Mr. R. Premkumar, 2026, 14:2
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

AI-Powered Pneumonia Detection Using Chest Ct Images

Mr. R. Premkumar(Ap/Cse), Mr.R.Eswara Prakash, Mr. M. Manoj, Mr. K. Janarthanan
Department of computer science & engineering, Kongunadu College of Engineering And Technology Tholurpatti (Po), Thottiam (Tk), Trichy (Dt) – 621 215

Abstract- The aim of this project is to classify CT Scan images of pa- tients with or without pneumonia. More specifically, we trained a Convolutional Neural Network(CNN) of differ- ent parameters with chest CT Scan images of children and the outcome classes are two: ”Pneumonia” or ”Non - pneumo- nia”. The findings follow in the next sections. The primary objective of this work is to develop an efficient and reliable model that can automatically classify chest CT scan images into two categories: pneumonia and non- pneumonia. For this purpose, Convolutional Neural Networks (CNNs) are utilized due to their proven effectiveness in image classification tasks. The model leverages transfer learning and fine-tuning techniques using pre-trained architectures, enabling better performance even with limited medical datasets. In addition, data augmentation methods such as rotation, zooming, flipping, and shifting are applied to enhance the diversity of the dataset and reduce overfitting. Various optimizers, including RMSprop and Adam, are implemented and compared to improve the training efficiency and accuracy of the model. Experimental results demonstrate that the proposed system achieves high accuracy and strong performance in distinguishing between infected and healthy lung images.

Keywords: Pneumonia Detection, Chest CT Scan Images, Convolutional Neural Network (CNN), Medical Image Classification, Deep Learning, Transfer Learning.

I. INTRODUCTION

Nowadays, the best available method for pneumonia diagnosis is chest CT Scan. Pneumonia is an inflammatory infection that affects one or both of the lungs and may be caused by a virus, bacteria, fungi or other germs. Pneumonia affects approximately 450 million people globally (7% of the population) and results in about 4 million deaths per year (Wikipedia).

We are going to use a CNN which seems the best option in order to solve this problem. What makes CNN special is its effectiveness and efficiency. CNNs have been highly successful for image recognition tasks whether we need to classify images or localize objects within images or to detect objects within images or even Pneumonia is a life-threatening respiratory disease that affects the lungs and can be caused by bacteria, viruses, or fungi. It remains one of the leading causes of death worldwide, particularly among children and elderly people. Early and accurate diagnosis is essential to ensure timely treatment and to reduce complications. In modern healthcare, chest CT scans and X-ray images are widely used for diagnosing pneumonia. However, the manual interpretation of these medical images by radiologists can be time-consuming and may sometimes lead to misdiagnosis due to human fatigue or limited expertise. segment objects in images. CNNs have proven to be highly successful in this With the rapid advancement of artificial intelligence, especially in the field of deep learning, there has been a significant improvement in medical image analysis. Convolutional Neural Networks (CNNs), a class of deep learning models, have shown remarkable performance in image classification and pattern recognition tasks. These models can automatically learn important features from medical images, making them highly suitable for detecting diseases like pneumonia.

This project focuses on developing an AI-powered pneumonia detection system using chest CT scan images. The proposed system utilizes CNN-based architectures along with transfer learning and fine-tuning techniques to improve classification accuracy. By leveraging pre-trained models, the system can efficiently learn from limited datasets while maintaining high performance. Additionally, data augmentation techniques are applied to enhance the robustness of the model and prevent overfitting.

he main objective of this project is to design a reliable and efficient automated system that can classify CT scan images into pneumonia and non-pneumonia categories. This system aims to assist healthcare professionals by providing faster and more accurate diagnostic support. Furthermore, such an automated solution can be highly beneficial in rural and underserved areas where access to skilled radiologists is limited.

Overall, this project demonstrates how artificial intelligence can be effectively integrated into the healthcare domain to improve diagnostic accuracy and support medical decision-making, ultimately contributing to better patient outcomes.

II. RELEATED WORKS

One of the most popular and first CNN was that of Krizheski, Sutskever and Hinton back in 2012 [4]. It is a very recent project and considered a pioneering paper. Krizhevsky, Sutskever and Hinton have trained a CNN over the complex dataset of imageNet LSRVC-2010. The dataset consists of over of 1.2 million images. Such a scale dataset is considered to be very large and consists of a lot of images. The CNN was able to identify what the object was with a very high level accuracy. Their CNN consisted of over 60 million different parameters including weights and biases and each one of them had six hundred fifty thousand neurons. It was a very big CNN which consist of convolutional layers, MaxPool layers, fully connected layers and a softmax final activation layer.

