

# 1 Description

<b>Dataset name</b>	CT Scans for COVID-19 Classification
<b>Link</b>	<a href="#">Kaggle</a>
<b>License</b>	<a href="#">CC BY 4.0</a>
<b>Medical discipline</b>	Radiology
<b>Medical procedure</b>	Computer tomography (CT)
<b>Multi-class problem</b>	✓
<b>Multi-label problem</b>	✗

Since December 2019, the outbreak of an unknown viral pneumonia has severely affected Wuhan, China. This virus was quickly identified and named by the World Health Organization (WHO) as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). To control the outbreak, early diagnosis of patients with COVID-19 pneumonia for timely treatment is critical, especially in epidemic regions.

Data were collected from 2 hospitals: Union Hospital (HUST-UH) and Liyuan hospital (HUST-LH) which is describes in detail [in this paper](#). The authors of this paper classified individual CT images into three types,

- 5705 non-informative CT (NiCT) images where lung parenchyma was not captured for any judgment,
- 4001 positive CT (pCT) images where imaging features associated with COVID-19 pneumonia could be unambiguously discerned, and
- 9979 negative CT (nCT) images where imaging features in both lungs were irrelevant to COVID-19 pneumonia.

So the first step of methodologies in the paper was to extract lung parenchyma from CT images. We extracted lung parenchyma by using the methods explained in the paper. These extracted images are in the preprocessed CT scans folder where the original images are in the original CT scans folder. All the images are resized to (512x512).

Classes that can be identified are

- NiCT
- pCT
- nCT

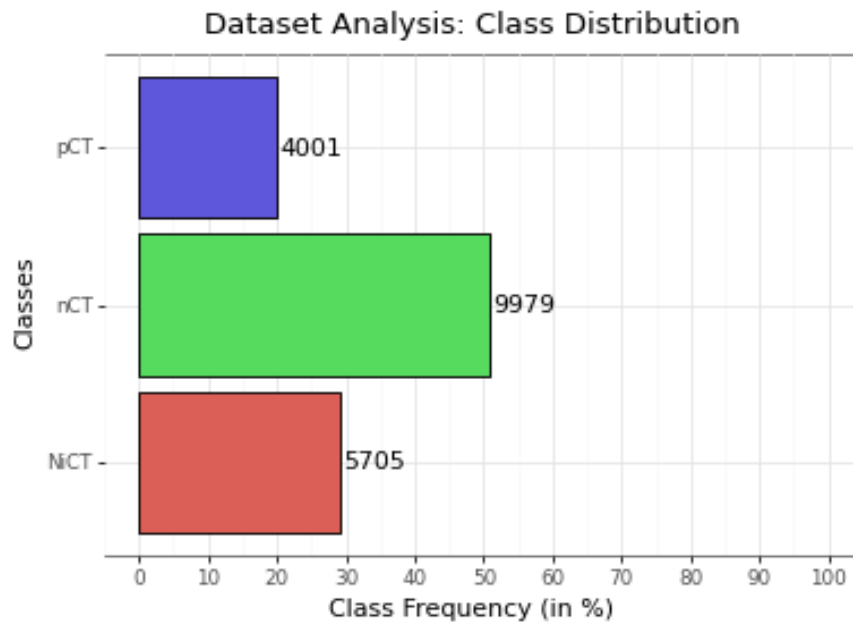


Figure 1: Class distribution: CT Scans for COVID-19 Classification

## 2 Pre-processing

The dataset consists of original and pre-processed CT scans. Only original images were considered. No further pre-processing was done.

