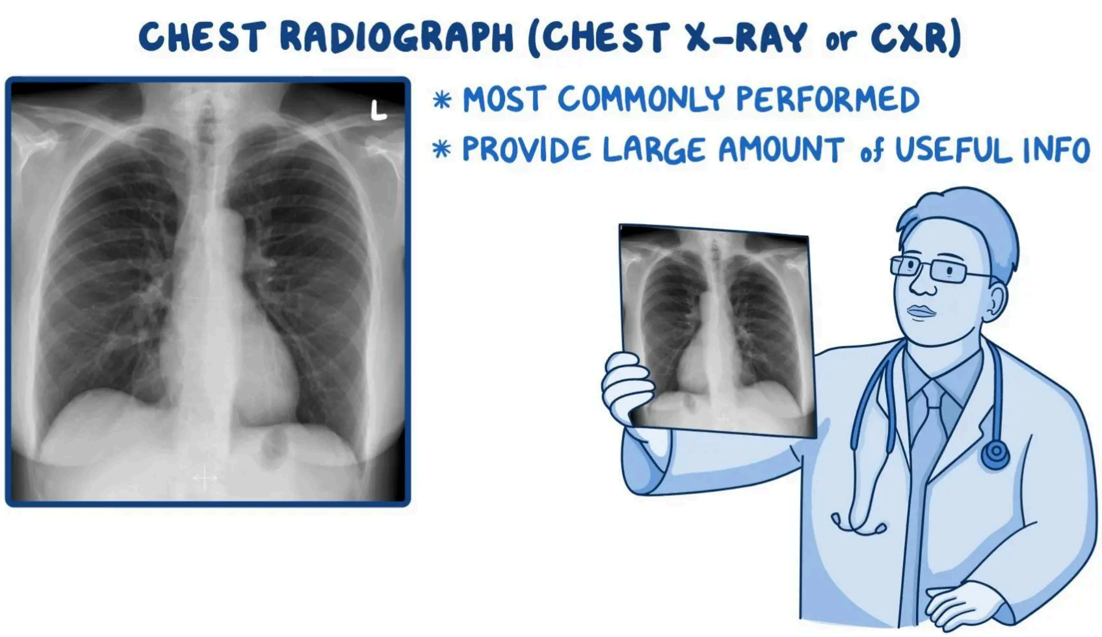
Medical Image Classification using Convolutional Neural Network



Minh Hien Ha(2060207)

Aleksandar Kirilov(2059920)

Beksultan Maatkulov(2056532)

Melis Zhorobaev(1992865)

ABSTRACT

This project involves the development of a Convolutional Neural Network (CNN) for classifying medical images into three categories: Pneumonia, Normal, and COVID-19. The dataset is preprocessed and augmented to enhance model performance, and the model is trained and evaluated using PyTorch. This report details the methods and outcomes of this classification task, providing insights into the model's training process, evaluation metrics, and overall performance.

KEYWORDS

Medical Imaging, Convolutional Neural Network, PyTorch, Data Augmentation, Image Classification

1. INTRODUCTION

The aim of this project is to build a CNN capable of classifying chest X-ray images into three categories: Pneumonia, Normal, and COVID-19. The project is structured into several stages including data preprocessing, data augmentation, model definition, training, and evaluation.

