Lung Nodule Classification and Localization

Suman Gunnala, Anil Luthra, Tony Reina, Kyle Shannon

UC San Diego
JACOBS SCHOOL OF ENGINEERING

UCSDCSE
Computer Science and Engineering

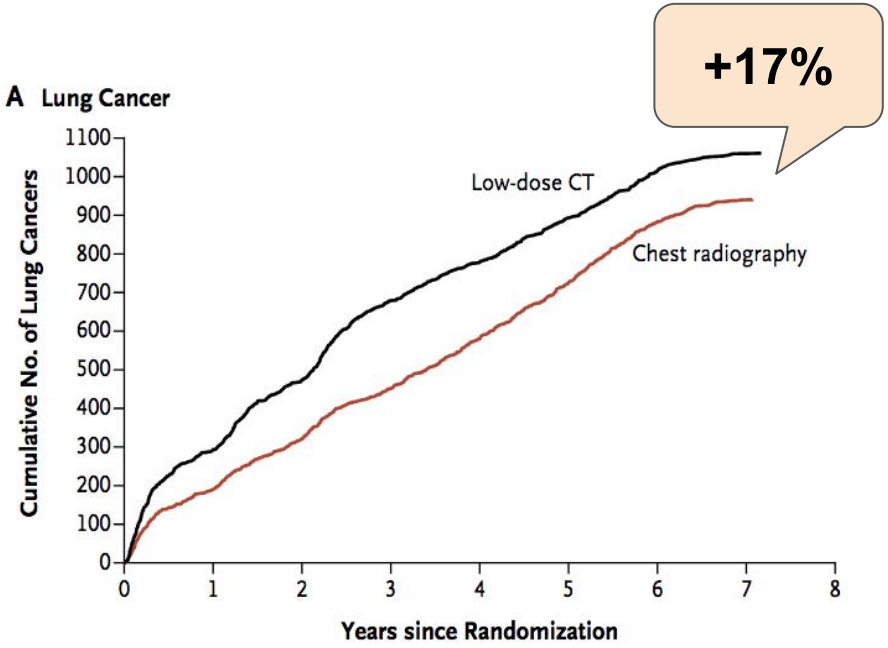
The NEW ENGLAND JOURNAL of MEDICINE

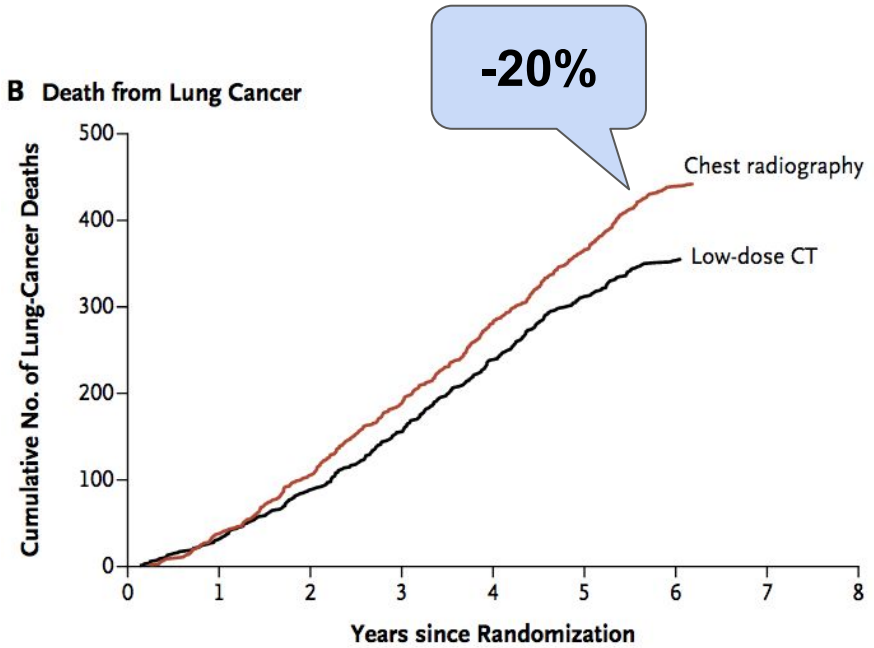
ESTABLISHED IN 1812 AUGUST 4, 2011 VOL. 365 NO. 5

Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening

The National Lung Screening Trial Research Team*

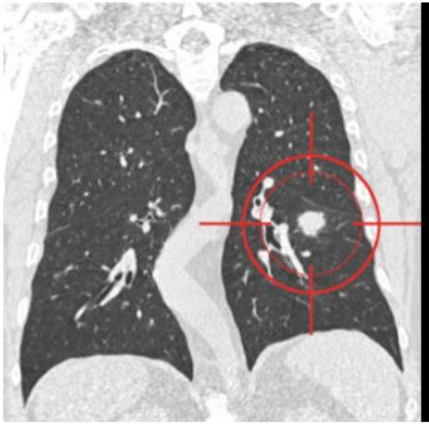
NEJM 2011; 365(5):395-409





Chest CT
Detects **MORE** Nodules

Chest CT
Saves **MORE** Lives



LUng Nodule Analysis 2016

Info

Join

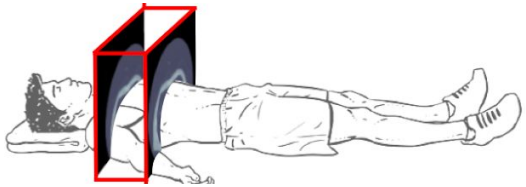
Home
Description
Procedure
Data
Evaluation
Results
Download
Submit
Tutorial
Email organizers

News

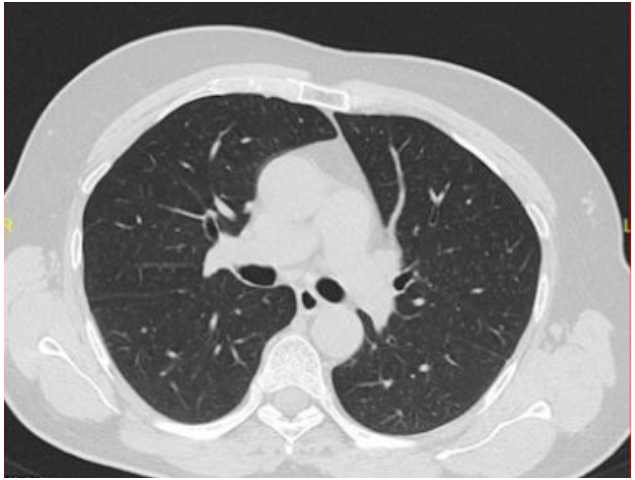
- January, 2018: We have decided to stop processing new LUNA16 submissions. Read more ...
- September, 2017: We have decided to stop processing new LUNA16 submissions without a clear description article. Read more ...
- June, 2017: The overview paper has been accepted for publication in Medical Image Analysis: https://doi.org/10.1016/j.media.2017.06.015
- May, 2017: Kaggle has held a competition that may be of interest for participants of LUNA16: https://www.kaggle.com/c/data-science-bowl-2017

LUng Nodule Analysis 2016

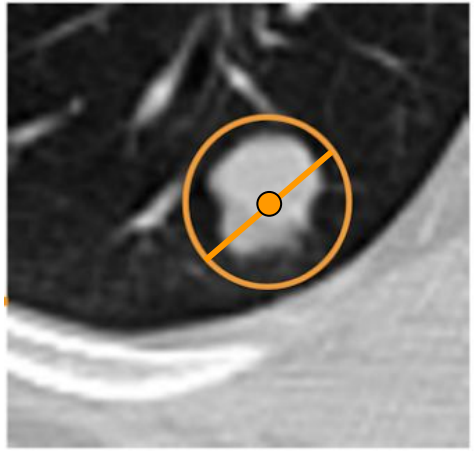
Lung cancer is the leading cause of cancer-related death worldwide. Screening high risk individuals for lung cancer with low-dose CT scans is now being implemented in the United States and other countries are expected to follow soon. In CT lung cancer screening, many



Chest CT Slice



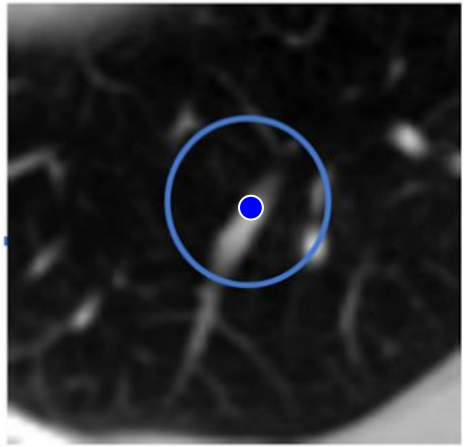
Class 1



(x,y,z)

(d mm.)

