PNEUMONIA DETECTION USING DEEP LEARNING AND TRANSFER LEARNING

Chandrahas Reddy Pola, Padhmanaban Mahendran, Satya Rellu, Suraj Salver, Venu Madhav Arikapudi

University of Calgary

ABSTRACT

Pneumonia is an infectious disease which caused due to the viruses, bacteria that causes inflammation in the heart, severe cough, fever chills and trouble breathing. The reasons to cause pneumonia is the inadequate medical conditions, Unhealthy environment, overcrowded etc., Pneumonia is a common disease in most of the developed and underdeveloped countries due to high pollution. The pneumonia leads to swelling of the lungs which leads to breathing infections. The pneumonia should be detected early to improve survival rates. Basically, the pneumonia is detected using the chest X-rays. Sometimes, the chest x-rays are challenging and prone to subjective variability. To overcome this problem, we have developed a computer-based diagnosis system which automatically detects the pneumonia by using the chest images. We designed the deep neural networks to handle the scarcity of available data and designed an assemble of convolutional neural networks [1][2].

1. INTRODUCTION

Pneumonia is an acute infection that is caused by bacteria, viruses and fungus which effects the lungs of a human by causing severe lung inflammations which fills the lungs with infectious fluids. This is caused due to the over pollution, over-crowding, unhygienic and worst environments, and poor medical resources. This has accounted more than 15-20% of deaths in infants and this number is more dominant in underdeveloped and back ward countries. The diagnosis of this decease is mostly done by using the techniques like Radiography (X-Rays), Computer Tomography scan (CT), or Magnetic resonance imaging (MRI). Early diagnosis had played a pivot role in preventing the decease from getting worst. The X-Ray imaging techniques is considered as an in expensive and non-invasive examination of the lungs. In this study, we developed a computer aided diagnosis system which uses the deep transfer learning models for classifying and studying the chest X-ray images and used Grad-Cam to visualize the predictions of the model [4].

2. ABOUT METHODOLOGY AND DEEP LEARNING

Convolutional neural networks (CNNs) model of the Deep learning framework has played a major role in solving many computational problems and mostly solved various image debugging and classification models. The main required factor for this classification model is larger data sets of the problem to work effectively. But, in the medical industry it becomes difficult to gather so much of labelled data to be computed and it is also a time consuming, expensive, and difficult task as we have limited data or information on certain problem under study. We have brought in the use of transfer learning to overcome issue of accuracy of the model dealing with less data, so we use a predeveloped classification model which is trained on larger data sets. We used a predefined datasets of chest x-ray images of high quality and resolution from Kaggle dataset and this data sets are used to train the CNN model and transfer learning model. Figure-1 shows the nominal level of the pneumonia infection in the human lungs x-ray[4].

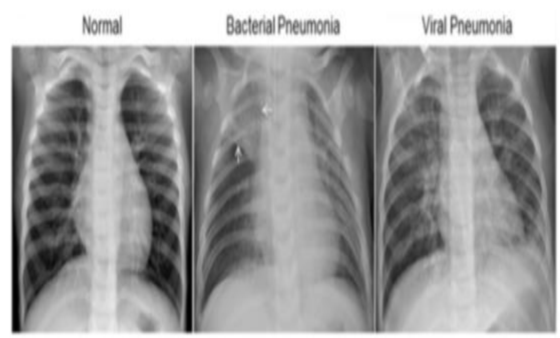


Fig. 1. Chest X-Ray Examples with and without pneumonia.

3. MOTIVATION AND SIGNIFICANCE:

As we discussed earlier, this pneumonia has affected a larger group of humans, most importantly children's age groups in underdeveloped countries with poor living standards and poor medical facilities where many other medical conditions are also predominant. The detection of pneumonia is mainly done by interpretation of the lungs X-ray by a radiologist and medically profound persons. With best in efforts, we developed a CNN model and transfer learning models which are trained to detect the pneumonia condition in provided data sets [5][6].