2. DATA PREPROCESSING

2.1 LOADING DATA

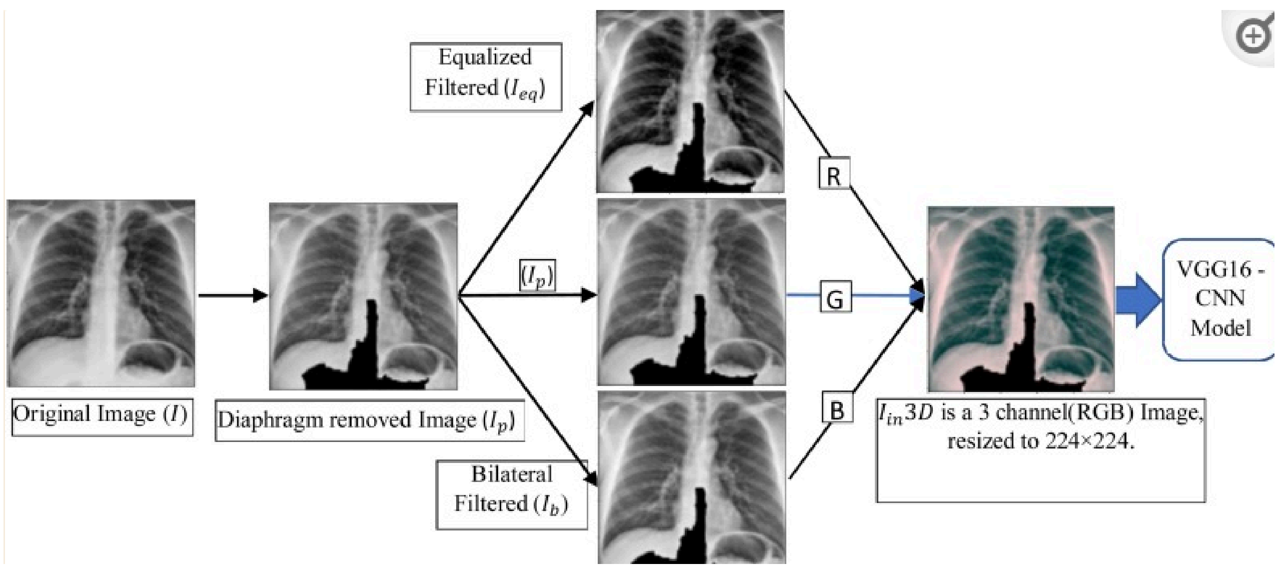
Data Loading and Preprocessing: The code loads and preprocesses image data from directories for three classes: 'pneumonia', 'normal', and 'covid'. Images are converted to grayscale and resized to 150x150 pixels and 224x224 pixels respectively to the different models.

Splitting Dataset: Dataset are splitted into training, validation, testing folders with 80/20 techniques

2.2 PREPROCESSING

Image Preprocessing: Includes diaphragm removal , which is an algorithm detects the maximum (the brightest - Vmax) and minimum (the darkest - Vmin) pixel value of the image, then uses a threshold T=Vmin + 0.9× (Vmax−Vmin) to segment the original image into a binary image then we remove it from the images. As the diaphragm region may have a negative impact on the result(4) .Then we split the image to three different channels and apply the bilateral filter and histogram equalization to it.

(see figure 2.1 and 2.2).



2.3 AUGMENTATION

Data augmentation was implemented to increase the used dataset's size by applying various image transformations, such as randomly zooming, flipping, shifting vertically and horizontally.

3. MODEL DEFINITION

3.1 CNN ARCHITECTURE

Two CNNs were created to correctly classify the X-ray Image, specifically:

1. Model 1: The CNN model has four Convolutional layers:

   64 filters, (3, 3) stride

   64 filters, (3, 3) stride

   128 filters, (3, 3) stride

   128 filters, (3, 3) stride

   Each convolutional layer is followed by Batch Normalization and a (2, 2) Max Pooling layer to reduce the spatial dimensions. After the convolutional layers, a Flatten layer is used, followed by two Dense layers: the first with 512 units and a Dropout for regularization, and the final layer with 3 SoftMax units for classification. This model uses a 224x224x1 input shape, ReLU (hidden layers) and SoftMax (output layer) activation functions, an Adam optimizer, and the Categorical Cross Entropy loss function.

