COVID-19 Medical Image Analysis

12/07/2020

Chengyuan Cai, Tiancheng Ren, Yuchen Sun, Yanwen Wang

Motivation

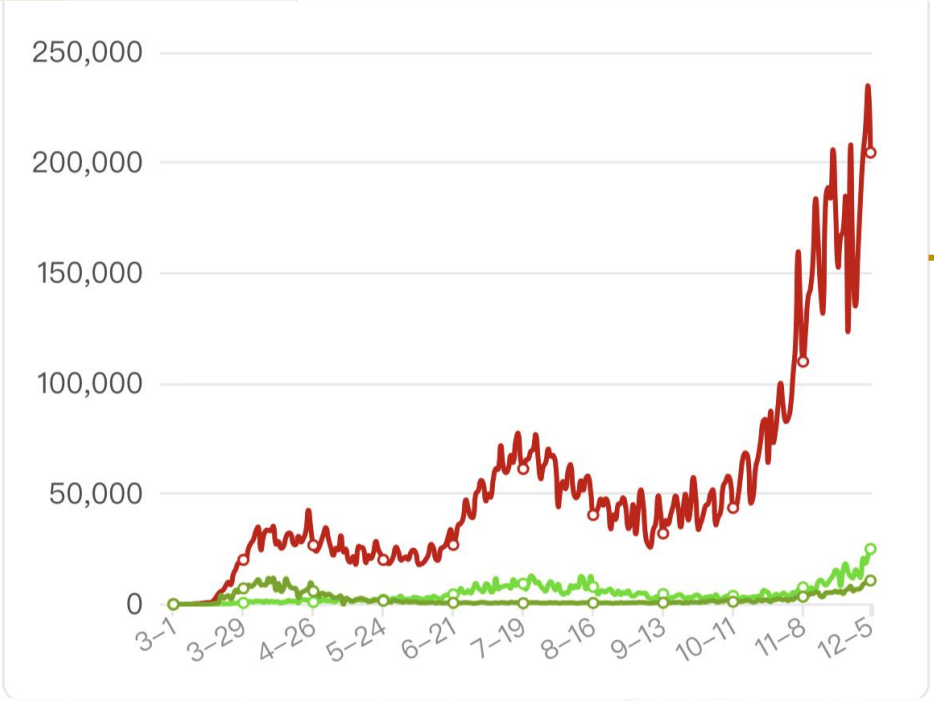
01 Help identify suspected Covid-19 cases

