BMI707 - Deep Learning for Biomedical Data
Group 1: Adèle Collin, Kyla Gabriel, Chuck Lin, Ritvik Raina, Sofía Rojas
5/4/24

CNN Model for COVID-19 Detection Using CT Scans

Abstract

In response to the limitations of PCR and rapid antigen tests for COVID-19, we developed a convolutional neural network (CNN) model that employs deep learning to analyze computed tomography (CT) scans for the detection of COVID-19. Our model incorporates advanced architectural enhancements specific to medical image analysis, such as increased convolutional layer depth and optimized filter sizes for enhanced feature extraction. Utilizing a dataset comprising 7,593 COVID-19 positive and 9,511 control images from multiple international sources, the model underwent a stringent training and validation protocol. It demonstrated superior diagnostic accuracy with a precision of 0.97 and recall of 0.98, outperforming traditional ML methods used as a baseline.

Introduction

The COVID-19 pandemic significantly impacted global health, exposing the limitations of traditional diagnostic methods like PCR and rapid antigen tests, which are either time-consuming or offer moderate sensitivity [1]. In response to these challenges, there has been a growing interest in leveraging advanced imaging techniques, such as computed tomography (CT) scans for the diagnosis of COVID-19. CT imaging offers the advantage of providing detailed anatomical information, allowing for the visualization of lung abnormalities associated with the virus, even in asymptomatic or early-stage cases [2]. However, manual interpretation of CT scans requires specialized expertise and can be subjective leading to variability in diagnosis. To address these limitations and enhance the efficiency and accuracy of COVID-19 diagnosis, there has been a surge in the development of artificial intelligence (AI) algorithms, particularly convolutional neural networks (CNNs). CNNs have shown promise in automating the analysis of medical images by extracting meaningful features and patterns indicative of the disease [3]. By using the power of deep learning, these models have the potential to revolutionize COVID-19 diagnostics that offer rapid and reliable detection and complement existing testing methods.

In this paper, we present a CNN-based approach designed to detect COVID-19 from CT scans with high accuracy and efficiency. Our model builds upon previous research in medical image analysis and incorporates novel architectural enhancements tailored specifically for COVID-19 detection. Through rigorous experimentation and validation on diverse datasets, we demonstrate the robustness and efficacy of our approach, highlighting its potential to streamline diagnostic workflows and improve patient outcomes in the ongoing battle against the pandemic. This project extends the research evidencing the efficacy of CT scans in identifying COVID-19-related anomalies. This study leverages advanced machine learning techniques to address the challenges posed by the scale of the pandemic and to enhance the diagnostic processes, to answer our main research question: how can deep learning be effectively applied to CT scans to develop a diagnostic tool for COVID-19? Specifically, this project aims to develop and evaluate a neural network architecture tailored for CT-based diagnosis of COVID-19, enhancing early detection capabilities crucial for controlling the spread of the virus.

Methods

Data Collection

A large lung CT scan dataset for COVID-19 diagnosis was compiled by Maftouni et al. by integrating data from seven public datasets. The dataset underwent quality control measures, including the removal of non-informative lung slices and images lacking clear class labels or patient information by the authors. It comprises 7,593 COVID-19 images from 466 patients, 6,893 normal images from 604 patients, and 2,618 images from community-acquired pneumonia (CAP) cases, with annotations provided for a subset of the CAP images. This dataset has been collected in multiple countries including Europe, the Middle-East, China and Japan (Figure 1). The average age of the patients is 53 years old, 51 for negative samples and 64 for COVID patients. About 59% of the patients are male, 32% female and 9% have an unknown gender. This formed the largest known COVID-19 lung CT dataset when it was collected in 2023. The utilization of CT scans in COVID-19 screening is acknowledged for its efficiency, prompting the creation of a dataset featuring lung CT scans from both normal individuals and those infected with COVID-19.

