PNEUMONIA DETECTION USING CNN

Ankit Kumar
Master of Science, Data Analytics
National College of Ireland
Dublin, Ireland

Abstract – **Pneumonia is a fatal disease which causes 2.5 million deaths each year globally out of which 24% are children. The total time required to provide a chest x-ray results to the patients may vary depending on the availability of the general practitioner. Mostly it can take up to 2 days. The literature review explores various deep learning approaches for pneumonia detection, emphasizing the effectiveness of CNNs in medical image analysis. Despite challenges such as data scarcity and model overfitting, CNNs remain a promising solution due to their ability to learn hierarchical features from raw data. With this, the paper focuses on developing a CNN architecture and comparing it with the pre-trained ResNet model to solve the business objective of reducing the time required for pneumonia diagnosis by automating x-ray analysis. The images were initially pre-processed, and the images were resized to 128 x 128 pixels for ResNet and 224 x 224 pixels for CNN. In this research, CNN model's recall was 82% and pre-trained ResNet model's recall was 77%, which is slightly less than the CNN model. This paper discusses the approach of achieving these results and then conclude the paper with future work.**

Keywords – ***Pneumonia, CNN architecture, pre-trained ResNet model***

I. INTRODUCTION

Pneumonia is inflammation and fluid in the lungs. They are caused by viral, fungal, or bacterial infection. It can make breathing difficult and may cause cough and fever (*Pneumonia* (n.d.). According to (Clinic Barcelona, 2021) pneumonia causes 2.5 million death each year globally. Out of this, 24% are children under the age of 5. According to a news article in one of the leading newspapers of India, 24% of global pneumonia cases are found in India (Sharma (2023)) with fatality rate between 14 to 30 percent. This is a raising cause for concern globally. The major cause of pneumonia among kids is the air pollution. When an x-ray of chest is performed, on an average, it takes about one to two days for the x-ray results to come out. Once out, the general practitioner (GP), according to their availability, looks at the result and informs the patient about the results. There can be some delays in the process as it requires the expertise of the GP and can only be interpreted when GP is available. Fig. 1 shows the image of a pneumonia x-ray and normal x-ray.

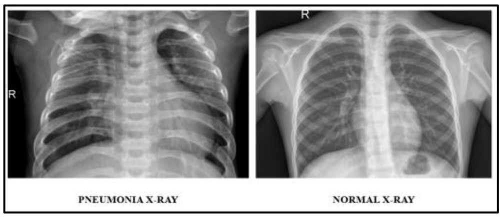

Fig. 1. Pneumonia vs normal x-ray

To a normal human (or non-experts), they both seem quite similar, and we won't be able to tell the difference between these x-ray images.

These limitations motivate the exploration of AI for automated detection of pneumonia from normal chest x-ray. Deep learning algorithms like Convolutional Neural Networks (CNNs) offer promising potential to address these challenges. This was also explored in the literature review where CNN was the widely used algorithm that was used by different researchers. Faster analysis by CNN models can significantly reduce turnaround time, leading to quicker treatment initiation.

With this business proposal of this project is as follows:

Reduce the time required for pneumonia diagnosis by automating X-ray analysis with a CNN model.

Working on this business proposal, we used a pre-trained ResNet model and proposed our CNN models with different convolutional layers. The ResNet model gave a recall of 77% and our proposed CNN model gave a recall of 82%.

In the subsequent section, this research will work its way to achieve this business goal and will be explained in deep. The layout of the paper is as follows. Firstly, the literature review will explore the research done on this topic and how they go about implementing similar research topic. The section ends with challenges and justification which critically assesses the techniques in the research with this paper. The data description gives a quick overview of the x-ray image dataset at hand. Post this, the papers talk about processing and modelling and finally concluding it with the results and interpretation leaving a room for future advancements.

II. LITERATURE REVIEW

Spiking Neural Network [SNN] proposed by Sivapriya et al. (2023) along with an enhancement technique called CLAHE was used for noise clearance and reshaped images were fed into SNN model to go through XGBoost and RF classifier for classification with obtaining 98% accuracy. Similarly, 3 architectural hierarchies were proposed by Widodo et al. (2022) to first build an architecture containing 7 convolution layers and 3 ANN layers using UBNetV1 to classify images and performed with 99.6% accuracy.

Detection of pneumonia and tuberculosis proposed by Ahmed et al. (2023) developed an ANN based on VGG16 with LBP, DWT and GLCM(LDG) reached an accuracy of 99.6%. Integrating these models with having higher accuracy are very helpful in medical system for better human healthy environment. Along with this a deep comparison of deep learning neural networks conducted by Khan et al. (2022) on medical image classification, used 3 neural networks and a sort of artificial neural network that is RNN was used, and it performed better.