In order to train this convolutional neural network the used two different GPUs (Nvidia), and they trained their network for 5-6 days. In addition, Hinton also contributed with a milestone overfitting procedure called "Dropout". In their paper which they demonstrate the aforementioned CNN, they used two CNNs and they train them in parallel in order to classify more than a million images into their corresponding labels. The VGGnet from Oxford is an idea of a deeper network with small filters and 7.3% error rate[6].It consists of 16 layers and the key idea was the very small filters (3,3). VGGnet is one of the most influential papers because it reinforced the notion that CNNs have to have a deep network of layers in order of the hierarchical representation of visual data to work.

The VGGnet from Oxford is an idea of a deeper network with small filters and 7.3% error rate[6].It consists of 16 layers and the key idea was the very small filters (3,3). VGGnet is one of the most influential papers because it reinforced the notion that CNNs have to have a deep network of layers in order of the hierarchical representation of visual data to work.

The rest of paper is organized as follows. In Section III details about the dataset, Section IV what is fine tuning and what is data augmentation. Section V we define the our model and the optimizers we used, Section VI experiments and results, Section VII summary of our work and achievements. Finally section VIII conclusion and Section IX future work. In recent years, deep learning techniques have significantly improved the performance of image classification tasks, especially in the medical domain. One of the pioneering works in this field was introduced by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton in 2012 through the development of a deep Convolutional Neural Network (CNN) trained on the ImageNet dataset. This model demonstrated that CNNs can effectively learn complex features from large- scale image data and achieve high classification accuracy.

Their work laid the foundation for applying deep learning models in various applications, including medical image analysis. Following this, more advanced ,architectures were proposed to improve performance and efficiency. The VGGNet model, introduced by Simonyan and Zisserman in 2014, emphasized the importance of deeper networks with smaller convolutional filters. This approach enabled better feature extraction and improved accuracy in image recognition tasks. VGGNet has since become one of the most widely used architectures for transfer learning in medical imaging applications due to its simplicity and effectiveness.

Several researchers have applied CNN-based models for pneumonia detection using chest X-ray and CT scan images. These studies highlight the ability of deep learning models to automatically extract relevant features such as lung texture, opacity, and abnormalities without manual intervention. Transfer learning has been widely adopted in these works, where pre-trained models are fine-tuned on medical datasets to overcome the challenge of limited data availability. In addition to model architectures, data augmentation techniques have also been explored to improve model generalization. Methods such as image rotation, flipping, zooming, and shifting have been proven effective in increasing dataset diversity and reducing overfitting. Perez and Wang demonstrated that data augmentation plays a crucial role in enhancing the performance of deep learning models, especially in image classification tasks with limited datasets.

Furthermore, optimization techniques such as Stochastic Gradient Descent (SGD), RMSprop, and Adam have been widely used to train deep neural networks. Among these, the Adam optimizer has shown better convergence speed and performance in many cases due to its adaptive learning rate mechanism.

Overall, previous research indicates that combining deep CNN architectures, transfer learning, data augmentation, and efficient optimization techniques can significantly improve the accuracy of pneumonia detection systems. These advancements have motivated the development of this project, which aims to build an effective and reliable AI-based solution for automated pneumonia diagnosis.

The application of deep learning in medical image analysis has gained significant attention in recent years, particularly for disease detection and classification. Early developments in this field were strongly influenced by the success of deep

Convolutional Neural Networks (CNNs) in large-scale image recognition tasks. A major breakthrough came with the introduction of AlexNet in 2012, which demonstrated the ability of CNNs to automatically learn hierarchical features from raw images. This success encouraged researchers to explore similar techniques for analyzing medical images such as CT scans and X- rays.

III. PROPOSED & METHODOLOGY

Our dataset consists of 1583 normal photos and 4273 pho- tos of pneumonia, split into train, test and validation sets (Figure 2). The dataset downloaded from the site of kaggle. The train set consists of 1341 normal and 3775 of pneumo- nia, while the test set 234 and 390 respectively.The valida- tion set has only 16 photos, of which 8 are pneumonia and 8 normal, which makes them really difficult to be classified correctly.