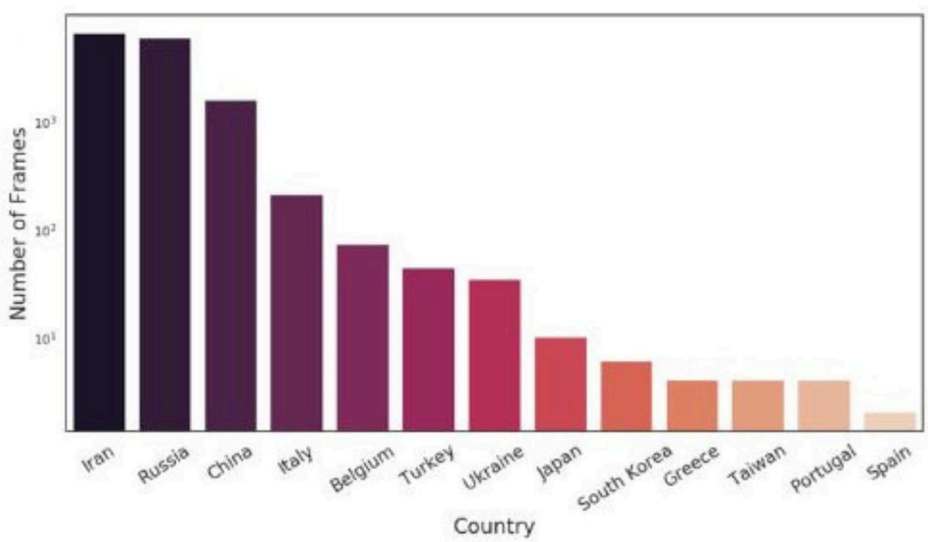


Figure 1. The sources of CT images by country (Maftoun et. al., 2023).

Deep Learning Approaches

A CNN model was constructed for the COVID diagnosis prediction. This CNN model consisted of convolutional layers, batch normalization layers, max-pooling layers, a flatten layer, a fully connected layer, and a dropout layer (Figure 2). We designed a convolutional block including a convolutional layer, a batch normalization layer, and a max-pooling layer. To increase the depth of the model architect, the convolutional block was stacked on top of each other for a total of four blocks. The filter size of the convolutional layer doubled in the next block, with 32 in the first block and 128 in the third and fourth blocks. Each convolutional layer had a kernel size of 5x5 and was also activated by RELU. Next, the output of the convolutional block was flattened to one dimension and passed through a dense layer with 100 nodes. A dropout layer was placed before the dense with a dropout rate of 0.2. Finally, the output layer generated the probability of each class using the sigmoid function. Consequently, our model contained 873,262 parameters in total. The model was then compiled using sparse categorical cross-entropy as the loss function and Adam as the optimizer. A sparse categorical cross-entropy is chosen to classify the integer class labels, and more stringent loss functions such as balanced cross-entropy and focal loss were not selected given the balanced class labels in the training and testing set. We trained the model using a batch size of 64 for 20 epochs. The model weights with the best validation accuracy were saved, which was then called back for use on the testing data for the final accuracy.

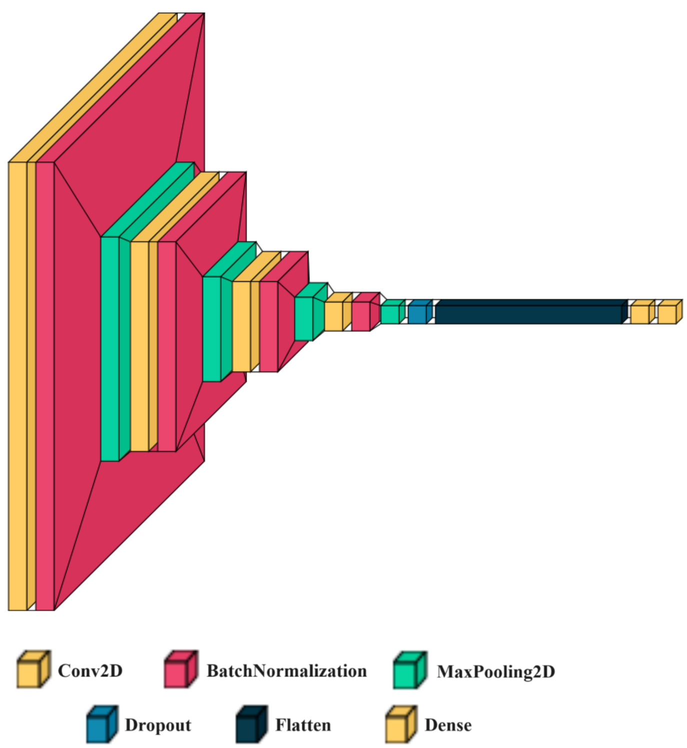


Figure 2. Summary of the Custom CNN Architecture.