In the current days, thanks to the application of computer-aided diagnostic (CAD) systems, especially for those based on machine learning (ML) and deep learning (DL) techniques, the future opens to several promising solutions regarding raising the accuracy of detecting pneumonia. Alenezi and Ludwig (2022) spotlighted that machine learning (ML) was considered of enormous importance for pneumonia detection in chest X-ray images, especially the use of custom convolutional neural network (CNN) models. Despite these benefits and justification of ML-based techniques, several issues are still there. Improvement for 2017 was needed in further optimization to increase accuracy. The release of Teaching Learning-Based Optimization (TLBO) is an outstanding approach to pneumonia detection, and combined with CNN, shows continued efforts to refine diagnostic accuracy. The model has exhibited notable improvements and highlights, indicating outperformance from all the former models, with high accuracy.

Similarly, Demir (2023) argues that there is a "need" for accurate and timely diagnosis of pneumonia, especially in paediatrics. The above capabilities were identified, and the CNN methods for transfer learning have been recommended as techniques best suited for accurate diagnosis. In this line, these good results have been produced, as with the use of the suggested CNN model, a test accuracy of 96.31% was yielded. These thus present a challenge in their interpretation, either from picture ambiguity or from errors of humankind and have therefore motivated researchers to investigate data augmentation and transfer learning as a way of improving CNN performance.

Alam et al. (2023) warn that these diseases of the lung are important across the world, which needs detection at an early stage to administer proper treatment. It developed RVCNet, a hybrid deep learning framework using various CNN models, and proved improvements in lung disease prediction from X-ray data. RVCNet showcases great accuracy in classification, as it shows the superiority of using ensemble DL models toward increasing diagnostic performance. The usefulness of RVCNet, on the other hand, is largely reliant on the ability to overcome the constraints on data in medical imaging.

Alenezi and Ludwig (2024) on the other hand, just indicated the robustness of TLBO-CNN for pneumonia detection and were able to underline its capability of improving accuracy in diagnosis in clinical surroundings. The better performance of TLBO-CNN requires an improvement to surpass common challenges faced during a process like pneumonia diagnosis, i.e., data scarcity and visual ambiguity for improved diagnostic results.

Taken together, these studies propose great potential for ML and DL techniques in improving the diagnosis of pneumonia from CXR images. Though each technique involves promising results, it needs to overcome the data constraint issues for robust diagnosis accuracy in practice.

Ayalew et al. (2022) uses YOLOv3 to check whether the input image is chest X-ray or not. The dataset that this paper uses has less training and test data and they have used image and data augmentation to increase the dataset size. Histogram equalization was used to increase the pixel brightness. Feature is extracted using HOG and after doing SVM classifier, the resultant model gives 99 percent training accuracy for DCC net and 100 percent accuracy for HOG. Even though the accuracy is impressive, it poses a threat of overfitting of the data. For medical science usage, relying on a model with 100 percent accuracy is not recommended as this may lead to life threatening results.

In another paper, Avola et al. (2022) is using transfer learning to detect pneumonia from chest x-ray images. After doing an impeccable job of comparing different pretrained model like Googlenet, Densenet, Resnet, Mobilenetv3 etc, the authors concluded that mobile net v3 is the best model for transfer learning because it gave accuracy of 84.36%.

However, this is not ideal when reproducibility is the key of a project. Pre trained models are trained on everything and they extract generic features. In order to have a tight model which extracts a feature from chest x-rays, it is important to create convolutional layers from scratch which this paper is focusing on.

Scaling image to 224,224 pixels is a key takeaway from research done by Verma, Gopal and Sharma (2023). RESNET was used which gave an accuracy of 99.96 percent which is a resultant of overfitting of the data. The reason is that the images are augmented because the initial database had less images. By building various models over CNN, Saparna and Mary (2023) in their research build KNN, Random Forest, Naïve Bayes, and SVM over CNN and calculate the model accuracy for each. SVM performs well by giving 90% accuracy. This is the inspiration for our paper and by training the model to extract features from the image, SVM will be used in the end to increase accuracy.

A. Challenges

Several studies, including those by Ayalew et al. (2022) and Avola et al. (2022), face challenges related to data scarcity in medical imaging. Limited datasets hinder model generalization and may lead to overfitting, while pre-trained models may extract generic features, impacting reproducibility. Verma, Gopal, and Sharma (2023) report overfitting issues due to image augmentation and a small initial database, resulting in inflated accuracy. Model overfitting poses a significant challenge in achieving robust and generalizable results.

Despite the varied approaches employed in the reviewed papers, the use of Convolutional Neural Networks (CNNs) emerges as a common thread. CNNs offer several advantages, including their ability to automatically learn hierarchical features from raw data, such as chest X-ray images. CNNs excel in capturing spatial dependencies in images, making them well-suited for medical image classification tasks.

However, while CNNs offer promising results, gaps exist in their implementation across the reviewed papers. Some studies may not fully exploit the potential of CNNs, such as relying solely on pre-trained models without fine-tuning for specific medical imaging tasks. This can result in non-reproducibility of the code, which is the main concern. Code reproducibility is code is one of the main focuses of this research.

