Application of Neural Networks in Pneumonia Detection

Guddu Kumar Shah
Faculty of Engineering,
Environment and Computing
Coventry University
Coventry, West Midlands,
United Kingdom
shahg4@uni.coventry.ac.uk

Abstract*— Pneumonia at present is one very commonly found chest infection in the world. It is basically the inflammation of the air sacs in the lungs which gets filled with fluids making people hard to breathe. In the UK, out of 1000 every 5-10 adults are likely to get affected by pneumonia. Babies, aged individuals above 65 and people having bad health history are at high risk of this chronic disease. Pneumonia in patients can be diagnosed through chest X-rays, where the images of internal tissues, organs and bones along with lungs are captured. However, it becomes quite challenging for radiologists to differentiate between the X-ray which is normal or X-ray having pneumonia because of either the disease is in its initial stage or part of the lungs not seen easily. With the increase of computer-aided diagnosis system, the pneumonia diagnosis accuracy has improved. This would facilitate the treatment of the disease by providing easy access to diagnosis to a broader range of the population. In this work, I will develop a computational model to detect pneumonia causing factors using the convolutional neural network to classify the pneumonia in individuals using chest X-ray images dataset from Kaggle. The model achieved an accuracy of 80% using Convolutional Neural which shows a great and effective approach for pneumonia detection.

Keywords*—Convolutional Neural Network, Image Classification, Chest X-ray (Pneumonia)

I. INTRODUCION

Pneumonia is currently the most common respiratory infection affecting millions of people in all parts of the world. Pneumonia is basically the inflammation in the lungs caused by the bacterial of virus infection and the basic symptoms for this contagious disease are coughing, difficulty in breathing increase in heartbeat and body temperature, chest pain and muscle pain (NHS, 2019). According to the WHO, pneumonia is the most infectious death causing disease in children globally accounting 14% of children under 5 years and 22% of death in children within the age group 1 to 5 (World Health Organization, 2021). This disease is also highly prone to people over 65 and having past health history of cancer, diabetes, smoking, alcohol addiction, HIV and drugs. In the UK, the Springer Link reports that over twenty hundred thousand patients dies every year diagnosed with pneumococcal disease and approximately 15% of the hospitalized patients dies within a month after admission (Chalmers et al., 2017). This disease can be controlled at it's early stage of development with medically prescribed drugs and antibiotics. . In this early stage of medication it reduces the future risks to mortality. Therefore, early medication is extremely important to control latter complications leading to death of patients. Chest X-ray is very popularly known method for diagnosing this disease (World Health Organization, 2001). Though it's the popularly known clinical method, pneumonia chechup from X-ray become a challenging task for experienced radiologists. This can be because of it's appearance, X-ray can be confusing, unclear and show the symptoms of other respiratory infections. This caused the variabilities in the decision makings of expert radiologists for diagnosis of pneumonia in children (Neuman et al., 2012).

There has been several developments in the computerized aided diagnosis system in the medical industry to help radiologists to diagnose pneumonia through chest X-rays. The study presented by Kosh et al. (2021) clearly explains one of the first developments in the deep learning industry using Convolutional Neural Network for checking pneumonia through image classification . This development showed a great improvement in pneumonia detection in patients. The main objective of using CNN in deep learning was to develop a neural networks model that could learn through backpropagation algorithm by making internal modifications in the parameters in computers making it to behave as human does and extract the most significant features from the entire image (Kosh et al., 2021). Researchers and scientists have developed deep neural networks using CNN which were capable to make object detection, localization and classification using computer vision system. Apart from computer vision recognition CNN has also achieved a great breakthrough in the medical industry in recognition of brain tumor MRI detection and breast cancer detection One of the main contributors to models improvement is using DenseNet-121 layer with transfer learning method (Antin et al., 2017). The 121-layer CNN on DesneNet also known as ChestNet trained the network using more than tens of thousands of chest X-rays with 14 diseases labels and ended achieving f1 score higher than the average of 4 radiologists.

