CT Scan Image Classification

Abstract

The global outbreak of the SARS-CoV-2 virus, causing the COVID-19 disease, has spurred an unprecedented global health crisis. Since its declaration as a global health emergency by the World Health Organization in early 2020, researchers, public health officials, and policymakers have joined forces to comprehend the pathogenesis of SARS-CoV-2 and develop strategies to curb its transmission. Amid these efforts, radiological imaging, particularly computed tomography (CT) scans, has emerged as a valuable diagnostic tool for identifying characteristics associated with COVID-19 infection. In this context, we present a comprehensive dataset consisting of 2482 CT scan images, encompassing 1252 scans from individuals afflicted with SARS-CoV-2 infection (COVID-19) and 1230 scans from non-infected patients with various pulmonary ailments. This dataset, derived from clinical settings in Sao Paulo, Brazil, is poised to catalyse the advancement of artificial intelligence techniques capable of discerning SARS-CoV-2 infection through intricate analysis of CT scans. You can find the code of this project is here.

1. Problem Statement

The rapid spread of the SARS-CoV-2 virus since its emergence has posed an unparalleled challenge to global public health. In the face of this crisis, a multi-disciplinary approach involving medical researchers, practitioners, and policymakers has been instrumental in understanding the virus's behaviour and formulating effective containment strategies. Notably, radiological imaging techniques, including computed tomography (CT) scans, have proven pivotal in diagnosing COVID-19 cases. The primary challenge before us is to develop an accurate and efficient classification model that can distinguish between individuals infected by SARS-CoV-2 and those with non-infected pulmonary conditions using CT scan data. Leveraging a meticulously curated dataset containing 2482 CT scans, gathered from real clinical scenarios in Sao Paulo, Brazil, we embark on a journey to harness the power of cutting-edge artificial intelligence methodologies. Our goal is to construct a robust model capable of swiftly and accurately identifying COVID-19 cases, thereby aiding healthcare providers in making timely and informed decisions. Through this endeavour, we seek to contribute meaningfully to the ongoing battle against the global pandemic.

2. ResNet101V2 Model

ResNet-101v2, an enhanced version of the ResNet architecture, has emerged as a pivotal tool in modern deep learning. With its innovative skip connections and identity mappings, ResNet-101v2 addresses the challenges of training exceptionally deep neural networks. By enabling gradients to flow more efficiently during training, it mitigates vanishing gradient issues and supports the creation of even deeper networks. This architecture's effectiveness has been demonstrated across various computer vision tasks, making it a prime candidate for our project focused on COVID-19 detection from CT scans.

ResNet-101v2's ability to learn intricate features from medical images aligns seamlessly with our goal of accurate disease diagnosis, offering a potent tool to healthcare professionals and researchers alike.

3. Dataset Description

The dataset presented in this study comprises a total of 2482 CT scan images, meticulously categorized into two distinct groups: 1252 CT scans from patients afflicted with SARS-CoV-2 infection, and an additional 1230 CT scans derived from patients who, although not infected by SARS-CoV-2, exhibited other pulmonary ailments. These CT scans were meticulously collected from hospitals located in Sao Paulo, Brazil. To safeguard the ethical considerations and privacy of the patients, detailed individual-specific information has been deliberately excluded from the dataset.