4. PROPOSED METHODOLOGY

The steps (a) to (d) are same for both the Convolutional neural network and transfer learning model.

a) Dataset:

Kaggle chest X-ray images data set has been used in this project. 5216 images were used for training the model and 500 images were used for testing the model. The dataset has two classes normal and pneumonia. Train and Test label distribution is shown in table 1.

Classes	Training	Validation	Testing
Normal (0)	1241	108	234
Pneumonia (1)	3623	260	390

Table. 1. Dataset class Distribution

Table-1 shows the distribution of images belonging to each class. The number of images belonging to class 1 are very high when compared to other classes. Hence the data set is imbalanced. To address the problem of data imbalance, class weights are defined for each class. By assigning appropriate weights to both the classes based on their class distribution, we were able to mitigate the problem of data imbalance to a certain extent and achieved better results. Without assigning class weights during the model training, the model was unable to predict images of normal (0) class.

b) Data Pre-processing:

Data pre-processing is done in a series of steps as mentioned below.
- The images used for training the model were loaded using flow from directory function. This function combines the image labels (i.e., normal and Pneumonia) present inside the dataset directory and labels them as 0 and 1 respectively.
- All the images are loaded with a target size of (150,150) and are converted to greyscale using the image data generator.
- The loaded images are normalized to 0 - 255 pixels using standardization
- Similar steps were followed for the images used for testing and validating the model.

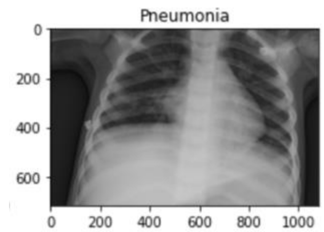


Fig. 1a. Before processing

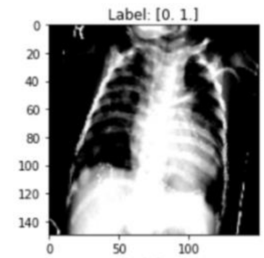
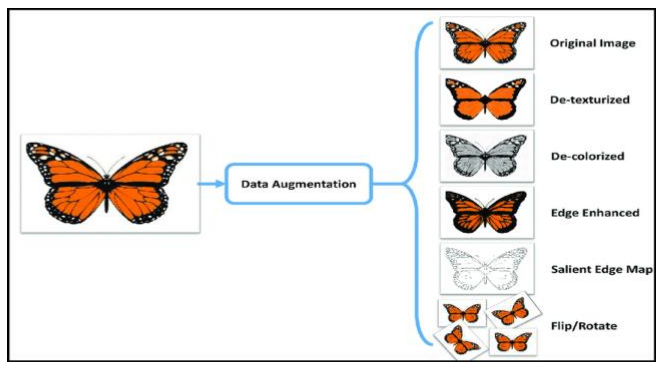


Fig. 1b. After processing

c) Data Augmentation:

Data augmentation is a technique to increase the diversity of dataset without an effort to collect any more real data but still help improve your model accuracy and prevent the model from overfitting.
We used Keras data augmentation method of image data generator. There are a few parameters we can set for data augmentation, the parameters used in our project are rotation, zoom, width shift, height shift, sheer shift, and horizontal flip. Data Augmentation is done on training, validation, and test data.



Reference@Media.com

Fig. 2. Data Augmentation Example

d) Grad-CAM:

Gradient-Weighed Class Activation Mapping (Grad-CAM) uses the gradients of any target concept, and a coarse localization map is produced by flowing into the final convolution layer highlighting the important regions in the image. This technique produces visual explanations for the decisions made by a large class of CNN-based models which makes them more transparent [7].
Grad-CAM retrieves the final convolution layer and the gradient of the model output with the final convolution layer is calculated, and they are pooled. The pooled gradient is iterated for the number of filters in the final convolution layer and the mean of it produces the heatmap which are standardized and plotted as the Grad-Cam.