Class 0

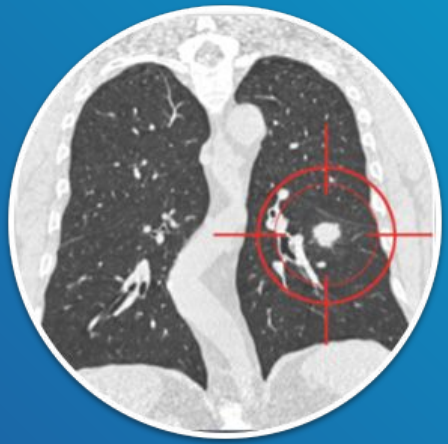


(x,y,z)

4 Radiologists annotated nodules > 3mm

Algorithm marked class 0s

10 Subsets for C.V.

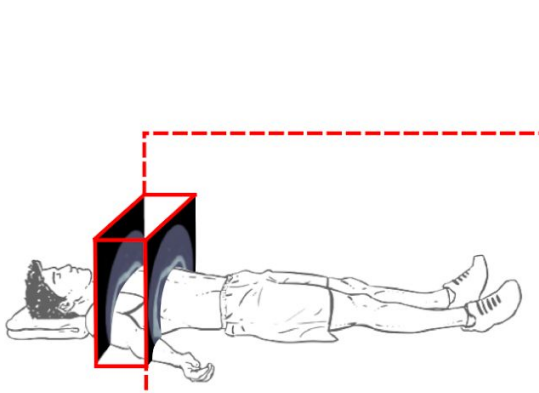


LUNA16 Challenge

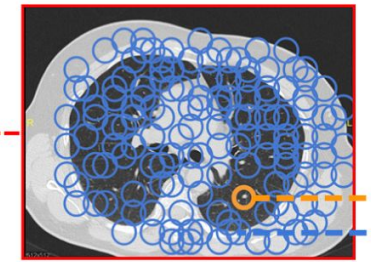
LUng **N**odule **A**nalysis 20**16**

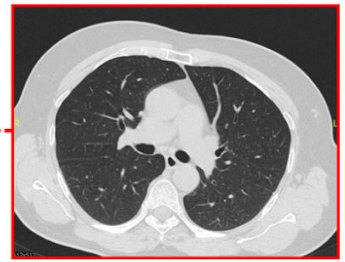
Med Image Analysis Dec 2017; Volume 42:1-1r3

888 CTs = ~300 GBs of DICOM/.mhd + Metadata



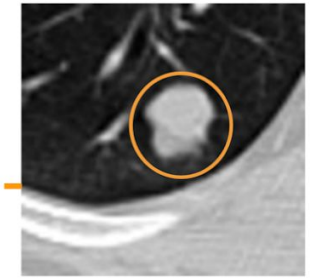
3D Chest CT

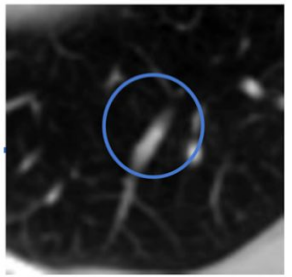




2D CT Slice

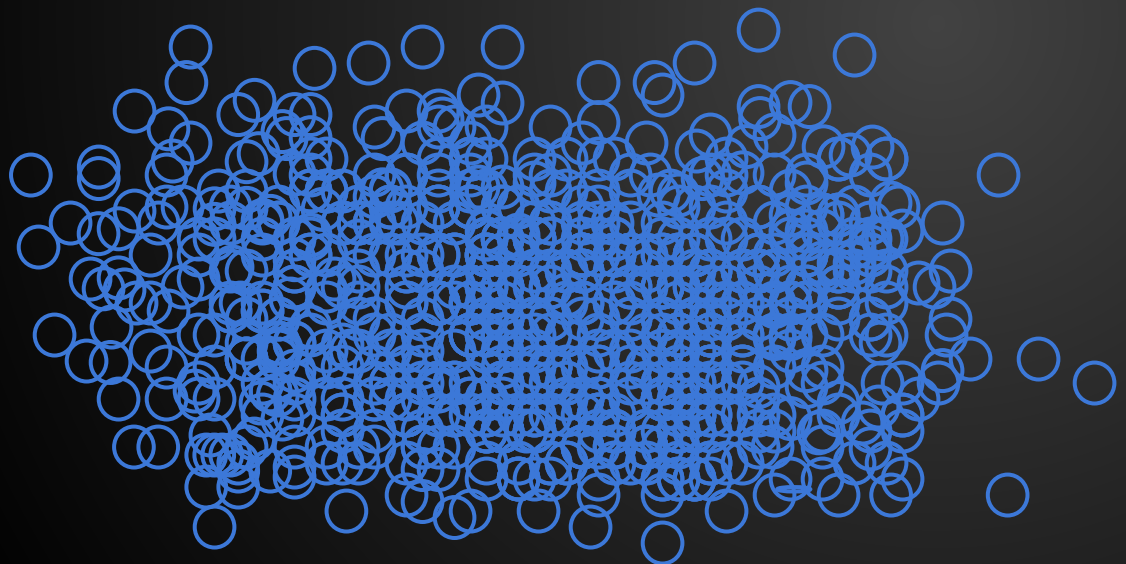
ROI: Class 1, ~1,200



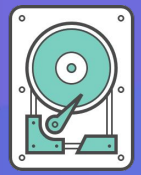


ROI: Class 0, ~550k

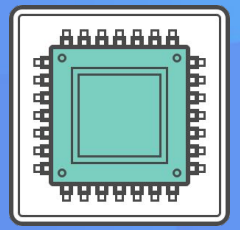
700 : 1



process/clean raw data



ready to train?

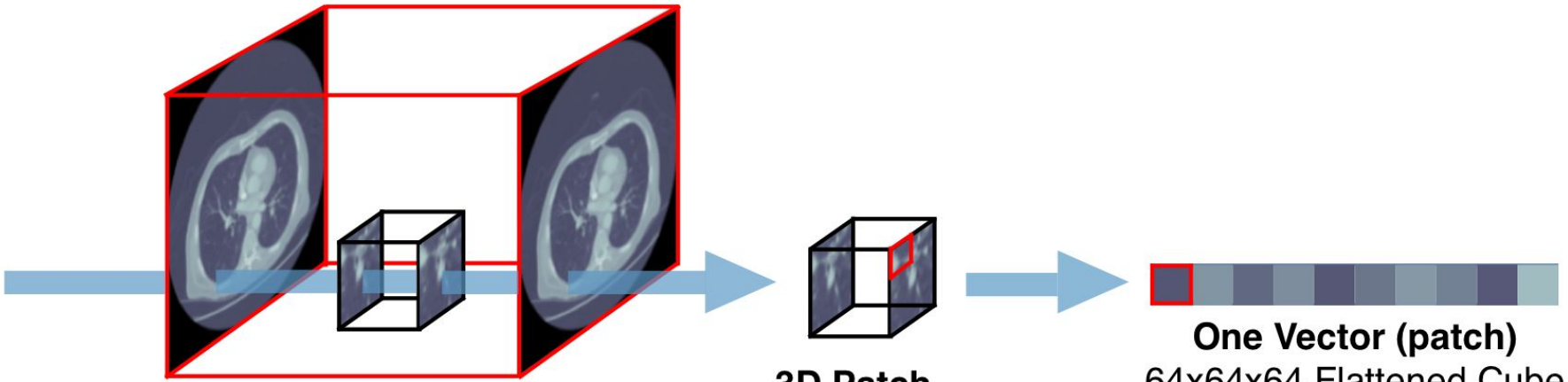


Building a Great Dataset!