## 3 Training

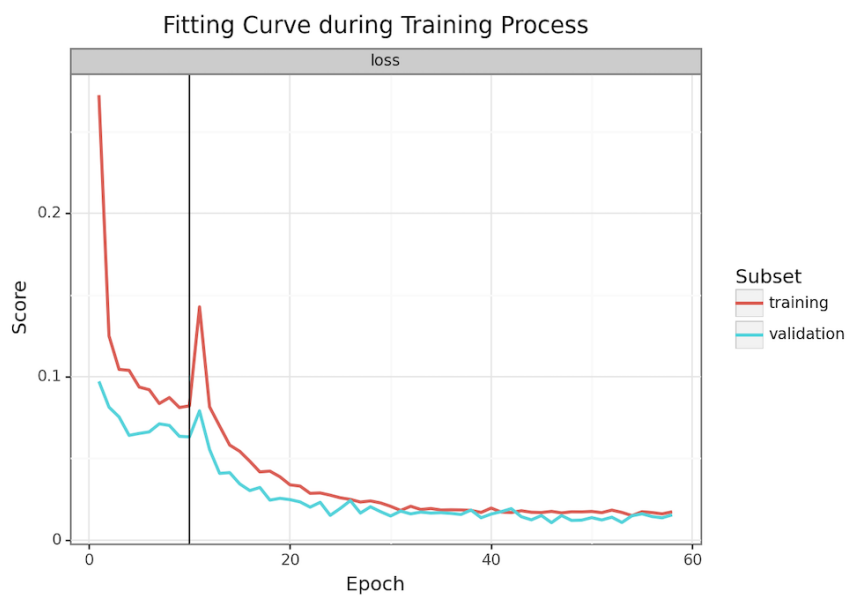


Figure 2: Training: CT Scans for COVID-19 Classification

Training and validation loss are decreasing beautifully. There is no indication for overfitting.

# 4 Results

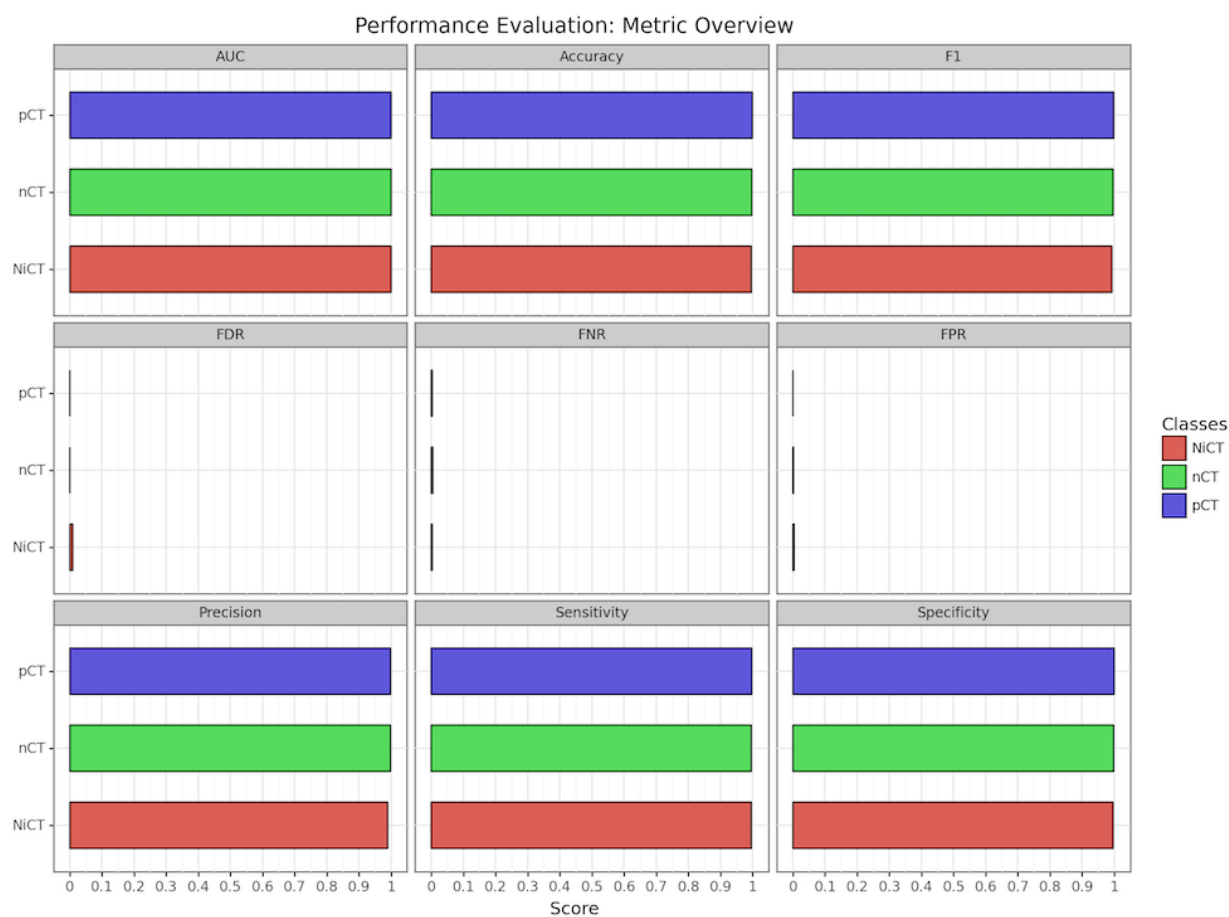


Figure 3: Metric overview: CT Scans for COVID-19 Classification

All calculated metrics attest to the model’s outstanding classification performance.