In addition to the custom CNN model, a ResNet50-based transfer learning model with ImageNet weight was also constructed. The output of the ResNet50 was removed and replaced with a flatten layer, a dropout layer with a 0.2 dropout rate, and a final dense layer whose number of filters was set to the number of classes. To increase the model performance, the final two layers from the ResNet base model were unfrozen and retrained. Besides the deep learning techniques, classical machine learning models including Logistic Regression and Random Forest model were also trained and evaluated on the same training and testing dataset. Model hyperparameters of the classical and deep machine learning model can be found in Table 1.

Table 1. Summary of Model Hyperparameters.

Model	**Hyperparameter**	**Value**
Custom CNN	Optimizer	*Adam*
	Learning rate	*0.001*
	Batch size	*64*
	Epochs	*20*
ResNet50	Optimizer	*Adam*
	Learning rate	*0.001*
	Batch size	*64*
	Epochs	*20*
Logistic Regression	Penalty	*l2*
	C	*1.0*
	Tol	*1e-4*
	Solver	*lbfgs*
	Max_iter	*100*
Random Forest	n_estimators	*50*
	Criterion	*gini*
	min_samples_split	*2*
	min_samples_leaf	*1*
	min_weight_fraction_leaf	*0.0*
	max_features	*sqrt*
	max_leaf_nodes	*None*
	min_impurity_decrease	*0.0*
	bootstrap	*True*

Evaluation Metrics

To ensure reliable model performance, multiple metrics were evaluated in this project. Firstly, the training and validation losses and accuracies were assessed to verify that the model was learning the x-ray pattern and that it could generalize well to the validation set. Secondly, the confusion matrices were produced to describe the false positives and false negatives of the predicted COVID and control cases. Lastly, the precision, recall, and F1 scores were calculated for each model using the confusion matrices and were used to evaluate and compare model performance. Additionally, support was included in the metrics to detect potential class imbalances in the predicted labels.

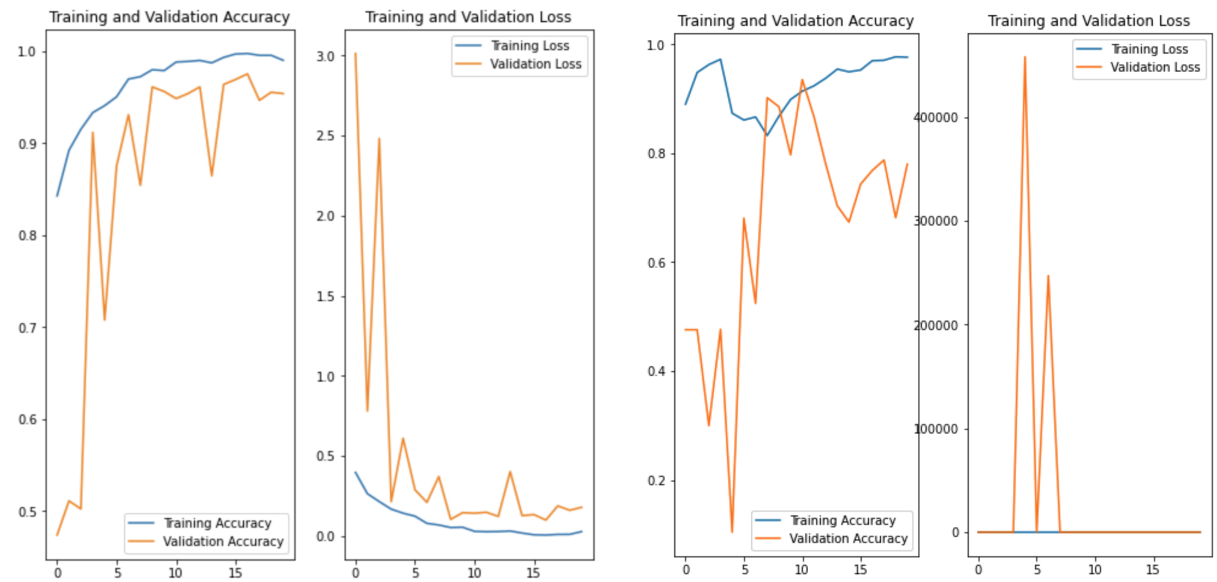
Results

Our primary goal for this project was to build and evaluate a convolutional neural network that is able to differentiate between individuals with COVID and without through lung CT-scans. As such, we used images from the dataset curated by Maftouni, Maede, et al, which had 14486 images with 6893 normal lung images and 7593 covid lung images. We split 10% of the images for our test set and 20% of our images for a validation set. Using this data split we trained and evaluated a shallow custom built CNN model and then benchmarked this shallow model's performance against a ResNet50 model initialized with imagenet weights (and data augmentations) as well as a Logistic Regression model and a Random Forest Model. Furthermore, for the purposes of ease of training and manipulation, all the images were converted into numpy tensors with the shape (128,128,1). To train the Logistic regression model and random forest model, the 3 dimensional image tensors were flattened to 1 dimensional arrays and then scaled with mean and standard deviation set to 0 and 1 respectively. The test accuracy, precision, recall, and F1 score for each model is shown in Table 2.

