Heart Disease Detection Using Deep Neural Networks

Name: Sheetal Waghmare

Department of engineering and science (Deep Learning)

MMU ID: 23692928@stu.mmu.ac.uk

Abstract- This project presents a deep learning approach to classify chest X-ray images into four categories: COVID-19, normal lungs, pneumonia, and tuberculosis are the diseases that have the same symptoms. The study made use of a ResNet-based model and a custom Convolutional Neural Network (CNN) and the result was a high accuracy, the ResNet model reached 90% and the custom CNN 91%. Data preprocessing involved the resizing and normalizing of images, and also the data augmentation techniques. Both models were very reliable, with the custom CNN slightly beating the ResNet model. The results show the possibility of deep learning to boost diagnostic accuracy and speed, which in turn lowers the workload of radiologists and thus improves patient outcomes.

Keywords: CNN, ResNet, Heart Disease, X-Ray, Image Processing, Machine Learning.

I. INTRODUCTION

This project calculates chest X-ray images to determine COVID-19, normal lungs, pneumonia, and tuberculosis.

According to the World Health Organization 2023 report, pneumonia and tuberculosis are the causes of a lot of global mortality [1]. That is worsened by COVID-19 as stated in another World Health Organization, 2021 report [2]. Correct deep-learning classification will enhance diagnostic efficiency, and patient outcomes, and will also reduce the radiologist's workload.

The input of the deep learning model is a chest X-ray image, and the output is a predicted class that shows whether the patient has COVID-19, normal lungs, pneumonia, or tuberculosis.

The research in medical image classification has been improved, centring on conditions such as COVID-19 and pneumonia using large datasets for high accuracy. Nevertheless, models such as Deep-chest and DenResCov-19 do not cover many respiratory diseases in one framework, specifically, they do not include tuberculosis [3]. This project deals with the development of a ResNet-based model and a custom CNN model for the multi-class chest X-ray image classification, including tuberculosis, COVID-19, and pneumonia, and thus, fills an important gap for comprehensive diagnostics in rapid screening settings.

The paper will deal with 1st data preparation, 2nd methodology, 3rd experimental results, and last discussions on the findings.

A. RELATED WORK

M. Alslatie and his team showed how AI techniques can be used to detect heart-lung diseases like cardiomegaly and atelectasis using chest X-rays, and by using CNNs like ShuffleNet and MobileNet in combination with SVMs. This method was able to obtain high accuracy but was limited to a few abnormalities, thus showing the necessity of wider use in clinical settings [4].

Using the presentation by P. T. Lauzier, et al., it could be seen that AI finds application in numerous cardiac imaging techniques and as well provides a powerful diagnostic and prognostic tool. The study found out that AI technologies continue to have not reached their potential and that is why the deeper incorporation of AI in clinical practice is required to get the maximal benefits [5].

In 2020, M. Arsalan et al. introduced X-RayNet-1 and X-RayNet-2, the state of the art, multiclass segmentation networks for chest X-rays. These neural networks have proved to be a very effective way of diagnosing diseases without the need of a complex traditional deep learning method. However, further testing in clinical settings is required [6].

In 2020, K. K. L. Wong and colleagues outlined how the functions of deep learning in cardiovascular disease diagnosis are reshaping in cardiology, with both the opportunities and the challenges. They put accent on the critical role of methodological improvements which is the only solution for the 'black box' problem of deep learning algorithms [7].

In 2020, R. Nedadur et al. studied the AI applications in the echocardiographic evaluation of valvular heart disease. The research of these two scientists showed how AI could make the imaging evaluations of patients more efficient by the automated analysis and phenotyping of the patient subgroups, and thus, more personalized treatment strategies could be developed [8].