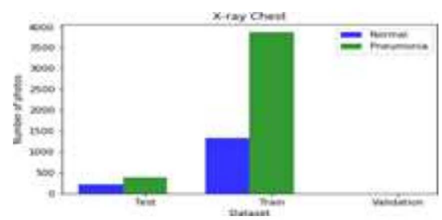
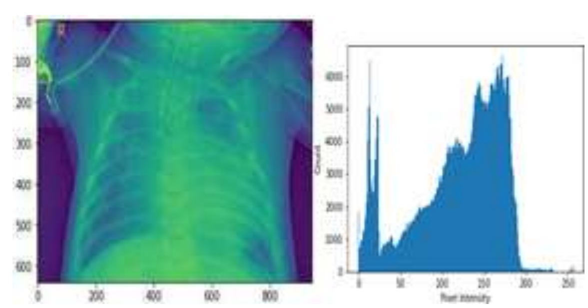


Figure 2: Images distribution of train, test and validation set

Figure 3 demonstrates a sample image and its corre- sponding pixel intensity histogram. The histogram repre- sents the flatten tonal distribution of the pixels that corre- spond to that image. From left to the right, the pixels are distributed from dark to lighter colors. A CNN takes an image as input, expressed as an array of numbers, applies a series of operations and at the end returns the probability that an object in the image belongs to a particular class of objects.Therefore, the distribution of the pixels is the char- acteristic of the specific image and could be considered as the first layer in the CNN.



(a) virus (b) pixel intensity

Figure 3: A sample image and its pixels intensity

Preprocess

Fine tuning DCNNs

Fine tuning is a process that takes a pretrained network model for a specific task and uses it for another similar task. Fine tuning is a very close term to transfer learn- ing and it occurs when we use knowledge that was gained from solving one problem and then apply it to a new but related problem. For instance, a knowledge gained from learning to recognize apples could potentiallybe applied in a oranges recognition.

In more details, fine tuning is a process that takes a model that has already been trained for a given task and then tun- ing or tweaking that model to make it perform a second or similar task. Using an artificial neural network that has already been designed and trained allows us to take advan- tage of what the model has already learned without having to develop it from scratch.

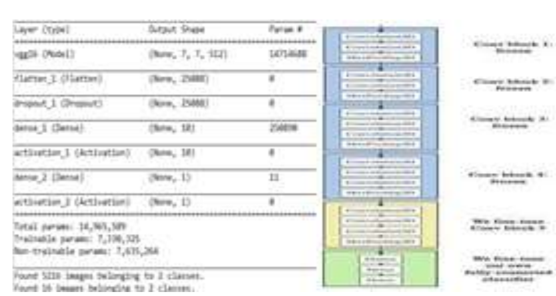
When building a model from scratch, we have to try many approaches and a type of trial and error. This process seems really difficult because we have to choose how many lay- ers we are using, what types of layers, what order to put the layers, how many nodes to include in each layer, how much regularization to use, what learning rate to use and etc. Validating our own model can be a huge task in its own right depending on what data we are training it on.

Like in our research, the two tasks are different and there will be information that the model has learned that may not apply to our new task. Or there may be new informa- tion that the model needs to learn from the data, regarding the new task that was not learned from the previous task. As we mention in the previous example, a model trained on apples is not going to have ever seen an orange before so this feature is something new the model would have to learn about. But this, already trained, model has learned to understand edges ,shapes and textures. All this previous knowledge will be benefit for our new model in order to learn to classify new images.

The implementation of this model is based on importing our first original model that was used on another task. Next for simplicity purposes, we remove the last layer of this model. The last has been used previously in order to clas- sify an object, for instance if it was an apple or not. After removing this, we add a new layer back whose purpose is to classify whether an image is an orange or not. In our case we want to remove more than the last single layer and we added more than just one layer. We made this change because our task was not similar with the task that the net- work was trained. The layers that we use seem to play important role. The first layers have learned more gen- eral features and the last layers more specific layers. After we have modified the structure of the existing network, we freeze that layers in our model that came from the origi- nal model. We use the term freezing, because those layers consider as non trainable. The weights of those layers will not be updated whenever we train the model on our new data for our new task. We update only the weights of our new or modified layers. Instead of training an entire net- work from scratch with random initial weights, we use the weights of a network that has been trained with a very large data set, on a similar task. In our experiments we used the weights from the ImageNet and in Figure 4a demonstrates the structure of our network and in Figure 4 the concept of the fine tuning.