B. Justification

Based on the critical analysis of the reviewed literature, CNN emerges as a suitable option for our paper due to several reasons. CNNs have demonstrated remarkable performance in capturing complex features and patterns from chest X-ray images, essential for accurate disease classification. The hierarchical nature of CNNs allows them to automatically learn and extract relevant features, reducing the need for manual feature engineering. CNNs offer flexibility in model architecture design, enabling customization to suit specific medical imaging tasks and datasets.

Despite challenges such as data scarcity and model overfitting, CNN remain a widely adopted and effective approach in medical image analysis, promising robust and generalizable results when appropriately implemented and fine-tuned. By leveraging the strengths of CNNs and addressing the challenges highlighted in the literature, our study can contribute to advancing the state-of-the-art in lung disease detection and classification, ultimately benefiting medical diagnosis and treatment outcomes.

III. DATA DESCRIPTION

The Pneumonia Detection dataset of chest X-rays from Kaggle (Mooney, P. (2018)) is accessible to the public. It comprises of 5856 X-Ray images. This data is organised into three folders namely train, test, and validation. Each of these folders have two sub-folders which contains normal and pneumonia chest x-ray images. Table I gives a detailed description of the dataset. Fig. 2 on the other hand shows the visual representation of the data distribution.

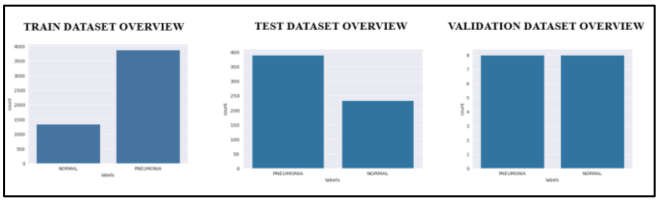

Fig. 2. Visual representation of the data distribution

TABLE I. DATASET DESCRIPTION

TEST	
Normal	234
Pneumonia	390
TRAIN	
Normal	1341
Pneumonia	3875
VALIDATION	
Normal	8
Pneumonia	8

The images have varied range of pixel which may cause irregularity in the processing the image. Hence, the pictures were brought down to 224 x 224 and 128 x 128 pixels to normalize the image data. 224 x 224 image was used in the proposed CNN model and 128 x 128 image was used in the pre-trained ResNet model.

IV. PRE-PROCESSING

Libraries were imported to support the model building. NumPy was used for data manipulation and data visualization. For numerical operations, pandas and matplotlib were used. Pyplot from matplotlib is extensively used for visualization and the operating system’s function were used for file operations. Apart from these libraries, various other libraries and modules were employed for deep learning, model building, optimization, and evaluation. For example, Seaborn module was used for statistical data visualization.

TensorFlow is an opensource machine learning framework that provides a set of tools that include different layers such as sequential, convolutional, pooling, and dense. To evaluate the model performance, confusion matrix, and classification reports were generated.

After importing the libraries, preprocessing of the images was done to enhance the model interpretability of the chest x-ray images. Three different paths are established for training, test, and validation of image dataset. The validation data is used during the training process to assess the model’s performance on data to tune hyperparameters and adjust model’s performance as needed. This helps in preventing overfitting by providing an independent dataset for model evaluation.

To organize the data, data frames for test, train and validation were created which stored the paths of images and their respective labels. A loop is implemented through the files in the test directory using the OS library and each file is assigned a full file path by concatenating the file directory with the file name.

The preprocessed images are converted into datasets using tensor flow-built function loaded in the gray scale formats as they have a single channel representing pixel intensity and they are used because of the less input dimensionality and specifies batch sizes for training, testing and validation data.

V. MODEL ARCHITECTURE

A. *Residual Network50V2 architecture*

ResNet, short for Residual Network, is a type of deep neural network architecture that is highly effective for deep learning tasks, particularly in computer vision applications like image classification and object detection. ResNet introduces skip connections that bypass one or more layers, allowing the input to be added to the output of deeper layers. This helps alleviate the vanishing gradient problem and enables training of much deeper networks and in return it ends up in performing an identity mapping. How identity mapping works is: If the input and output of a residual block have the same dimensions, the skip connection simply adds the input to the output. If the dimensions differ, the skip connection includes a 1x1 convolutional layer to match the dimensions.

ResNet is used for pneumonia detection because of its efficiency in dealing with the limitations presented by deep neural networks. Pneumonia identification requires effectively biasing between normal and abnormal (pneumonia-infected) chest X-ray pictures, which can be difficult due to variations in image quality, patient demographics, and the different types of pneumonia symptoms in some circumstances. ResNet's architecture allows for the training of very deep neural networks, which is critical for extracting detailed characteristics from input pictures.

Additionally, ResNet architectures, mainly the ones which are pre-trained on big datasets like as ImageNet, are effective for feature extraction and collecting layered representations of visual information that may be used for pneumonia detection. Fine-tuning a pre-trained ResNet model using pneumonia-specific data enables effective training even with limited labelled data, exploiting the generalisation skills acquired from massive volumes of varied picture data. Overall, using ResNet for pneumonia identification improves the model's ability to learn discriminative characteristics from chest X-ray pictures, resulting in more accurate and reliable diagnosis of pneumonia cases.