02 Improve efficiency of diagnosis of patients

03 Reduce the occurence of misdiagnosis

04 Develop useful new algorithm

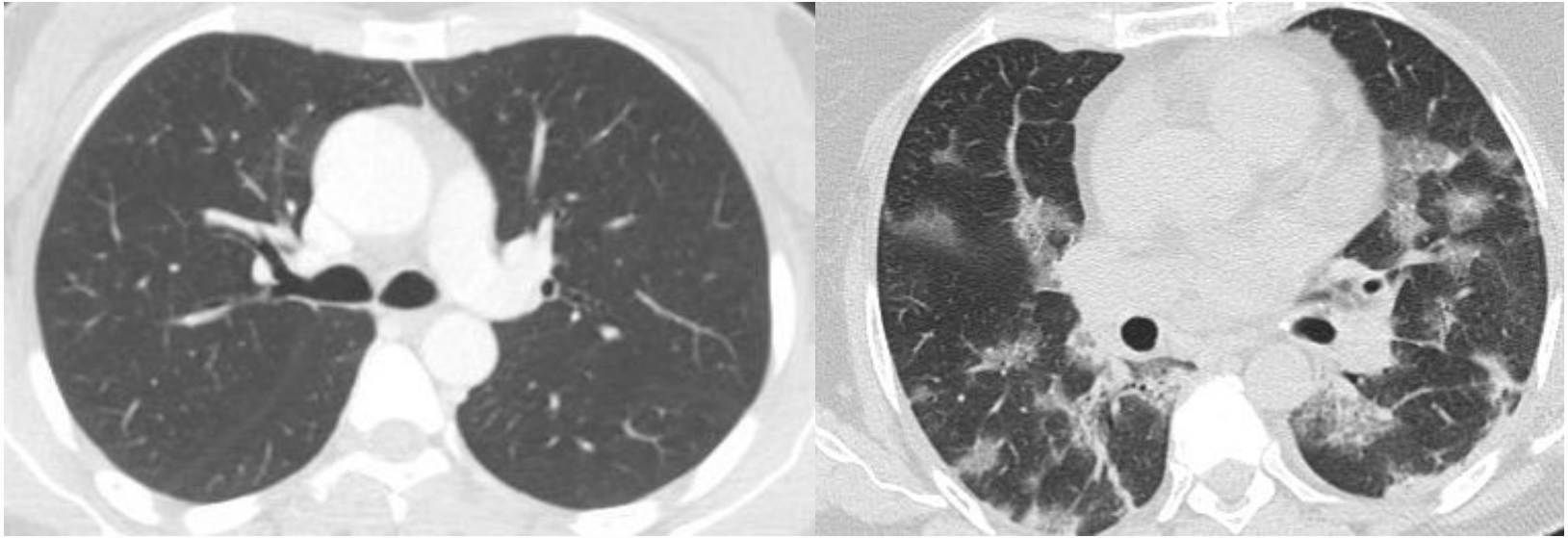
Background



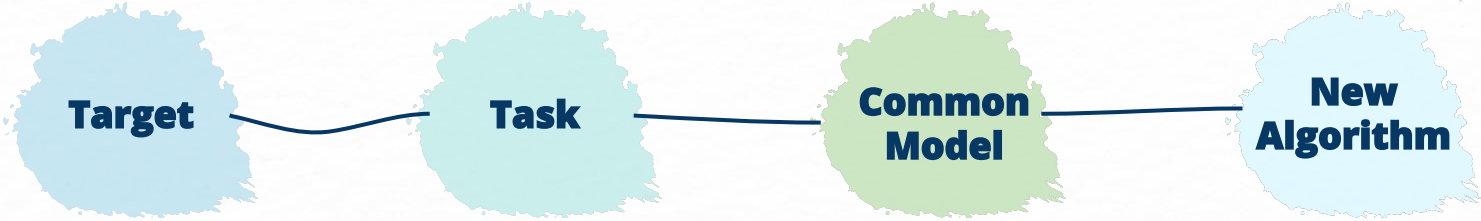
DAILY NEW CASES IN UNITED STATES

Developing machine-learning based programs to aid the diagnosis of COVID-19 and other lung diseases in combination with computed tomography