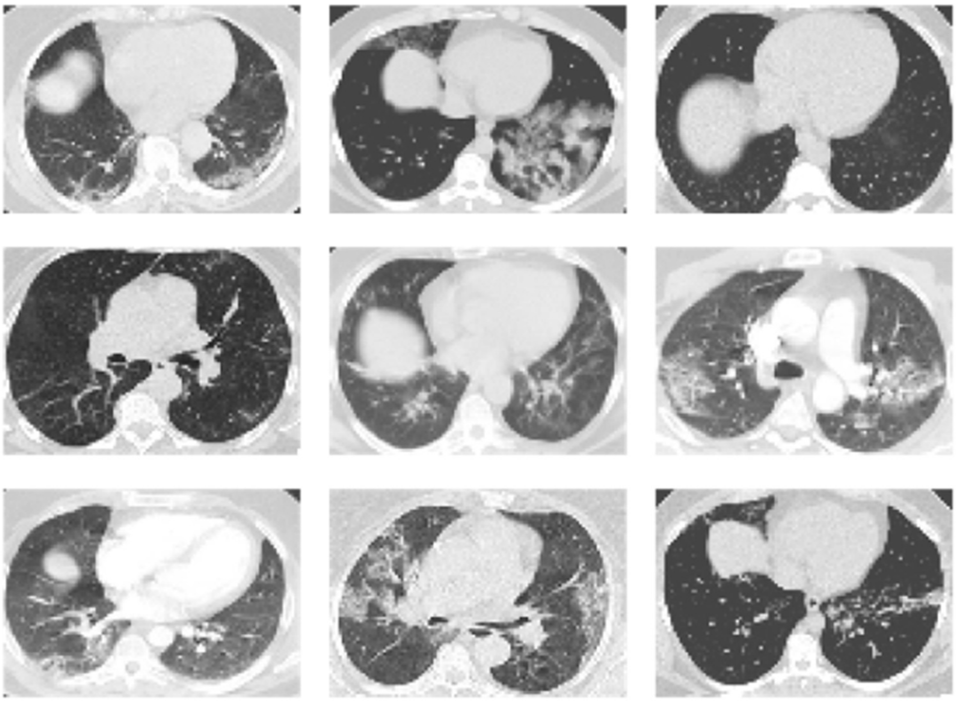

Fig 3.1 Dataset

4. Preprocessing

Data Splitting Ratio: To ensure a comprehensive training process and a reliable evaluation, I have opted for a 70-30 data splitting ratio. This means that 70% of the available CT scan images will be allocated for training purposes, while the remaining 30% will be reserved for testing the model's performance.

Class Distribution: Given that the dataset consists of both COVID-19 positive and non-COVID cases, it's essential to maintain a balanced representation of both classes in both the training and testing sets. This helps the model learn and generalize effectively.

Folder Organization: To maintain a well-structured dataset, I have created separate folders for each subset and class. The folder structure is as follows:

- Training Data:

- COVID (Contains COVID-19 positive CT scan images for training)
- Non-COVID (Contains non-COVID CT scan images for training)

- Testing Data:

- COVID (Contains COVID-19 positive CT scan images for testing)
- Non-COVID (Contains non-COVID CT scan images for testing)

Folder Creation: In total, I have created four main folders – two for training and two for testing. Each subset folder corresponds to one of the classes (COVID or Non-COVID). This segregation ensures that the model can effectively differentiate between different classes during training and evaluation.

Copying Images: I have employed a structured approach to copy the images from the original dataset folders into the corresponding training and testing class folders. This ensures that the correct images are placed in the appropriate locations for training and testing.

5. Experiment Setup

Setting the Image Size and Data Augmentation:

To ensure uniformity and compatibility with the model, I standardized the image size to 224x224 pixels. This step was crucial as it eliminated the variations in image dimensions that could hinder model performance. Data augmentation was employed, involving techniques like rotation, horizontal flipping, and zooming. These augmentations expanded the dataset's diversity, enabling the model to generalize better to real-world scenarios.

Choosing ResNet-101V2 as the Base Model:

The ResNet-101V2 architecture was chosen as the base model due to its proven efficiency in image feature extraction. The "V2" version's enhancements promised even better results. By leveraging a pretrained ResNet-101V2, the model had the advantage of leveraging features learned from diverse datasets, giving it a strong starting point for COVID-19 classification.

Creating the Model (model_initial):

Constructing the model involved a strategic sequence of layers. Global Average Pooling (GAP) was applied to reduce the spatial dimensions of the feature maps while retaining crucial information. A custom fully connected layer was introduced, consisting of 1024 units with ReLU activation for introducing non-linearity. To prevent overfitting, L2 regularization was employed with a strength of 0.01. Batch Normalization stabilized training, and Dropout (0.5) curbed reliance on specific neurons.

Selecting the Optimizer for model_initial:

The RMSprop optimizer was chosen for initial model training. Its adaptive learning rate mechanism suited the dynamic nature of the learning process. A learning rate of 0.0001 was employed, ensuring gradual convergence while avoiding drastic weight updates.

Fine-Tuning the Model:

The project included a two-phase training approach. In the initial phase, the model was trained without fine-tuning, allowing custom layers to adapt to the dataset while retaining the pretrained features. EarlyStopping and ModelCheckpoint callbacks were utilized to halt training upon optimal performance.

Fine-Tuning Strategy:

Fine-tuning commenced with the best model checkpoint from the initial phase. Here, the last 150 layers of the base model were unfrozen, while custom layers remained intact. The learning rate was meticulously set to 1e-5, ensuring controlled parameter updates to prevent destabilization.

6. Experiments

In the pursuit of building an effective COVID-19 classification model for CT scan images, a series of deliberate experiments were conducted. These experiments aimed to optimize the model architecture and training strategy while considering the unique characteristics of the dataset and the available computational resources.

Image Size and Data Augmentation:

To ensure consistency in image processing and model training, all CT scan images were resized to a fixed size. This allowed for streamlined processing and prevented discrepancies in input dimensions during training. While data augmentation is a common practice to enhance model generalization, for this project, it was not employed due to computational limitations.