The proposed architecture includes creating a neural network model based on the ResNet50V2 architecture (fig. 3), specified for binary classification tasks. The model is implemented by loading the pre-trained ResNet50V2 model from the keras module integrated to use pre-trained weights on the image dataset. Notably, the model will be allowed for custom classification layers to be appended as all layers in the pre-trained model will be frozen to reserve the learned features during training. A pooling layer is added to the output layer of the model to facilitate feature consideration and normalization. In continuation to this, an activation layer ReLu was added to capture higher features and finally, dense layer with sigmoid function was added for binary classification. Binary cross entropy loss, Adam optimizer, and accuracy metrics are used for compiling the model.

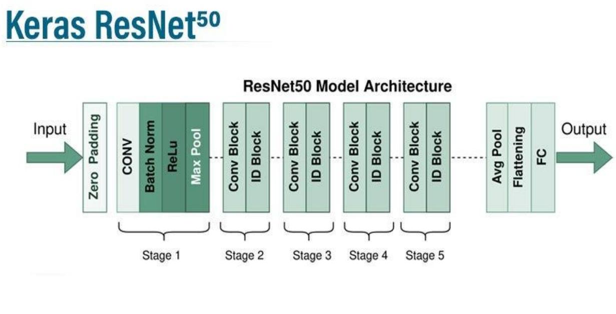

Fig. 3. Keras ResNet50 (Serej, 2022)

B. *CNN architecture*

CNN are a class of deep neural networks which are designed to process and analyze visual data such as images or videos. The kernels or filters are applied on input images to extract features. It has multiple layers such as convolutional layers, pooling layers, and fully connected layers to perform feature extraction and detecting complex patterns, textures, edges, and shapes to gain insights of the data for model performance. CNN was integrated in this study, as the dataset was image data and using neural networks was apt for making predictions. By critically analyzing the literature review, CNN emerged as a widely utilized model. CNN provides a strong architecture suited for processing visual data. The hierarchical structure of the model allows the extraction of features by capturing the basic features like edges and texture in the initial layers and by extracting complex patterns in the later layers.

CNN provides the flexibility of reorganizing the image sizes (it is important to resize as working on large dataset is challenging) by allowing spatial dimension reduction and accounting to reduce the computational time. It can generalize unseen data and variations within the same class. CNNs have showed remarkable performance on various computer vision tasks, that includes image classification, object detection, image segmentation and image generation which have achieve better results in complex and challenging datasets and have set benchmarks.

