Pneumonia Detection Using Novel Deep Learning Techniques

Chandrashekhara K T, Thungamani M

1Research Scholar, Computer Science and Engineering, BMS Institute of Technology and Management, Affiliated Visvesvaraya Technological University, Belagavi, Karnataka, India
2Assistant Professor of Computer Science, College of Horticulture, UHS Campus, Karnataka, India
Email: chandru@bmsit.in
DOI: 10.47750/pnr.2022.13.S09.131

Abstract

Pneumonia is one of the leading infectious diseases that can kill children and elderly people around the world. The development of an automated system to detect pneumonia would be advantageous, especially to enable treatment of this disease in remote areas without much delay. Pneumonia is a lung infectious disease which mainly affects the small air sacs known as alveoli. The main symptoms of pneumonia include cough, fever, and breathing problems. Aged people, children, and persons who have medical problems are the main victims of this disease. Around 450 million people are affected by this disease on an average each and every year. The most commonly used technique for diagnosing this disease is chest X-ray imaging. Chest X-ray examination is a complex procedure to detect the disease because it involves lots of vulnerabilities. With latest advances in technologies we can use deep learning algorithms to detect the disease using chest X-ray images. To deal with the scarcity of data, we used Deep Transfer Learning and designed the Hybrid Algorithms. The images of the chest X-rays were fed into the individual algorithms for training purposes. Parallel Deep Feature Extractors are used in conjunction with various algorithms. For classifying chest X-ray images into normal and pneumonia, we are proposing an hybrid model based on VGG16, VGG19, CNN, and MobileNet networks. Individual image classification algorithms were combined to form a hybrid model. In comparison to individual algorithms, the new Hybrid Model with Deep Learning achieved higher accuracy than existing methods.

Keywords: VGG16, VGG19, Mobile-NetConvolution Neural Network

1. Introduction

Pneumonia is a severe disease caused by infections inside the lungs. It is an important respiratory organ. These infections, which can occur in one or both lungs, infection is mainly by viruses, fungi, and bacteria. The infection can block air sacs inside lungs preventing them from receiving oxygen-rich air. Hence the patient's breathing would be difficult and dangerous. Pneumonia is a severe disease that causes the air sacs in one or both lungs to become inflamed. Every year, it kills more children under the age of five than any other infectious disease, including HIV infection, malaria, and polio. One or more visual and/or auditory stimuli or tuberculosis may be provided by the pneumonia screening device. The model we are proposing here has achieved 99.47 accuracy, it is good enough compared with existing models. Pneumonia is mainly caused by a variety of bacteria, viruses, and fungi, the most familiar of which is Streptococcus pneumonia. Pneumonia is an Infectious disease that easily spreads among the members of the society. Other characteristics that a pneumonia victim experiences include cough, chest pain, and fever. Based on the criticality, the person's age, immunity factors, the effects of this disease can vary from moderate level to organ failure.
Our goal is to use Deep Learning to create a hybrid framework for early diagnosis of pneumonia using chest X-ray images. Pneumonia is a dangerous disease that can affect one or both lungs and is typically caused by viruses, fungi, or bacteria. Our goal is to use Deep Learning to create a hybrid algorithm for diagnosis of pneumonia using chest X-ray images. The x-rays we have will help us detect this lung disease. The Chest X-rays dataset is from Kaggle and contains various x-ray images divided into two categories: "Pneumonia" and "Normal." We will be developing a Hybrid deep learning model that will tell us whether the person has pneumonia or not. The model we are proposing achieves an accuracy of 99.47. It is better performing model compared with existing models.

2. Related Work

Methods for Detecting Pneumonia Using Computer-Aided Diagnosis (IJETER July 2021). With latest improvements in computer based technologies, we use computer based diagnosis to detect the disease. The article investigates techniques used to detect the disease. The major intention of this article is to study the present work, issues left and future issues in this area. This disease leads the cause of deaths especially in aged people and kids below 5 years across the world. As per the statistics of the WHO, "it consistently kills around 1.4 million children under the age of five," but this is a highly treatable disease.