4.1. Convolutional Neural Network

The Network used in our project is a Deep Convolutional Neural Network which is capable to discover subtle local features without human intervention for manual annotation of individual X-ray images. The network was trained on classification task which gives a continuous output ranging between 0-1. This score represents the classifiers confidence in detecting presence of pneumonia in the images.
The model is built from scratch and the layers are finalized after testing the model with different convolutional layers, max pool layers, dense layers and drop out layers. The final Convolutional Neural Network architecture consists of one input and output layers, 4 convolutional layers with filters 32,32,64 and 128 respectively, two max pool layers and two drop out layers. The output layer consists of 2 neurons with sigmoid as activation layer. Figure-3 shows the developed convolutional neural network model architecture.

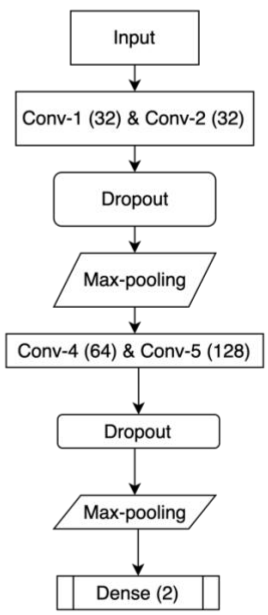


Fig. 3. CNN architecture

4.2. Transfer Learning

Transfer learning is one of the popular approaches in deep learning where a model developed for one task is reused as a starting point for another similar task. For example, a model trained to recognize passenger flights can also be used to predict the fighter jets. The ImageNet challenge has numerous images with many classes and many deep neural network architectures are trained upon the huge dataset. Those pre-trained models can be used for other problems.
In our project, we have used seven architectures namely VGG16, VGG19, DenseNet201, DenseNet169 and Inception. Using transfer learning, we were able to load and use the weights of the above-mentioned architectures and train the model. The models are loaded with the top layers since we need to use the Grad-CAM to visualize the last convolutional network. Due to the inclusion of the top layer, the input image target size has been changed based on the model architecture. The input target size has been changed to (224,224) for VGG-16, VGG-19, DenseNet169 and DenseNet201 and (299,299) for Inception.

5. RESULTS AND DISCUSSION

5.1. Metrics Evaluation:

The accuracies for the test data are computed for all the models and table-2 displays the same. The VGG-19 model has the highest accuracy when compared with other models and CNN model has the second highest accuracy.

S. No	Model Architecture	Accuracy (%)
1.	CNN model	85.2
2.	VGG-16	83.16
3.	VGG-19	87.6
4.	DenseNet201	77.8
5.	DenseNet169	78.5
6.	Inception	68.16

Table.2. Accuracy of Models

5.2. Train and Validation data metric analysis:

The Train and Validation accuracy increases from the starting epochs, and it stabilizes around 98% and 94% respectively. Whereas the train and validation loss decrease from the initial epoch and reads the same at 0.04 and 0.2 respectively in the case of CNN model. Likewise, when the accuracy curves are in the incrementing stage and the loss curves are in the decrementing stage, the model is said to be learning from the features. This curve also determines the stages of overfitting, underfitting and good fit of a model.
The transfer learning model also exhibits the similar behavior.
Below are the loss and accuracy curves for the Train and Validation data of CNN and transfer learning models.

5.1.2.CNN:

