to build a classifier that classifies pneumonia in chest x-ray images. In recent years, the artificial neural network made a wide impact in the fields of medical image processing from detecting tumors in brains to finding the cure to several diseases using different types of neural network architectures. A few main medical fields where neural networks are widely used are diagnostic systems, biomedical analysis, image analysis, and drug development. The concept of developing neural networks for medical purposes started with processing techniques used for reducing noises and blurs that are indispensable. Secondly, these are also used for Magnetic Resonance Images (MRIs), where these could produce high-quality images of x-ray scans, however, with little pixel variation. Neural Networks are used on these MRIs for semantic segmentation of tumors and blood clots.

In our architecture, we shall be training our CNN extensively to extract precise output images wherein Pneumonia detection statistics are plotted. We feed a set of chest x-rays, and the model trains on the images given and outputs the labels associated with each image.

1. Background

Most of the classification in which pneumonia is categorized is based on the country average low income and high income [1] which was presented by the World Health Organization. There are a few guidelines that classify the clinical signs of pneumonia which describe the threshold of risk for children. In this work, we aim to work on the chest radiology images which give a complete report of the threshold data as well as pneumonia classified images. We’ve identified that almost 832 out of 16,031 children die by pneumonia which is 5% of the whole. When it comes to teenagers, out of 11,788, 321 die due to pneumonia, which when summed up, boils down to 3% [1]. Most of the diagnosis that is carried out in the intensive care configuration is highly unreliable; this includes the diagnosis of the chest radiology manually. There are a few advanced techniques in neural networks dedicated to finding several abnormalities in chest X-Rays. Few include ChexNet [2], a 121-layer neural network architecture that is used to find about 14 abnormalities that also include screening, diagnosis, critical segmentation, and pneumonia classification, and achieved classification accuracy of 0.7680 on the pathology pneumonia. The one main disadvantage is that they take a lot of time to be implemented in rural areas and less populated countries due to minimal availability of the required data. The current practices in detecting pneumonia include manual identification by the radiologist using the weights of the deep neural networks (less precise) for classification, and by using the parameters with respective to the fields of microbiology, etiology, radiology, etc, calculating the entropy to measure the difference between normal respiratory system that’s all fine with the one affected by pathology, and comprehending the snoring patterns to detect for irregularities that might be the causes for pneumonia.

Along with the mentioned ones, a lot more techniques have forayed trying to improvise the mundane ways of detecting Pneumonia. It started with the work[8] proposed. It involved pulmonary segmentation and extraction of characteristics using an ANN. The other work [9] includes detecting the pulmonary nodules from a set of CT images.

Pattern recognition in machine learning includes four main steps: data acquisition, preprocessing, feature extraction by Convolutional neural networks and classification. Convolutional Neural Networks are the most predominant neural networks used for the multidimensional signal preprocessing in the fields of deep learning. These are proposed by YanLeCunn et al.[4] which is named the Alex net. It is an eight-layered convolutional neural network consisting of five convolutional kernels and three fully connected layers. The activation used in Alex Net is “relu” activation that happens at the end of every convolutional layer. The main goal of Alex Net is to classify the images, but not limited to any specific set of images.

1. Artificial Neural Networks

Artificial Neural Networks are inspired by biological neurons inside the human brain. They form network-like structures to solve complex problems where the basic computational units are neurons. The data travels through this artificial structure and then the network finds the patterns and connections between them and improves overtime to make the output more accurate and learning process more efficient. These artificial neural networks are organized in three layers. First, the input layers wherein we pass all the features of the output, the second are hidden layers where all the patterns and weights of the network are updated and last are the output layers where the data is classified or predicted. Neural Network weights, biases, and activation functions are the most important parameters to train the network, they iterate overall the training process in order to return the best prediction. Below is a diagram of a basic artificial neural network wherein the inputs are passed through the input layer and multiplied with the weights (the connection between the neurons) and resulting in output layer traversing through activation function in the hidden layers.

****

Figure 1. The basic intuition of artificial feedforward neural network with two hidden layers.

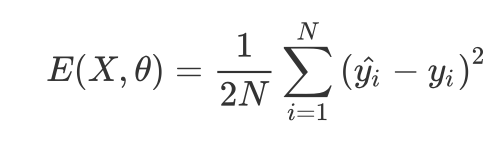
* 1. Hyperparameters

The neural network can only be defined by its hyperparameters, which include the number of Layers, Neuron Weights, Biases and Activation Function, Loss Function, Optimizer. Once the network is defined with all these hyperparameters we call it a model. The model will change based on the data that is given as the input, the dimensions of the data are the most considered the total number of input and output features a model can have. In this case of pneumonia classification, the model we take in will be a binary model as the network returns us a binary value. Prediction as 1 if it has a positive pneumonia response and 0 if the pneumonia test response is negative. Activation Function -Theyintroduce non-linear properties to our Network, Loss Function- Machines learn by means of a loss function. It’s a procedure of finding how well an algorithm models the given data, Optimizers- they update the weight parameters thereby to minimize the loss function.