The aim of our work [2] is to develop a hybrid model for detecting and classifying pneumonia using chest X-rays in the following ways: First hybrid model is obtained by combining VGG16, VGG19 and Convolution neural network (CNN). Second hybrid model is obtained by combining VGG16, VGG19 and Mobile Net models.

This article [4] is about Convolution networks are at the heart of the majority of cutting-edge computer vision techniques for a wide range of problems. Here, they're looking into ways to scale up networks while utilizing the extra computation as efficiently as possible through appropriately factorized convolutions and aggressive regularization.

Image processing has achieved great strides in these years and in turn attracted the minds of the global research fraternity, according to this detailed and comprehensive article [5]. Text analysis is a catch-all term for AI-powered methods that aid in the extraction of meaningful information from unstructured data. The CNN framework has won the popular ILSVRC competition and the innovations they introduced are detailed in this review paper. The article investigates on different CNN flavors such as V1, V2, V3, V4, ResNet, VGG16 etc. Another important practice to follow before making business decisions is text classification. This also investigates on optimizing SVM, LR, Random Forest (RF), and Boosting Regression Tree hyper parameters (BRT). We investigated parameters for fine-tuning machine learning models' performance.

Efficient Convolution Neural Networks for Mobile Vision Applications: Alex Net (Research Gate Dec 2017). Alex Net is a streamlined model which constructs light weight neural networks using depth-wise convolutions. Here we propose two simple parameters to balance latency and accuracy. As per the limitations, the parameters makes the model maker to choose the correct sized model for the application. We do variety of experiments on resource, accuracy and exhibit strong performance on Image Net model compared with respect to other models. The effectiveness of Alex Net is exhibited on many applications including object detection.
Figure 2. Existing Alex Net Model

Deep Learning for Pneumonia Classification from Chest X-Ray Images Throughout COVID19 (Springer Link Jan 2021). The outbreak of the novel corona virus disease (COVID-19) in 2019 has created a problems across the world. Symptoms and physical examination were used to make a diagnosis. Dated March 11, 2020, the WHO announced the disease a pandemic. As on 10 October 2020, the disease has affected nearly 200 nations, as many as 370 lakhs of people affected by this disease and nearly 10 lakh deaths across the world. The standard procedure for detecting COVID-19 disease is RT-PCR). This has number of limitations such as incorrect reports, costly and needs expertise people to test. As the disease spreads rapidly there needs an immediate attention to device a suitable accurate techniques to detect disease accurately.

Transfer learning with VGG-16 and Deep Convolution Neural Network for Image Classification is described in paper [10]. (IJSRP July 2019). Traditionally, data mining and machine learning algorithms have been designed to approach problems in isolation. These algorithms are used to separate the model on a specific feature space and distribution. A model is trained for a specific task using a machine learning algorithm, depending on the business case.

3. Methodology

Here we are proposing two hybrid models for detection of pneumonia. First one is by combining VGG16,Vgg19 and convolution neural network(CNN) model. Second model is obtained by combining VGG16,VGG19 and Mobile Net models.

We have proposed two novel hybrid neural network algorithms for detection of pneumonia. The algorithms are as follows

1. Hybrid (VGG16 + VGG19 + CNN) Architecture
2. Hybrid (VGG16 + VGG19 + MobileNet) Architecture

![Diagram of Hybrid Model](Image)
Data Sets As previously stated, the collection of X-ray images for pneumonia disease. We combined two real data sets to increase the number of samples for our work. The initial dataset was obtained from Kaggle. We obtained a dataset for pneumonia and normal chest X-ray images.