Table 2: Model Performances and comparisons

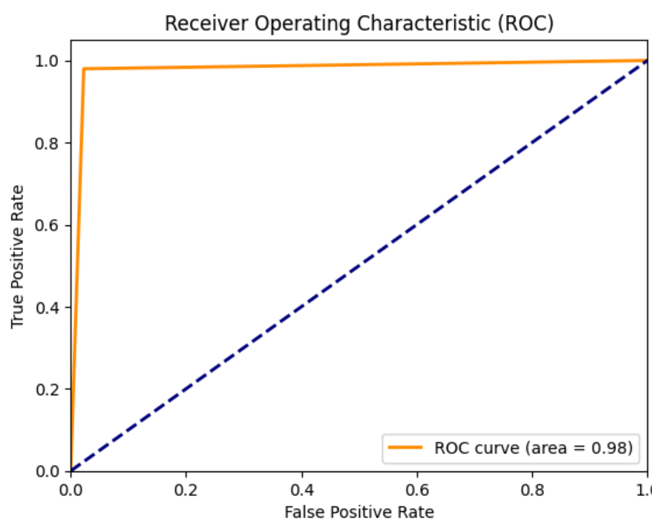
Model	Test Accuracy	Precision	Recall	F1 Score
Custom CNN	0.97	0.97	0.98	0.97
ResNet50	0.51	0.48	0.26	0.34
Logistic Regression	0.93	0.94	0.92	0.93
Random Forest	0.95	0.98	0.93	0.96

We trained the Custom neural network and the ResNet50 model with a batch size of 64 and 20 epochs, with data augmentations added to the images used to train the ResNet50 model through the use of Image Generators. In Figure 3 and Figure 4, the accuracy and loss curves over epoch range can be seen.



(a) (b)

Figure 3: Accuray and Loss curves for (a) Custom CNN Model and (b) ResNet50 Model



(c)

Figure 4: AUROC curve for the custom CNN model

While training we added in model checkpoints to save the model weights with the best validation accuracy and used these checkpoints to test the models. The confusion matrices for our deep learning models and the non-deep learning models are below in Figure 5. In Figure 6, we have added some sample images with the ground truth labels and the custom CNN model's prediction on those images.

