Lung Cancer Image Classification and Bounding Box Predictions

By Daniel Sun

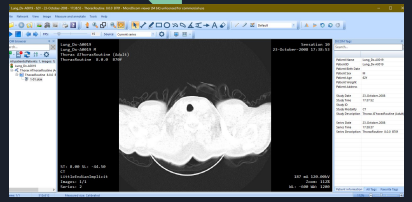
The Dataset

- Taken from Cancer Imaging Archive: A Large-Scale CT and PET/CT Dataset for Lung Cancer Diagnosis

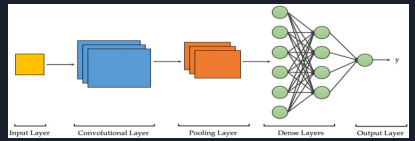
- Entire dataset:
  - 127gb Dicom files
  - 4 classes of lung cancer
  - Hand-annotated bounding boxes in separate XML files

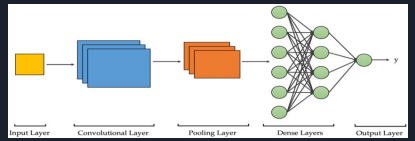
- My subset:
  - ~ 15gb
  - ~1000 images used for training

General Workflow



Training





Evaluate final models

Pull out predictions

The metric IOU (Intersection Over Union)

- IOU = Area of Overlap / Area of Union：

Initial Statistics of Bounding Box Predictor Model Performance

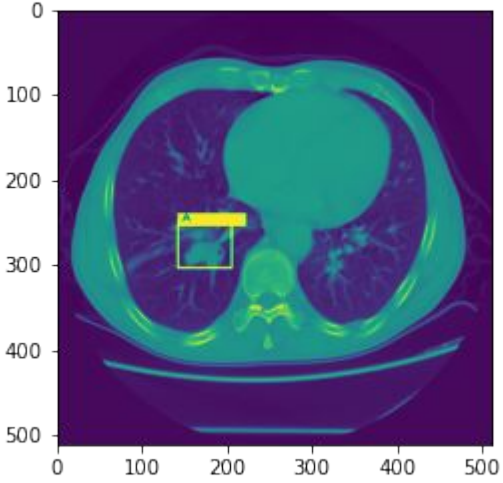
- Ratio of predicted Boxes *with* overlap: 0.772
  - Total testing length: 232
  - Total non-zero IOU: 179