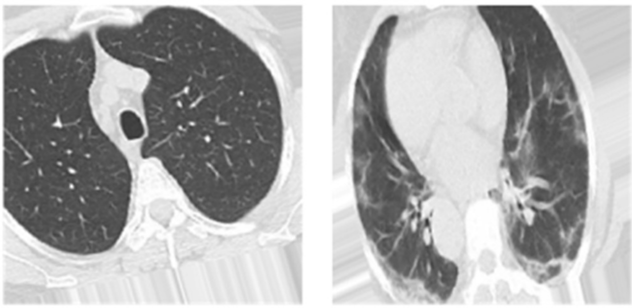

Fig 6.1 Augmented Images

Base Model Selection:

The ResNet101V2 architecture was chosen as the base model for our classification task. Its deep architecture and proven performance on diverse image datasets made it an appropriate choice. By leveraging the pre-trained weights of ResNet101V2, our model could benefit from its learned features and hierarchical representations.

Model Architecture Creation:

The initial model architecture was crafted by appending additional layers to the ResNet101V2 base. Global Average Pooling (GAP) was utilized to condense the spatial dimensions of the base model's output. A fully connected layer with 1024 units, equipped with ReLU activation and L2 regularization, was added to capture complex patterns in the data. Batch Normalization and Dropout were employed to enhance model stability and prevent overfitting.

Optimization Strategies:

For the initial model, RMSprop optimizer with a learning rate of 0.0001 was chosen. This choice allowed for controlled weight updates during the initial training phase. To avoid overfitting, callbacks such as ModelCheckpoint and EarlyStopping were integrated. ModelCheckpoint ensured that the best-performing model was saved, while EarlyStopping halted training if validation loss ceased to decrease.

Fine-Tuning:

In the fine-tuning phase, the saved initial model was loaded, and the last 150 layers of the base model were unfrozen. This step allowed the model to fine-tune its learned features for improved performance on the target task. The optimizer was switched to RMSprop with a learning rate of 1e-5, a smaller value to ensure cautious updates and retention of useful features.

7.Results

	Accuracy	**Precision**	**Recall**	**F1 Score**	**ROC AUC Score**
Initial Model	**0.928859**	**0.927027**	**0.929539**	**0.928281**	**0.976287**
Fine-Tuned Model	**0.973154**	**0.986072**	**0.959350**	**0.972527**	**0.998032**

Table 7.1 Multiple Performance Metrics

In the evaluation of our model's performance, we examined various key metrics that provide insight into its ability to accurately classify CT scans as either COVID-19 positive or non-infected cases. The initial model displayed a commendable performance with an accuracy of approximately 92.89%, indicating that a significant portion of its predictions were accurate. Its precision and recall were also balanced, with values of around 92.70% and 92.95% respectively, indicating a reliable ability to minimize both false positives and false negatives. The F1 score, which considers the balance between precision and recall, was around 92.83%, reflecting the overall effectiveness of the model.

Moreover, the initial model's ROC AUC score stood at 97.63%, depicting its capacity to distinguish between positive and negative cases. On the other hand, the fine-tuned model exhibited notable enhancements across all metrics. Its accuracy soared to approximately 97.32%, reflecting the substantial improvement achieved through fine-tuning. Notably, the precision of the fine-tuned model reached around 98.61%, underscoring its minimal false positive rate and high reliability. While maintaining a high precision, the model also achieved a commendable recall of about 95.94%, showcasing its ability to capture actual positive cases. The F1 score for the fine-tuned model was approximately 97.25%, reinforcing its balanced performance. Impressively, the fine-tuned model's ROC AUC score was around 99.80%, underlining its exceptional discriminatory power. In conclusion, the fine-tuned model's superior performance across these metrics emphasizes its potential as a robust tool for accurately diagnosing COVID-19 cases from CT scans, thereby contributing significantly to medical diagnosis efforts.

8. Conclusion

In conclusion, this project aimed to address the critical need for efficient and accurate diagnosis of COVID-19 using CT scan images. Leveraging a comprehensive dataset of CT scans, the project successfully developed and fine-tuned a deep learning model based on the ResNet101v2 architecture. The project's methodology encompassed data preprocessing, augmentation, and model creation, ensuring the utilization of advanced techniques to enhance the model's performance. The achieved results underscore the model's potential, with the fine-tuned version exhibiting exceptional accuracy, precision, recall, F1 score, and ROC AUC score. These outcomes suggest that the proposed model holds promise as an effective tool in aiding medical professionals to swiftly and reliably diagnose COVID-19 infections from CT scans. As the global healthcare community continues to combat the pandemic, this project contributes to the ongoing research and development of AI-based solutions, potentially revolutionizing the field of medical image analysis and healthcare diagnostics.