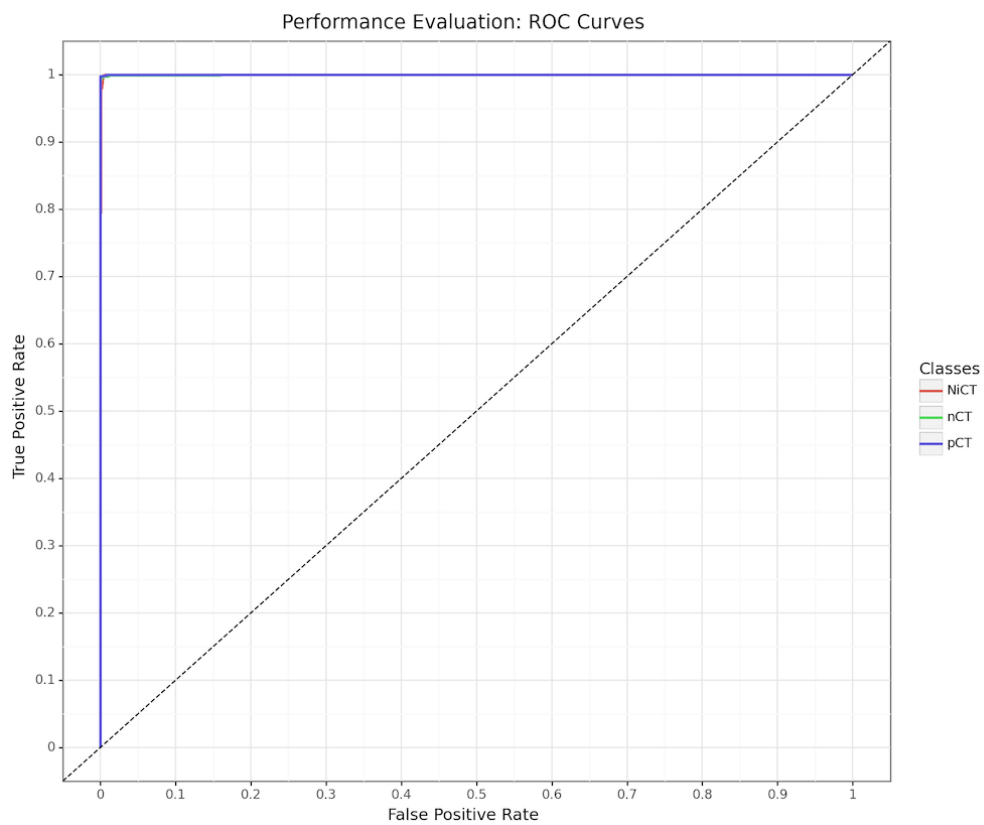


Figure 4: ROC curve: CT Scans for COVID-19 Classification

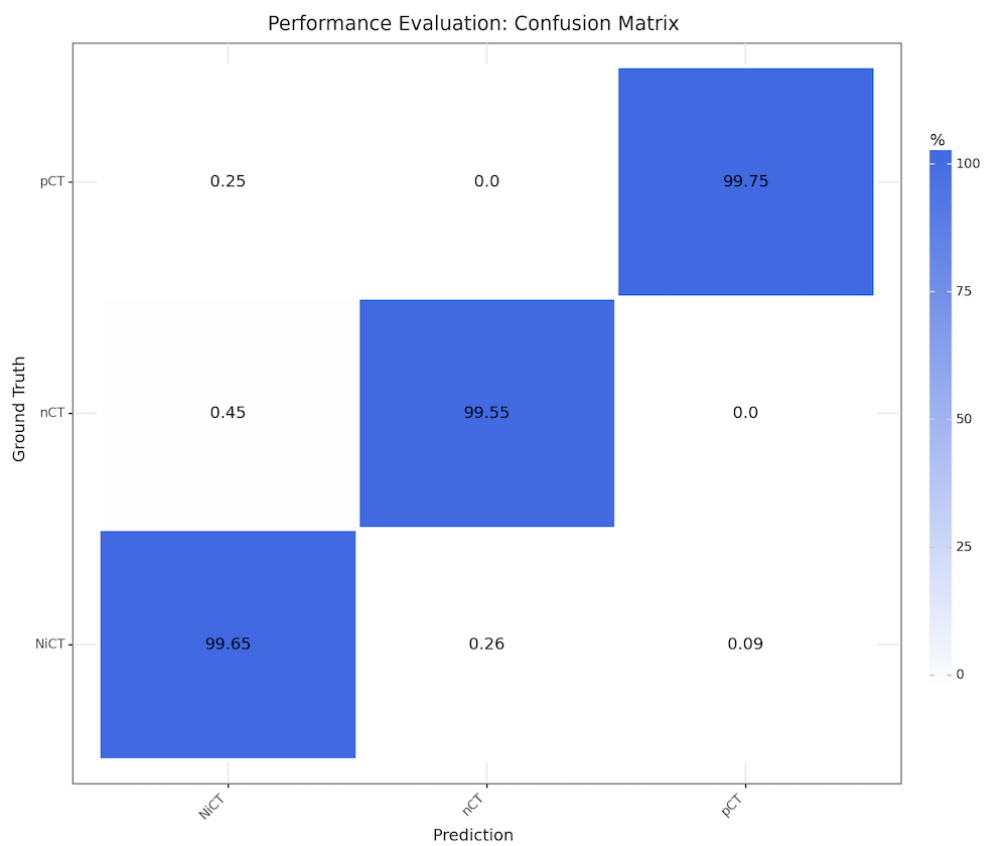


Figure 5: Confusion matrix: CT Scans for COVID-19 Classification

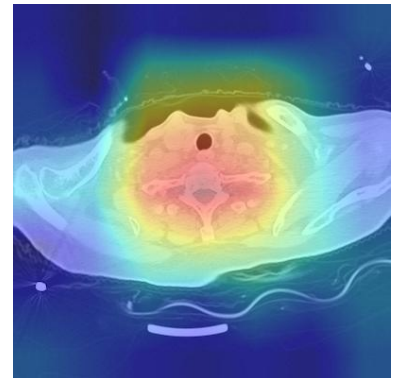
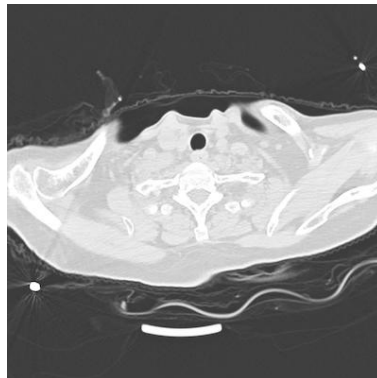
## 5 XAI

Non-informative CT means, that there is no gas-exchange-relevant lung tissue on the image. For sample **NiCT4128** we see that our model is marking the center of the scan in search of lung parenchyma. For the classification of sample **pCT474**, which does show a COVID-19 infection, our model focuses