# Advanced Multimodal COVID-19 Detection Using Chest X-Ray Imaging

Hongdu WEI
*PROJ-H402 PROJ-H419

***Abstract*—This study introduces an advanced approach to COVID-19 detection leveraging deep learning to classify chest X-ray (CXR) images into COVID-19, Viral Pneumonia, and Normal categories. Addressing the limitations of RT-PCR tests, our work utilizes CXR imaging for its potential in rapid diagnostics. Our methodology integrates Convolutional Neural Networks (CNNs), Fully Connected Networks (FCNs), and a comparative analysis of traditional machine learning models, establishing a robust diagnostic framework. The dataset, encompassing images from three categories, underwent extensive preprocessing to ensure model robustness. Our evaluation framework, using metrics such as accuracy, precision, recall, and F1 score, indicates superior performance of CNNs, with FCNs further enhancing classification accuracy. A comparative analysis of machine learning models offers insights into the most effective methodologies for CXR image analysis. Our project not only marks a significant advancement in medical AI but also sets the stage for future exploration and innovation within AI and healthcare, aiming to develop a universally applicable framework that can adapt to new pathogens and respiratory illnesses. By inviting collaboration and sharing our findings, we seek to leverage AI's full potential in healthcare diagnostics, highlighting the project's impact on the global health landscape.**

***Index Terms*—COVID-19 detection, chest X-ray imaging, deep learning, comparative analysis, healthcare impact.**

## I. Introduction

The COVID-19 global pandemic, originating from the novel SARS-CoV-2 virus, has imposed unprecedented challenges upon worldwide public health infrastructure, economic stability, and healthcare delivery systems. The urgent demand for swift and precise diagnosis is paramount in stemming the spread of the virus and ensuring effective patient management. Although reverse transcription-polymerase chain reaction (RT-PCR) tests have emerged as the primary modality for COVID-19 detection, their widespread utilization is hampered by limitations such as restricted availability, prolonged turnaround times, and logistical complexities. Consequently, there exists a pressing necessity for alternative diagnostic modalities that offer expeditious, reliable, and scalable solutions. Radiographic imaging, notably chest X-rays (CXR), emerges as a promising avenue due to its pervasive accessibility and the potential for prompt disease identification.

In this context, we present a pioneering deep learning-based methodology tailored for the multifaceted classification of CXR images into three distinct categories: COVID-19, Viral Pneumonia, and Normal. Leveraging the inherent capabilities of deep learning algorithms, our approach seeks to harness the intricate patterns embedded within CXR images to facilitate accurate disease discrimination and diagnosis. By capitalizing on the extensive availability of CXR imaging data, we endeavor to furnish healthcare professionals with a powerful tool for expeditious and precise COVID-19 detection.

Through the integration of advanced machine learning techniques, our methodology aims to bridge the gap between conventional diagnostic modalities and contemporary AI-driven approaches. By employing a combination of Convolutional Neural Networks (CNNs) and Fully Connected Networks (FCNs), our framework endeavors to extract and elucidate intricate features embedded within CXR images, thereby facilitating comprehensive disease classification. Furthermore, our approach encompasses a meticulous comparative analysis of traditional machine learning models, including Logistic Regression, Decision Tree Classifiers, and Random Forest Classifiers, to discern the most effective diagnostic methodologies in the context of COVID-19 detection.

The comprehensive evaluation of our diagnostic framework encompasses an array of performance metrics, including accuracy, precision, recall, and F1 score, to ascertain the efficacy and reliability of our proposed methodology. Preliminary findings indicate the superior performance of our deep learning models, particularly CNNs, in facilitating accurate disease classification. Additionally, our comparative analysis sheds light on the inherent strengths and limitations of traditional machine learning models, thereby guiding future research endeavors and model selection processes.

Beyond immediate diagnostic applications, our research underscores the transformative potential of AI-driven approaches in revolutionizing medical diagnostics. By fostering an open-source ecosystem and facilitating collaboration within the global research community, we endeavor to accelerate the development and deployment of AI-driven diagnostic tools, thereby enhancing healthcare delivery and mitigating the impact of global health crises.