The Convolution properties of images have been extracted to represent information in a way, where CNN model used to understand the properties of the images. First we have performed domain transformation on X-ray images. Every model used to consider patient past data to make predictions out of that data. The hybrid mode proposed here achieves an accuracy of 99.4%. The new algorithms gives better results among existing ones. The image decomposition into respective coefficients is done by high and low pass filters. The approximate coefficients contain most information about input images of four coefficients. These chest radiograph coefficients were used as information in this study. The approximate coefficients of a sample image can be calculated using

$$cA_{i+1}(x,y) = \frac{1}{2} \sum_{k} \sum_{l} d_k d_l cA_{i,(2x+k, 2y+l)}$$
Additional options for training are:

1. **Brightness** - Helps the model adapt to the light while providing images of various light during training.
2. **Shear** - Adjust the shear angle.

![Image Processing Diagram](image_url)

**Figure 5.** Image Processing

Whenever we apply decomposition the image size reduces to half of the original image size. The DT technique gives us the samples of size 128*128 pixels. Now images are fed into Convolution model which extracts the properties of the image. After the above process number of kernels are obtained and comparisons are made to each part of the image in the convolution layer.

To train any CNN Model it requires massive amount of computational power and time resources. Transformational Learning provides techniques to reduce computational power and time. Transformation Learning models are ideal to start with for many jobs. In this paper, we have built 2 hybrid models First one is a combination of VGG16, VGG19 and CNN, the second hybrid model is a combination of VGG16, VGG19 and Mobile Net. The Chest x-ray images were passed to individual algorithms such as VGG16, VGG19, CNN and mobile net for training purposes. These algorithms considers past data for prediction purposes. The first proposed model accords an accuracy of 94.8 percent and second proposed model accords an accuracy of 99.4 percent. The above hybrid algorithms gives better accuracy than individual algorithms. Here we have achieved the goal of improving the accuracy compared to individual algorithms. Using these hybrid models we have classified the images into either normal or pneumonia. To deal with the scarcity of available data, we used the deep transfer learning method. Using the concepts of a widely used image recognition model, a hybrid deep learning model is created.
We proposed two hybrid deep learning models in this section based on variations in the use of CNN VGG16, VGG19, and MobileNet in deep learning layers. Figure 6 depicts the architecture of these hybrid models, and the details are discussed below.

**Input Chest X-Ray Image:** The dataset was initially obtained from Kaggle. We obtained a database of pneumonia and normal chest X-ray images. This dataset consists of 1,345 pneumonia and 1,341 normal images. The images of the chest X-rays were fed into the individual algorithms for training purposes. Parallel Deep Feature Extractors are used in conjunction with various algorithms.

**Convolution Layers:** Convolution layers are helpful for extracting features from images because they concentrate on spatial redundancy through weight sharing. Since we move on in-depth into network the attributes of the images will be very informative and redundancy decreases. It is mainly because of the use of cascaded convolutions and information condensation through sampling. Once the redundancy is decreased we obtain the condensed attributes of the image, which represents the image contents. Output layers are now in charge of mapping this attributes to the appropriate output categories. Because the entire feature vector is required to take appropriate decision, this mapping function no longer requires weight sharing. The convolution layers convolutionize the data before transforming it with the rectified linear unit (ReLU) function. It has three convolution layers, each with nine filters that produce nine feature maps. Each kernel's operation is defined by the following equation:

\[ X^i = \phi(W^i \ast X^{i-1} + \beta^i) \]

**Pooling Layers:** The main outcome of use of pooling layer is to produce characteristics maps, which are very helpful because any minute changes in features are detected by convolution layer the input. The model's invariance to local translation is the capability added by pooling. The average pooling function is defined by the following equation:

\[ X^i = \text{AvgP}(X^{i-1}) \]

**SoftMax:** The softmax function, also known as normalised function., This can be most commonly used for the activation of neural network. The softmax normalizes a vector \( z \) of \( K \) real numbers into a probability distribution with \( K \) probabilities proportional to the exponentials of the input numbers.

\[ \sigma(z)_i = \frac{e^{z_i}}{\sum_{k=1}^{K} e^{z_k}} \text{ for } i = 1, \ldots, K \text{ and } z = (z_1, \ldots, z_K) \in \mathbb{R}^K \]
multi-class classification problems makes use of softmax for activation purpose where class membership on more than two class labels is required.