EDA

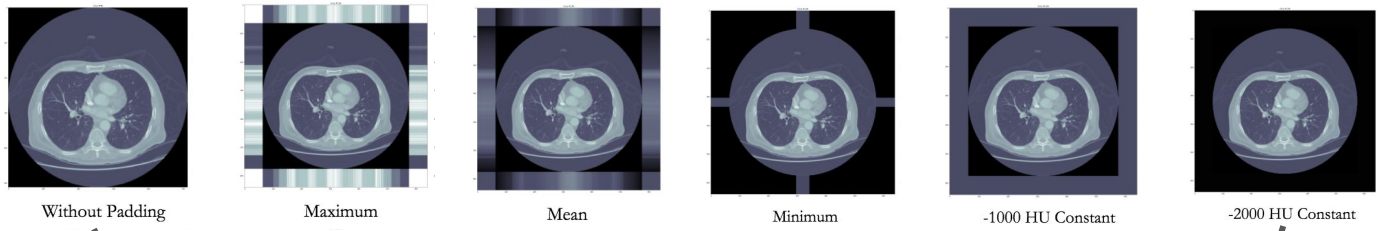
data storage

Processing Raw Data

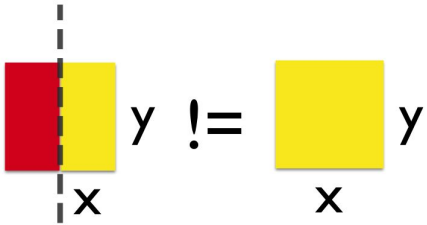


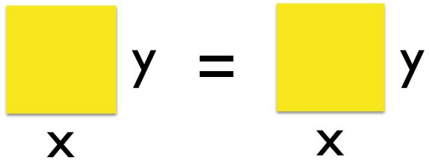
Patch Creation Challenges (2D shown)

Possible Solutions

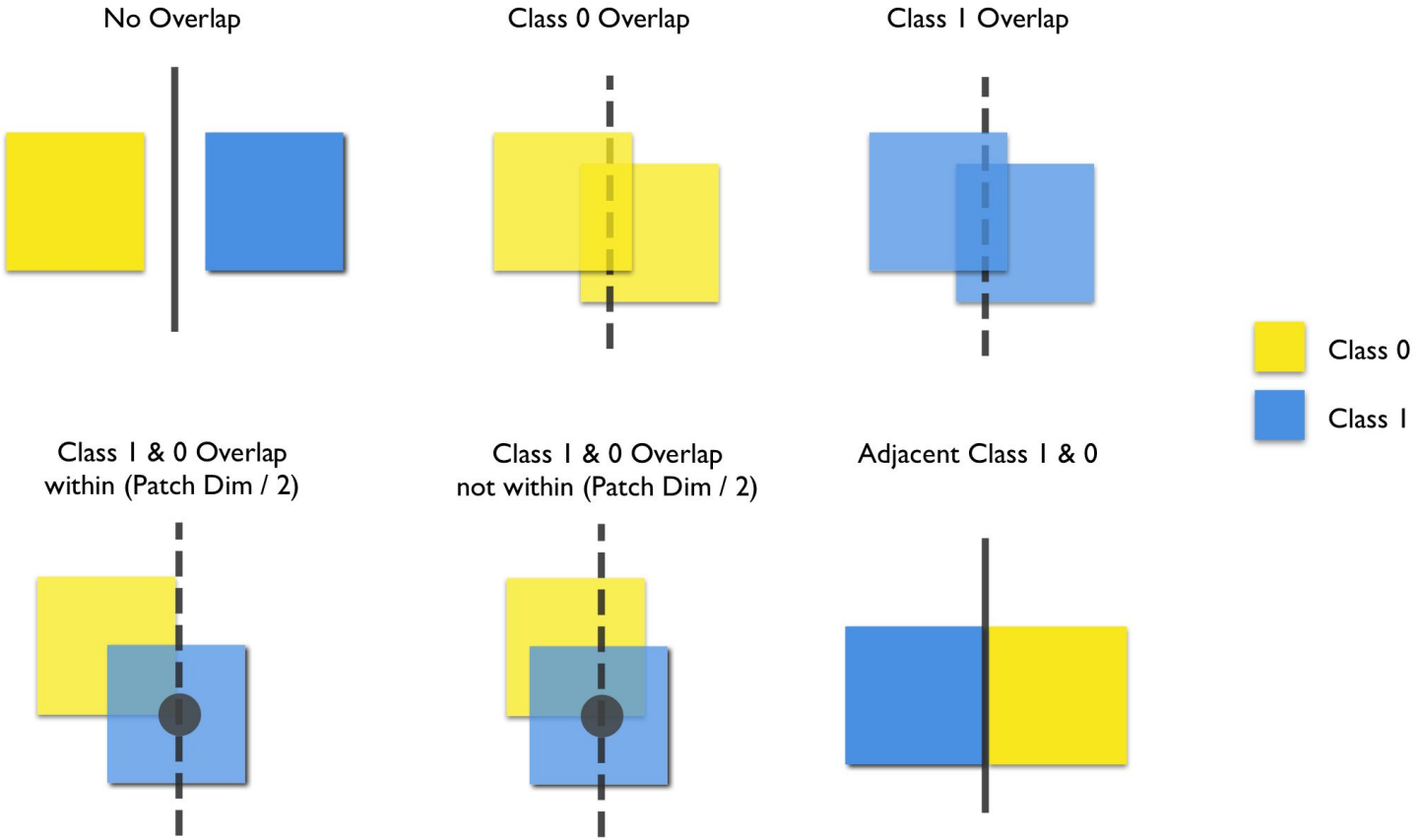


Best Solution:
-2000 HU Constant

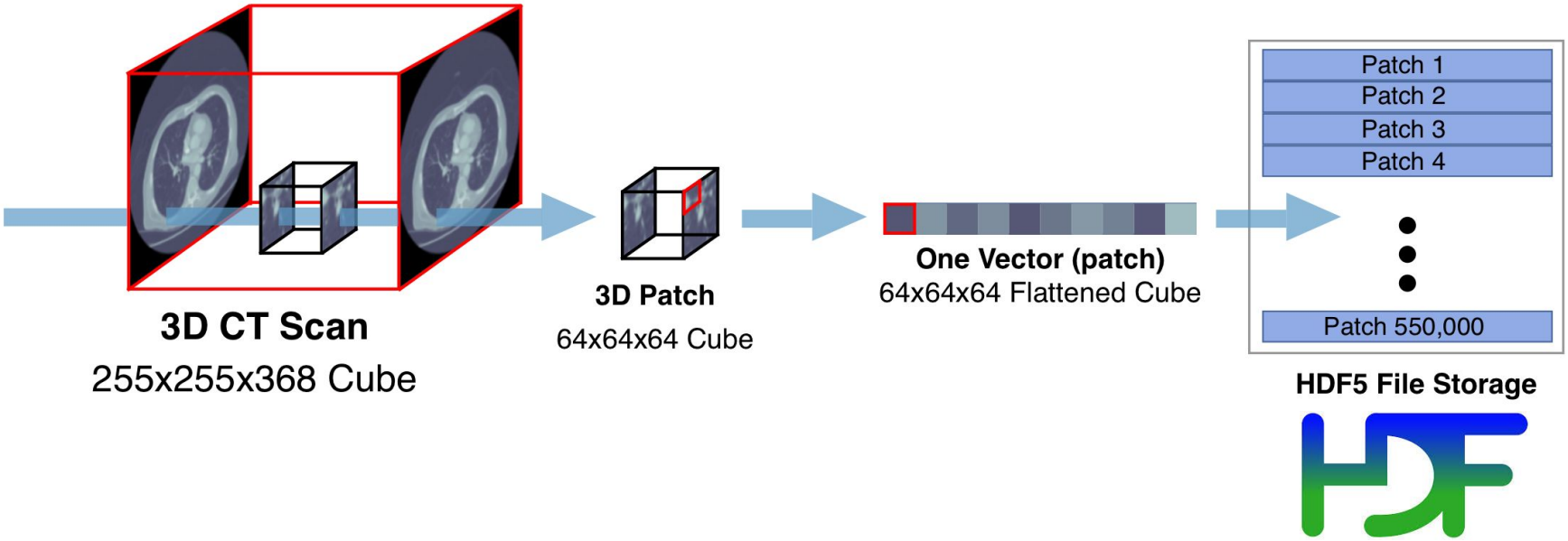




Patch Creation Challenges (2D shown)