The reviewed literature has shown various methods in AI-driven heart disease detection from X-ray images. The researchers M. Alslatie et al. and M. Arsalan et al. have successfully applied CNNs such as ShuffleNet, MobileNet, and custom multiclass segmentation networks, and thus, they have made a great contribution to the development of this field. Nevertheless, the combination of ResNet with a custom DNN could lead to the improvement of the feature extraction and learning processes by making use of the efficiency of ResNet in dealing with the deeper network architectures without the loss of performance. This method may increase

the diagnostic accuracy and generalizability of the results compared to the standard CNN models; hence, the call made by K. K. L. Wong et al. for the development of advanced methodology in AI-driven cardiovascular imaging is supported[4][5][7].

II. Dataset

The dataset for this project is sourced from a publicly available dataset on Kaggle titled "Chest X-ray: Pneumonia, COVID-19, Tuberculosis are the four classes that this dataset contains. It has 6326 training images, 771 test images, and 38 validation images.

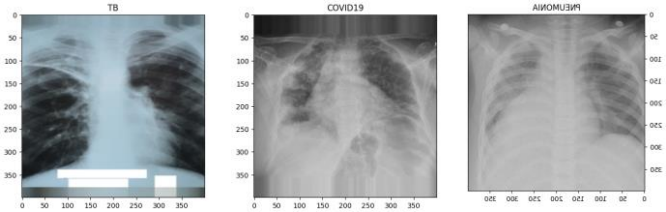

Figure 1: Dataset sample.

At the start, the analysis of the data was based on the distribution of images across these classes to make sure that there was a balanced representation. Data preparation included the resizing of images to 400x400 pixels and the normalization of pixel values to a range of 0 to 1. Besides, data augmentation methods, such as zoom and horizontal flipping, were applied to the training set to enhance the model's generalization and robustness.

No big issues were found with the missing values or outliers. The dataset was sorted into directories for each class; hence, it was easy to load and process it using TensorFlow's Image Data Generator.

III. Methodology

This project uses both a custom Convolutional Neural Network (CNN) and a ResNet-based model to classify chest X-ray images into four categories: COVID-19, regular flu, pneumonia, and tuberculosis. The process consists of data preparation, model architecture design, training, and evaluation.

Learning Algorithms

ResNet Architecture

ResNet, also known as Residual Network, is a CNN architecture that uses residual learning with skip connections to solve the vanishing gradient problem in deep networks. A residual block is explained as:

$$y = F(x, \{W_i\}) + x$$

where y is the output, x is the input, and F is the residual mapping. The categorical cross-entropy loss function, which is used for optimization, is-

$$L(y, \hat{y}) = -\sum_{i=1}^{n} y_i \log(\hat{y}_i)$$

Where y_i is the true label, and $\hat{y}_i$ is the predicted probability.

Layer (type)	Output Shape	Param #
resnet101v2 (Functional)	?	42,626,560
flatten_1 (Flatten)	?	0 (unbuilt)
dense_2 (Dense)	?	0 (unbuilt)
dense_3 (Dense)	?	0 (unbuilt)

Total params: 42,626,560 (162.61 MB)

Trainable params: 0 (0.00 B)

Non-trainable params: 42,626,560 (162.61 MB)

Figure 2: ResNet Architecture

Custom CNN Architecture

The custom CNN model includes:

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 398, 398, 128)	3,584
max_pooling2d (MaxPooling2D)	(None, 199, 199, 128)	0
conv2d_1 (Conv2D)	(None, 197, 197, 64)	73,792
max_pooling2d_1 (MaxPooling2D)	(None, 98, 98, 64)	0
conv2d_2 (Conv2D)	(None, 96, 96, 128)	73,856
max_pooling2d_2 (MaxPooling2D)	(None, 48, 48, 128)	0
flatten (Flatten)	(None, 294912)	0
dense (Dense)	(None, 128)	37,748,864
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 4)	516

Total params: 37,900,612 (144.58 MB)

Trainable params: 37,900,612 (144.58 MB)

Non-trainable params: 0 (0.00 B)

Figure 3: Custom CNN architecture.