Healthy VS. Infected



Target Task & Proposed Solution



Identify possible COVID-19 infections by using machine learning approach

Analyze model results and compare each of them by accuracies and speed

Apply five common models

Develop a brand new algorithm which based on weight function

Common Model Summary

MLP

Multi-Layer Perceptron

CNN

Convolutional Neural Network

ResNet

Residual Network

KNN

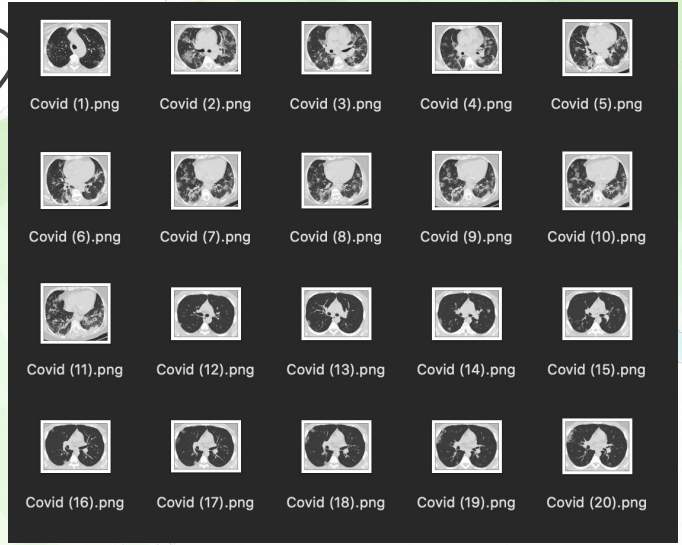
K-Nearest Neighbors

SVM

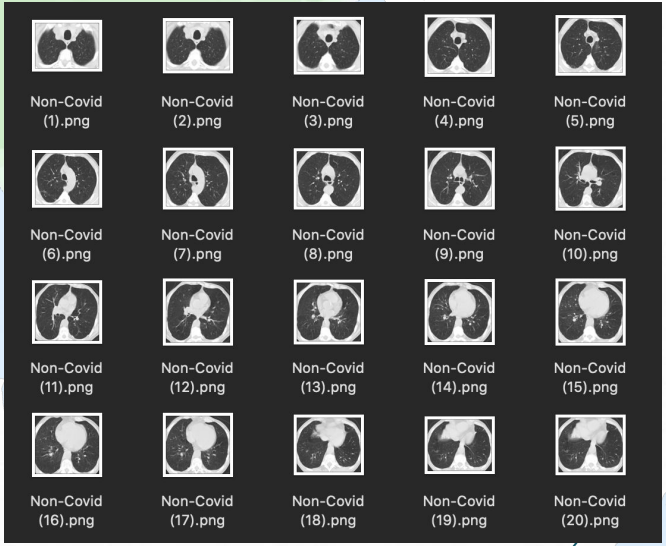
Support Vector Machine

Dataset

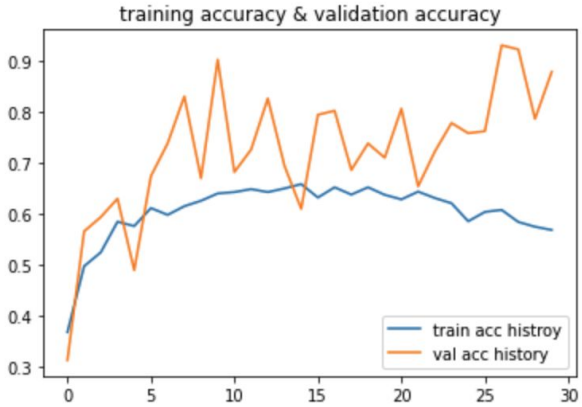
1252 COVID infected lung scans



1230 healthy lung scans



Implementation-MLP

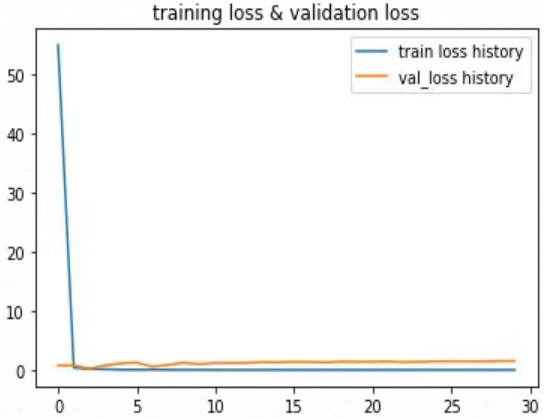


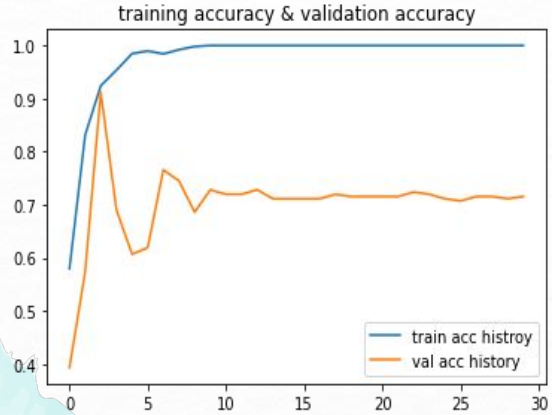

```
'learning_rate': 0.001,
'batch_size': 64,
'epochs': 30,
'validation_split': 0.1
```