Data Storage Pipeline



888 Total
3D CT Scans

550,000
3D Patches

550,000
Vectors (patches)

Single ~1.0 TB
HDF5 File Storage

Computer Vision Primer

Classification

Vehicle

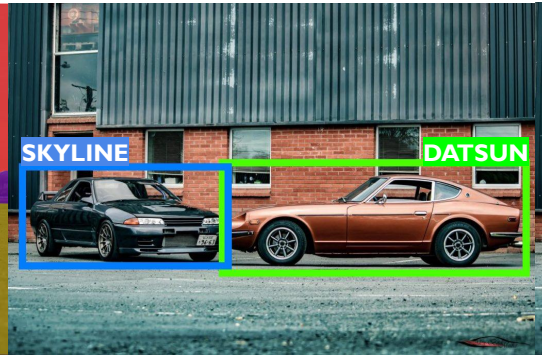
One object

Semantic Segmentation

- Vehicle
- Sky
- Forest
- Road

Each pixel classified, providing location

Object Detection



- Datsun 280z
- Nissan Skyline

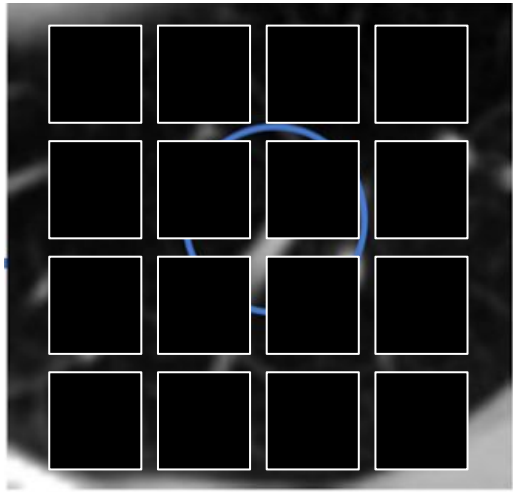
Multiple objects, located

Instance Segmentation, new!

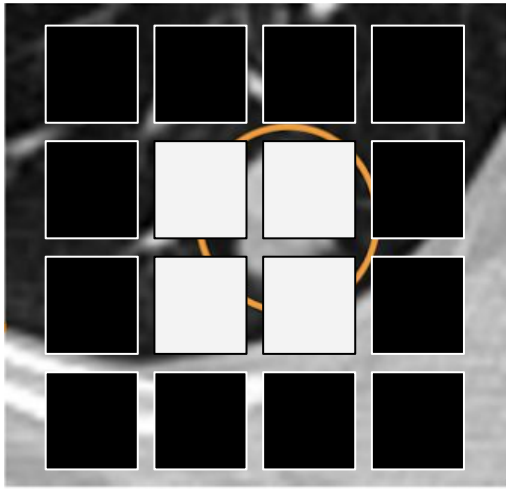
- Datsun 280z
- Nissan Skyline

What's Our Goal?

Semantic Segmentation

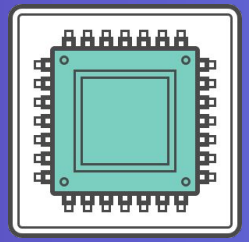


Class 0

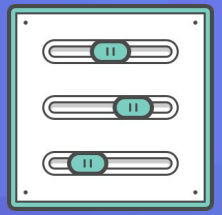


Class 1

Training a Model!