## II. Dataset and Preprocessing

Our meticulously curated dataset encompasses a rich array of CXR images spanning three pivotal categories: COVID-19, Viral Pneumonia, and Normal as demonstrated in Fig.1. To fortify our model's resilience and tackle the intricacies stemming from diverse imaging conditions, we have embarked on an exhaustive preprocessing journey. This includes but is not limited to image resizing, normalization, and augmentation techniques such as rotation, zooming, and flipping. In order to study the characteristics of the dataset in more depth, we also performed a visualization analysis(Fig.2), which allowed us to better understand the diversity and challenging nature of the dataset by plotting sample images for each category. This step provided us with valuable insights into the internal structure of the dataset, which guided the optimization of model training. By looking at these visualized images, we can clearly see that the images for each category exhibit significant variability and diversity. This diversity is very beneficial for modeling because the significant variability between the different categories helps the model to better distinguish between them. These meticulous efforts are geared towards bolstering model robustness, ensuring its adeptness in navigating the complexities inherent in medical imaging datasets. At the end of this section we get the enhanced image as shown in Fig.3.

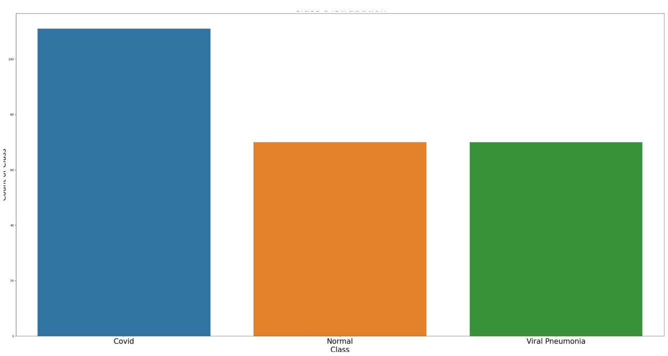


Fig. 1. Class distribution of dataset

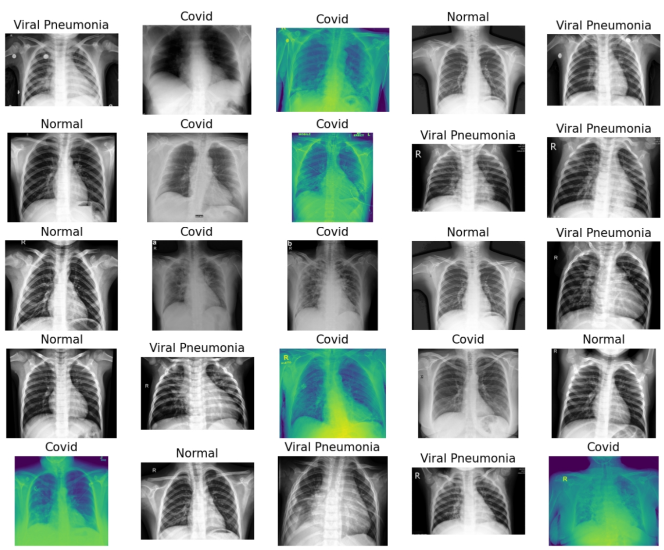


Fig. 2. Visualization analyse

## III. Methodological Overview

Our approach synergistically integrates Convolutional Neural Networks (CNNs) and Fully Connected Networks (FCNs), complemented by a thorough investigation into traditional machine learning models, to forge a resilient diagnostic framework. Before we start to introduce mentioned three ways of deep learning, we would like to processed the image features by principal component analysis (PCA) dimensionality reduction technique to explore feature redundancy and optimize feature dimensionality in the dataset. By analyzing the variance explained graph, we found that the variance explained is close to $100\%$ at about 100 principal components (features). This result indicates that there is significant feature redundancy in

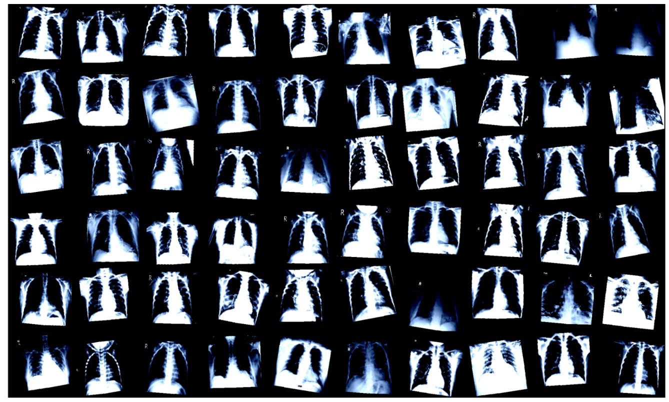

Fig. 3. Images with augments

the image data, and many of these features have limited ability to explain the data. The variance interpretability plot is an effective tool for assessing the importance of features in the data as well as determining the appropriate dimensionality for dimensionality reduction. In our case, this curve reveals that not only are the features in the dataset redundant, but also that the variance explainability is able to remain high even after the dimensionality is rapidly reduced, suggesting that there are some major features that capture most of the information in the data. This finding is important for the training of deep learning models. High dimensional features may lead to overfitting of the model, increasing computational complexity, reducing training efficiency, and possibly leading to performance degradation. Therefore, with PCA dimensionality reduction, we can choose to keep relatively few features and still retain most of the information, thus reducing the burden on the model(Fig.4).