Relative High Validation Accuracy: around **80**%

Medium Training Accuracy: around **60**%

Implementation-CNN






```
'learning_rate': 0.0001,
'batch_size': 64,
'epochs': 30,
'validation_split': 0.1
```

Optimizer: Adam

Highest training Accuracy: **100**%

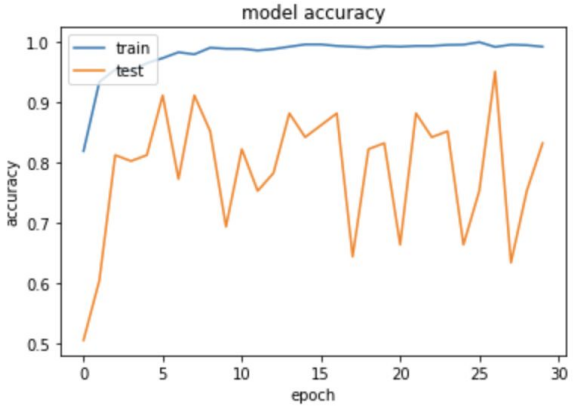
Relatively High validation accuracy: **70**%

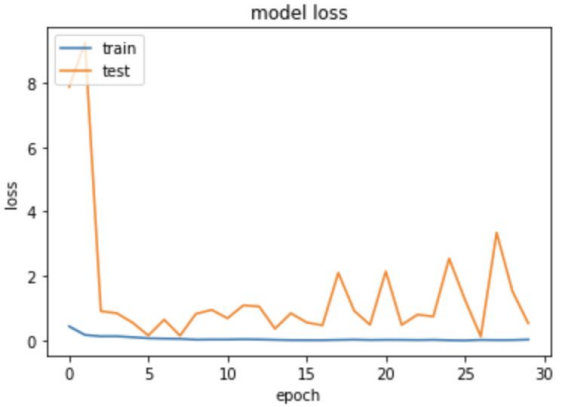
Relatively High test accuracy: around **80**%

Implementation-ResNet

01 Use transfer learning to create ResNet50, ResNet101, and ResNet152 models and fine tune each model.

02 Compare the validation accuracy of each model: ResNet152 is less stable and ResNet50 has relatively lower accuracy compared with the other two.

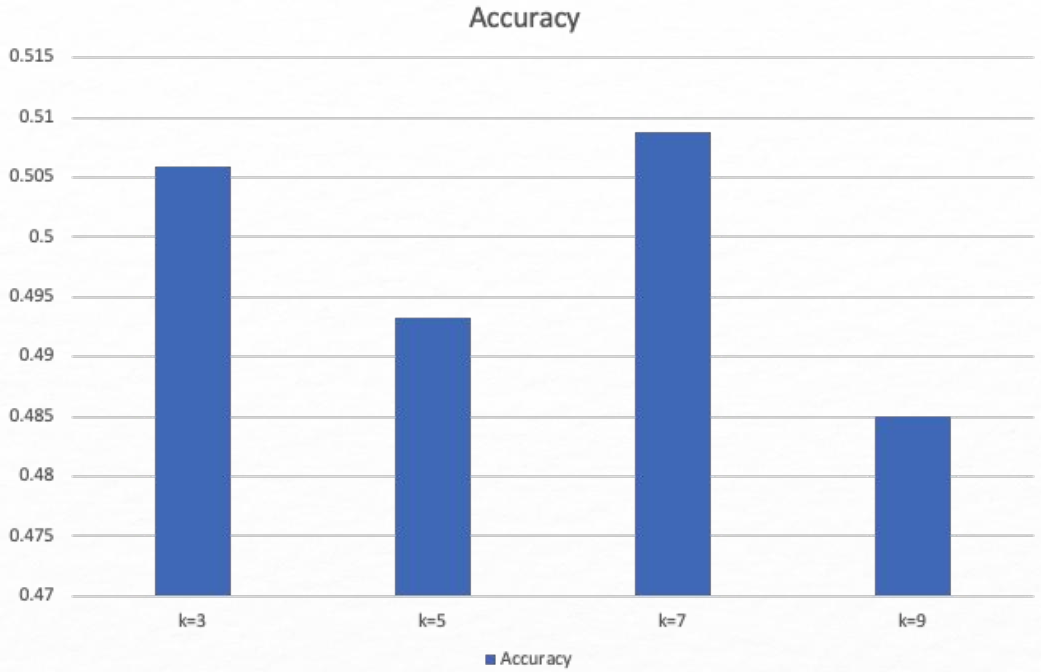





```
history2.model.evaluate(x_test_re_stack, y_test_re)

4/4 [==============================] - 0s 77ms/step - loss: 0.1294 - accuracy: 0.9505
[0.1294102668762207, 0.9504950642585754]
```