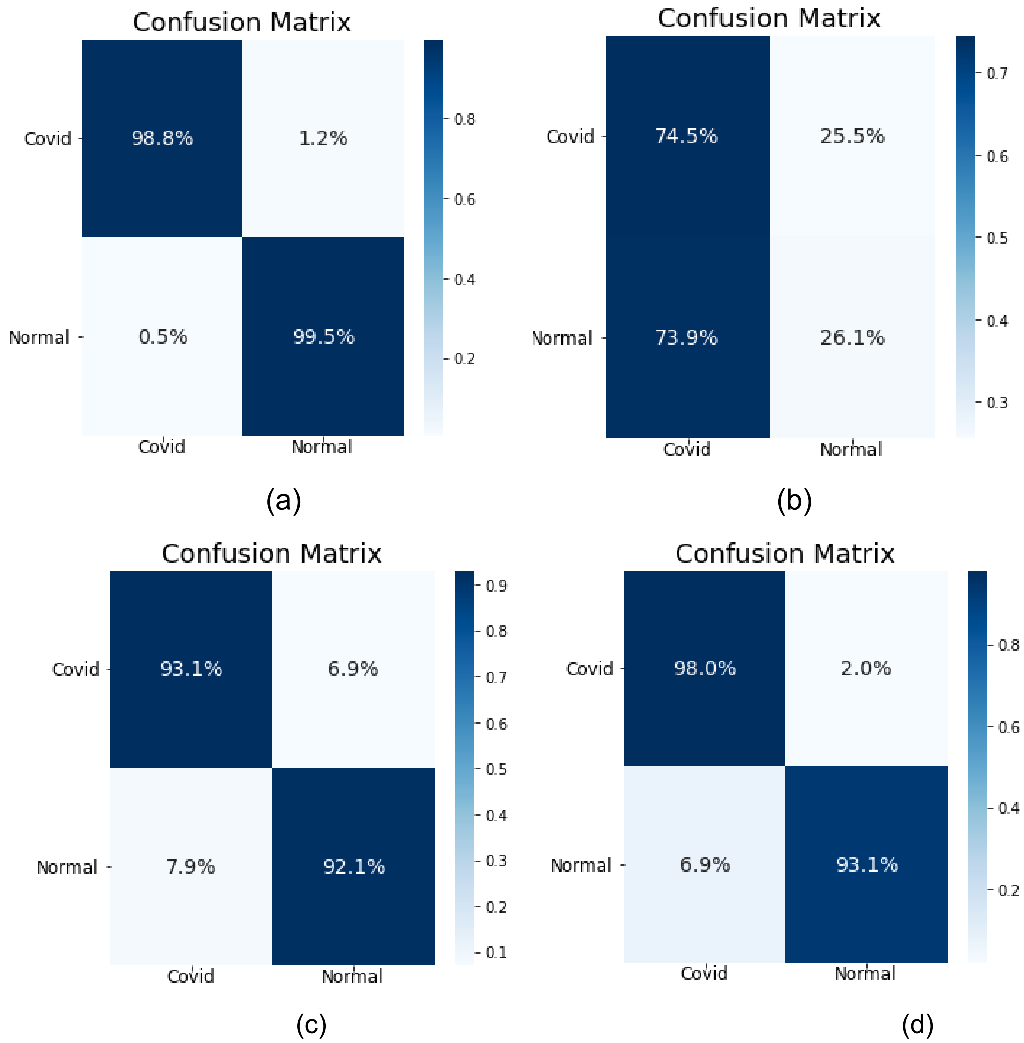


Figure 5: Confusion Matrices for (a) Custom CNN, (b) ResNet50, (c) Logistic Regression, and (d) Random Forest

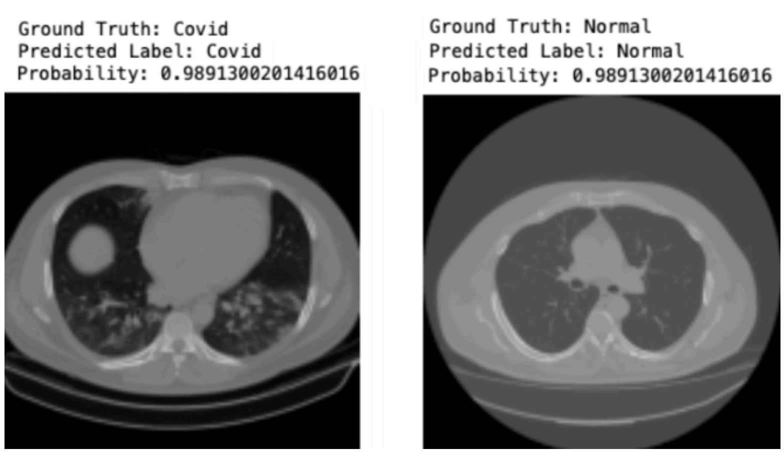


Figure 6: Custom CNN model predictions with COVID-19 (left) and normal (right) examples

Discussion

The findings from our study carry significant implications for the field of COVID-19 diagnostics and the broader application of machine learning in healthcare. Firstly, our results underscore the efficacy of CNNs in accurately differentiating between individuals with COVID-19 and those without using lung CT scans. The superior performance of our custom CNN model, with a test accuracy, recall, and F1 score of 0.97, 0.98, and 0.97 respectively, highlights the potential of tailored deep learning architectures in achieving precise and reliable diagnosis of the disease. Moreover, the comparison with established models such as ResNet50, Logistic Regression, and Random Forest provides valuable insights into the relative strengths and limitations of different machine learning techniques in this context.

The ResNet50 model performed the worst, likely due to the small amount of images in the dataset. This complex model likely did not have enough information to perform well and capture the underlying patterns in the images. However, the exploration of traditional machine learning algorithms, like Logistic Regression, demonstrates the feasibility of leveraging simpler models for COVID-19 detection, particularly in scenarios where computational resources are limited or interpretability is paramount. Additionally, the observation that the Random Forest model achieved the best precision score suggests its potential utility in scenarios where minimizing false positives is critical, such as screening programs or triage systems. This highlights the importance of considering not only overall accuracy but also specific performance metrics tailored to the clinical context when evaluating diagnostic models. While the Random Forest model showcased competitive performance, our custom CNN model outperformed it likely due to its ability to optimize the data with the applied hyperparameter tuning capabilities.

