Enhancing Pneumonia Detection in Chest X-ray Images Using Neural Networks with Fourier Transformation Mask Optimization

Zhixiang Yao

Department of Biomedical Engineering, Duke University, Durham 27708, USA; zy154@duke.edu

Introduction

Pneumonia is a severe respiratory condition characterized by the inflammation of lung tissue. It continues to be a major barrier in the healthcare industry. Swift and accurate diagnosis of the condition is essential, particularly given its varied manifestations in chest X-rays.[1] Traditional diagnostic methods often rely on the expertise of radiologists, despite facing challenges such as subjective and inconsistent interpretation.[2] Consequently, there is a growing inclination to use computational techniques, particularly neural networks, in order to enhance the precision of detection.

Recent advancements in deep learning have shown considerable promise in automating the analysis of medical images. Neural networks have proven to be highly effective in the identification and interpretation of chest X-ray images for the purpose of pneumonia detection.[1] The Fourier Transformation is a prominent technique for enhancing these networks. It is a mathematical method that translates images into the frequency domain. This modification enhances the ability to amplify certain traits and diminish the visibility of others, hence enhancing the network's ability to detect patterns that signal the existence of pneumonia.[2]

Furthermore, recent studies have examined the integration of trainable Fourier transformation masks into neural networks. This innovative approach allows the network to adaptively focus on relevant frequencies, hence potentially improving the accuracy of pneumonia classification from chest X-rays.[1] This project investigates the application of neural networks augmented with Fourier Transformation for the detection of pneumonia, highlighting its efficiency and capacity to revolutionize diagnostic techniques in the biomedical field.

Figure 1: An example of chest X-ray Image.

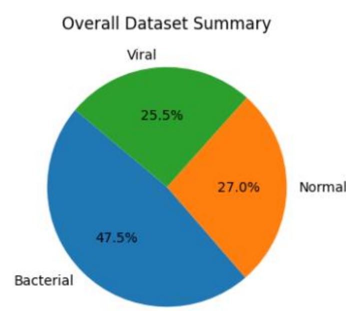

Figure 2: Summary of dataset.

The main variables considered in this project are the chest X-ray images and their corresponding labels. The input variable is the X-ray image, a digital depiction of the chest, that is processed by the neural network. Images undergo pre-processing and resizing to ensure uniform dimensions, ensuring uniformity. The output variable represents the categorization outcome, assigning each image to one of three categories: normal, viral pneumonia, or bacterial pneumonia. The neural network's objective is to examine the input image and generate a classification that shows the most probable condition portrayed in the X-ray.

The dataset used in this project consists of 5,856 chest X-ray images, which are divided into three categories: normal lungs, viral pneumonia, and bacterial pneumonia.[3] An example of chest X-ray image is presented as Figure 1. The dataset comprises 1,583 images illustrating healthy lungs, 1,493 images representing instances of viral pneumonia, and 2,780 images showcasing cases of bacterial pneumonia (Figure 2). The diverse selection enables the application of neural network models in classification tasks, providing a comprehensive basis for training and assessing the model's ability to differentiate between these three essential lung diseases. The dataset's wide variety and extensive size make it a valuable resource for the advancement and assessment of advanced image processing algorithms in the field of biomedical engineering.

Methods

Data Preprocessing

The project started by conducting preprocessing on the dataset of 5,856 chest X-ray images. Every image is uniformly downsized to a dimension of 256x256 pixels in order to maintain consistency. Standardizing the input data is an essential preparatory step for neural network training.

Neural Network Architecture

The basal neural network architecture, SimpleANN, forms the foundation of the analysis. This architecture comprises a sequence of linear layers with ReLU activation functions, with the network output tailored for binary or multi-class classification. The network is optimized using the Adam optimizer, with a learning rate of 0.0004. For binary classification tasks, a binary cross-entropy loss function (BCELoss) is used, while multi-class tasks utilize a cross-entropy loss function (CrossEntropyLoss).