For image classification tasks, Convolutional Neural Networks (CNN) is the mostly used neural networks. CNN has shown a remarkable improvement in image recognition in the past few years. In this work, I am using the CNN based neural network to perform classification on the X-ray images for pneumonia detection by performing the model building tasks on the Chest X-rays dataset present at Kaggle. The model gave the accuracy of 81% with precision and f1 score of 78% and 86% The further analysis work is described as follows. Section defines the dataset in more details, Section III defines the methodologies used. Experiments and results are demonstrated in Section IV and V. Section VI is social, ethical considerations and Section VII is Discussion and conclusion.

II. DATASET DESCRIPTION AND METHODOLOGY

A. DATASET DESCRIPTION

The dataset is composed of 5863 X-ray pictures (JPEG) separately kept in train, test and val directories with subdirectories having pictures NORMAL and PNEUMONIA. These data were extracted from the pediatric patients from Guangzhou Women and Children's Medical Center, Guangzhou as a for routinely check. There are in total 1583 normal images and 4273 pneumonia images in this dataset. A brief description of each category distributed in training, testing and validating phases shown in *Table 1*. Here normal image is represented as 0 and pneumonia image is represented as 1. The sample of Normal case and Pneumonia case is presented in *Figure 1*and *Figure 2*

Table 1: X-rays images data

Category	**Train**	**Test**	**Val**
Normal	1341	234	8
Pneumonia	3875	390	8
Total	5216	624	16

Figure 1: X-ray image of Normal case

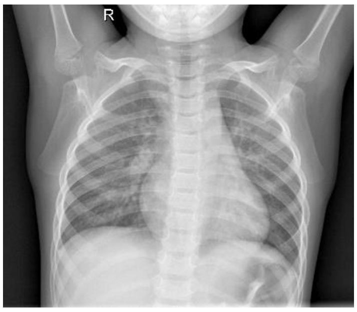
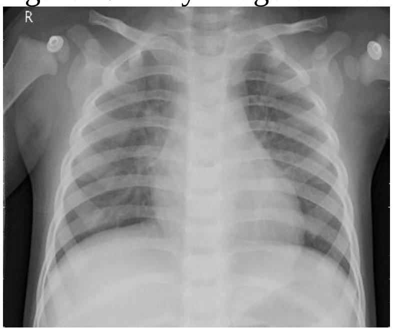


Figure 2: X-ray image of Pneumonia case



B. DATA AUGMENTATION

Deep learning requires large number of data to obtain great results. The existing data is not sufficient enough to train the deep neural network. In solving of medical problems it is very costly in obtain large the data. This becomes very difficult to build a powerful classifier. However, the solution to this existing problem is data augmentation. Data augmentation basically avoids the overfitting in the model and increases the accuracy (Ayan & Unver, 2018). From Bissoto et al. (2018) and Perez et al. (2018) the studies clearly suggested that for reducing the overfitting the best strategies during data augmentation are vertical and horizontal flip, shear, scaling and random rotation. Both the studies also stated using color variations and brightness will drastically reduce the accuracy. In this work, Keras ImageGenerator API is used to perform image augmentation. This is automatically label the label the images data in the normal folder and the pneumonia folder with their corresponding names. Several augmentation methods such as zooming, shearing, random at 40-degree angles, shifts and horizontal flips were used for creating random images for building cnn classifier.. Similarly, this operation was performed in the test folder and val folder.

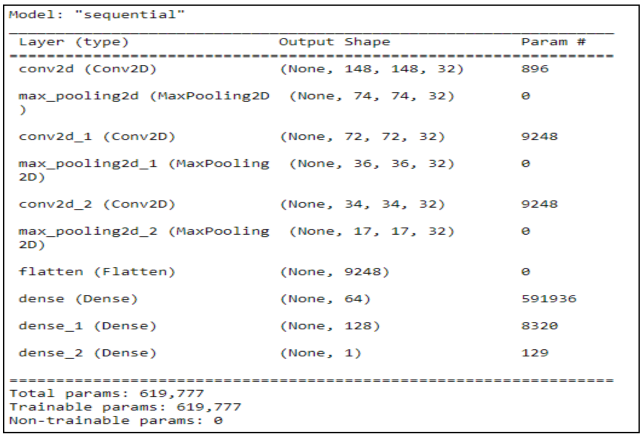
C. CONVOLUTIONAL NEURAL NETWORK ARCHITECTURE