* 1. Backpropagation

The backpropagation algorithm [5] can update the weights efficiently and is used in conjunction with the Gradient descent optimization method. The weights are enhanced resourcefully by means of Backpropagation that calculates the gradient and loss function in comparison with the weights of the network for single input-output, and for changing weights in multi-layer networks gradient methods is the best choice. It updates the weights repeatedly to minimize the errors between the actual output vector and the desired output vector. As a part of this weight adjustment, internal nodes which are neither inputs or outputs play a significant role in a task domain [6]. It involves finding the derivative of the cost function and then backpropagating the errors to update the weights. The cost function in most cases is calculated using Mean Squared Error (MSE) to find the deviations of predicted output over the actual output in case of Supervised Machine Learning algorithms.

Backpropagation tries to minimize the following error function by calculating for each weight $w^k\_{ij}$, weight for node j in layer k for an incoming node i.



1. Convolutional Neural Networks For Medical Image Classification

Convolutional Neural Networks is one of the widely used neural networks for image classification and image semantic segmentation. The one main advantage of Convolutional Neural Networks or ConvNets is that they can recognize the features if the dimension of the data is very high. They operate over volumes and assume that the input is a multi-channeled image. ConvNet arranges its neurons into the following three dimensions first the width of the medical image, second the height and the depth of the medical image to avoid a number of neurons. ConvNets have several layers based on the features to be predicted in the case of medical images few most used ConvNet architecture is U-Net. This U-Net architecture is proposed by Olaf Ronneberger et al for bio-medical image purposes. It has a convex path with a symmetrical expanding and contracting path, including deconvolutional layers. The deconvolutional layers are used for making the convoluted image to a high resolution which outperforms the prior best method (a sliding-window convolutional problem). This architecture entirely segments the nerves and body organs of humans.

* 1. Convolutional Neural Network Architecture

The convolutional neural network mainly consists of convolutional layers which are used for finding the edges of features in a given image. Adding more convolutional layers to your neural network results in finding more number of features as well as overfitting. In order to make your network more efficient, we need to track down the feature extraction process. Max Pooling and Average pooling operations are applied once the convolutional operation is applied to the images. These results in finding feature highlighting and finding the import features given in the image. RELU activation is used for making the network non-linear. In the last layer preceding the output the activation function used is a sigmoid activation function. This activation results in which class the image belongs to, in this case, pneumonia or any segmentation task. Below is the image of AlexNet the first convolutional neural network used to classify the digits.



Figure 2. Convolution pipeline with convolution, pooling and fully connected layers

* 1. data

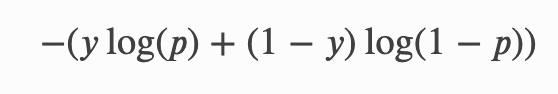
We use our convolutional model for pneumonia classification on an open-source dataset proposed by Daniel Kermany et al. The dataset trained on consists of OCT and Chest X-Ray images described and analyzed in "Deep learning-based classification and referral of treatable human diseases". The images in the dataset are already labeled based on a few features. The labels present in the dataset are disease which is almost 14 abnormalities, randomized patient ID (to make the learning more efficient and unbiased, and the image number by the patient. The dataset consists of almost 5,500 images of patients of two categories pneumonia and normal patients. We label the images into two categories present in the dataset the first as positive where the pneumonia is detected for the patient which we assign as one and the other as pneumonia negative which is assigned to zero. For the pneumonia detection task, we randomly split the dataset into training (~5000 patients, 5800images), and test (~800 patients, 1000 images). There is no patient overlap between the sets.

|  |
| --- |
|  |

Figure 3. Open-Source Dataset proposed by Daniel Kermany et al.

* 1. Problem Formulation

As discussed in the previous section we consider this problem to be a binary classifier; initially, the inputs to our proposed models should be frontal-scans of chest radiography. The labels are also passed into the network in terms of a binary value which returns 1 when pneumonia is positive and 0 if the pneumonia is negative. For every iteration of the input image, the weights are updated in the network by using the binary cross-entropy loss which is mathematically stated as,



where y is the actual class label that can be a 0 or 1 in case of a binary classifier (correct classification) and p is the predicted probability. Since there a number of data points that will be fed into the model, this loss iterates with respect to every data point.