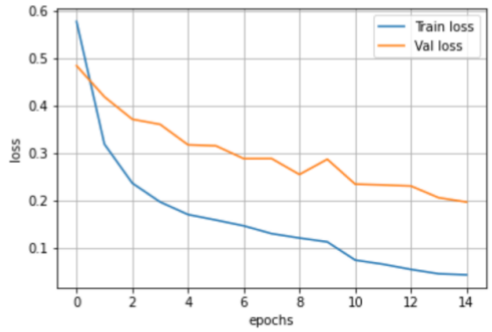


Fig. 4a. Train and Validation loss of CNN model

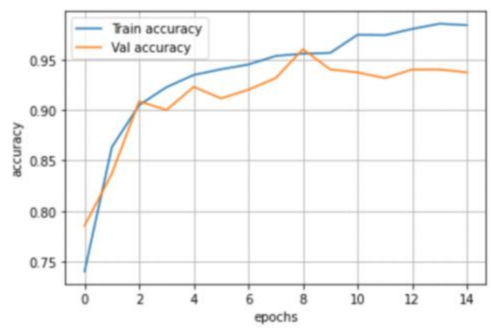


Fig. 4b. Train and Validation Accuracy of CNN model

5.1.2.VGG-16:

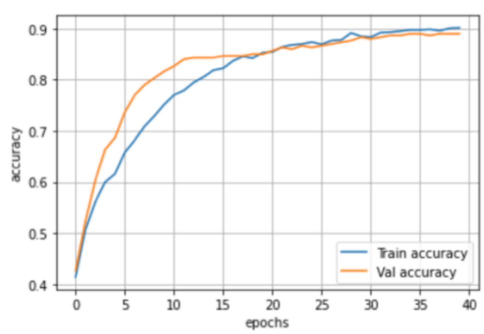


Fig. 5a. Train and Validation loss of VGG-16 model

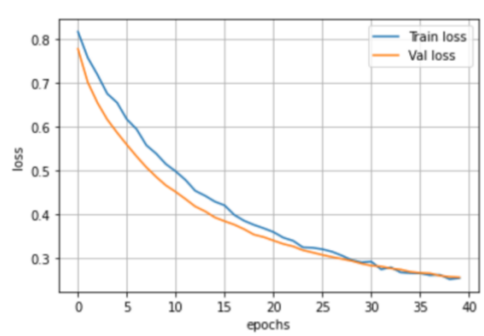


Fig. 5b. Train and Validation Accuracy of VGG-16 model

5.1.2. DenseNet201:

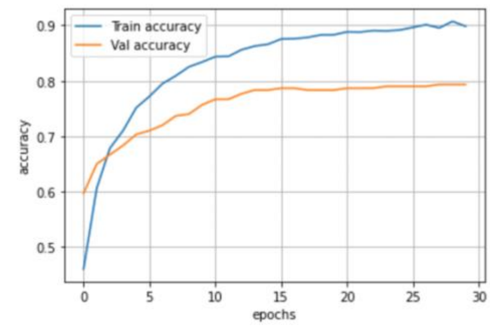


Fig-6a: Train and Validation accuracy of DenseNet201

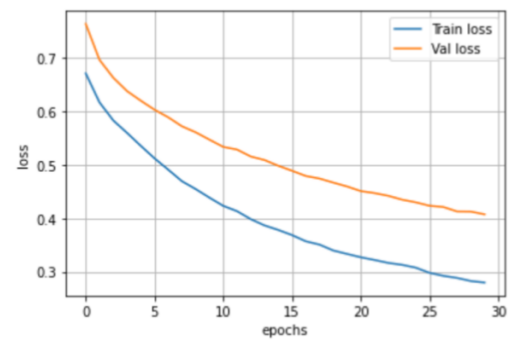


Fig-6b: Train and Validation loss of DenseNet201 model

5.1.2. DenseNet169:

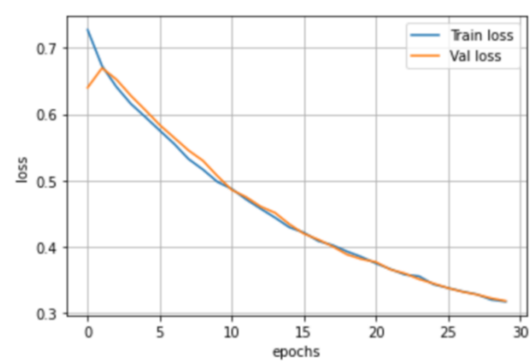


Fig. 7a. Train and Validation loss of DenseNet169

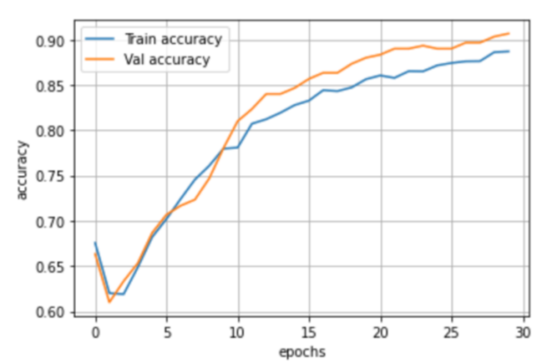


Fig. 7b. Train and Validation accuracy of DenseNet169

5.1.2.VGG-19:

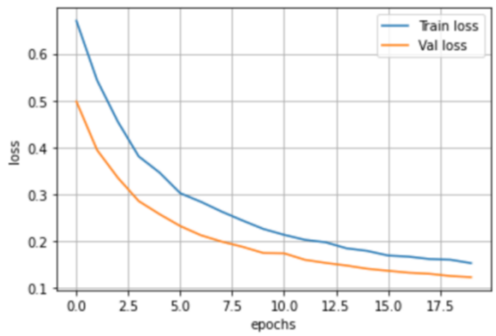


Fig. 8a. Train and Validation loss of VGG-19 model

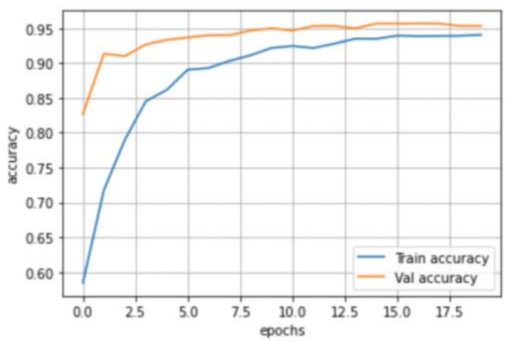