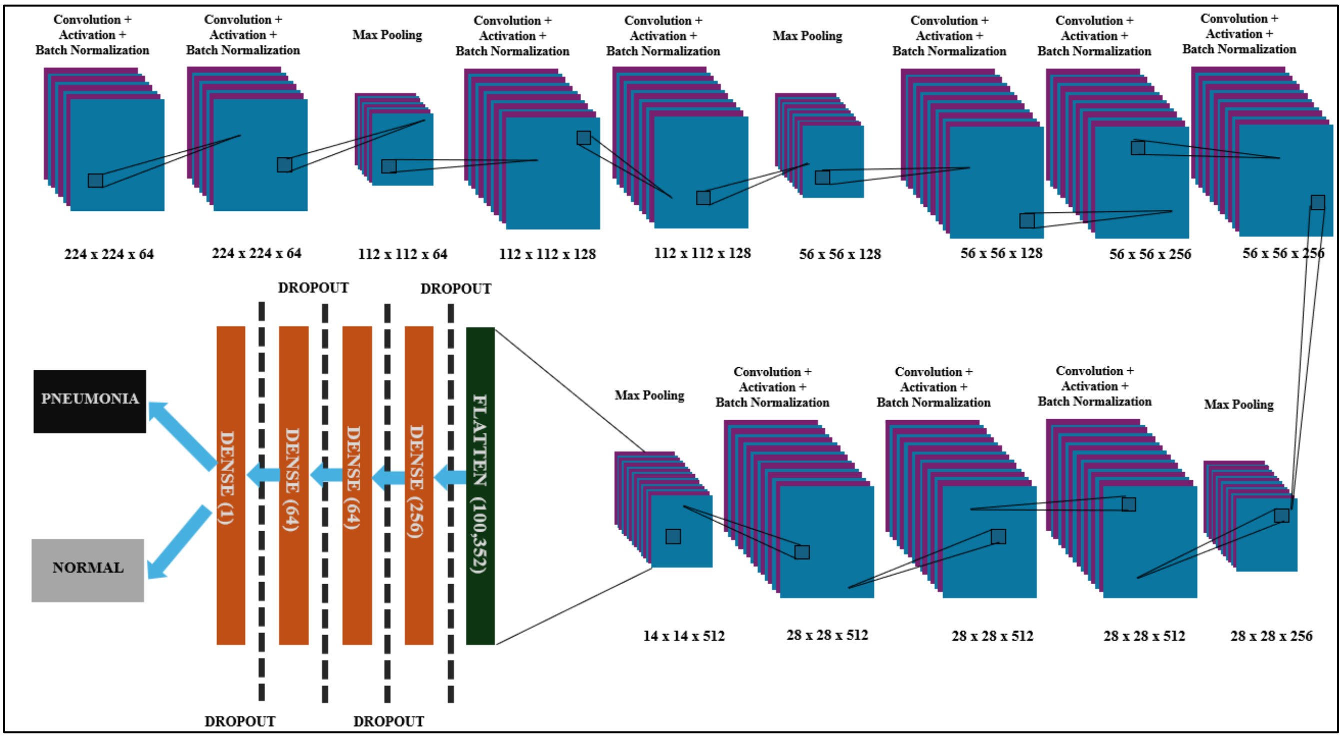

Fig. 4. Proposed CNN model architecture

VI. MODELLING

A. *Resnet Model*

The initial methodology involved developing a ResNet model to better understand its behaviour and performance in pneumonia identification. The ResNet model was pre-trained on the ImageNet dataset and fine-tuned using pneumonia-specific information. This strategy helped to take advantage of ResNet's solid feature extraction capabilities, which are very useful for picture classification applications. ResNet model was trained on the pneumonia dataset to see how well it could learn discriminative characteristics from chest X-ray pictures and reliably diagnose pneumonia cases.

Initial stages of data preparation were carried out and the images were loaded using image data generator and resized to 128 x 128 pixels. Core model development is the transfer learning approach using ResNet50V2 architecture. Additionally, layers were added on the top of the pre-trained model for binary classification and the output layer had one neuron along with sigmoid function for predicting the YES or NO of pneumonia. A training accuracy of 0.617 and with the testing accuracy of 0.562 was achieved. Fig. 5 shows the Resnet model's layer summary.

After evaluating the ResNet model, convolutional neural network (CNN) model was implemented from scratch. A variety of variables contributed to the choice to create a customised CNN model most importantly being the ability to reproduce the data and to have complete control over the network architecture. Furthermore, while the ResNet model offered a solid foundation, there were several limitations that required the examination of alternate methodologies. Despite its effectiveness in feature extraction, the ResNet model performed poorly in several aspects of pneumonia identification. ResNet model struggled to effectively categorise many pneumonia patients, especially those with small or unclear symptoms on chest X-ray images. This shortcoming underlined the need for a more specialised and finely tuned model capable of capturing all the minute aspects associated with pneumonia.

```
res_model.summary()
Model: "functional_1"
```

Layer (type)	Output Shape	Param #	Connected to
input_layer (InputLayer)	(None, 128, 128, 3)	0	-
conv1_pad (ZeroPadding2D)	(None, 134, 134, 3)	0	input_layer[0][0]
conv1_conv (Conv2D)	(None, 64, 64, 64)	9,472	conv1_pad[0][0]
pool1_pad (ZeroPadding2D)	(None, 66, 66, 64)	0	conv1_conv[0][0]
pool1_pool (MaxPooling2D)	(None, 32, 32, 64)	0	pool1_pad[0][0]
conv2_block1_preact_bn (BatchNormalization)	(None, 32, 32, 64)	256	pool1_pool[0][0]
conv2_block1_preact_relu (Activation)	(None, 32, 32, 64)	0	conv2_block1_preact_bn[0]…
conv2_block1_1_conv (Conv2D)	(None, 32, 32, 64)	4,096	conv2_block1_preact_relu[…
conv2_block1_1_bn (BatchNormalization)	(None, 32, 32, 64)	256	conv2_block1_1_conv[0][0]
conv2_block1_1_relu (Activation)	(None, 32, 32, 64)	0	conv2_block1_1_bn[0][0]
conv2_block1_2_pad (ZeroPadding2D)	(None, 34, 34, 64)	0	conv2_block1_1_relu[0][0]
conv2_block1_2_conv (Conv2D)	(None, 32, 32, 64)	36,864	conv2_block1_2_pad[0][0]
conv2_block1_2_bn (BatchNormalization)	(None, 32, 32, 64)	256	conv2_block1_2_conv[0][0]
conv2_block1_2_relu (Activation)	(None, 32, 32, 64)	0	conv2_block1_2_bn[0][0]
conv2_block1_0_conv (Conv2D)	(None, 32, 32, 256)	16,640	conv2_block1_preact_relu[…
conv2_block1_3_conv (Conv2D)	(None, 32, 32, 256)	16,640	conv2_block1_2_relu[0][0]

Fig. 5. Resnet model layer summary snapshot

B. *CNN Model*

As a result, by creating own CNN model, we hoped to overcome and improve the model's performance for pneumonia identification. We aimed to improve the model's sensitivity and precision by iterative testing and fine-tuning of the CNN architecture. This technique enabled to carefully analyse several modelling approaches, find areas for improvement, and tweak the CNN model to better match the unique needs of detecting pneumonia from chest X-ray pictures.