Fourier Transform Integration

An exceptional feature of this project is the incorporation of Fourier Transform methodologies. The neural network incorporates a trainable layer that applies Fourier Transforms to the input images. The purpose of this layer is to convert the images into the frequency domain, enabling the network to concentrate on specific frequency components that are essential for accurately categorizing pneumonia. The Fourier Transform layer is flexible, as its settings are acquired through the training process. This project investigates various Fourier Transform layers, specifically focusing on trainable phase variation, low-pass filtering (LPF), and high-pass filtering (HPF). The layers are analyzed and contrasted.

Training and Validation

The dataset is partitioned into training, validation, and testing sets at a 70:15:15 ratio. The model undergoes training using the training set, during which it acquires the ability to identify patterns that are linked to each category. The validation set is employed to optimize the model's hyperparameters and mitigate overfitting. The testing set offers an impartial assessment of the model's effectiveness.

Five unique machine learning experiments are carried out:

1. Binary Classification (Healthy vs. Unhealthy Lungs): The base Artificial Neural Network (ANN) is employed to categorize lungs as either healthy or unhealthy.

2. The baseline artificial neural network (ANN) categorizes images into three distinct classifications: bacterial pneumonia, normal, and viral pneumonia.

3. The incorporation of trainable phase masks in the frequency domain into a modified artificial neural network (ANN) that utilizes Fourier Transformation.

4. The implementation of a Low-Pass Filter (LPF) involves enhancing the Artificial Neural Network (ANN) with a filter that operates in the Fourier domain.

5. The implementation of a High-Pass Filter (HPF) enhances the Artificial Neural Network (ANN) by using it in the Fourier domain.

Results and Discussion

Each model was quantitatively analyzed using standard classification metrics, and the confusion matrices provided insights into their specific strengths and weaknesses.

The ANN achieved high performance in distinguishing healthy from unhealthy lungs, with an accuracy of 94.43%, precision of 95.91%, recall of 96.07%, and an F1 score of 95.99%. However, the three-category classification is more challenging. The ANN demonstrated an overall accuracy of 77%, with precision, recall, and F1 scores varying across categories. The model was most effective in classifying bacterial pneumonia. Additionally, the inclusion of a trainable phase mask in the frequency domain yielded an overall accuracy of 76%, indicating the potential of frequency domain transformations in enhancing classification accuracy. The application of an LPF showed an accuracy of 78%. It was particularly effective in identifying bacterial pneumonia, showcasing the utility of frequency domain filtering. Yet, the HPF model achieved an accuracy of 65%, indicating that while useful, high-frequency components alone might not be sufficient for robust classification in this context.

The Fourier Transform's role in enhancing feature extraction and classification was evident, especially in models with the Fourier phase mask and LPF/HPF implementations. These models showed a nuanced ability to focus on specific frequency components, which is crucial for classifying pneumonia in chest X-rays. Examples of the Fourier filtering in a 9-plot manner on 3 different cases where Fourier Transformation composed a layer of the neural network (Figure 3). Examples of prediction results are provided in Figure 4.

Conclusion

The project demonstrates the efficacy of neural networks in classifying chest X-ray images, with a notable improvement observed when incorporating Fourier Transform techniques. The exploration of different Fourier Transform applications (phase masks, LPF, and HPF) highlighted their potential in enhancing diagnostic capabilities in biomedical imaging. Further research could explore the combination of these techniques or delve into more complex network architectures for improved performance.