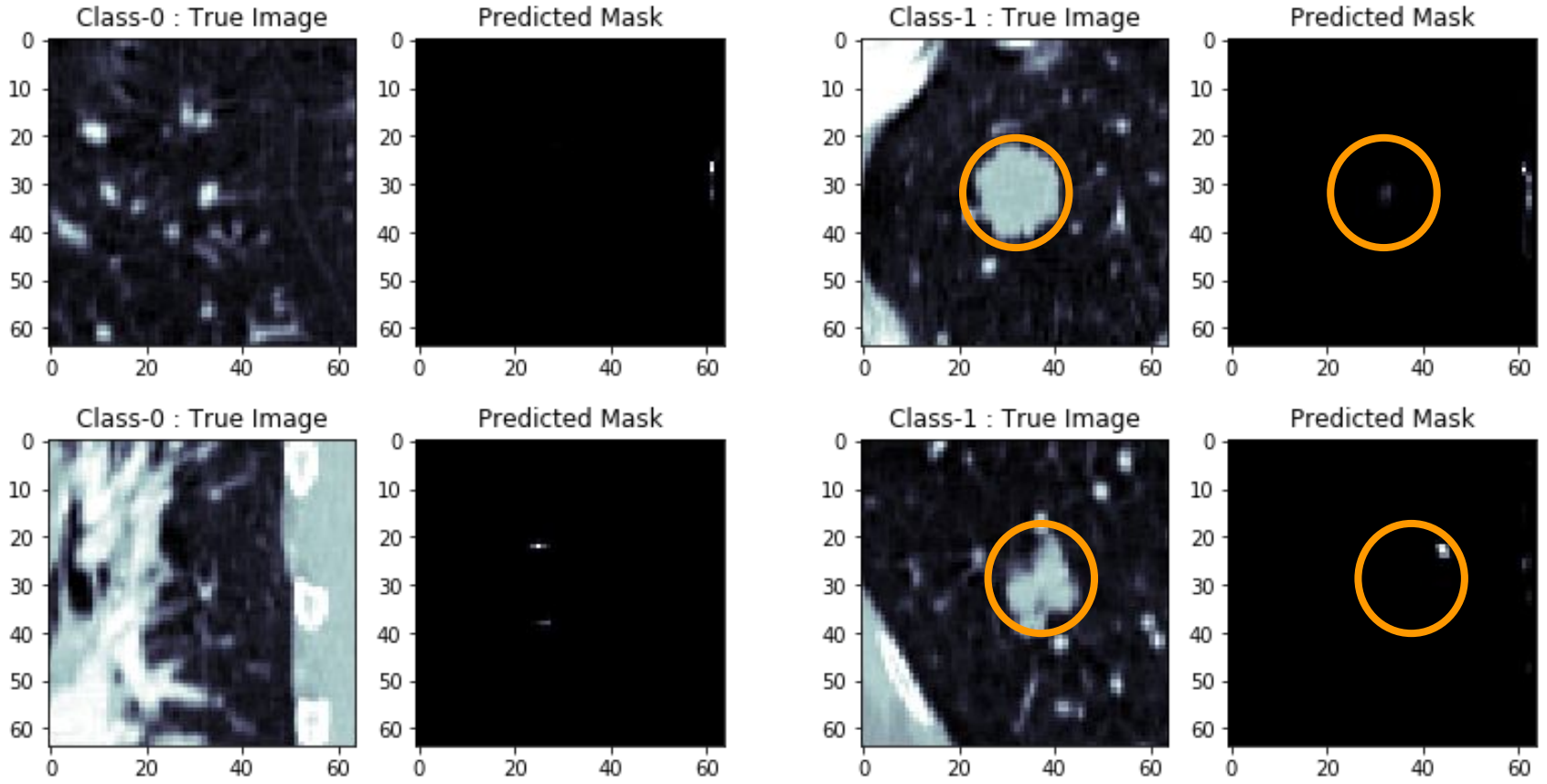
preprocess
more data… :/

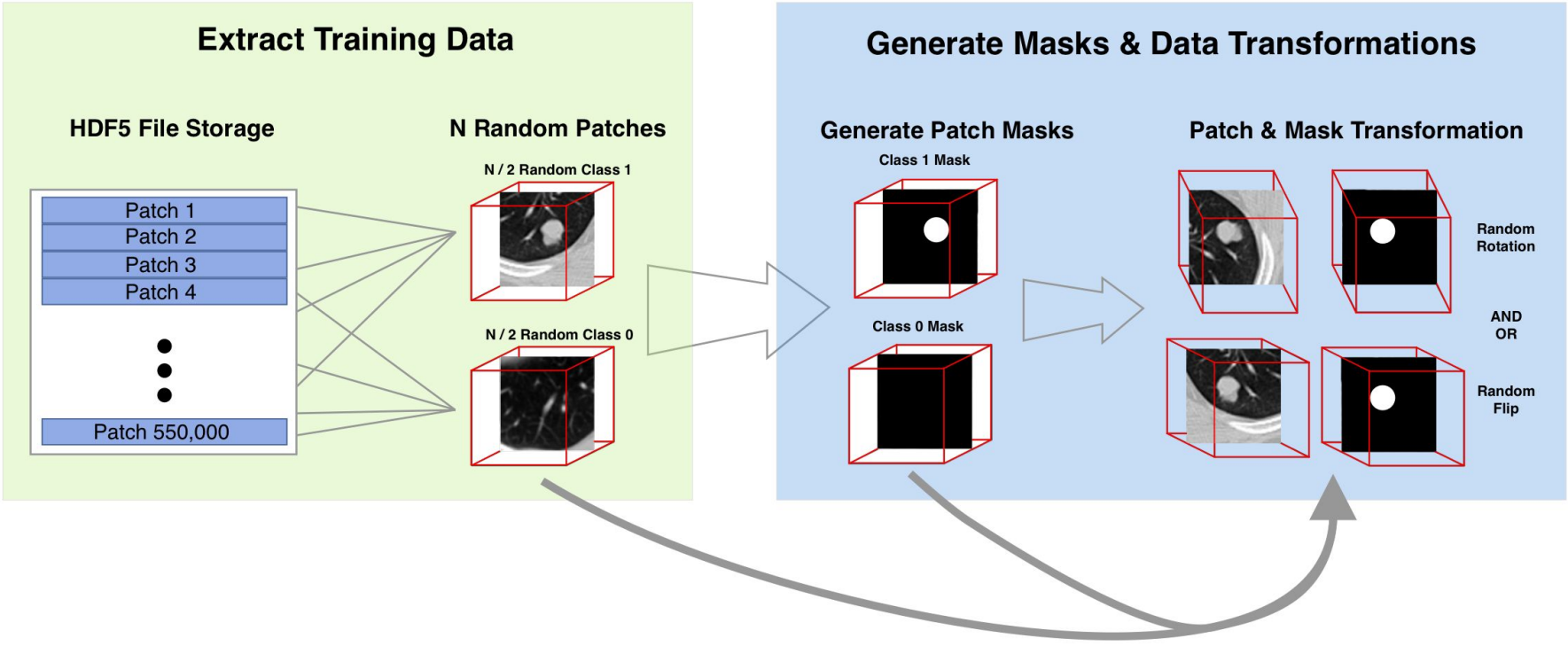


HyperParam
Tuning… :/

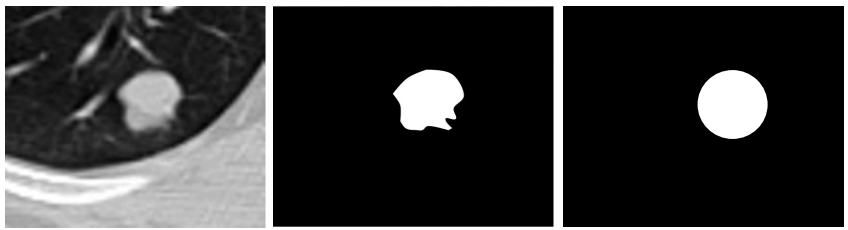
Waiting… :/

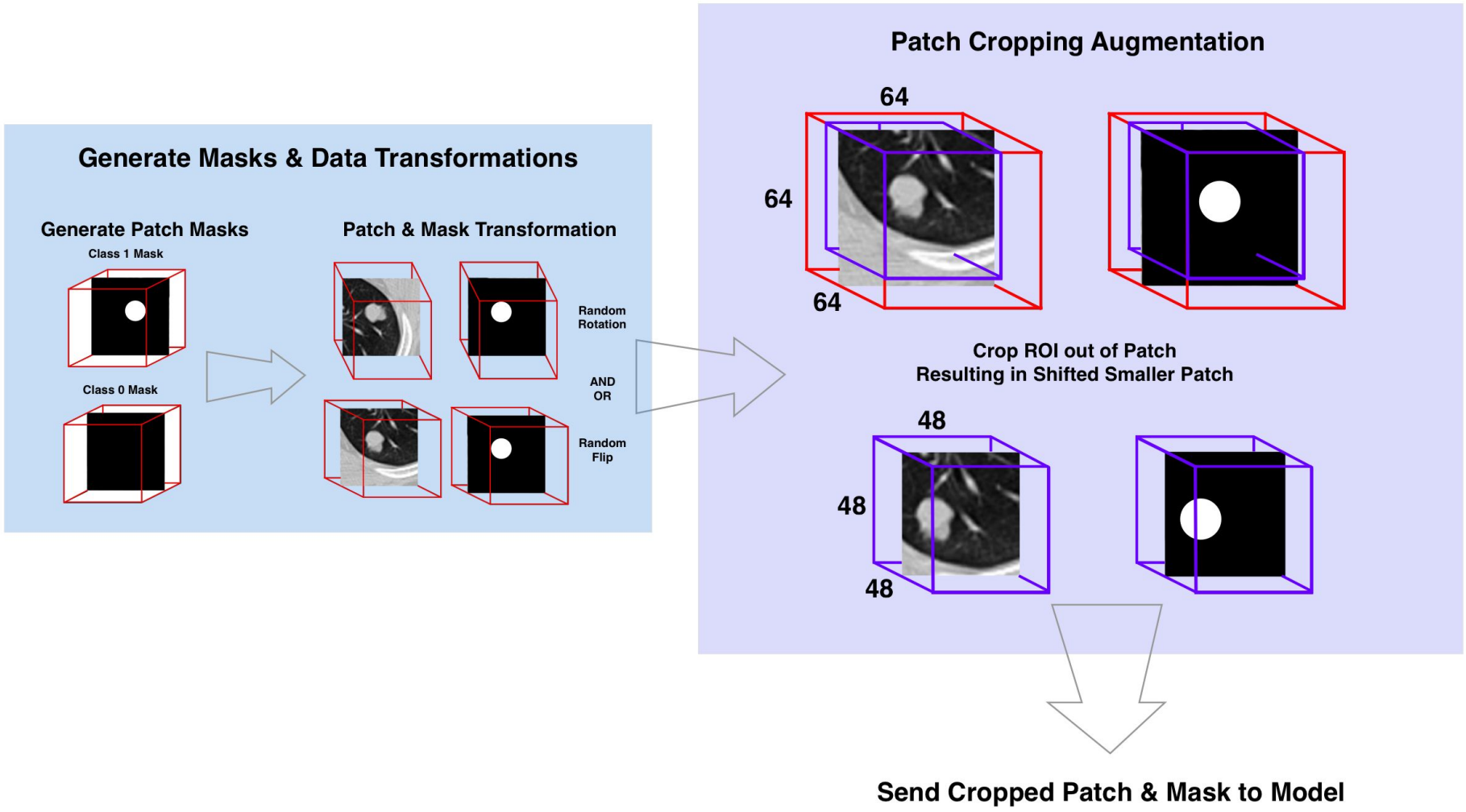
Initial Trial Run UNet Results

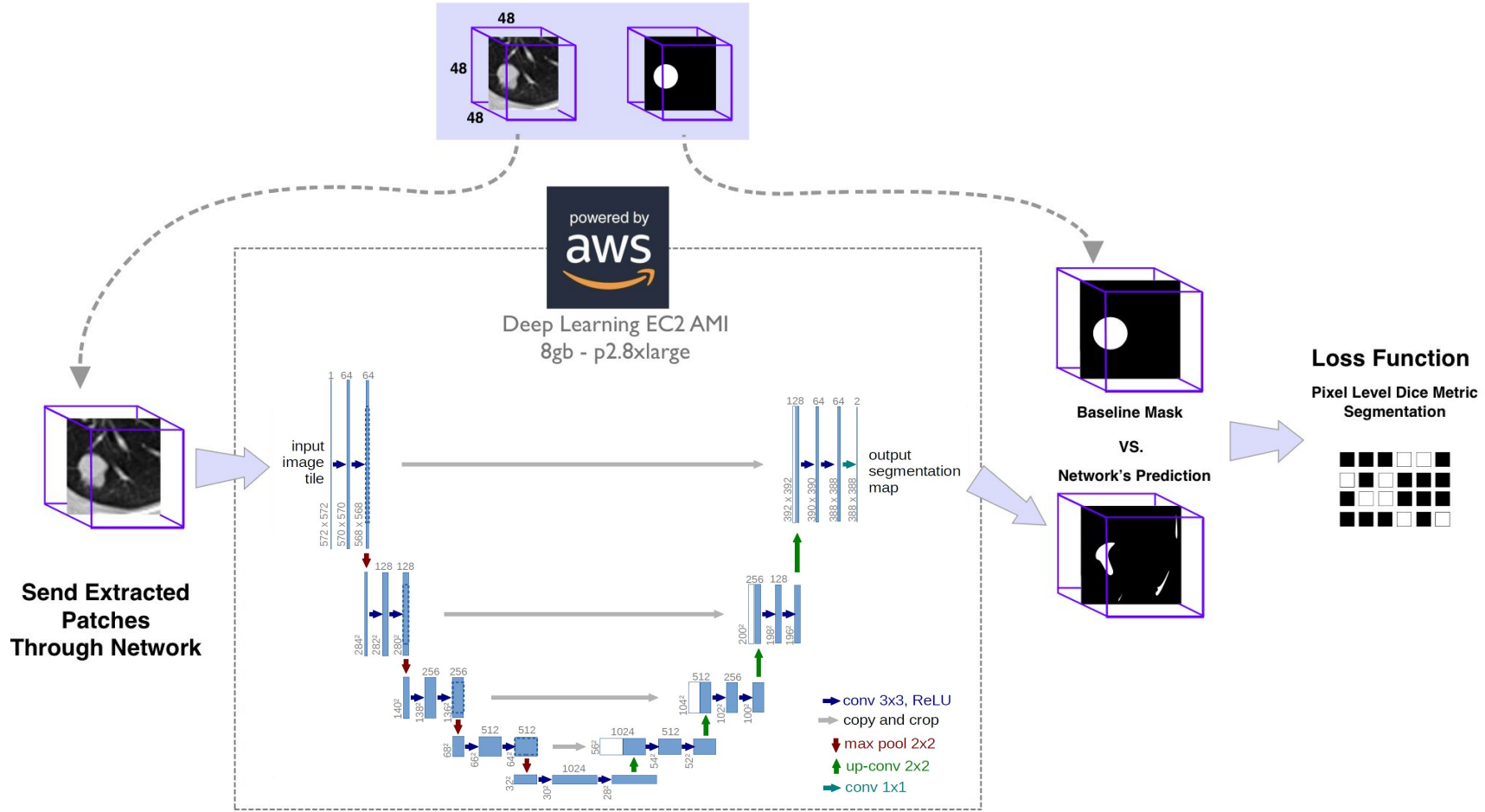




Generating Patch Masks:



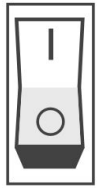

Generate Masks & Data Transformations
Generate Patch Masks
Class 1 Mask
Class 0 Mask
Patch & Mask Transformation
Random Rotation
AND OR
Random Flip
Patch Cropping Augmentation
64
64
64
Crop ROI out of Patch
Resulting in Shifted Smaller Patch
48
48
48
Send Cropped Patch & Mask to Model

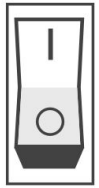