CNN model is built using tensor flow keras API. The image sizes are specified as 224 x 224 pixels and channels are set 1 as greyscale images are 1D. The architecture of CNN model is sequential model, and each layer is added one by one in order. A 2D CNN layer is used to perform convolution operations on input data to extract spatial features from images with 32 filter (Kernels) for determining the output feature maps of size 3*3 (size of kernel) with the ReLu activation function. ReLu function is used for its ability to introduce the non-linearities to the model decision boundaries. It clearly determines the information that should be permitted and the information that should be blocked. This aids in the identification of complex patterns and relationships in the data. By offering multiple processing options and preventing unnecessary information, ReLu helps optimize the overall process.

Max pooling is used to divide the picture into smaller sizes of squares or rectangles to analyze each segment independently. It identifies the brightest spot within each segment, retaining only the important information and discarding the rest. The max value information is converted into a new smaller picture with less pixels and makes it easier for neural networks to process which leads in reducing the computational load and prevent overfitting of data. Flatten layer gathers and prepares a piece of information for CNN layers for analysis and dense layer looks for patterns in it and connects information with different neurons, which looks for different patterns or combinations.

The 2nd dense layer makes decisions based on the patterns found and soft max function helps in making output decision which has 2 neurons and each representing either one of the options. Finally, a summary of the model offers insights into the model's architecture by providing details on its layers, output shapes, and number of parameters. Finding the optimal solution for the model is paramount. To achieve this, a learning rate scheduler was employed that starts the learning rate by a smaller value of 0.001. The learning rate is adjusted at each checkpoint and is decreased by 10 %, which is helpful in finding optimal solution. This helps fine-tune the model and enhance its performance. Fig. 6. Shows the CNN layer summary.

Next section discusses the result and conclusion of this project and compares the two models implemented for Pneumonia detection.

```
Model: "sequential"
```

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 224, 224, 64)	640
batch_normalization (Batch Normalization)	(None, 224, 224, 64)	256
conv2d_1 (Conv2D)	(None, 224, 224, 64)	36928
batch_normalization_1 (Bat chNormalization)	(None, 224, 224, 64)	256
max_pooling2d (MaxPooling2 D)	(None, 112, 112, 64)	0
conv2d_2 (Conv2D)	(None, 112, 112, 128)	73856
batch_normalization_2 (Bat chNormalization)	(None, 112, 112, 128)	512
conv2d_3 (Conv2D)	(None, 112, 112, 128)	147584
batch_normalization_3 (Bat chNormalization)	(None, 112, 112, 128)	512
max_pooling2d_1 (MaxPoolin	(None, 56, 56, 128)	0

Fig. 6. CNN model layer summary snapshot

VII. RESULTS AND EVALUATION

To correctly classify the chest x-ray image into pneumonia and normal, ResNet and CNN models were used and run on Google Colab. The CNN model has 5 epochs i.e. completes 5 iterations through the training dataset to train the model thoroughly. After each epoch, model performance is evaluated on the training data against validation model and checked for overfitting and underfitting along with metrics such as loss and accuracy. Finally, predictions were made on the test dataset and the predicted class for each sample is stored. These reports provide insights into the model's precision, F1 score, recall, and support score. Confusion matrix is generated to visualize the model's prediction on the test dataset, highlighting the areas where the model may struggle in its prediction. All this helps in gauging the model's accuracy and performance. The obtained training accuracy and testing accuracy were 0.973 and 0.796 respectively. Table II shows the comparison between total parameters of pretrained model and proposed model.

TABLE II MODEL PARAMETERS

PRE-TRAINED RESNET MODEL	
Total Parameters	23,827,201
Trainable parameters	262,401
Non-trainable parameters	23,564,800
PROPOSED CNN MODEL	
Total parameters	33,355,970
Trainable parameters	33350594
Non-trainable parameters	5376

The proposed CNN model has a higher total parameter count as compared to the pre-trained ResNet model. This suggests a higher model complexity and capacity to learn intricate patterns in the data. Most parameters are non-trainable in the ResNet model which is inversely proportional to the proposed model. As initially guessed and discussed in the literature review, pre trained model relies heavily on the fixed features and extracts generic features from an image. In this case, ResNet extracts fixed features like edges, curves etc which are generic without understanding the image itself. On the contrary, the higher parameter count in the proposed CNN model implies higher model complexity which could lead to better performance on the task at hand i.e. pneumonia detection.

Fig.7. depicts the training and validation loss and training and validation accuracy. The figure shows the best epoch for model is epoch 5 where the losses decreased on training set and at the same time, the validation accuracy increased which indicates that the model is performing well on the unseen data.