2. Model 2: The CNN has four Convolutional layers:

   32 filters, (3, 3) kernel

   64 filters, (3, 3) kernel

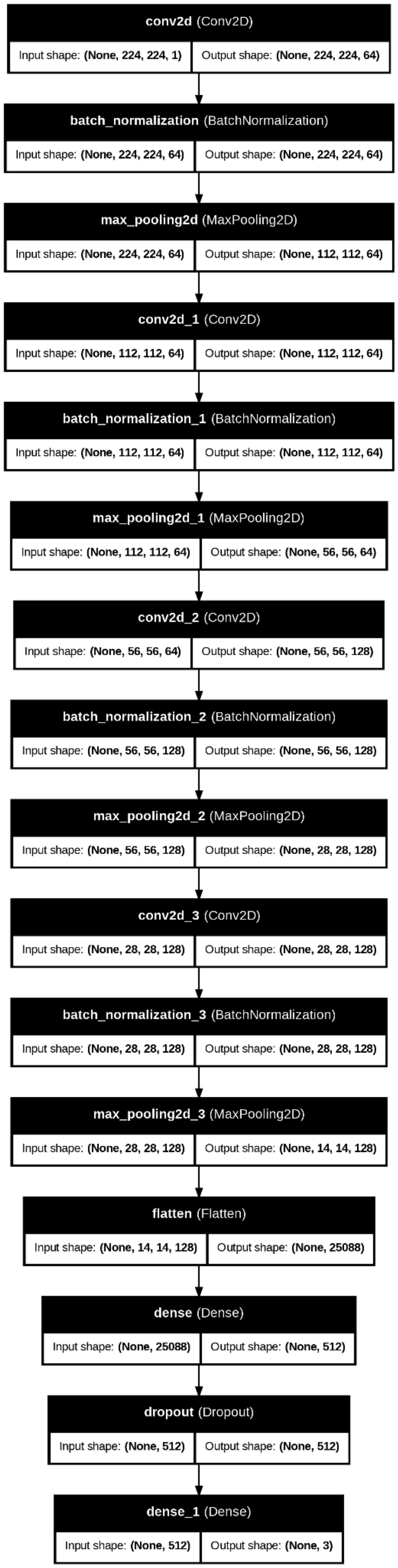
   128 filters, (3, 3) kernel

   256 filters, (3, 3) kernel

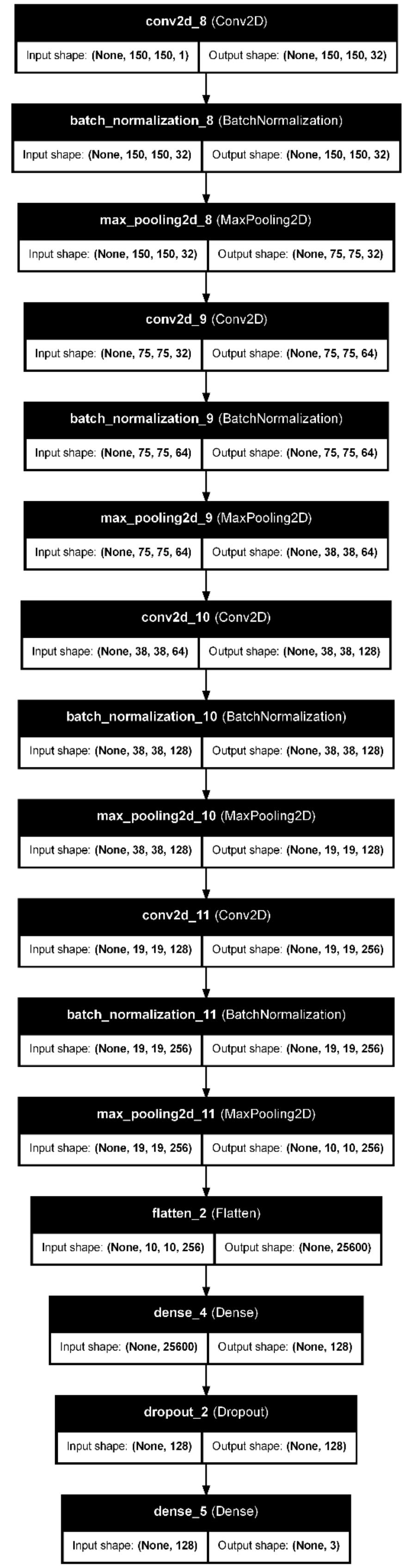
   Each Conv layer is followed by Batch Normalization and (2, 2) Max Pooling layers.

   The network ends with a Flatten layer, a Dense layer with 128 units, and a Dropout layer (0.5), followed by a SoftMax layer with 3 units for classification.

   This model uses a 150x150x1 input shape, ReLU (hidden layers) and SoftMax (output layer) as activation functions, an RMSprop optimizer, and the Sparse Categorical Cross Entropy loss function.



Model 1



Model 2

4. TRAINING AND EVALUATION

4.1 TRAINING AND HYPERPARAMETER

Hyperparameters values for the CNN were:

Model 1 version 1 (using normal dataset):

- 50 epochs
- Batch size 32
- Using Learning Rate Reduction with 0.0010 as the starting learning rate

Model 1 version 2 (using preprocessed dataset) :