Overall, our findings contribute to the growing body of evidence supporting the integration of AI-driven diagnostic tools into clinical practice, offering the promise of more efficient and accurate identification of COVID-19 cases from CT imaging. Moving forward, further research and validation in larger and more diverse cohorts will be essential to fully realize the clinical utility of these models and ensure their seamless integration into healthcare workflows.

Conclusion

In conclusion, our study presents a significant step forward in the development of a CNN-based approach for the detection of COVID-19 from CT scans. Despite the challenges faced, such as initial model training constraints due to computational limitations, we successfully mitigated these obstacles by preprocessing images into Numpy arrays. However, it's crucial to acknowledge the limitations inherent in our study. Potential biases may exist due to variations in image acquisition protocols, and the dataset used may not fully capture the diverse manifestations of COVID-19 across different demographic populations. Moving forward, future research endeavors should aim to address these limitations by conducting further validation studies. We propose performing additional validation by testing our model on datasets collected by Shakouri et al., which could provide valuable insights into the generalizability of our approach. Furthermore, we intend to enhance the robustness of our model by further validating transfer learning results with a shallower architecture ensuring its effectiveness across various settings and populations. By addressing these challenges and limitations, we can continue to refine and optimize our model, ultimately advancing the field of COVID-19 diagnostics and contributing to improved patient care and public health outcomes

Group Member Contributions

- Chuck Lin: Methodology and analysis of results
- Kyla Gabriel: Interpretation of results and discussion
- Adele Collin: Data description and analysis of results
- Sofia Rojas: Introduction and analysis of results
- Ritvik Raina: Building and evaluating the custom CNN and analysis of results; also assisted with other models

While we did have some tasks divided, we primarily worked on the project together with everyone's contributions overlapping across tasks. We discussed results as a team and all members contributed to the final write-up.

References

1. Li, Yan, and Liming Xia. "Coronavirus Disease 2019 (COVID-19): Role of Chest CT in Diagnosis and Management." *AJR. American Journal of Roentgenology*, vol. 214, no. 6, June 2020, pp. 1280–86.
2. Li, L. ,et al. (2020). Using Artificial Intelligence to Detect COVID-19 and Community-acquired Pneumonia Based on Pulmonary CT: Evaluation of the Diagnostic Accuracy. Radiology, 296(2), E65-E71.
3. Wang, S., et al. (2020). A deep learning algorithm using CT images to screen for CoronaVirus Disease (COVID-19). European Radiology, 30(10), 1-9.
4. Maftouni, Maede, et al. "A Mask-Guided Attention Deep Learning Model for COVID-19 Diagnosis Based on an Integrated CT Scan Images Database". *IISE Transactions on Healthcare Systems Engineering*, vol. 13, nº 2, avril 2023, p. 132‑49. *DOI.org (Crossref)*, https://doi.org/10.1080/24725579.2022.2142866.
5. Misztal, Krzysztof, et al. "The Importance of Standardisation - COVID-19 CT & Radiograph Image Data Stock for Deep Learning Purpose." *Computers in Biology and Medicine*, vol. 127, Dec. 2020, p. 104092.
6. Peña, Mónica, et al. "Performance of SARS-CoV-2 Rapid Antigen Test Compared with Real-Time RT-PCR in Asymptomatic Individuals." *International Journal of Infectious Diseases: IJID: Official Publication of the International Society for Infectious Diseases*, vol. 107, June 2021, pp. 201–04.
7. Shakouri, Shokouh, et al. "COVID19-CT-Dataset: An Open-Access Chest CT Image Repository of 1000+ Patients with Confirmed COVID-19 Diagnosis". *BMC Research Notes*, vol. 14, 1, Decembre 2021, p. 178. *DOI.org (Crossref)*, https://doi.org/10.1186/s13104-021-05592-x.

Supplemental Information

1. Code 🗀 BMI707_CODE

   https://drive.google.com/drive/folders/1NM1DQv_0oKzRrH17AY3aMtXiMZiEhLrZ?usp=sharing

2. Metrics: 🗀 Metrics

   https://drive.google.com/drive/folders/1Pjl85vbemPKZZi6BBPYeRFFIAQnrzrPG?usp=drive_link