The initial weights for pneumonia classification are updated from the VGG16 ImageNet model for the first four convolutional layers in the proposed model.

1. Proposed Model and Separable Convolutions

Most of the image classification techniques in medical fields include transfer learning, while the trick here for achieving the highest accuracy is by applying separable convolutional layers than normal convolutional and pooling layers. We will apply partial transfer learning to the model and the rest of the entire convolutional model is built from scratch. We used VGG16 [8] weights for initial layers for finding a few features like colors, shapes, blur, and noise in the given medical radiographic image. The model is entirely developed based on the python Machine Learning framework called Keras. Once the initial weights of the ImageNet model are added to the proposed model the second step is to fine-tune the model by initializing the training process. Now that the training starts the error and weights get better and better while we pass in the pneumonia dataset images. The entire network is trained using an end-to-end Adam Optimizer using the standard parameters. The standard parameters are beta1 which is the first constant is set to 0.9 and second beta2 to 0.999 same with reference to the ChexNEt training model. Initially, the learning rate of the model is considered to be 0.01 but with the less change in the values of the cost function, we considered that to be a learning rate of 0.0001, that has been delayed factor of to 1e-5 after every epoch.

Every epoch here defines as the one complete forward propagation and one complete backward propagation of the network to validate and update the loss. The total number of epochs that are considered to the proposed model is 20 with a batch size equal to 60.

* 1. Defining the Model

The proposed model only has two convolutional layers which are initialized with the weights of the VGG16 network. VGG16 is a pretrained convolutional neural network model presented for Keras. This dataset contains nearly 14 million images supporting 1000 variant classes.The two layers are followed by max-pooling operations which returns the main features and shapes in the given medical image. Once the basic features are extracted now we tune the network by using the two separable two-dimensional convolutional layers, the difference between the normal and separable convolutional layers is that separable convolutional layers take the column, the features and multiply with the convolutional kernel normally we take the first row of the feature matrix and multiply with the first column in the convolutional kernel. Separable convolutions are used to reduce the processing time, since they divide the convolutions depth wise so that the operation can be done with fewer number of multiplications in comparison with the regular convolutions that increases the no. of multiplications there by the processing time. The first two separable convolutional layers are followed by a max-pooling operation and again succeeded by the third separable convolutional layer. Now the number of features that goes into the next layer will be approximately 256 and the resolution of the images entirely shirks down where we add a batch normalization layer in order to streamline the total number of input features that can be passed to the next separable convolutional layer to find out more number of features. This set of three separable layers and batch normalization is repeated with a max-pooling operation making a total number of 18 layers neural network architecture. Now the input image totally loses it's dimensioned its now time to classify the extracted features here where we add a flatten layer that reshapes and extracts the feature map of the input image. This flattens layers might overfit the trained model, Hence we remove 0.7 neurons with dropout layers out of every iteration where we could solve the problem of overfitting the extracted pneumonia features. Since the output layer should return a binary value the last layer is a dense layer containing only two neurons followed by the softmax function. The advantage of using a softmax function is that it returns a binary value based on the training and the classification that the model could predict. All the convolutional, max-pooling, dense and separable two-dimensional convolutional layers are followed by a linear rectifier activation function making it non- linear and a classification model at every layer and epoch.



Figure 4. Proposed CNN architecture that includes Separable convolution layers to classify Chest X rays as Pneumonia Positive or Pneumonia negative

* 1. Training the Model

The proposed neural network architecture for pneumonia classification is trained on over 6000 frontal chest radiology scans as discussed in the Data Section. The labels of the pathologies are considered to be positive and negative with binary values and weights of first four layers same as the VGG16 network. We randomly split the dataset into another thirty percent for validating our training loss. Since the network has a number of parameters the network is added with batch normalization and dropout layers in order to reduce the overfitting of the data in our model. The size of the image is normalized to have 224x244 pixels and a 3-band making the image processed based on the input features of the network architecture. Interpolation and binomial filters are used to maintain the image quality even after it is being down sampled. The hyperparameters considered while training the model are a number of epochs to be 10 and the batch size considered is 16 images for the input. The transfer learning rate was applied at a patience rate of 5 and a learning rate as 0001, and the method of testing the efficiency is by the loss and the calculated accuracy for each iteration.