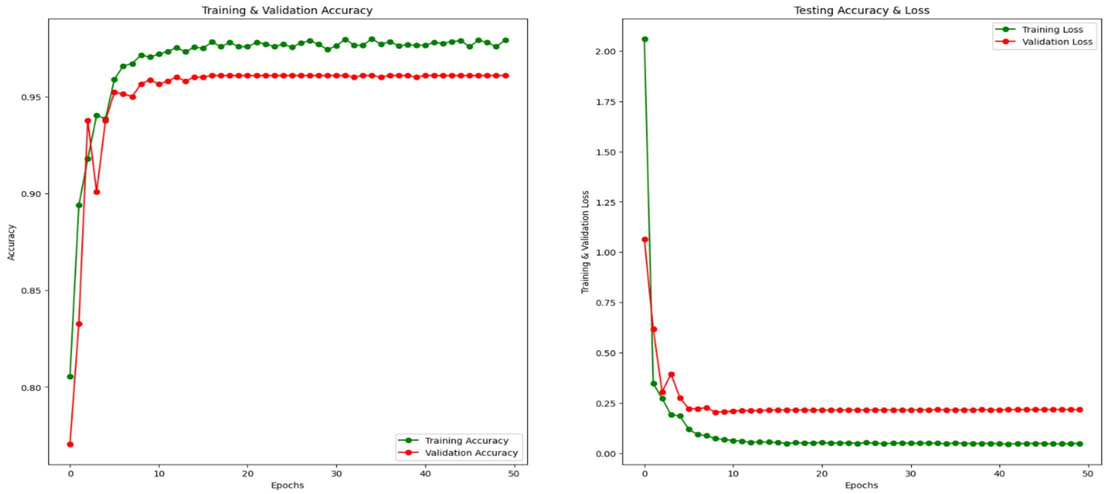
- 50 epochs
- Batch size 32
- Using Learning Rate Reduction with 0.0010 as the starting learning rate
- Early stopping was implemented to prevent overfitting.

Model 2:

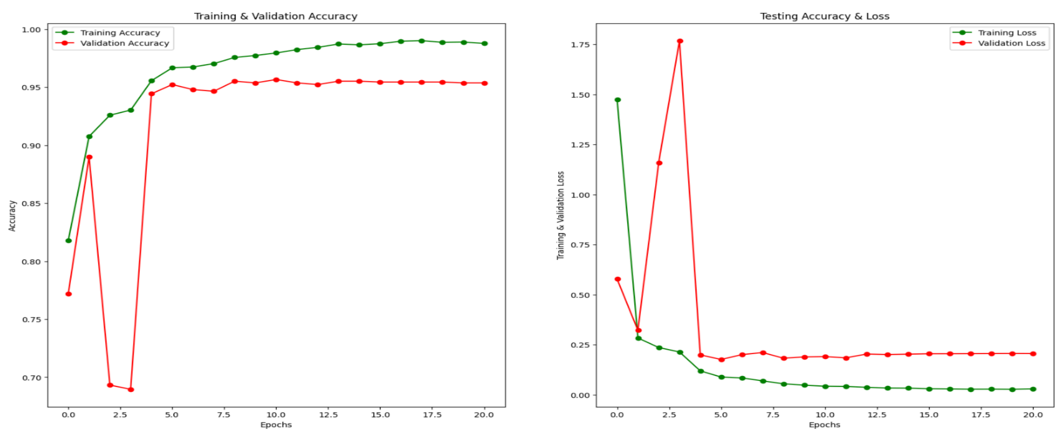
- 50 epochs
- Batch size 32
- Using Learning Rate Reduction with 0.0010 as the starting learning rate
- Early stopping and learning rate reduction were implemented to prevent overfitting.

4.2 TESTING AND EVALUATION

- For the first model first version, a maximum of 99.55% accuracy during training with loss function of 0.0179 and 95.68% accuracy on validation with loss function of 0.265 were achieved.
- For the first model second version, *surprisingly*, a maximum of 97.65% accuracy during training with loss function of 0.0556 and 95.51% accuracy on validation with loss function of 0.1825 were achieved.
- For the second model, a maximum of 87.60% accuracy during training with loss function of 0.3894% and 85.01% accuracy on validation with loss function of 0.4332 were achieved.



Model 1 version 1



Model 1 version 2

Model 2

5. CONCLUSION

This project demonstrates a comprehensive workflow for medical image classification using a CNN. Key steps include data preprocessing, data augmentation, defining the CNN architecture, and training and evaluating the model. The project highlights the importance of proper data handling, model design, and performance evaluation in building effective deep learning models for medical image analysis.

LINK AND PAPERS THAT HELPED

1) https://www.kaggle.com/code/madz2000/pneumonia-detection-using-cnn-92-6-accuracy#Data-Augmentation
2) https://www.kaggle.com/datasets/amanullahasraf/covid19-pneumonia-normal-chest-xray-pa-dataset
3) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8982897/
4) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7510591/