the left lung. The Grad-CAM mark-up for the negative sample **nCT2535** however seems to highlight the heart, instead of any lung tissue.

**Image:** NiCT4128

**Class:** NiCT

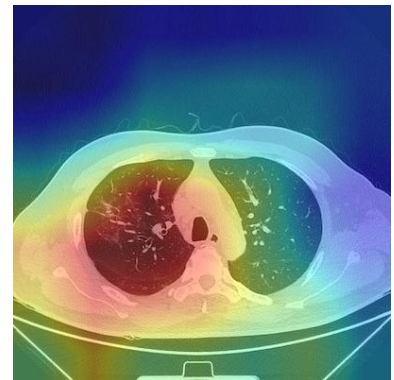
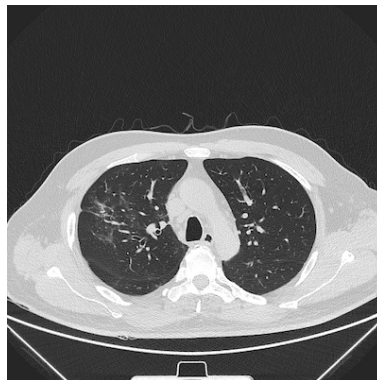
**Classified as:** NiCT (97.2 %)



**Image:** pCT474

**Class:** pCT

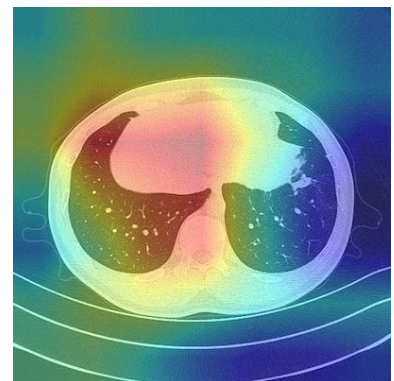
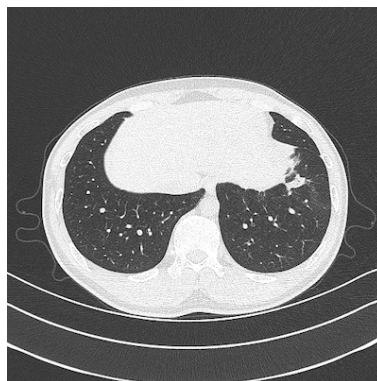
**Classified as:** pCT (98.3 %)



**Image:** nCT2535

**Class:** nCT

**Classified as:** nCT (100.0 %)



## 6 Summary

Our model masters the classification task reliably and with high confidence.