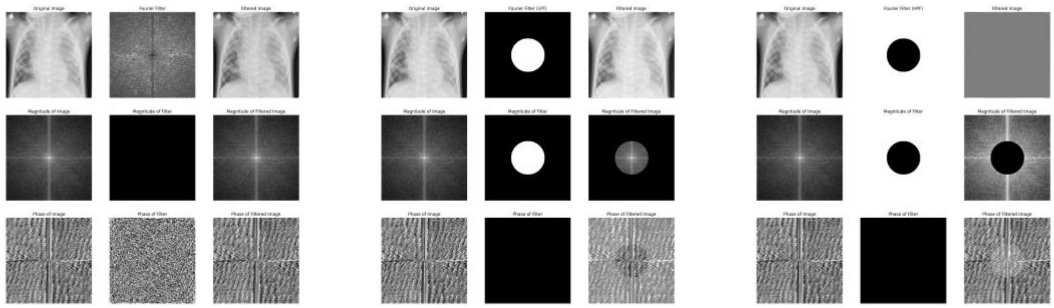

Figure 3: Examples of (a) implementing Fourier filter with trainable phase, (b) implementing Fourier filter with LPF, and (c) implementing Fourier filter with HPF.

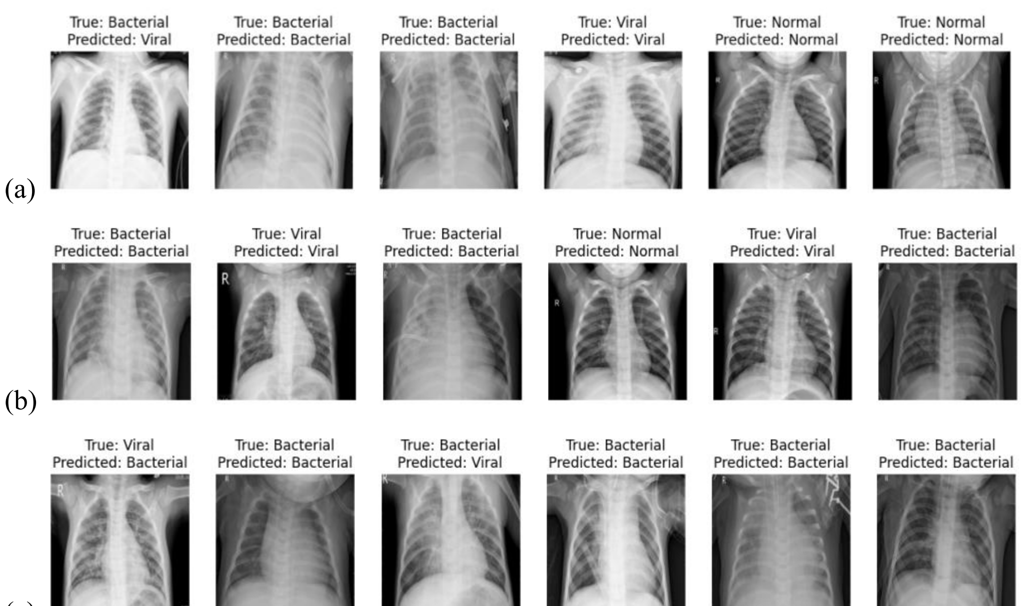

Figure 4: Example of predictions with (a) trainable phase mask, (b) LPF implementation, and (c) HPF implementation

References

1. Mujahid, M., Rustam, F., Álvarez, R., Luis Vidal Mazón, J., Díez, I. de, & Ashraf, I. (2022). Pneumonia classification from X-ray images with inception-V3 and Convolutional Neural Network. *Diagnostics*, *12*(5), 1280. https://doi.org/10.3390/diagnostics12051280

2. Alapat, D. J., Menon, M. V., & Ashok, S. (2022). A review on detection of pneumonia in chest X-ray images using neural networks. Journal of Biomedical Physics and Engineering, 12(6). https://doi.org/10.31661/jbpe.v0i0.2202-1461

3. Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C. S., Liang, H., Baxter, S. L., McKeown, A., Yang, G., Wu, X., Yan, F., Dong, J., Prasadha, M. K., Pei, J., Ting, M. Y. L., Zhu, J., Li, C., Hewett, S., Dong, J., Ziyar, I., … Zhang, K. (2018). Identifying medical diagnoses and treatable diseases by image-based Deep Learning. Cell, 172(5). https://doi.org/10.1016/j.cell.2018.02.010