The Convolutional Neural Network is a specialised type of neural network, ideal for data that can be represented as a grid. CNN is mostly used for image recognition tasks since this input can be transcribed as a 2D grid of pixels. The CNN are neural networks that use at least one of their layers the convolution operation. CNN's can handle larger inputs due to their unique characteristics. The convolution kernel is smaller than the input size, meaning that many input units interact with the kernel at the same time. Common NN's have a one by one interaction, which results in thousands or millions of interactions if we consider an image processing task. CNN's also makes use of parameter sharing, i.e. instead of learning the weights between every element of the network, it learns a set of parameters that can be used across all input nodes.

The next step in the process is the pooling where the features extracted from the convolution layer transformed into the features maps get reduced dimensionally. The output from pooling layer is shrinked version of the input. The detailed information regarding pooling is explained in (Kosh et al., 2021). The series of convolution and pooling is applied on the input data until it is reduced to the size being efficient in cost of processing for the machine.

In the work, the CNN architectures basically has convolutional layer with ReLU(helps in making the model non-linear), pooling layer, fully connected layer and a loss layer. In the first step, the Sequential model is initialized. The CNN model has a convolutional layer which is also the core block applies filter on the input data creating a feature map as explained in. The feature map summarises the detected features in the input during feature extraction. As the input data progress to become smaller the network gets deeper with the feature map. Here, I added one Conv2D layer with the kernel size 3x3 along with ReLU activation function to set the negative values in the feature map as 0 and initializing the input_shape to be (150,150,3). The pooling layer in placed between the two adjacent convolutional layers for the dimensionality reduction of the convoluted features over the network. Maxpooling is the mostly used method for size reduction pooling the most significant features. A simple explanation on the differentiation between different types of pooling techniques in Suárez-Paniagua & Segura-Bedmar (2018) concludes Maxpooling performs the best amongst other techniques. Maxpool2D with pool size of 2x2 is used for pooling. The convolutional and pooling operations are repeated further 3 consecutive times. After that the Flatten layer is introduced to convert the image data into 1D array. The flatten data is passed over to the two dense layers and an output dense layer with one neuron using the activation function sigmoid for classification. The final output gives the image data belongs to class 0 (Normal) or class 1(Pneumonia). Since the validation data is significantly smaller in comparison to the train set, I will set validation_data = test_generator for testing the valid training. As there are 624 image data in the test set, validation step = 624/32, where 32 is the batch size of the data. *Figure 3* represents the summary of CNN Model and *Figure 4* represents the CNN Architecture.

Figure 3: Summary of CNN Model

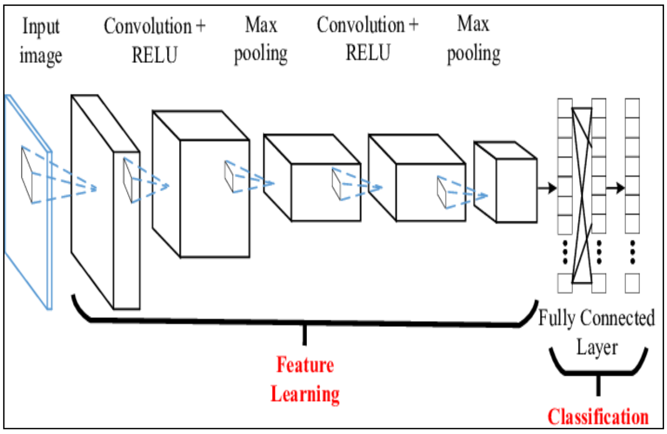


Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 148, 148, 32)	896
max_pooling2d (MaxPooling2D)	(None, 74, 74, 32)	0
conv2d_1 (Conv2D)	(None, 72, 72, 32)	9248
max_pooling2d_1 (MaxPooling2D)	(None, 36, 36, 32)	0
conv2d_2 (Conv2D)	(None, 34, 34, 32)	9248
max_pooling2d_2 (MaxPooling2D)	(None, 17, 17, 32)	0
flatten (Flatten)	(None, 9248)	0
dense (Dense)	(None, 64)	591936
dense_1 (Dense)	(None, 128)	8320
dense_2 (Dense)	(None, 1)	129