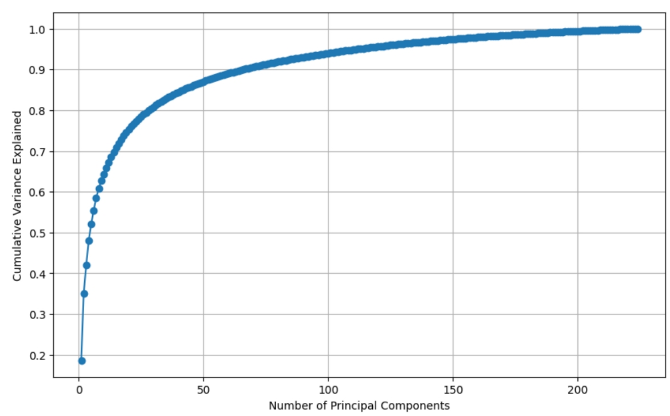


Fig. 4. Images with augments

### A. Comparative Analysis of Machine Learning Models

Acknowledging the diverse spectrum of AI techniques, we embarked on an exhaustive comparison of various machine learning models, encompassing Logistic Regression(C), Decision Tree Classifiers (max depth), and Random Forest Classifiers (n estimators). This meticulous evaluation not only serves to benchmark the performance of deep learning models against conventional algorithms but also furnishes invaluable insights into the most efficacious methodologies for CXR image analysis in the realm of COVID-19 detection.

When comparing the performance of the three models, logistic regression, decision tree and random forest, we mainly based on the evaluation metrics including Accuracy, F1 Score, Recall and Precision as shown in Fig.5. The logistic regression model performs best on these evaluation metrics, with an Accuracy of 82.35%, an F1 Score of 81.02%, a Recall of 79.50%, and a Precision of 85.69%. This indicates that the logistic regression model exhibits high accuracy and stability in balancing positive and negative sample classification.

In contrast, the decision tree model has only 56.86% accuracy, 53.44% F1 score, 53.21% recall and 54.32% precision. This result points to the poor performance of the decision tree model in dealing with this dataset, probably because the decision tree is easy to overfit and is more sensitive to small fluctuations in the training data, which leads to a lack of generalization ability on unknown data.

The random forest model has an accuracy of 58.82%, an F1 score of 45.98%, a recall of 52.80%, and a precision of 41.27%. Although Random Forest is slightly higher than the decision tree model in terms of accuracy, its F1 score and precision are lower, which may be due to the fact that the performance of Random Forest suffers if there is a large amount of noise when combining the prediction results of multiple decision trees.

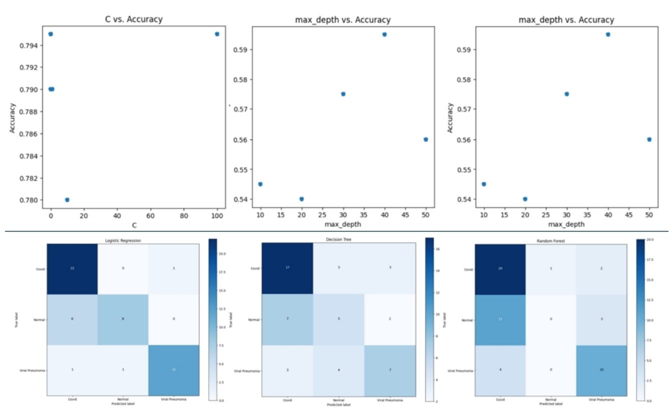


Fig. 5. Comparison of different machine learning methods

### B. Exploration of Fully Connected Networks (FCNs)

Our endeavor delves into the efficacy of FCNs for image classification tasks. These networks, characterized by dense layers that intricately interconnect every neuron across layers, play a pivotal role in refining classification based on features gleaned by CNNs.

In a performance evaluation of three different configurations of fully connected neural networks (FNNs), we compared their accuracy, precision, recall, F1 score, and average classification time. The three configurations are FNN1 (with three layers and 128 neurons per layer), FNN2 (with three layers and 256 neurons per layer), and FNN3 (with three layers and 512 neurons per layer) as illustrated in Fig.6. By comparison, we found that both FNN1 and FNN2 achieved 83.31% in accuracy, while FNN3 had a slightly lower accuracy of 82.35%. In terms of precision, recall and F1 score, FNN2 shows slightly better performance than FNN1 and FNN3 with 84.41% precision, 82.82% recall and 83.39% F1 score. It is worth noting that FNN3 did not significantly outperform FNN1 and FNN2 despite having more neurons, which may indicate that increasing the complexity of the network to some extent does not always result in improved performance.