Fig. 8b. Train and Validation accuracy of VGG-19 model

5.1.2. Inception:

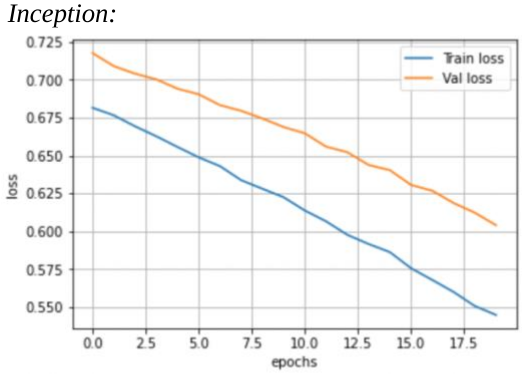


Fig. 9a. Train and Validation loss of Inception model

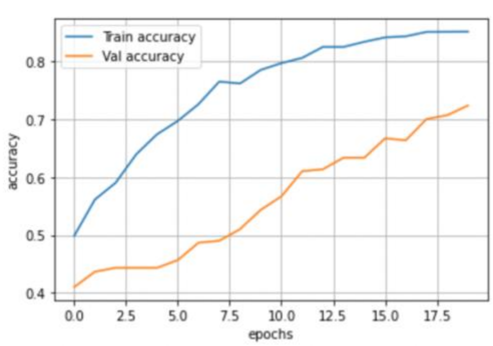


Fig. 9b. Train and Validation accuracy of Inception model

5.3. Classification Report:

The classification report is run based on the prediction of the model. The prediction returns the class which has highest score among the two classes.
The precision provides the percentage of positive predictions which is given by the ratio of true positive to the sum of a True and False Positives. This value must be high for both the classes.
Recall is the fraction of true positives to the sum of true positives and the false negatives.
F1-score is calculated using the precision and recall. It is used to compute the classifier models. The best score is considered as 1 and the worst score is 0.
Support is the number of instances of each class.