Total params: 619,777
Trainable params: 619,777
Non-trainable params: 0

Figure 4: CNN Architecture



III. EXPERIMENTAL SETUP

The experiment was performed using Python programming language along with importing of scientific libraries that contains predefined functions. The code is tested by using Jupyter Notebook (interactive computer environment) that supports IPython command shell for Python as well as other programming languages like R. The project runs Python 3.10.5 on Jupyter Notebook in version 6.4.8 installed on Windows 10 – 64bit operating system (OS). The whole package is installed by downloading Anaconda Distribution framework version 4.12.0. Prior to launching the experimental environment, the OS required the installation of additional libraries for running the experiments. The main part of the project (classification algorithms) uses a scikit-learn library that provides required functions for Python programming language. Libraries like tensorflow and keras (runs on the top of tensorflow) API's are also imported to perform deep learning tasks in this project. The chest X-ray dataset is loaded into the notebook using os module (used for creating, deleting, changing and accessing directory on the top of operating system) in python as presented in *Appendix A.*

The pre-processing operation was started with fetching all the directories of the image data from the os. The images in the dataset present with hazy patches also known as ground-glass opacity resembles pneumonia in patients and the other resembles normal patients. During this pre-processing stage it was found that there was a class imbalance problem showing pneumonia cases thrice the normal cases present at *Appendix B*. This became an advantage as the CNN model will perform better for the major class pneumonia. The FP (Normal predicted as Pneumonia) will tend to have higher data which will help in building a robust model rather than FN (Pneumonia predicted as Normal) which would then create a big overhead. Therefore, this type of positive class imbalance favors the model accuracy. After the pre-processing and organization of image data, data augmentation is performed on the images present in train, test and validation set by creating images through multiple processing like rotation, shear, flips to remove the condition of overfitting and boost the performance of the deep neural networks and building a robust classifier by using few training data. The ImageDataGenerator() will help in labelling the images created through multiple processing. Since the validation set has only 16 image data, this is significantly small in from of 5216 iamge data. I decided to use the training as a validate set. After data augmentation the input image for training is resized to (150, 150) with batch size of 32. Here, the class_mode is set to binary since I am using the binary_crossentropy loss. After this step, an open-source library called cv2 is imported for performing the operations of dividing the image data in normal and pneumonia sets along with labelling it as present in the code snippet below in *Appendix A.*

Image data after the pre-processing and augmentation is ready building deep learning model, CNN is one of the most popular deep learning techniques in healthcare for image recognition because of it's ability to identify and capture the most significant features from the data. The input data (150, 150,3) is sent over the CNN architectures comparised of convolutional layer, pooling layer, fully connected layer and output. At this stage, the convolutional layer apply filter to the input and create a feature map. Feature map reduces the image dimension by extracting the most intact features. A Conv2D layer with kernel size (3x3) and ReLU activation function was added to the model architecture along with batch size of 32, the output size was (148,148,32).