Implementation-KNN

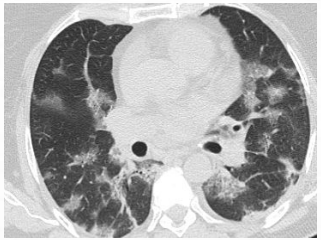


K Value	Accuracy
3	0.5058626465661641
5	0.4933061349108978
7	0.5087981714593032
9	0.48494773197089125

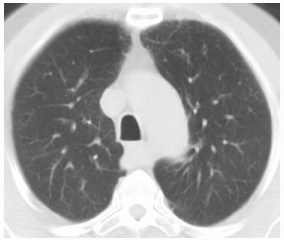
Implementation-SVM

01 Read in CT image files as numpy arrays

02 Create SVM matrix for training



COVID



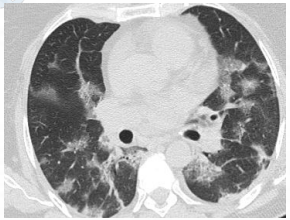
Non-COVID

pixel1	pixel2	...	pixel40000	label
198	184	...	254	1
68	68	...	228	0

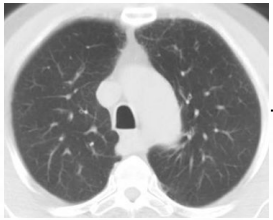
Implementation-Our Model

1. Read in CT image files as numpy arrays
2. Generate the weight of all images, determine the potential lung cavity region
3. Train the model
    a. Random select pixels, expand to 50x50 block
    b. determine if the block represents lung cavity
    c. determine parameters
4. Use the model to predict test samples
    a. Calculate scores for COVID and NON-COVID cases
    b. Compare the score to determine the result