1. Results and Model Interpretation

In order to interpret the model the considered metrics were heat maps that are generated by the network where the pneumonia is identified and the confusion matrix that we could generate once the training process is done. The recall of the model is found to be 0.93 and the precision of the model is found to be 98 %. The loss table is tabulated below for every iteration.

|  |  |  |  |
| --- | --- | --- | --- |
| **Model**  **Accuracy** | **Model Loss** | **Validation Accuracy** | **Validation Loss** |
| 0.5474722385 | 0.75 | 0.02738 | 0.98322 |
| 0.2586314082 | 0.875 | 0.02566 | 0.98504 |
| 0.06607481092 | 1 | 0.02245 | 0.98792 |
| 0.06071493775 | 0.9375 | 0.02147 | 0.98945 |
| 0.632596314 | 0.8125 | 0.01357 | 0.99118 |
| 0.743275404 | 0.6875 | 0.02145 | 0.98696 |
| 0.0728430748 | 0.9375 | 0.02027 | 0.98888 |
| 0.02700931393 | 1 | 0.01535 | 0.99271 |
| 0.1509871036 | 0.875 | 0.01506 | 0.99022 |
| 1.122514844 | 0.625 | 0.01346 | 0.99309 |
| 0.3206430376 | 0.875 | 0.01183 | 0.99213 |
| 0.2556969225 | 0.9375 | 0.01978 | 0.99252 |
| 0.8750720024 | 0.875 | 0.00925 | 0.99271 |

Tabel 1. The Table above shows the Model accuracy and loss and Validation accuracy and loss. Model loss changed for the first four epochs it was because of the activation and the pooling layers. Once the features are highlighted by the first five layers using the weights of VGG model, the architecture slowly got adjusted and reduced the loss. Generated below is the graphical review of the Model accuracy and loss and Validation accuracy and loss with the number of epochs.



Figure 5. Epochs Vs Value Loss

The accuracy of the model increased for the first 10 epochs and tend to decrease because of overfitting. Once the model is hit with the dropouts the accuracy increases with the increase in the total number of epochs. At last, the model is settled with an average accuracy of 98% after 12 complete epochs without changing the batch size. Below are the critical metrics of the model accuracy and the loss metric which is generated by the developed model.



Figure 6. Epochs Vs Model Accuracy



Figure 7. Epochs Vs Value Accuracy

* 1. Model Output:

The display of images below show the chest radiographs that are classified by the proposed architecture as “Pneumonia” or “Normal”.



Fig 8: Predicted Labels

1. Conclusion

Pneumonia is one of the most identified diseases in India as well as many other countries identified. It is most commonly found in children and reported by the radiologists. In this work, we have proposed a new neural network model that can classify a chest radiography image having pneumonia positive or negative using a separable convolutional neural network and weights of a pre-trained model called VGG16. We usually use X-rays as the most common imaging tools in order to give a brief analysis of the pneumonia diagnosis, radiologists are not always precise in terms of analysis with the naked eye, Hence we added an additional tool with accuracy almost 97% that can find label pneumonia positive and negative based on the given lung X-Ray scan.Execution of the future work would accomplish an efficient medical image analysis for digital chest radiographs. Research work by means of Deep Learning for medical image analysis and consequent diagnosis needs to be more vividly done in Indian framework.

References

[1] Agweyu A, Lilford RJ, English M. Appropriateness of clinical severity classification of new WHO childhood pneumonia guidance: a multi-hospital, retrospective, cohort study. Lancet Glob Health 2018;6:e74–e83.

[2] Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, et al. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning; 2017 [cited 1 July 2018]. Preprint. Available from: <https://arxiv.org/abs/1711.05225>.

[3]Causes of Pneumonia, https://www.lung.org/lung-health-and-diseases/lung-disease-lookup/pneumonia/what-causes-pneumonia.html.

[4] Yann LeCun, L Bottou, Y Bengio and P Heffner Gradient-Based Learning Applied to Document Recognition, Proc. of the IEEE, November 1998.

[5] Learning representations by back-propagating errors David E. Rumelhart, Geoffrey E. Hinton & Ronald J. Williams, Nature Volume 323, pages 533–536 (09 October 1986).

[6] Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification

https://data.mendeley.com/datasets/rscbjbr9sj/2

[7] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In CVPR09, 2009.

[8] Khobragade, S., Tiwari, A., Patil, C. Y., and Narke, V. (2016). Automatic detection of major lung diseases using chest radiographs and classification by a feed-forward artificial neural network. In 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), pages 1–5.

[9] Paing, M. P. and Choomchuay, S. (2017). A computer-aided diagnosis system for the detection of lung nodules from series of ct slices. In *2017 14th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, pages 302–305.