When compared to logistic regression, the previous best performing machine learning model, the fully-connected neural network FNN2 showed a slight advantage in all major evaluation metrics. The accuracy of the logistic regression model was 82.35% compared to 83.31% for FNN2; the accuracy of logistic regression was 85.69% compared to 84.41% for FNN2; and in terms of F1 scores, logistic regression was 81.02% compared to 83.39% for FNN2. This comparison shows that although the logistic regression model provides very competitive performance in some cases, fully connected neural networks, especially the more complex configurations, can provide better accuracy and overall performance.

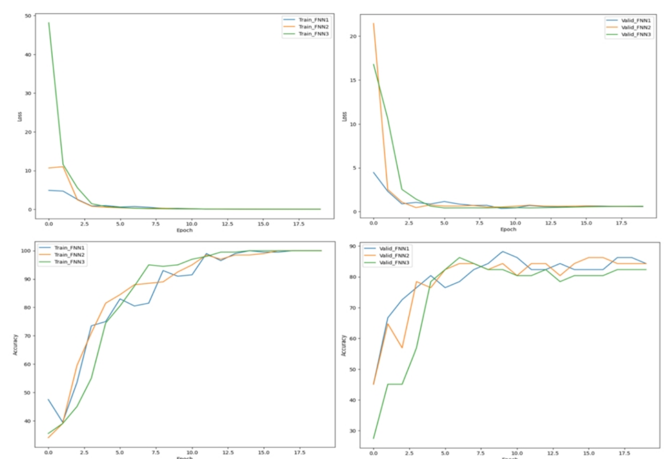

Fig. 6. Comparison of different FCNs configurations

### C. Development of Convolutional Neural Networks (CNNs)

Tailored CNN architectures were meticulously crafted for CXR image analysis, featuring layers meticulously engineered to extract and discern intricate patterns germane to COVID-19 and other respiratory conditions, thereby optimizing detection efficacy. In our evaluation, we scrutinized the performance disparities among three CNN architectures: the base CNN (NN1), LeNet (CNN2), and AlexNet (CNN3). Notably, as the models progress from the foundational CNN to the more intricate AlexNet, discernible performance enhancements are evident, underscoring the prowess of deep learning models in navigating intricate datasets. Thus, opting for a sophisticated CNN model such as AlexNet emerges as the superior choice for tasks mandating heightened accuracy and nuanced recognition capabilities.

We evaluate the difference in performance between three CNN architectures: base CNN (CNN1), LeNet (CNN2), and AlexNet (CNN3). The evaluation metrics include accuracy, precision, recall, F1 score, and average classification time. The results show a general improvement in performance as the model complexity increases. The results can be seen directly in Fig.7.

CNN1, as the most basic CNN model, achieves $86.27\%$ in accuracy, $87.19\%$ in precision, $85.20\%$ in recall, $85.72\%$ in F1 score, and 0.0045 seconds in average classification time.CNN2, which uses the LeNet architecture, improves its performance compared to CNN1, with an increase in accuracy to $88.24\%$, precision to $88.46\%$, recall of $86.65\%$, F1 score of $87.33\%$, and a slight reduction in average classification time to 0.0042 seconds. This indicates that the LeNet architecture, through its specific network layer design, is able to capture image features more efficiently, thus improving the overall performance of the model.

CNN3, which utilizes the more complex AlexNet architecture, has the best performance with $92.16\%$ accuracy, $92.59\%$ precision, $92.34\%$ recall, $91.75\%$ F1 score, and its average classification time is maintained at 0.0042 s. This performance enhancement of AlexNet is attributed to its deeper network structure and more sophisticated feature extraction capabilities to more accurately process and classify image data.

Compared to the previously analyzed Fully Connected Neural Network (FNN) models, advanced CNN models such as AlexNet show significant advantages in accuracy, precision, recall, and F1 scores, while maintaining similar fast classification times. This result highlights the importance and superiority of convolutional neural networks in processing images or tasks requiring advanced feature extraction. AlexNet, in particular, has shown great accuracy and efficiency in applications in the field of image recognition.

To summarize, as the model transitions from the basic CNN to the more complex AlexNet, we can observe significant performance gains, which highlights the power of deep learning models when dealing with complex datasets. Therefore, choosing an advanced CNN model such as AlexNet would be a superior choice for tasks that require high accuracy and fine-grained recognition.