Data augmentation

Data augmentation is a popular technique used in many machine learning tasks, such as image classification, which aims in virtually enlargement of the training dataset size and avoid overfitting. Traditional data augmentation tech- niques for image classification tasks create new samples from the original training data by, for example, flipping,



Our Model Fine Tuning

Data augmentation settings

zoomrange	1, 1.25
widthshiftrange	0.1
heightshiftrange	0.1
fillmode	’reflect’

Data augmentation settings

Settings

_zoom range	1, 1.25
__width shift range	0.1
__height shift range	0.1
_shear range	0.2
_horizontal flip	True
_rotation range	40
_fill mode	’reflect’

Model and Optimizers

A model mlp is based on a collection of connected units nodes or neurons which are able to capture and represent input/output relationships. The knowledge of neural net- work is stored within inter-neuron connection which are known as weights. The architecture of mlp seems to play very important for the behavior of our network. There is not actually formula which can give us an quick answer about the number of nodes in the hidden layer or how many hidden layers we must use. Usually the number of hidden units can be smaller than the number of inputs. In addition when the number of training data are huge , we use mul- tiple hidden units, when we have little data 2 hidden units works best. When we have a simpler task on hidden unit.

Nowadays, research in deep neural network architectures has shown that many hidden layers are the best option for difficult object like handwritten character and face recogni- tion problems.In our approach we used a simple mlp Table 3, in order to fully understand the architecture we display it also in 4a. By using only the mlp and set other layers as non-trainable we would

distorting, adding a small amount of noise to, or have done a method called file- transfer. But since we select cropping a patch from an original image.Figure 5 the last layer to be trainable we implement Fine-Tuning displays an exam- ple of an augmented image of our

data set using the 2. In the following experiments we tried various augmentation techniques, such as enlargement of the zoom range, width shifting, rotation, height shifting and etc that are proposed in the paper of Jason Wang and Luis Perez [5]. Tables 1 and 2, show the features and their corresponding settings that we used in the experiment.

MLP

layers	Values
Dropout	0.4

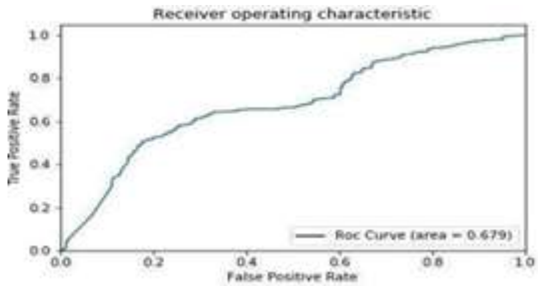
IV. RESULT AND DISCUSSION

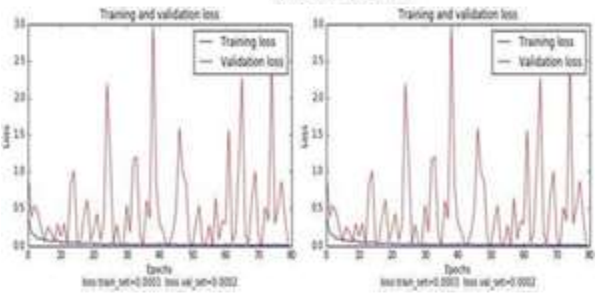
Early Experiment Alexnet

Our first experiment is a simple CNN that was based on the paper [4]. The network has kernel size (2,2) a stride of 1 we use as an optimizer the SGD(stochastic gradient descent). This network has been applied earlier to a hy- drangea dataset which was able to recognize specific flow- ers in the nature photos and it has achieved an accuracy on the test dataset 95% and loss 0.1919. The data augmen- tation that we used for this experiment are demonstrated in the Table 1.

The learning rate that we used it’s higher from our next experiments we used this kind of learning rate because we performed callbacks with reducing learn- ing rate. This is a keras operation in which every 20 epochs the learning rate decreases by a step we have defined, if the validation has not increased. We got better results using a more advance architecture of vgg16 more informations about the vgg16 you can find in [4].

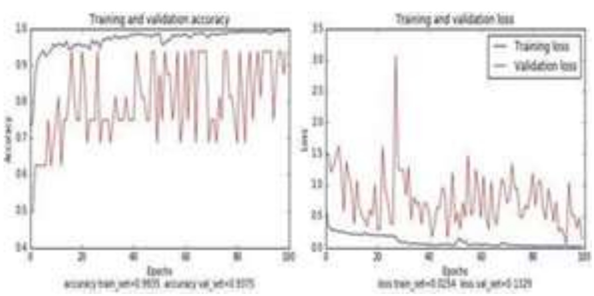
The result using our dataset was not satisfactory, since we achieved 81.7% accuracy in our validation dataset, but in the test set the accuracy was 63.46% and the loss 0.7770 (Table 4). In more details it recognized 226 normal photos





dataset compared to the training which seems to converge smoothly.The loss is the number of errors that has been done in training and validation set. After each iteration we