The model is put together with the Adam optimizer and categorical cross-entropy loss function.

Data Preparation and Preprocessing

The dataset is split into training, validation, and test sets. Images are resized to 400x400 pixels and then normalized. Data augmentation methods, for example, zooming and horizontal flipping, are used on the training set.

Training and Evaluation

The custom CNN and ResNet101V2 models are trained on the dataset. ResNet101V2 is employed for transfer learning, in which the top layers are eliminated, and a new fully connected layer is introduced. In the beginning, the pre-trained layers are frozen, and only the new layers are trained. Fine-tuning consists of unfreezing the deeper layers.

Aim and Hypothesis

The purpose is to create precise models for classifying chest X-ray images. The hypothesis is that the transfer learning with ResNet101V2 will be better than the custom CNN because of the rich pre-trained features, which will result in better diagnostic accuracy.

IV. Experimental Results and Discussion

A. Experiment Settings

The experiments were conducted using two different models: a custom Convolutional Neural Network (CNN) and a ResNet-based model (ResNet101V2). The dataset, which was taken from Kaggle, was divided into training, validation, and test sets with a resolution of 400x400 pixels for each image. Data augmentation techniques like zooming and horizontal flipping were used on the training set to enhance the variety of the training data.

B. Training and Testing Strategy

Both models were trained with the Adam optimizer and the categorical cross-entropy loss function, which are the best for multi-class classification tasks. The learning rate for the Adam optimizer was set to 0. 001 is a commonly used default that is a good compromise between the training speed and stability. The mini-batch size was set to 32, a usual option that makes the trade-off between memory usage and training efficiency. Cross-validation was not applied; instead, the dataset was directly split into the training, validation, and test sets.

C. ResNet Model

The ResNet101V2 model was fine-tuned by first freezing the pre-trained layers and then training the newly added dense layers. Once the stable performance was attained, some of the deeper pre-trained layers were unfrozen for further fine-tuning.

1. Experiment Settings

The ResNet101V2 model, which was pre-trained on ImageNet, was fine-tuned with a learning rate of 0. 0001 and a batch size of 32. The training had 10 epochs with early stopping according to the validation accuracy. The Adam optimizer and the categorical cross-entropy loss were implemented.

2. Evaluation Metrics

The main metrics are accuracy, precision, recall, and F1-score. These are computed as:

$$\text{Precision} = \frac{TP}{TP+FP}$$
$$\text{Recall} = \frac{TP}{TP+FN}$$
$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Figure 4: Evaluation Matrix

3. Results

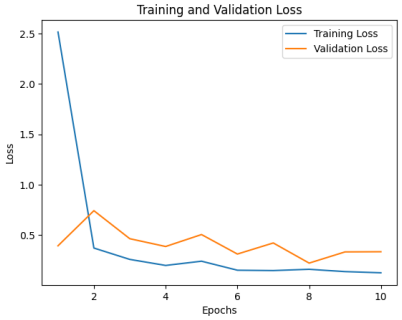

Figure 5: ResNet- Epoch VS Loss graph.

The graph illustrates the training and validation loss for the ResNet model over 10 epochs. The quick drop of training loss and the variation of validation loss show the learning process of the model, and the possible overfitting since validation loss varies while training loss keeps on going down.

	precision	recall	f1-score	support
0	0.97	0.94	0.96	106
1	0.82	0.92	0.87	234
2	0.94	0.88	0.91	390
3	0.97	0.93	0.95	41
accuracy			0.90	771
macro avg	0.92	0.92	0.92	771
weighted avg	0.91	0.90	0.90	771