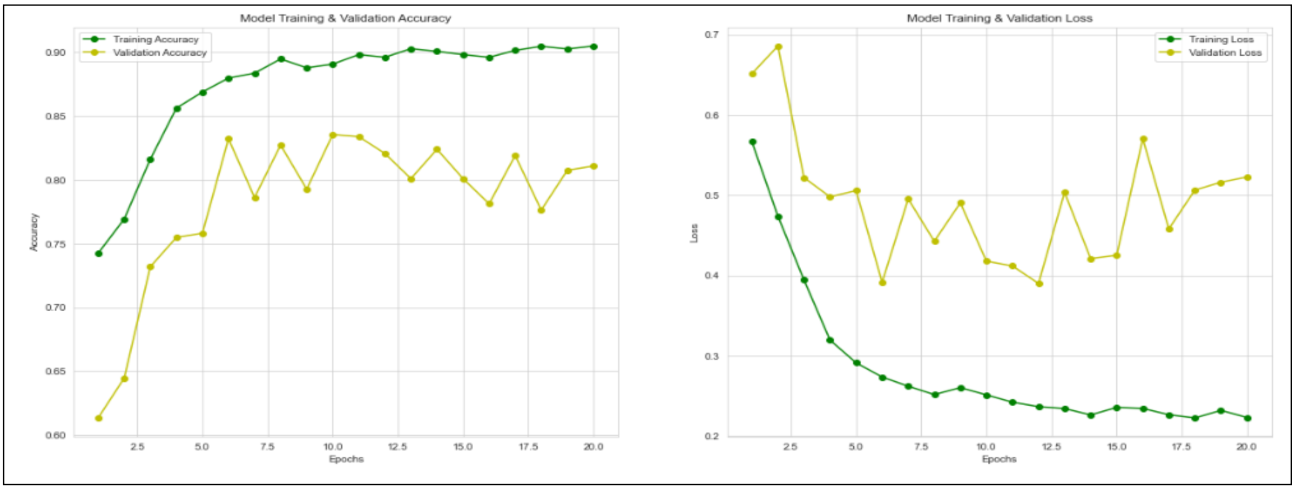
After that I used MaxPool2D to reduce the size of (2x2) and got the reduced output (74,74,32). The convolution layer and pooling layer is repeated converting the image size by 3 times till the output shape of the image was (17,17,32). The Flatten layer is added for converting the data into 1D array form of output shape 9248. Furthermore, 3 dense layers were added to the model with 64, 128 and 1 neurons simultaneously with activation function as sigmoid. The model was compiled using Adam optimizer which is used to improve the performance of the neural network. The fitting is applied on the model using the validation_data = test_generator, validation_steps = 624/32 where 32 is the batch size and epochs = 20. After 20 epochs the accuracy of the model was set to 81 %.

The final step is the model evaluation where the cnn_train_acc, cnn_train_loss, cnn_val_acc, cnn_val_loss is plotted and checked for CNN model training and validation accuracy. The confusion matrix showed a great result where TP (Normal predicted as Normal) was 167 and FN (Pneumonia predicted as pneumonia) was 346. This shows the model is robust.

IV. RESULTS

Here, methodologies regarding training and testing results are set. For the model the image data are resized according to their dimensions before the training phase. Because, The image was converted to the dimension of 150x150x3 so that this becomes computationally efficient for the network to handle. During the training of the model, epoch size = 20, categorical cross entropy is set to loss function, Adam is used as the optimizer, learning = 1e-4, and batch size = 32. Data augmentation approach was used for avoiding overfitting. In the foremost step, after every convolutional batch normalization is appliedlayer. *Figure 5* shows the model accuracy and loss plottings. The selected networks are implemented by using Keras framework. In this work, Training time was calculated to be 82 minutes 54 seconds. The estimated time per image is 0.018. The model prediction on the test set shows the accuracy of 81 % according to the classification_report.

Figure 5: Model accuracy and loss plotting



The model was evaluated by making the use of performance metrices such as accuracy, recall, precision and f1-score (Ayan & Unver, 2018). These metrices can be expressed in formulas as

Accuracy = (TP +TN) / (TP + TN + FP + FN)
Recall = TP / (TP + FN)
Precision = TP / (TP + FP)
F1-score = TP / (TP + (FP + FN) / 2)

where TP = True Positive, TN = True Negative, FP = False Positive and FP = False Positive.
The test image set contains 234 normal and 390 pneumonia cases. *Table 2* presents case base precision, recall and f1 scores of CNN network. In addition, *Figure 6*shows the confusion matrix from CNN model. *Figure* 7 shows model pneumonia classification and the summary of all the scores in displayed in the tabular format in *Figure 8*.