Ronneberger, Olaf et al. "U-Net: Convolutional Networks for Biomedical Image Segmentation." ArXiv abs/1505.04597 (2015)

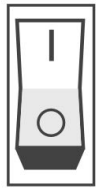
Model HyperParameter Tuning

Epochs 4

Batch Norm  Yes

HU Norm  No

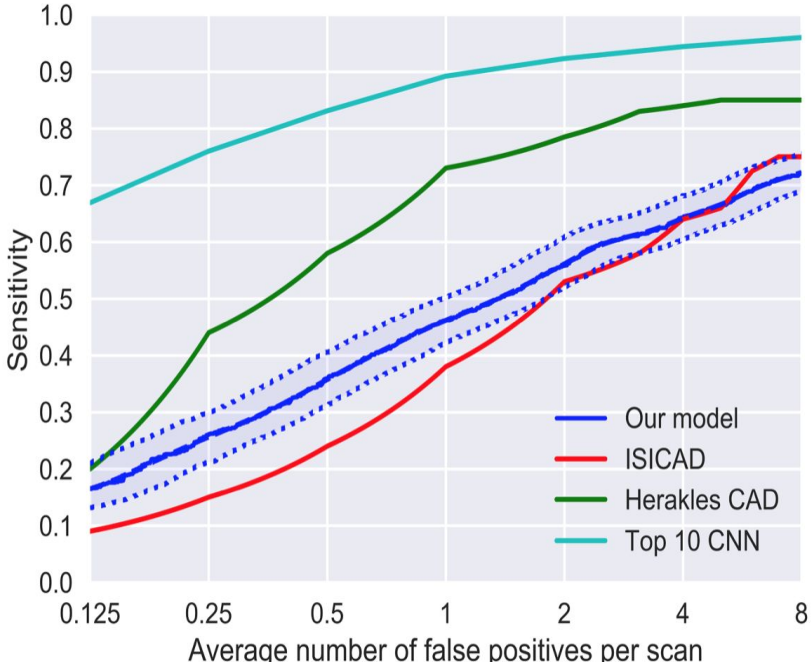
Learning Rate .009

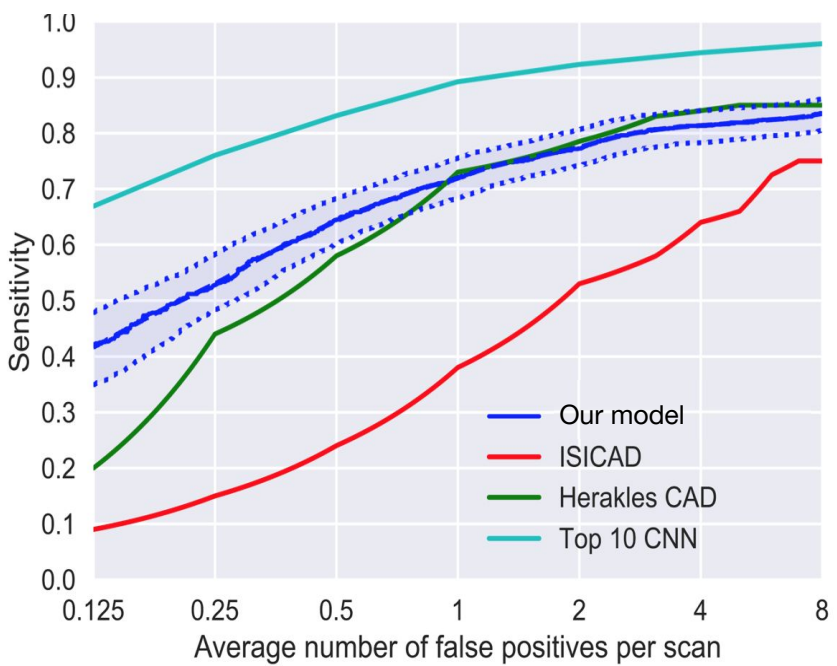
Weighted Loss Function  No

Resolution Layers 4

Stratified Sampling 3:1

FROC (Free-Response Operating Characteristic) Assessment Results Comparison (3 vs 64 Slices)



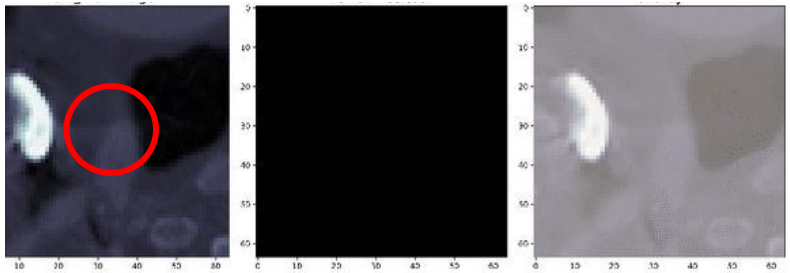


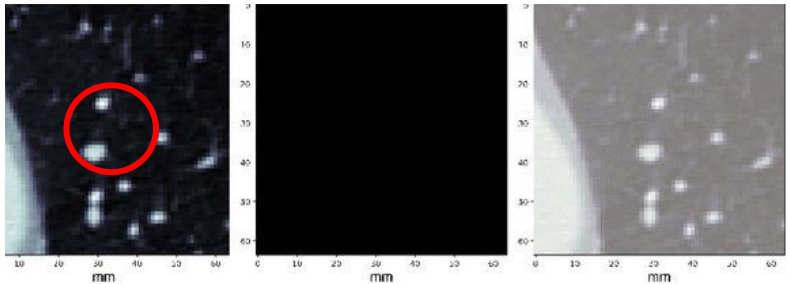
Bandos, A. I., et al. (2009). Area under the free-response ROC curve (FROC) Biometrics, 65(1), 247-256.