Model	Class	Precision	Recall	f1-score	support
CNN	0	0.80	0.75	0.77	178
	1	0.86	0.91	0.89	322
VGG-16	0	0.82	0.70	0.75	225
	1	0.82	0.96	0.88	375
DenseNet 201	0	0.68	0.45	0.54	225
	1	0.91	0.71	0.80	375
DenseNet 169	0	0.73	0.66	0.69	221
	1	0.73	0.95	0.82	379
VGG-19	0	0.96	0.47	0.63	225
	1	0.91	0.94	0.93	375
Inception	0	0.37	1.00	0.54	221
	1	0.64	1.00	0.78	379

Table. 3. Classification report of models

5.4. Confusion Matrix:

The confusion matrix provides the following ways to know the accuracy of the predictions.
1.TN / True Negative: the case was negative and predicted negative
2.TP / True Positive: the case was positive and predicted positive
3.FN / False Negative: the case was positive but predicted negative
4.FP / False Positive: the case was negative but predicted positive.

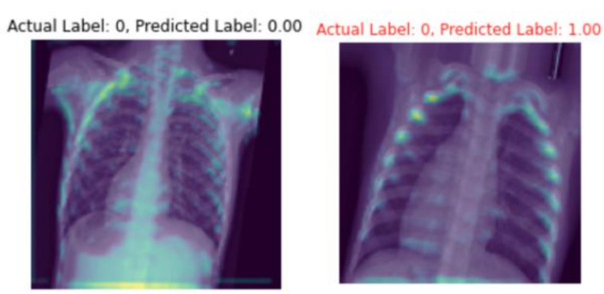
Models	Positive	Negative	
CNN	133(TP)	45(FP)	Positive
	29(FN)	293(TN)	Negative
VGG-16	147(TP)	78(FP)	Positive
	23(FN)	352(TN)	Negative
DenseNet201	162(TP)	63(FP)	Positive
	70(FN)	305(TN)	Negative
DenseNet169	108(TP)	113(FP)	Positive
	16(FN)	363(TN)	Negative
VGG-19	160(TP)	65(FP)	Positive
	9(FN)	366(TN)	Negative
Inception	164(TP)	57(FP)	Positive
	134(FN)	245(TN)	Negative

Table. 4. Confusion Matrix of models

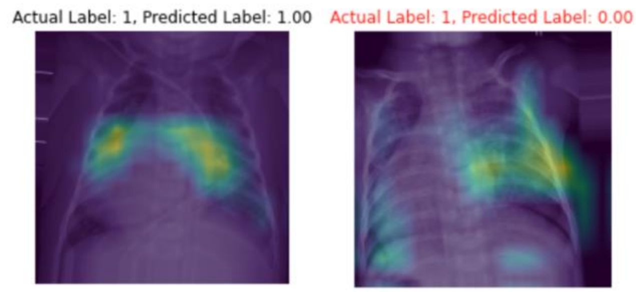
5.4. Heatmap:

The heatmap visualizes the sections which are used for the prediction. Following pair of images contains samples which are predicted as per the label and wrongly predicted. Incorrectly predicted images have the bright spot focused on the parts apart from the white infiltrates which is usually an indication of pneumonia.The correctly predicted images have the bright spots focused on the pneumonia infected areas predicting the samples whether it has pneumonia.