**Figure 7.** Hybrid Model: Vgg16+Vgg19+MobileNet Architecture

**Flatten:** The flatten function reduces multidimensional input tensors to a single dimension, allowing us to build hybrid neural network algorithms, then effectively send data into each neuron.

Equation: Assuming an NMP tensor is flattened, we take NMP affine combinations of all tensor elements such that for each combination, exactly one unique coefficient is 1 and all other coefficients are zeros. Flattening is a method which converts data to 1-D array and sends as input to the next layer. Combine all pixel data into a single line and connect it to the final layer.

**Figure 8.** Flattening Layer

**Concatenate:** This layer concatenates a collection of inputs. It takes a collection of tensors with the same shape as the concatenation axis as input and produces a single tensor that is the combination of many inputs. The concatenate function generates 62720 layers in the given hybrid architecture. Concatenation is simply a stacking operation. The x1 layer, for example, has 256 channels, as does the x2 layer. The proposed new algorithm, which is used in our work, has complex structures, with less parameters. If we concatenate these two layers channel by channel, the output will have 512 channels. We can perform all available operations on the resulting tensor as we would on any other tensor. Concatenation means keeping all information in the same place.
Because you don't change the data, this is what should be done by default. That, however, is not free. You must remember that, so it may be beneficial to avoid it as much as possible.

Figure 9. Concatenation Feature

**Weighted Classifier:** A weight (Wk) corresponding to each model was estimated in this module of the proposed methodology. Wk is defined as belief in the kth model, with k equal to 5 because this paper used 5 pre-trained models. The weighted sum of all these prediction arrays was computed as follows:

\[ P_1 W_1 + P_2 W_2 + P_3 W_3 + \cdots + P_K W_K = P_f \]

\[ W_1 + W_2 + W_3 + \cdots + W_K = 1 \]

\[ \text{Loss} = \frac{1}{N} \sum_{i=1}^{N} y \times \log(p) + (1 - y) \times \log(1 - p) \]

Because these weights are fixed (i.e., they are not tuning parameters like W), we can minimise a weighted classification cost precisely as we would any other, for example, using a local optimization scheme like gradient descent or Newton's method.

Figure 10. Weighted Classifier

**Dense:** This is the most commonly used neural network layer. It does the following operations as shown below.

**Equation:**

\[ \text{output} = \text{activation} \left( \text{dot} \left( \text{input}, \text{kernel} \right) + \text{bias} \right) \]

where,

- input indicates the input data.
- kernel indicates the weight data.
- bias indicates a biased value used in machine learning to optimize the model.
- activation indicates the activation function.
Figure 11. Dense Layer

**Dropout:** Dropout is one kind of technique where arbitrarily selected neurons are neglected, which are "dropped out" at random. It indicates contribution to downstream neuron activation is detached temporally on the forward pass, and any weight updates are not applied to the neuron during backward pass.

Figure 12. Drop Out Layers

We can generalize the model easily using Dropout and if the probability is set to 25% between two models, the chance of overfitting is very less. This helps in managing the unregularized network easily and this manages any kind of network. Validation accuracy is increased for Dropout and this normalizes the computation accuracy.

**CNN (Convolution Neural Network):** AI has grown significantly in connecting gap between human beings and machines. Convolution neural networks are designed to exercise on un-structured and semi-structured data such as images and videos. These models are good in performing operations on videos and images. This has made the CNN the best technology for computer vision problems.

**VGG16:** VGG16 is a CNN model. It overtake the Alex Net model by making use of huge kernel-sized filters 5 and 11, in the first and second convolution layers with multiple 33 kernel-sized filters one after the other. VGG16 had been training for weeks on NVIDIA Titan Black GPUs.

**VGG19:** The network received fixed size of RGB picture as input, matrix has been shaped to (224,224,3). During preprocessing only mean RGB of every pixel is obtained. We have covered the entire set of images using the kernel of size 3*3 using a stride size of 1 pixel. We are able to maintain image resolution using padding techniques.