Implementation - Our Model

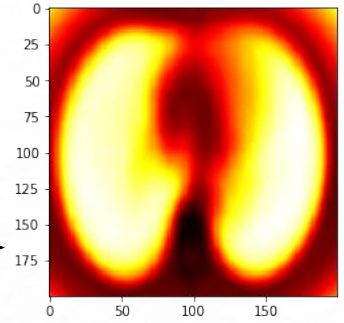


COVID

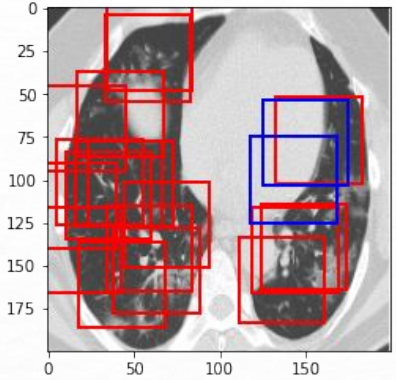


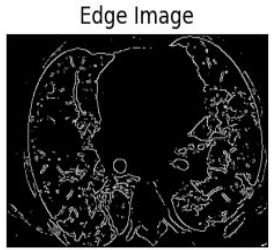
Non-COVID

pixel1	pixel2	…	pixel40000	label
198	184	…	254	1
68	68	…	228	0



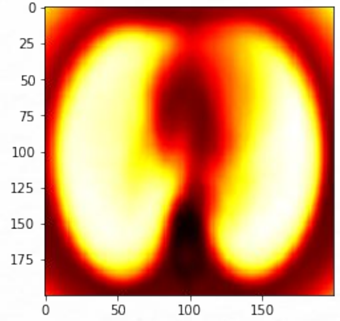
COVID Score: xxx
Non-COVID Score: xxx

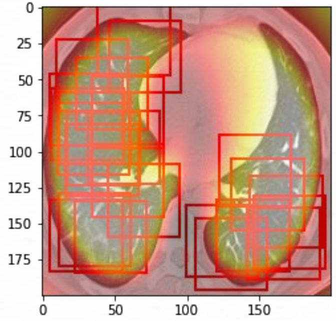


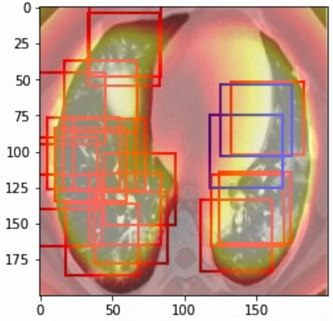


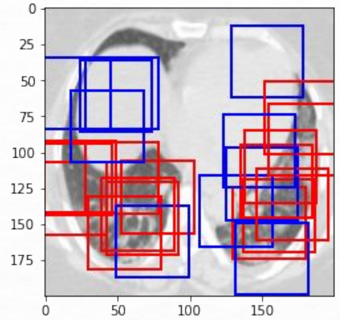

```
[ 184.92345317  122.13717546  174.48917112   49.18585586 7579.13316852]
[ 140.36916401  113.84491405  176.84756763   47.74327912 7305.1220815 ]
```

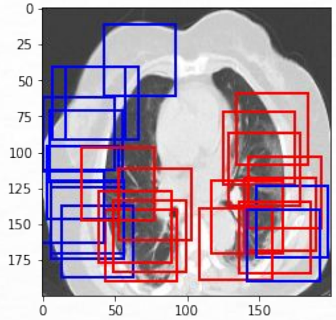
Implementation - Our Model

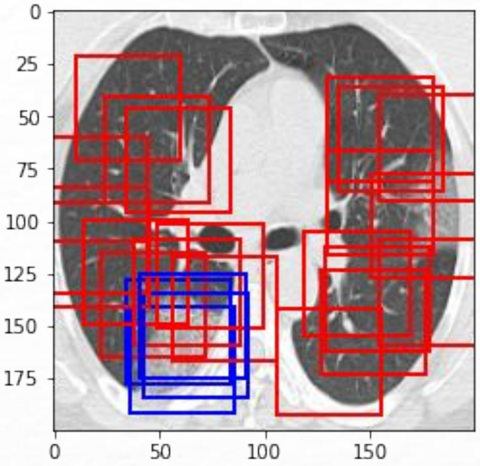




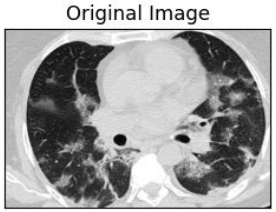


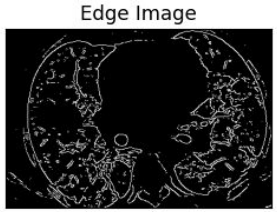


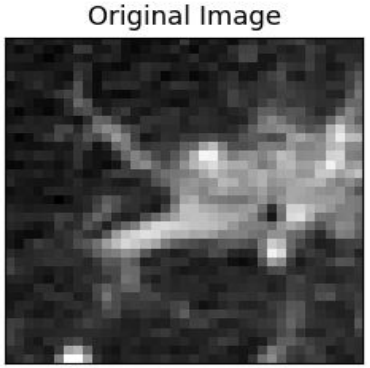


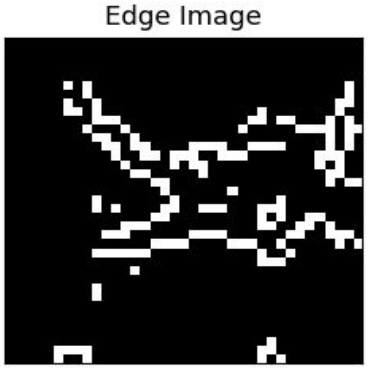


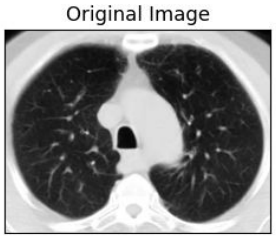
Implementation - Our Model

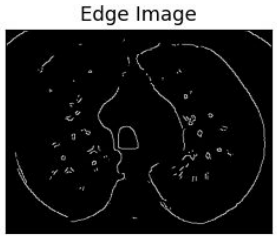












Contributions

- Find one model with the highest accuracy
- Provide doctors with auxiliary judgment on patients' conditions
- Provide a stable and rapid way for detecting covid ct-scans