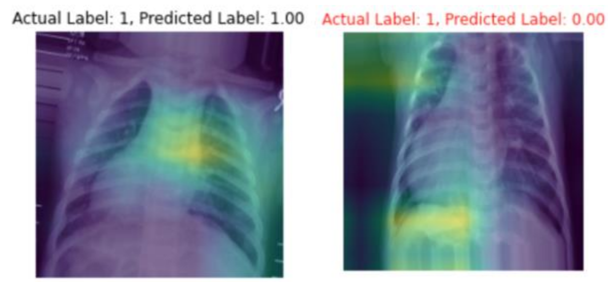
5.4.1.CNN:



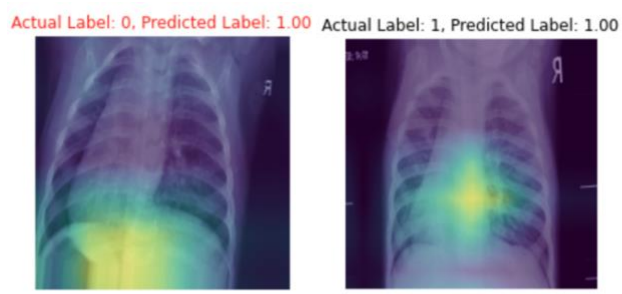
5.4.2. VGG-16:



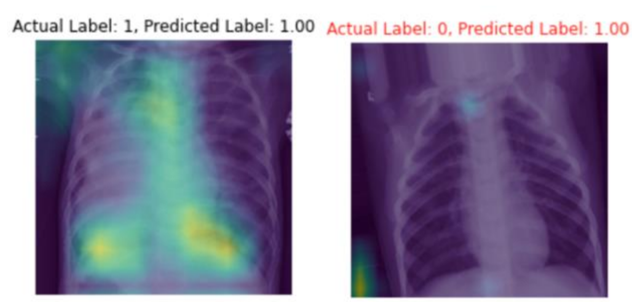
5.4.3. DenseNet201:



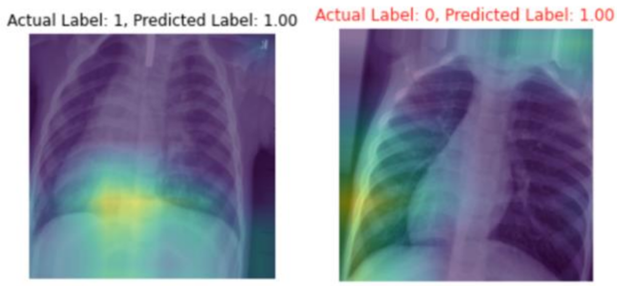
5.4.4. DenseNet169:



5.4.5. VGG-19:



5.4.6. Inception:



5.5. Importance of Data Augmentation:

To test the need of Data augmentation, CNN models and the transfer learning models are trained without the data augmentation process and found out that the accuracy of CNN model dropped by 5-7% and the accuracies of all the transfer learning model dropped by 3-5%. Thus, it was concluded that the Data augmentation procedure increases the accuracy of the model by modifying the input data for training.

6. CONCLUSION AND FUTURE WORK

To reduce the adverse effects of pneumonia, an early detection of it is important, so that the person receives the treatment as early as possible and reduce the deaths caused due to it. For pneumonia detection the chest radiography is used mostly, and medically trained people would be detecting the existence of pneumonia by prediction. The CNN and some of the transfer learning models have an accuracy rate of more than 80%. The accuracy of the model can be increased by increasing the size of the dataset. As we could see that there is still scope for developing a more accurate framework and cost-effective frameworks in future.

7. REFERENCES

[1] Kundu R, Das R, Geem ZW, Han G-T, Sarkar R (2021) "Pneumonia detection in chest X-ray images using an ensemble of deep learning models". PLoS ONE 16(9): e0256630. https://doi.org/10.1371/journal.pone.0256630

[2] WHO Pneumonia. *World Healh Organization*. (2019), https://www.who.int/news-room/fact-sheets/pneumonia

[3] Albahli S., Rauf H., Algosaibi A. & Balas V. "AI-driven deep CNN approach for multi-label pathology classification using chest X-Rays". PeerJ Computer Science. 7 pp. e495 (2021) pmid:33977135

[4] Jaiswal A., Tiwari P., Kumar S., Gupta D., Khanna A. & Rodrigues J. "Identifying pneumonia in chest X-rays: A deep learning approach". Measurement. 145 pp. 511–518 (2019)