Dice Coefficient (0.0 - 1.0 Scale)

True Positive

Original Img | Tumor Prediction | Overlay



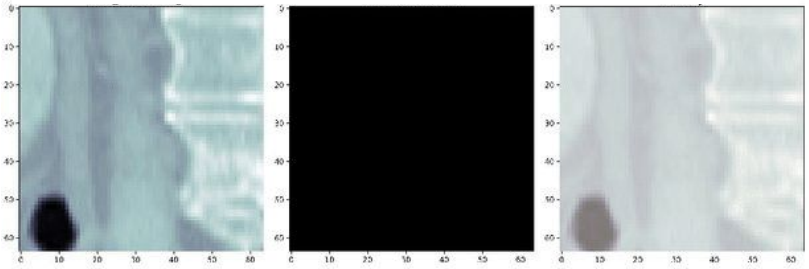


Look for the nodule in the circle

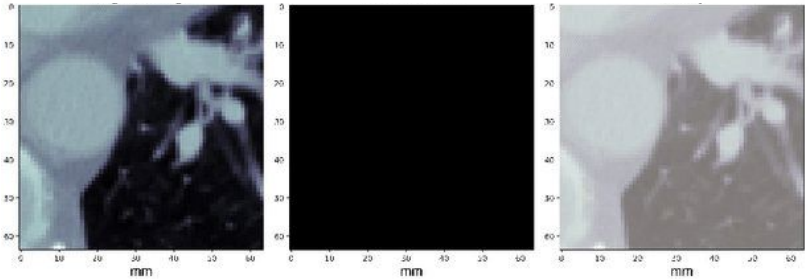
True Negative

Original Img | Tumor Prediction | Overlay

Sagittal Plane

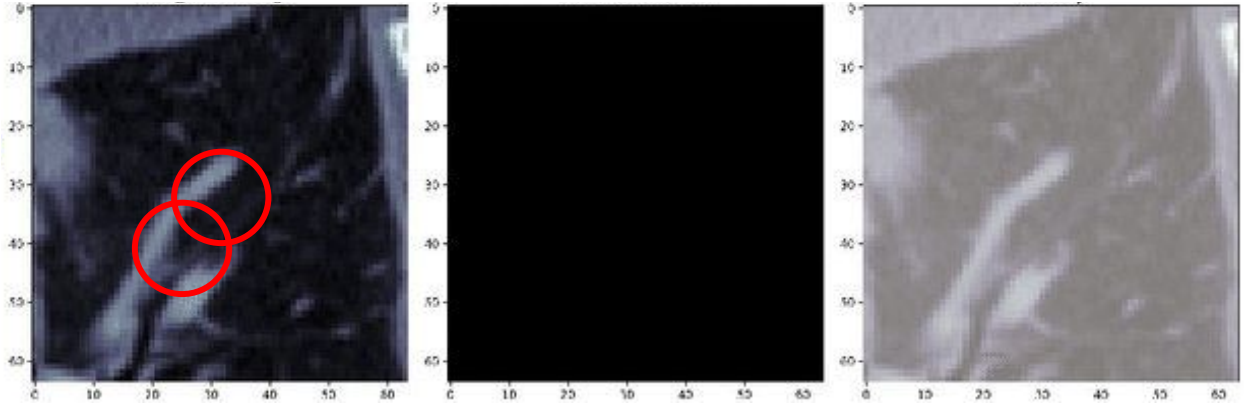


Transverse Plane



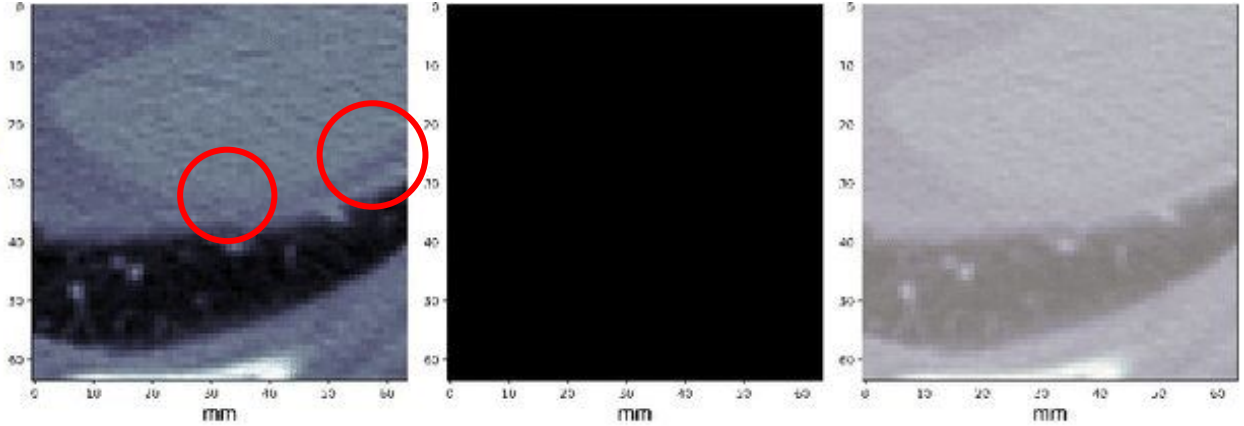
3 True Positives

Sagittal Plane



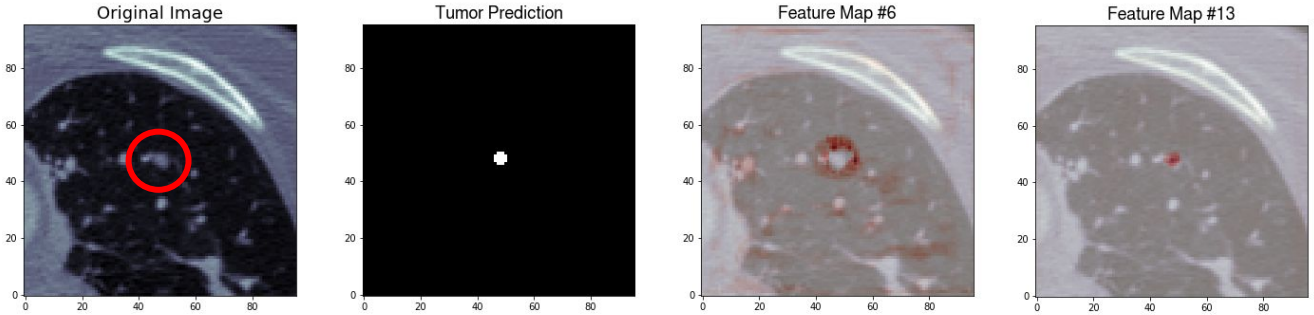
Original Img | Tumor Prediction | Overlay