## IV. Evaluation and Impact

In our evaluation framework, we employ a multifaceted approach, utilizing metrics such as accuracy, precision, recall, F1 score, and a comprehensive examination of the confusion matrix. This thorough assessment provides a nuanced understanding of each model's performance, offering insights into their strengths and areas for improvement. Initial observations highlight the superior performance of CNNs in tackling image-based classification tasks, while the integration of FCNs further refines classification accuracy through advanced feature extraction techniques. Moreover, our comparative analysis of machine learning models elucidates the diverse landscape of AI techniques, guiding future research endeavors and aiding in the selection of optimal models for specific diagnostic tasks.

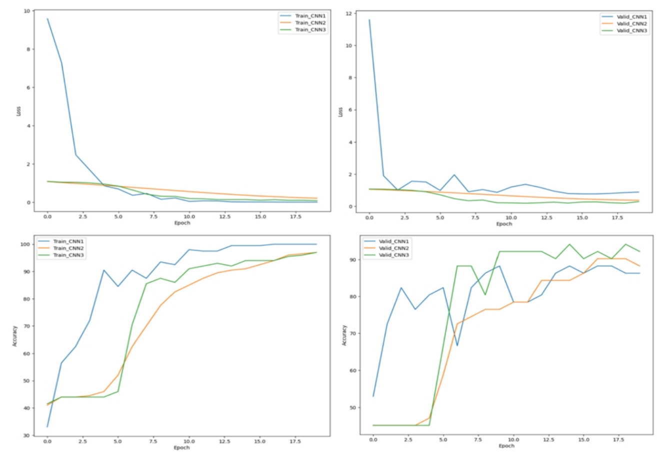

Fig. 7. Comparison of different CCNs architectures

Beyond its immediate application in COVID-19 detection, this project carries profound implications for the field of medical diagnostics at large. By releasing an open-source dataset and disseminating our findings, we aim to catalyze collaboration and innovation within the AI and healthcare communities. This collaborative effort fosters an environment conducive to continuous exploration and development, paving the way for novel advancements in AI-driven healthcare solutions. Ultimately, our endeavor seeks to leverage the power of AI to revolutionize medical diagnostics, enhancing accessibility, accuracy, and efficacy across diverse healthcare settings. Our evaluation test with the optimal model is demonstrated as in Fig.8.

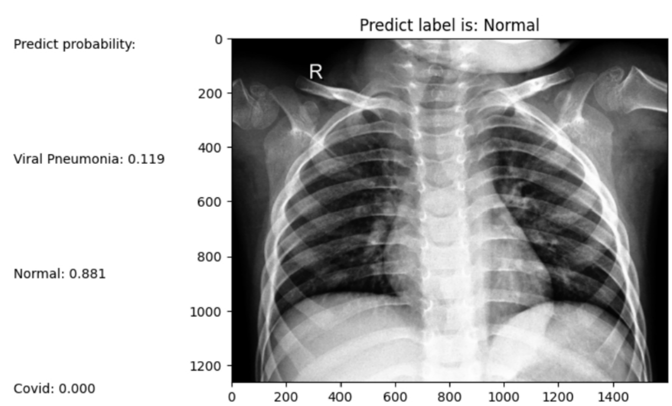


Fig. 8. Test results with CNN3

## V. Conclusion and Future Directions

Our endeavor to pioneer AI-driven diagnostic tools for the rapid and accurate detection of COVID-19 through CXR imaging represents a monumental leap forward in the realm of medical artificial intelligence. By amalgamating a diverse array of AI methodologies, including Convolutional Neural Networks (CNNs), Fully Connected Networks (FCNs), and traditional machine learning models, our project presents a sophisticated and nuanced approach to disease detection, promising unparalleled effectiveness and precision. As we embark on a journey of continual refinement and expansion, we extend an open invitation to the global research community for collaboration and contribution. Together, we aspire to harness the full potential of AI in tackling not only the current challenges posed by COVID-19 but also the myriad healthcare challenges of the future.

Looking ahead, our vision encompasses a multifaceted strategy aimed at elevating our research to new heights. We intend to broaden our dataset, incorporating a more diverse and expansive array of CXR images to enhance the robustness and generalizability of our models. Additionally, we seek to explore the integration of supplementary data modalities, such as clinical data, to further enrich the diagnostic process and improve overall accuracy. Furthermore, our roadmap includes a concerted effort to delve into the realm of advanced neural network architectures. By pushing the boundaries of innovation, we endeavor to develop cutting-edge frameworks capable of adapting to the dynamic landscape of emerging pathogens and respiratory illnesses. Our ultimate aspiration is to forge a universally applicable framework that not only mitigates the impact of existing diseases but also fortifies global health security against future threats.