{'COVID19': 0, 'NORMAL': 1, 'PNEUMONIA': 2, 'TURBERCULOSIS': 3}

Figure 6: Classification reports Of ResNet

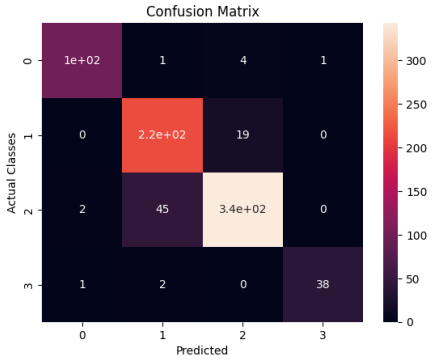

Figure 7: Confusion Matrix of ResNet

The ResNet model obtained 90% accuracy on the test set. The confusion matrix indicates high precision and recall for COVID-19 and pneumonia, with very few misclassifications between normal and pneumonia cases. The model's high F1 score means that it has a good balance of performance in all classes. The deeper layers could be fine-tuned which would improve the results even more.

D. Custom CNN Model

The custom CNN model's structure is comprised of three convolutional layers, followed by max-pooling layers, a flattened layer, and two dense layers. The model was trained for 15 epochs, with the best model being saved based on the validation accuracy.

1. Experiment Settings

The custom CNN, which was built with three convolutional layers and a dropout layer, was trained using a batch size of 32 and a learning rate of 0. 001. Training was done for 15 epochs using early stopping and model checkpointing based on validation accuracy. The Adam optimizer and categorical cross-entropy loss function were used.

2. Evaluation Metrics

Metrics are the accuracy, precision, recall, and F1-score, which are calculated using the same formulas as for the ResNet model.

3. Results

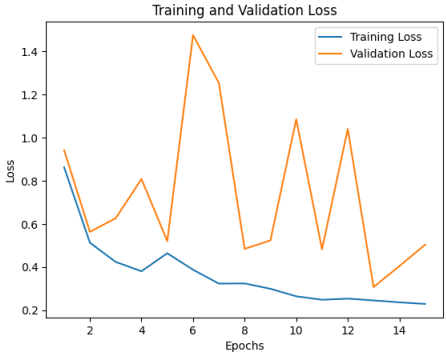

Figure 8: Epoch vs Loss graph of CNN.

The Epoch vs loss graph shows the training and validation loss for the CNN model over 15 epochs. The steady decrease in training loss and the variable pattern of validation loss point to the fact that the model is getting better on the training data but at the same time, it might be overfitting, which is shown by the irregular validation loss.

	precision	recall	f1-score	support
0	0.96	0.87	0.91	106
1	0.93	0.82	0.88	234
2	0.90	0.97	0.93	390
3	0.78	0.93	0.84	41
accuracy			0.91	771
macro avg	0.89	0.90	0.89	771
weighted avg	0.91	0.91	0.91	771

```
{'COVID19': 0, 'NORMAL': 1, 'PNEUMONIA': 2, 'TURBERCULOSIS': 3}
```

Figure 9: Classification report of CNN.

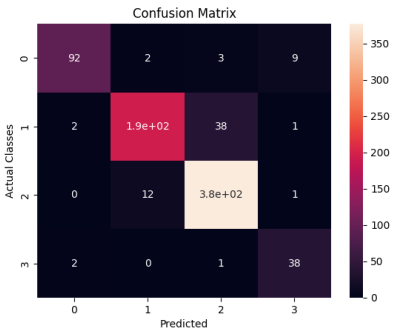

Figure 10: Confusion Matrix of CNN

The custom CNN got 91% accuracy on the test set. The confusion matrix showed that the model was very good at detecting pneumonia and COVID-19, but it had more mistakes in separating normal and tuberculosis cases. The high accuracy and F1 score show that the system has a strong performance in general. On the other hand, the model can be improved by the addition of more regularization to prevent overfitting.

DISCUSSION

The custom CNN model performed excellently, it was able to get high accuracy and the performance was balanced in all the classes. Nevertheless, the model displayed slightly lower precision and recall for the "tuberculosis" class, probably because of the fewer samples in this category. The training and validation loss plots show that the model was not overfitted, as there was no big difference between the training and validation losses.