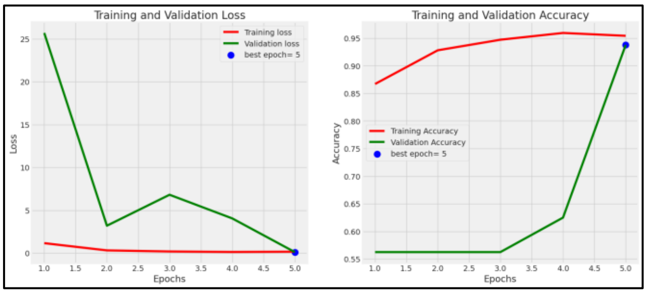

Fig. Validation loss VS Validation accuracy

Fig.8. shows the confusion matrix of the CNN. With True positive as 321, it suggests that it correctly identifies a chest x-ray as pneumonia. This is the crucial step as it is important to identify a pneumonia x-ray accurately. With false negative at 69, model incorrectly classifies a pneumonia chest x-ray as normal.

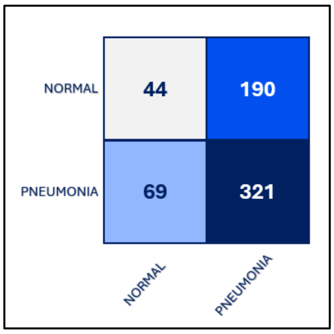

Fig. 8. Confusion matrix of CNN

The overall accuracy of the CNN model is 0.68 which means that model correctly classify 68% of the chest x-rays. On the other hand, precision being 0.63 suggests that 63 percent of the time, pneumonia was accurately classified.

Refer Table III for class imbalance. It is important to address the class imbalance in this scenario as both in test and train, Pneumonia x-rays are higher than the normal ones. Validation has a 50/50 split. In these scenarios, it is important to consider recall rather than the model accuracy itself as high/low accuracy can be misleading. The model's recall is 82% which indicates that the model correctly classifies 82% of the pneumonia cases. Fig.9 shows the model accuracy, recall and precision for the proposed CNN model.

TABLE III CLASS IMBALANCE

Pneumonia	
Train	74%
Test	62%
Normal	
Train	26%
Test	38%

When looking at the pretrained ResNet model, the confusion matrix is shown in fig. 10. True positive rate is 323 and overall recall for the ResNet is 77%, slightly less than the proposed CNN model. The reason as explained in the sections above lies in the fact that the CNN model is trained only on the x-ray images that extract features related to them instead of extracting generic features, as done by ResNet. Fig. 11 Shows the model accuracy with f1 scores, precision, and recall for ResNet.

	precision	recall	f1-score	support
NORMAL	0.39	0.19	0.25	234
PNEUMONIA	0.63	0.82	0.71	390
accuracy			0.58	624
macro avg	0.51	0.51	0.48	624
weighted avg	0.54	0.58	0.54	624

Fig. 9 CNN model performance metrics

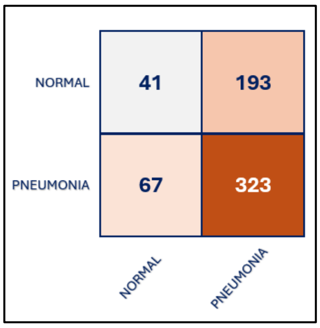

Fig. 10. Confusion matrix of ResNet

Classification Report:

	precision	recall	f1-score	support
0	0.36	0.22	0.27	234
1	0.62	0.77	0.69	390
accuracy			0.56	624
macro avg	0.49	0.49	0.48	624
weighted avg	0.52	0.56	0.53	624

Fig. 11. ResNet performance metrics

VIII. CONCLUSION

The study investigated the effectiveness of CNN for pneumonia detection using chest x-ray images. A custom CNN model was developed that had a testing accuracy of 79.6%. This paper talked about how recall is more important than the accuracy of the model as in medical imaging, it is important to correctly identify number of the pneumonia cases. Compared to the pretrained model, the proposed CNN model performs better as it is trained on chest x-ray images only.

A. Limitations

The study encountered two major limitations. The data imbalance exhibited an imbalance between normal and pneumonia x-rays where it was evident that pneumonia cases were more than the normal ones. This can eventually skew the model metrics like accuracy towards the majority class. Publicly available data is different from clinical or real-world data. It limits the model's generalizability to the unseen cases.

For future, exploring the data augmentation techniques can be useful to tackle class imbalances and can improve accuracy and recall of the overall model.

IX. Future Work

There is a vast avenue for future research when taking chest x-ray pneumonia detection into consideration. Firstly, Data augmentation can be implemented over this which can flip, rotate, scale the image to increase the size of the dataset artificially. This could be helpful as the real-world data is not simple and might involve variations in the image. This technique may bring the dataset close to the real-world data and can increase the model generalizability.

Tuning hyperparameters like learning rate, optimizer, number of epochs etc. can further be an effective technique and improve the model's performance. Techniques like grid search or randomized search can be employed to explore a wider range of hyperparameter combinations.

More diverse and large data containing chest x-rays from various demographics can be important for a robust performance of the model. Datasets can be sourced from different publicly available websites and then pulled into one data frame and augmented to increase the data size further to make it more diverse and closer to the real-world data.