Table 2: CNN model precision, recall and f1-score result

	precision	recall	f1-score
Normal	0.88	0.56	0.69
Pneumonia	0.78	0.95	0.86

Figure 6: Confusion Matrix

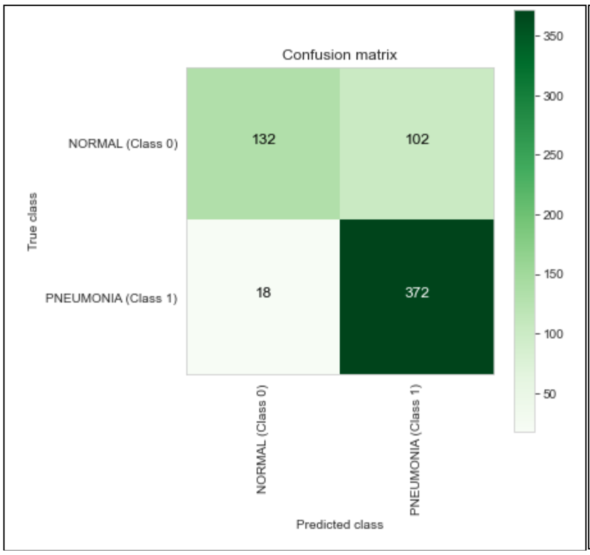


Figure 7: Model pneumonia classification

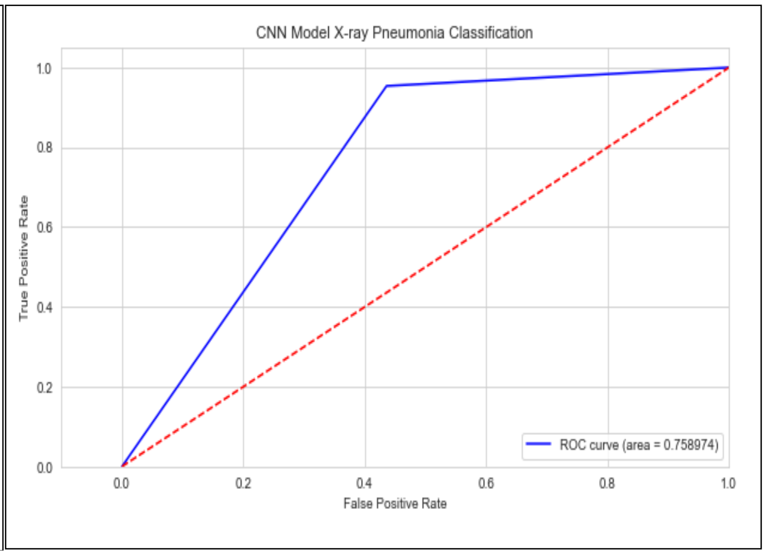


Figure 8: Score Summary Table

	Model	Accuracy	Precision	Recall	F1 Score	AUC
0	Convolutional Neural Network Model	0.81	0.83	0.76	0.77	0.76

V. SOCIAL, ETHICAL AND PROFESSIONAL CONSIDERATIONS

Though this technology has helped expert radiologists to detect pneumonia chest X-ray images, it can also be misused which can be a matter of concern. The accuracy doesn't of the deep learning or neural networks doesn't always helps in accurately classifying the classes. There are two most occurring errors when building a deep learning model i.e., FP (False Positive) and TN (True Negative). The FP can be seen as Normal predicted as Pneumonia, here the neural networks lacks to classify and makes a wrong decision leading to unnecessary medication to the patients The TN represents Pneumonia predicted as Normal, this prediction by the model fails here, which will make the pneumonia to get worse risking the life patients to death as it's predicted as normal. The. The other concern is the hidden biases raising the condition of overfitting where the model becomes biased because of the limited training data. During this condition the model starts wrong prediction decreasing the performance of the model. The third concern would be the algorithm corruption or the computation failure. With the development in the technology, the model should be updated as the functionalities of the old algorithms gets outdated with time, if this not done the data in the algorithm of the deep learning networks starts to get corrupted causing computational failure. This condition might hamper the confidence of the model in future. Although the deep learning shows the 99 % accuracy the model doesn't guarantee the replacement of the doctor instead marketed as a supplementary tool for helping in decision making. This indicates that model doesn't guarantee decision making for the radiologists affecting life or death though it's accuracy is high. The final concern will be professional issue where big organizations might twist the algorithms building new products for the business growth with caring it's consequences over the people.