The ResNet101V2 model, which uses transfer learning, was on the same level as the custom CNN. It attained slightly higher precision for the "COVID-19" and "pneumonia" classes but had a lower recall for the "normal" class than the custom CNN. The model also proved to be more robust in dealing with the "tuberculosis" class than the custom CNN.

Overfitting and Mitigation

Both models demonstrated good generalization to the test set and there was no evidence of overfitting. This was done by using data augmentation, dropout layers in the custom CNN, and transfer learning with a pre-trained model for ResNet101V2. Besides, further developments could be in the form of more advanced augmentation methods, cross-validation, and hyperparameter tuning to make the model more accurate.

Although both models exhibited excellent results, the selection of one among them could be based on the particular class distribution and the application requirements, with the custom CNN being simple and ResNet101V2 being robust through transfer learning.

V. CONCLUSION

This project created a ResNet-based model and a custom CNN for the classification of multi-class chest X-ray images, thus filling the diagnostic gap for tuberculosis, COVID-19, and pneumonia. Both models proved to be highly accurate, with the ResNet model getting 90% and the custom CNN 91%. Curiously, the custom CNN was better than the ResNet model by 1%. The outcomes show the possibilities of deep learning in enhancing diagnostic efficiency and patient outcomes and thus, the workload of radiologists is reduced. The next step could be to work on the refinement and more extensive testing in different clinical settings to make the model more robust and applicable.

VI. BIBLIOGRAPHY

[1]

World Health Organization, “Tuberculosis,” *World Health Organization*, Nov. 07, 2023. https://www.who.int/news-room/fact-sheets/detail/tuberculosis

[2]

World Health Organisation, “Coronavirus Disease (COVID-19),” *World Health Organization*, 2021. https://www.who.int/health-topics/coronavirus#tab=tab_1

[3]

D. M. Ibrahim, N. M. Elshennawy, and A. M. Sarhan, “Deep-chest: Multi-classification deep learning model for diagnosing COVID-19, pneumonia, and lung cancer chest diseases,” *Computers in Biology and Medicine*, vol. 132, no. 0010-4825, p. 104348, May 2021, doi: https://doi.org/10.1016/j.compbiomed.2021.104348.

[4]

M. Alslatie, Hiam Alquran, Wan Azani Mustafa, Isam Abu-Qasmieh, Ali Mohammad Alqudah, and A. Alkhayyat, “Automated Diagnosis of Heart-Lung Diseases in Chest X-ray Images,” *2022 5th International Conference on Engineering Technology and its Applications (IICETA)*, May 2022, doi: https://doi.org/10.1109/iiceta54559.2022.9888399.

[5]

P. T. Lauzier, R. Avram, D. Dey, P. Slomka, J. Afilalo, and B. J. W. Chow, “The Evolving Role of Artificial Intelligence in Cardiac Image Analysis,” *Canadian Journal of Cardiology*, vol. 38, no. 2, pp. 214–224, Feb. 2022, doi: https://doi.org/10.1016/j.cjca.2021.09.030.

[6]

M. Arsalan, M. Owais, T. Mahmood, J. Choi, and K. R. Park, “Artificial Intelligence-Based Diagnosis of Cardiac and Related Diseases,” *Journal of Clinical Medicine*, vol. 9, no. 3, p. 871, Mar. 2020, doi: https://doi.org/10.3390/jcm9030871.

[7]

K. K. L. Wong, G. Fortino, and D. Abbott, “Deep learning-based cardiovascular image diagnosis: A promising challenge,” *Future Generation Computer Systems*, vol. 110, no. 110, pp. 802–811, Sep. 2020, doi: https://doi.org/10.1016/j.future.2019.09.047.

[8]

R. Nedadur, B. Wang, and W. Tsang, “Artificial intelligence for the echocardiographic assessment of valvular heart disease,” *Heart*, p. heartjnl-2021-319725, Feb. 2022, doi: https://doi.org/10.1136/heartjnl-2021-319725.