Data Summary

	Best Test Accuracy	Average Time (min)*
MLP	0.7 ~ 0.8	10
CNN	0.78	5
ResNet101	0.95	25
KNN	0.73	10
SVM	0.88	20
Our Model	**0.85**	**1**

* Without the time of loading data

Experimental Result and Conclusion

MLP, CNN, ResNet101

- Higher average accuracy
- Blackbox
- ResNet101 has highest accuracy
  - Unstable (acc 64~96)
  - Time consuming (~25 min)

KNN, SVM

- Lower average accuracy
- Transparent
- Considers all pixels in an image

KNN, SVM

Initially considers all pixels
Faster speed

Evaluate lung cavity only

High Accuracy

Our Model

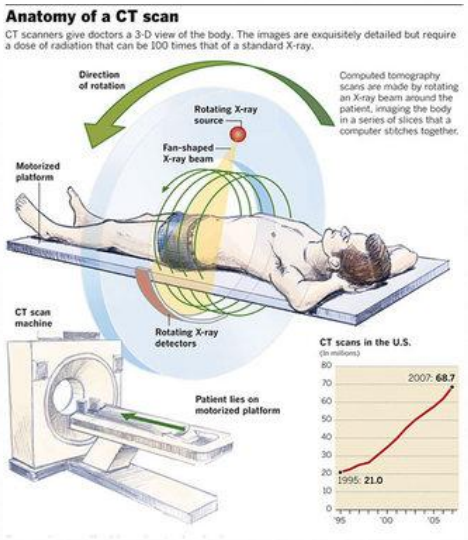
High speed

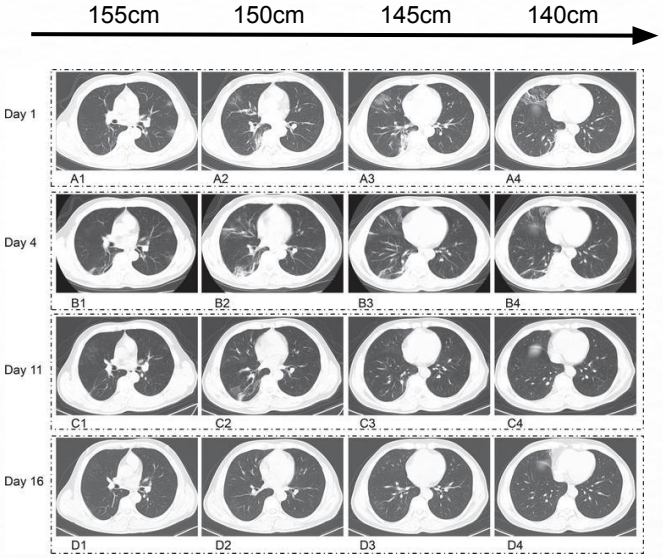
MLP, CNN, ResNet101

Consider key regions with higher weight

Calculate transparent scores

Experimental Analysis





Future Work

- Get more training data that includes detailed information about lung cavity and try to improve the performance of our model
- Find CT scans dataset that has already been segmented by radiologists and try other models such as YOLO to determine the anatomical location of the lesion

References

- https://blog.paperspace.com/fighting-coronavirus-with-ai-building-covid-19-classifier/
- https://docs.opencv.org/master/
- https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_imgproc/py_canny/py_canny.html
- https://arxiv.org/abs/1512.03385
- https://www.medmastery.com/guide/covid-19-clinical-guide/how-identify-covid-19-chest-ct-scan

THANKS