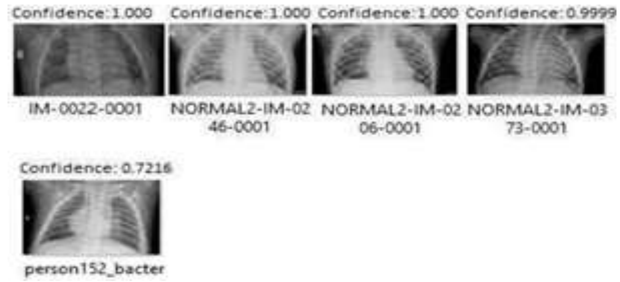
as pneumonia out of 234 and 2 as normal out of 390 pneu- monia. The performance of our classifier is demonstrated



Accuracy Loss

Most misclassified images

After ending the process of training and evaluation we printed the most misclassified CT Scan chest photos of our network. In Figure 11 and 12 you can observe the con- fidence and the most misclassified normal as pneumonia CT Scan images in the upper row and in the last the one mis- classified pneumonia. In order to make things more clear we display the actually title of the images above. The CT Scan normal images which classified as pneumonia was 57 and the pneumonia which classified as normal was 5 im- ages for the model of 11.



improve the accuracy and reduce the loss on the test set. We achieved the best performance by using smaller batch size. The actual accuracy on the test set was 90.71% and the loss 0.2867. We consider this score as a huge achieve- ment because the relative score of the top contributor using VGG16 was 90.01 score page. The ROC curve plot and the AUC score depicts that our classifier performance is good enough. At last a different approach was tried: adding more train- able layers.It seems more difficult to train the network and the layers to start learning to identify our images. The accu- racy and the loss that we obtained were not as satisfactory as we expected.

The approach of fine-tuning is a good option when the amount of data is small because, we are not training the entire network .In addition, a big part of the network is pre- trained and we do not need to trained it from scratch. The parameters that we needed to update are less so we do not need to consume a lot of time in order to have satisfactory results however since the network was pretrained fro dif- ferent tasks in our case too much was needed to adjust it in order to have results.At last we can train our network with- out using GPU and only with CPU. In our case we used GPU.

REFERENCES

1. Franc¸ois Chollet. Xception: Deep learning with depth- wise separable convolutions. CoRR, abs/1610.02357, 2016.
2. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pat- tern Recognition (CVPR), pages 770–778, 2016.

3. Andrej Karpathy and Li Fei-Fei. Deep visual-semantic alignments for generating image descriptions. 39, 12 2014.
4. Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hin- ton. Imagenet classification with deep convolutional neural networks. In Proceedings of the 25th Interna- tional Conference on Neural Information Processing Systems - Volume 1, NIPS'12, pages 1097–1105, USA, 2012. Curran Associates Inc.
5. Luis Perez and Jason Wang. The effectiveness of data augmentation in image classification using deep learn- ing. CoRR, abs/1712.04621, 2017.
6. Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recogni- tion. CoRR, abs/1409.1556, 2014.
7. Q. An, W. Chen, and W. Shao, "A Deep Convolutional Neural Network for Pneumonia Detection with Attention Ensemble," Diagnostics, vol. 14, no. 4, 2024.
8. L. Wu et al., "Pneumonia Detection Based on RSNA Dataset and Anchor-Free Deep Learning Detector," Scientific Reports, 2024.
9. D. Li, "Attention-Enhanced Architecture for Improved Pneumonia Detection in Chest X- ray Images," BMC Medical Imaging, 2024.
10. R. Kailasam and S. Balasubramanian, "Deep Learning for Pneumonia Detection Using CNN and YOLO," Human-Centric Intelligent Systems, 2025
11. W. Akbar et al., "Pneumonia Detection Using Diverse Neural Network Architectures," International Journal of Applied Mathematics and Computer Science, 2024.
12. High-Accuracy CNN-Based Pneumonia Detection with Explainability," Biomedical Signal Processing and Control, 2025.
13. S. Aljawarneh and R. Al-Quraan, "Enhanced CNN Model for Pneumonia Detection," Big Data Journal, 2025.
14. A. Buriboev et al., "CNN-Based Pneumonia Detection Using Fuzzy-Enhanced Dataset," Sensors, 2024.