VI. DISCUSSION AND CONCLUSION

In this work, I just considered using CNN architecture for chest X-ray images. This is because CNN is considered as the best technique in the deep learning for image classification. The data in the dataset were divided into , train, test and validation. Since the images in the dataset were of several dimensions, they were resized to 150x150 before the training for improving the performance. In the data augmentation part a combination of multiple processing is used for generating images where ImageDataGenerator API from keras was used. The batch size is taken as 32. Furthermore, the image separation was performed using cv2 library which is very popularly used in image recognition, obstacle detection. After this the image data is ready for training over the CNN architecture.

In this study, we evaluated the CNN networks performance on the pneumonia examination from X-ray images. In order to avoid overfitting , data augmentation techniques were used. The training is done through CNN model where the input image data passes through 3 Conv2D layers and 3 MaxPooling2D layer. The Adam optimizer and binary-crossentropy are used during the training for enhancing the model performance making it computationally efficient. After the training phase the testing done on the test set where the accuracy of the model was 0.81%, pneumonia precision was 0.78%, recall was 0.95% and f1-score = 0.86%. According to the experiment result and confusion matrix the model developed showed great results in pneumonia detection on the image data.

REFERENCES

- Antin, B., Kravitz, J., & Martayan, E. (2017). Detecting pneumonia in chest X-Rays with supervised learning. *Semanticscholar. org*.
- Ayan, E., & Ünver, H. M. (2018, April). Data augmentation importance for classification of skin lesions via deep learning. In *2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT)* (pp. 1-4). IEEE.
- Bissoto, A., Perez, F., Ribeiro, V., Fornaciali, M., Avila, S., & Valle, E. (2018). Deep-learning ensembles for skin-lesion segmentation, analysis, classification: RECOD titans at ISIC challenge 2018. *arXiv preprint arXiv:1808.08480*.
- Chalmers, J., Campling, J., Ellsbury, G., Hawkey, P. M., Madhava, H., & Slack, M. (2017). Community-acquired pneumonia in the United Kingdom: a call to action. *Pneumonia*, *9*(1), 1-6. 10.1186/s41479-017-0039-9
- Kose, U., Deperlioglu, O., & Hemanth, D. J. (Eds.). (2021). *Deep Learning for Biomedical Applications*. CRC Pres
- Neuman, M. I., Lee, E. Y., Bixby, S., Diperna, S., Hellinger, J., Markowitz, R., ... & Shah, S. S. (2012). Variability in the interpretation of chest radiographs for the diagnosis of pneumonia in children. *Journal of hospital medicine*, *7*(4), 294-298. https://doi.org/10.1002/jhm.955
- NHS. (2019, June 30). *Overview Pneumonia*. https://www.nhs.uk/conditions/pneumonia/
- Perez, F., Vasconcelos, C., Avila, S., & Valle, E. (2018). Data augmentation for skin lesion analysis. In *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis* (pp. 303-311). Springer, Cham.
- Suárez-Paniagua, V., & Segura-Bedmar, I. (2018). Evaluation of pooling operations in convolutional architectures for drug-drug interaction extraction. *BMC bioinformatics*, *19*(8), 39-47.
- World Health Organization. (2001). *Standardization of interpretation of chest radiographs for the diagnosis of pneumonia in children* (No. WHO/V&B/01.35). World Health Organization
- World Health Organization. (2021, November 11). *Pneumonia.* https://www.who.int/news-room/fact-sheets/detail/pneumonia

APPENDICES

A. Code Link:

https://drive.google.com/file/d/157kifXABtmEInyU1oCfrAl0O4sboPNf4/view?usp=sharing

B. Class Imbalance data