- Average overlap of non-zero IOUS: 0.377
  - Average overlap of ALL: 0.291

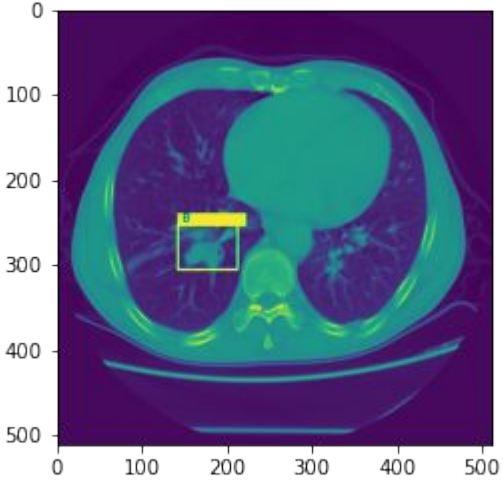
Best prediction

IOU score: 0.870

Model Predicted



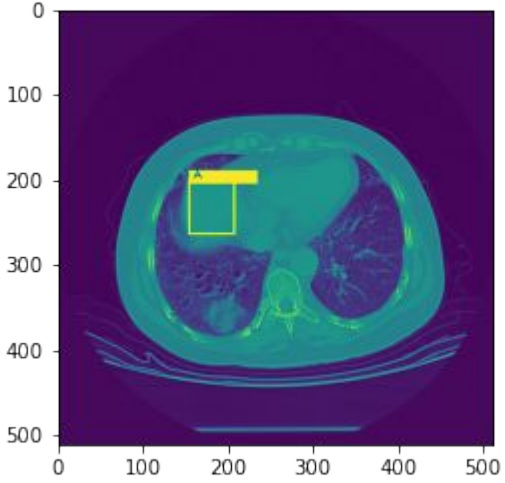
Actual



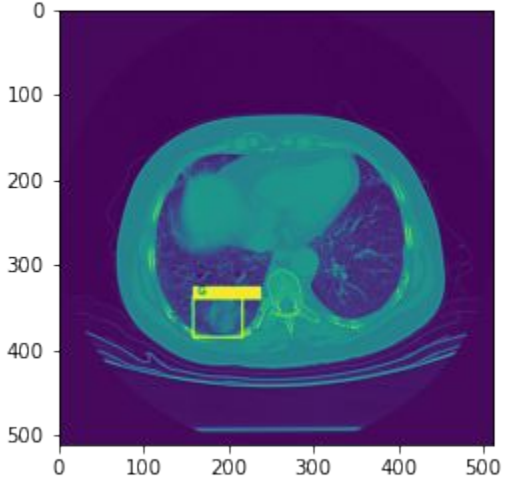
Worst Prediction

IOU score: 0

Model Predicted



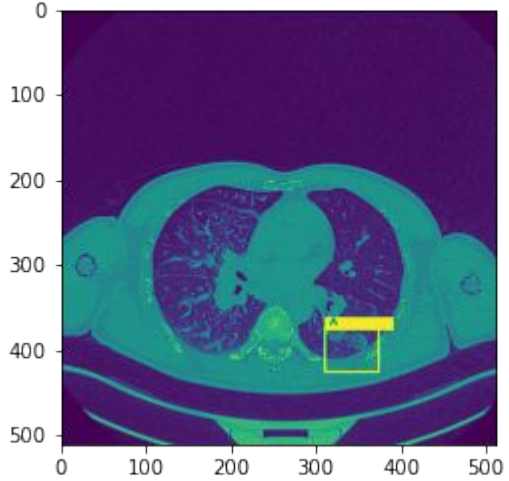
Actual



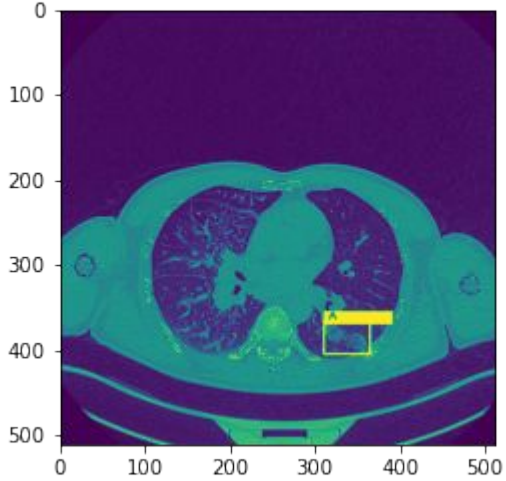
The Average Prediction

IOU score: 0.455

Model Predicted

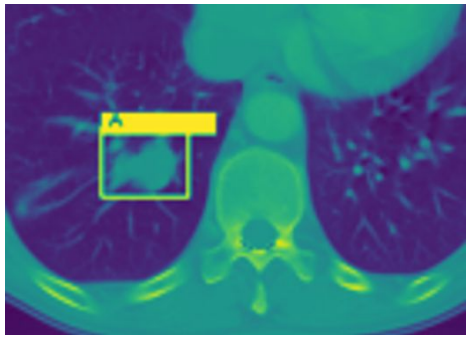


Actual



Classification using Bounding Box versus Image Data

Bounding Box Coord → Predicts: A



Actually: B

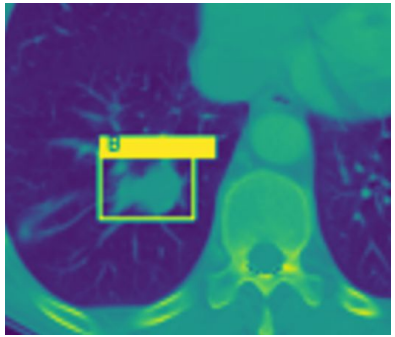
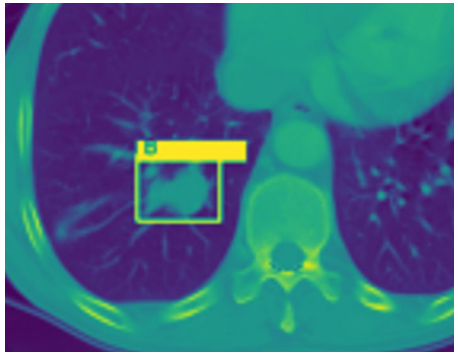


Image Data → Predicts: B



Classification Accuracy

Accuracy using Bounding Box Coordinates to classify			
A	B	E	G
0.830	0.667	0.513	0.671
Weighted F1 Score		0.704	
Accuracy using Image Data to classify			
0.959	0.959	0.889	0.940
Weighted F1 Score		0.947	

Testing on a new set of images

Accuracy using Bounding Box Coordinates to classify			
A	B	E	G
0.349	0.139	0.386	0.468
Weighted F1 Score		0.402	
Accuracy using Image Data to classify			
0.583	0.421	0.107	0.431
Weighted F1 Score		0.452	

- Ratio of predicted Boxes *with* overlap: 0.215
  - Total testing length: 1861
  - Total non-zero IOU: 401
- Average overlap of non-zero IOUS: 0.214
  - Average overlap of ALL: 0.105

Conclusion

- Model is not ready for production: Accuracy and IOUs need to be much more accurate on testing sets

- Model is niche: trained on images where it is given that there *is* a cancerous mass
  - Cannot say ‘there is no cancerous mass’ within an image

- Predicting class by total image slice may be more viable than by location of tumor within image slice.

Future Steps

- Upgrade hardware to be able to handle more complex models
  - Perhaps utilize cloud gpus for offloading memory pressure
- Begin with a pre-trained neural net like EfficientNet or ResNet50V2
- Utilize a larger training set or the entirety if possible