[5] Gabruseva T., Poplavskiy D. & Kalinin A. "Deep learning for automatic pneumonia detection". Proceedings Of The IEEE/CVF Conference on Computer Vision And Pattern Recognition Workshops. pp. 350-351 (2020)

[6] Kadry S., Nam Y., Rauf H., Rajinikanth V. & Lawal I. "Automated Detection of Brain Abnormality using Deep-Learning-Scheme: A Study". 2021 Seventh International Conference on Bio Signals, Images, And Instrumentation (ICBSII). pp. 1-5 (2021)

[7] Selvaraju R., Cogswell M., Das A., Vedantam R., Parikh D. & Batra D. "Grad-cam: Visual explanations from deep networks via gradient-based localization". Proceedings Of the IEEE International Conference on Computer Vision. pp. 618-626 (2017)

[8] Mahmud T., Rahman M. & Fattah S. "CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization". Computers In Biology and Medicine. 122 pp. 103869 (2020) pmid:32658740

[9] He K., Zhang X., Ren S. & Sun J. "Deep residual learning for image recognition". Proceedings Of the IEEE Conference on Computer Vision and Pattern Recognition.

[10] Yue H., Yu Q., Liu C., Huang Y., Jiang Z., Shao C., et al. & Others "Machine learning-based CT radiomics method for predicting hospital stay in patients with pneumonia associated with SARS-CoV-2 infection: a multicenter study". Annals Of Translational Medicine. 8 (2020) pmid:32793703

[11] Ranjan, A., Kumar, C., Gupta, R.K., Misra, R. (2022). "Transfer Learning Based Approach for Pneumonia Detection Using Customized VGG16 Deep Learning Model". In: Misra, R., Kesswani, N., Rajarajan, M., Veeravalli, B., Patel, A. (eds) Internet of Things and Connected Technologies. ICIoTCT 2021. Lecture Notes in Networks and Systems, vol 340. Springer, Cham. https://doi-org.ezproxy.lib.ucalgary.ca/10.1007/978-3-030-94507-7_2

[12] Fayemiwo, M.A. et al. (2022). "Comparative Study and Detection of COVID-19 and Related Viral Pneumonia Using Fine-Tuned Deep Transfer Learning". In: Mehta, M., Fournier-Viger, P., Patel, M., Lin, J.CW. (eds) Tracking and Preventing Diseases with Artificial Intelligence.Intelligent Systems Reference Library, vol 206.Springer.https://doi-org.ezproxy.lib.ucalgary.ca/10.1007/978-3-030-76732-7_2

[13] Sirazitdinov I et al (2019) "Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database." Comput Electr Eng 78:388–399

[14] Bhandary A et al (2020) "Deep-learning framework to detect lung abnormality–A study with chest X-Ray and lung CT scan images". Pattern Recogn Lett 129:271–278

Link to Code GitHub Repository:

https://github.com/chandu-55/ENEL645.git

Work done Table:

S. No	Name of the Student	UCID	Contribution	Score
1.	Chandrahas Reddy Pola	30140068	Effectively worked in group activities, idea sharing, project development and gathered data required for the project and was helped everyone work together as a team.	3
2.	Padhmanaban Mahendran	30117430	Contributed to making the project work and did logic breaking tasks and contributed greatly to brain storming sessions, documentation and supported all his teammates and helped his teammates understood when struck.	3
3.	Satya Rellu	30143162	Worked collectively on various portions of the project, brain storming sessions, and collectively sharing ideas and knowledge for making things work when going out of bound. Also contributed towards making a better documentation for the project.	3
4.	Suraj Salver	30124245	Worked as a team for making the project and helped in forming the crust to the idea. Provided write-up for project presentation and report.	3
5.	Venu Madhav Arikapudi	30147578	Collectively worked in understanding the essence of the idea behind the project and worked towards it. Helped in forming the notes for documentation purpose and worked for making united decisions.	3