Finally, clinical validation of these models with real patients will play a crucial role in assessing the suitability of such models in the clinical trials. A trained physician can be involved and then the total time can be recorded for the machine and physician to see whether there is a significant difference between the time taken to read a report by a machine and a human.

REFERENCES

Alam, F. B., Podder, P., & Mondal, M. R. H. (2023). RVCNet: A hybrid deep neural network framework for the diagnosis of lung diseases. *PloS One*, *18*(12), e0293125. https://doi.org/10.1371/journal.pone.0293125

Alenezi, R., & Ludwig, S. A. (2022). Detecting Pneumonia Based On Chest X-Ray Images Using Reinforcement Learning. *2022 IEEE Fifth International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*. https://doi.org/10.1109/aike55402.2022.00018

Alenezi, R., & Ludwig, S. A. (2024). Using Teaching Learning-Based Optimization with Convolutional Neural Network to Detect Pneumonia Based On Chest X-Ray Images. *2024 International Conference on Artificial Intelligence, Computer, Data Sciences and Applications (ACDSA)*. https://doi.org/10.1109/acdsa59508.2024.10467902

Ahmed, I. A., Senan, E. M., Shatnawi, H. S. A., Alkhraisha, Z. M., & Al-Azzam, M. M. A. (n.d.). Multi-Techniques for Analyzing X-ray Images for Early Detection and Differentiation of Pneumonia and Tuberculosis Based on Hybrid Features. Diagnostics, 13(4), 814. https://doi.org/10.3390/diagnostics13040814

Avola, D. et al. (2022) 'Study on transfer learning capabilities for pneumonia classification in chest-x-rays images,' Computer Methods and Programs in Biomedicine, 221, p. 106833. https://doi.org/10.1016/j.cmpb.2022.106833

Ayalew, A.M. et al. (2022) 'Detection and classification of COVID-19 disease from X-ray images using convolutional neural networks and histogram of oriented gradients,' Biomedical Signal Processing and Control, 74, p. 103530. https://doi.org/10.1016/j.bspc.2022.103530

Clínic Barcelona (2021) *Pneumonia causes 2.5 million deaths around the world each year*, *Clínic Barcelona*. Available at: https://www.clinicbarcelona.org/en/news/pneumonia-causes-2-5-million-deaths-around-the-world-each-year (Accessed: 08 May 2024)

Demir, Y. (2023). Detection of pneumonia from pediatric chest X-Ray images by transfer learning. *Sigma MüHendislik Ve Fen Bilimleri Dergisi*, 1157–1176. https://doi.org/10.14744/sigma.2021.00074

Khan, M. F., Kumar, R. N. S., Patil, T., Reddy, A., Mane, V., & Santhoshkumar, S. (2022). Neural Network Optimized Medical Image Classification with a Deep Comparison. *2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)*. https://doi.org/10.1109/icaiss55157.2022.10011109

Mooney, P. (2018) Chest X-Ray Images (Pneumonia). Available at: https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia (Accessed 2-Mar-2024)

Pneumonia (no date) *HSE.ie*. Available at: https://www2.hse.ie/conditions/pneumonia/ (Accessed: 01 May 2024)

Serej, A.D. (2022) ResNet-50, Medium. Available at: https://medium.com/@arashserej/resnet-50-83b3ff33be7d (Accessed: 01 May 2024)

Sharma, K. (2023) Why pneumonia cases are on rise and how you can identify it on time, The Times of India. Available at: https://timesofindia.indiatimes.com/life-style/health-fitness/health-news/why-pneumonia-cases-are-on-rise-symptoms-prevention/photostory/97242937.cms. (Accessed: 08 May 2024)

Sivapriya, G., Gowri, P., Govarthanan, S., Hareeshkumar, K. S., & Buvana, S. (2023). Diagnostic Classification of Lung Disease with Chest X-Ray Images Using SNN. 2023 2nd International Conference on Edge Computing and Applications (ICECAA). https://doi.org/10.1109/icecaa58104.2023.10212244

S. P and A. V. A. Mary (2023) "A Comprehensive study on the Detection of Pneumonia using Machine Learning and Deep Learning Approaches," 2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT), Gharuan, India, 2023, pp. 233-237, doi: 10.1109/InCACCT57535.2023.10141758

Verma, S., Gopal, D. and Sharma, P.K. (2023) 'An Optimized CNN and Transfer Learning Approach for Pneumonia Detection,' IEEE Conference ID: 59847 [Preprint]. https://doi.org/10.1109/ictacs59847.2023.10390240

Widodo, C. S., Naba, A., Mahasin, M. M., Yueniwati, Y., Putranto, T. A., & Patra, P. I. (2022). UBNet: Deep learning-based approach for automatic X-ray image detection of pneumonia and COVID-19 patients. Journal of X-ray Science and Technology, 30(1), 57–71. https://doi.org/10.3233/xst-211005