**Mobile Net:** It is type of light weighted deep learning model. Mobile Net uses depth wise convolutions, It does the convolution of single channel instead of merging 3 channels and flattens into a worth 2D CNN model. This gives significant improvement in performance because of depth wise convolutions also because of division of features into 2 layers. This module works well for all computer vision problems.
4. Results

**Evaluation Metrics:** The results of our proposed model were evaluated using accuracy, precision, recall, and F1 score:

\[
\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)}
\]

where TP represents true positive, TN represents true negative, FP represents false positive, and FN false negative. Precision and recall were defined; thus,

\[
\text{Precision} = \frac{TP}{(TP+FP)}
\]

\[
\text{Recall} = \frac{TP}{(TP+FN)}
\]

F1 score is defined as

\[
\text{F1 score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}
\]

The cells in the diagonal position in configuration matrix represents TP and TN. Other cells represents FP and FN. Observations are recorded in every cell. From the above observations we can say that misclassification rate is very less for all the proposed hybrid algorithms.

![Figure13.Model Results](image)

**Figure13.Model Results**

**Accuracy Comparison of Different Models:** This can be evident from the below table that, Hybrid neural network algorithms yields good accuracy in classifying the chest X-Ray images and detect whether it's pneumonia or normal with better overall accuracy.

<table>
<thead>
<tr>
<th></th>
<th>CNN</th>
<th>VGG16</th>
<th>VGG19</th>
<th>MobileNet</th>
<th>Hybrid VGG16+VGG19+CNN</th>
<th>Hybrid VGG16+MobileNet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss</td>
<td>14.6</td>
<td>29.18</td>
<td>25.18</td>
<td>29.19</td>
<td>23.18</td>
<td>11.69</td>
</tr>
<tr>
<td>Preci</td>
<td>75.0</td>
<td>81.9</td>
<td>78.23</td>
<td>64.32</td>
<td>91.50</td>
<td>95.92</td>
</tr>
</tbody>
</table>
Table 1. Accuracy of Different Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Recall</th>
<th>Precision</th>
<th>Accuracy</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN</td>
<td>74.9</td>
<td>82.2</td>
<td>80.25</td>
<td>65.7</td>
</tr>
<tr>
<td>VGG16</td>
<td>81.5</td>
<td>91.13</td>
<td>95.53</td>
<td></td>
</tr>
<tr>
<td>VGG19</td>
<td>83.8</td>
<td>94.47</td>
<td>99.47</td>
<td></td>
</tr>
<tr>
<td>MobileNet</td>
<td>79.1</td>
<td>87.7</td>
<td>85.45</td>
<td>73.56</td>
</tr>
<tr>
<td>First Hybrid Algorithm</td>
<td>94.47</td>
<td>99.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second Hybrid Algorithm</td>
<td>99.4</td>
<td>99.51</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Figure 14. Comparison of Different Models Used |

Above table 1 indicates Loss, precision, Recall, Accuracy and AUC for various models used in our work. CNN model achieves an accuracy of 70.4 percent, VGG16 achieves an accuracy of 82.1, VGG19 achieves an accuracy of 83.8 and MobileNet achieves an accuracy of 70.4 percent. The first hybrid algorithm achieved an accuracy of 94.47 and second hybrid algorithm achieved an accuracy of 99.4 percent while doing the classification of chest x-ray images.

5. Conclusion

In this paper, we presented two hybrid neural network algorithms/models. The first hybrid model proposed above is a combination of VGG16, VGG19, and CNN and has achieved an accuracy of 94.47 and second hybrid model proposed, is a combination of VGG16, VGG19, and MobileNet has achieved an accuracy of 99.4 percent while classifying chest x-ray images into either normal or pneumonia. We have taken dataset from kaggle which contains 3883 images of pneumonia patients and 1385 images from normal people.

It can be understood from the tables that the concatenated network performs better in classifying the chest X-Ray images and detect whether it’s pneumonia or normal with better overall accuracy.

References