Transverse Plane

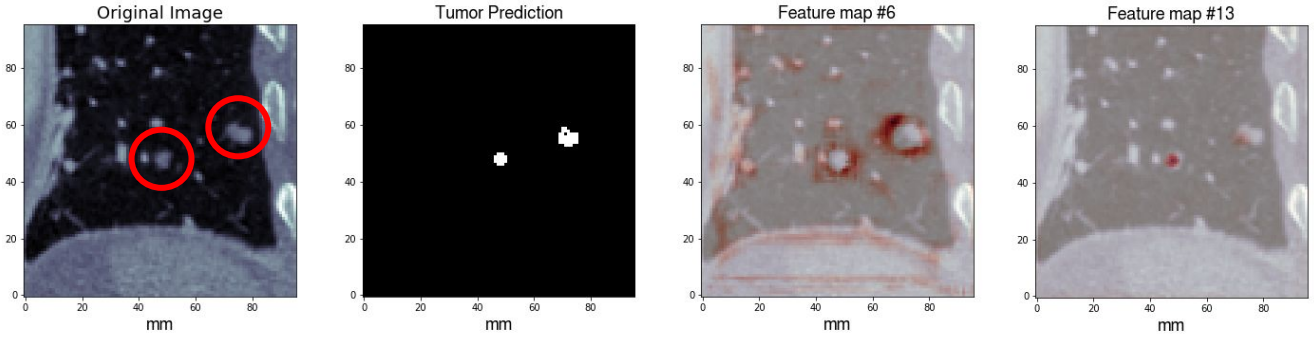


Look for the nodule in the circle

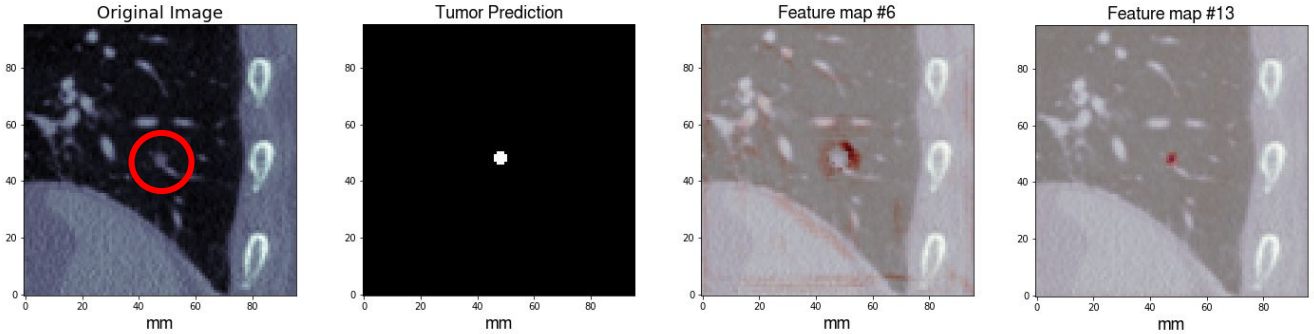
Transverse Plane



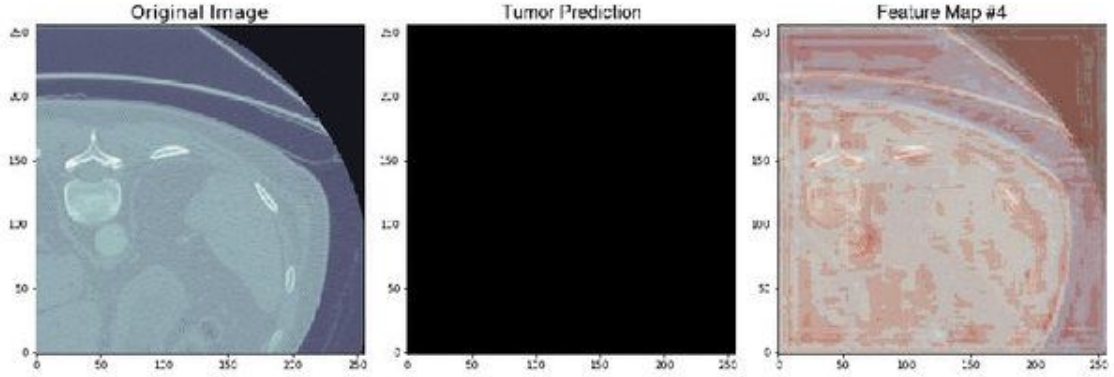
Coronal Plane



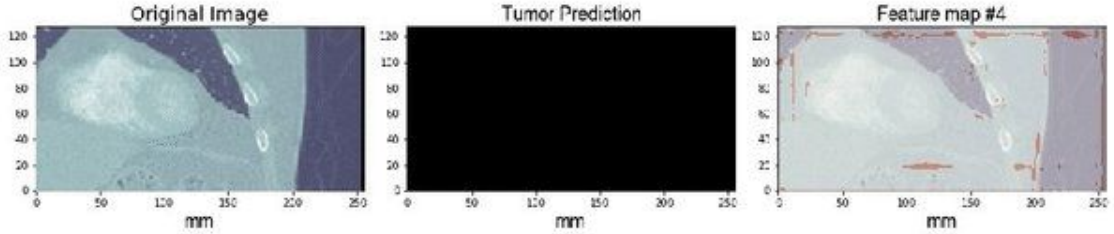
Sagittal Plane



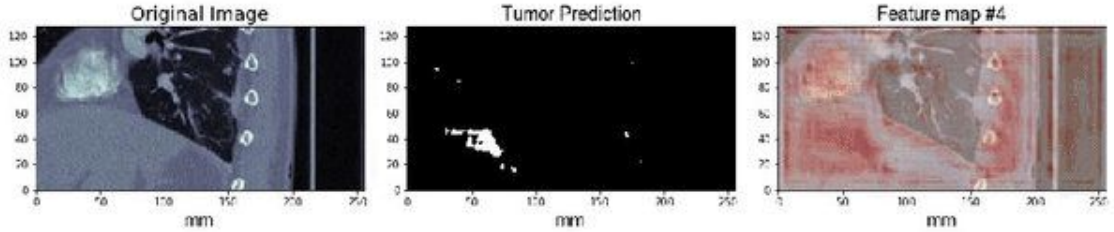
Transverse



Coronal



Sagittal



hand marked nodules
in red for easier spotting

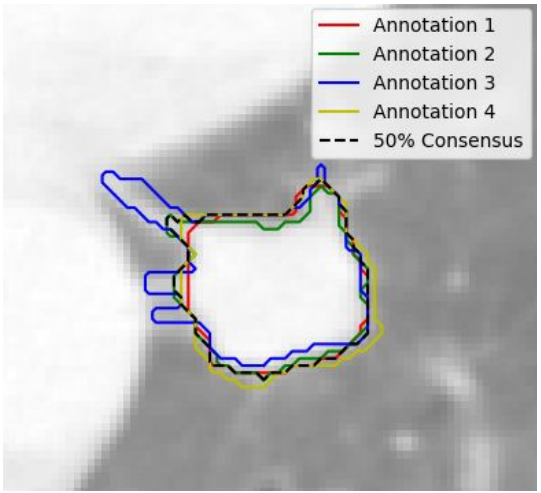
our model missed a nodule

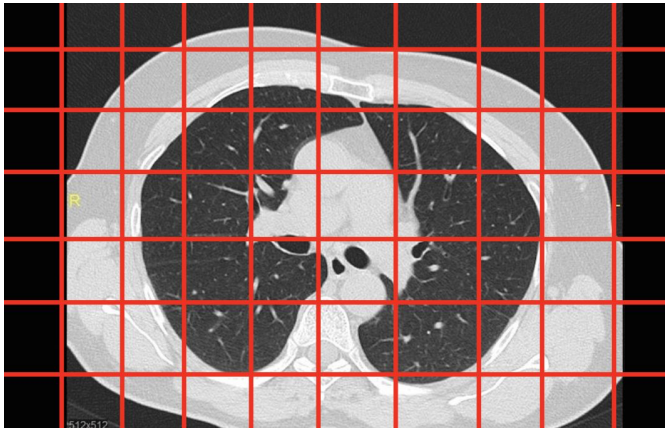
"Everybody has a plan until they get punched in the mouth."

Iron Mike inking by Rob Stinogle

How to Improve this System?

How to Improve this System?

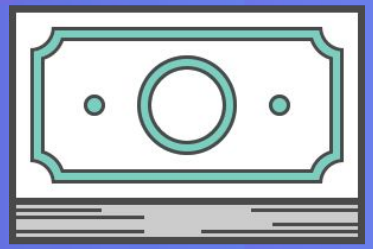




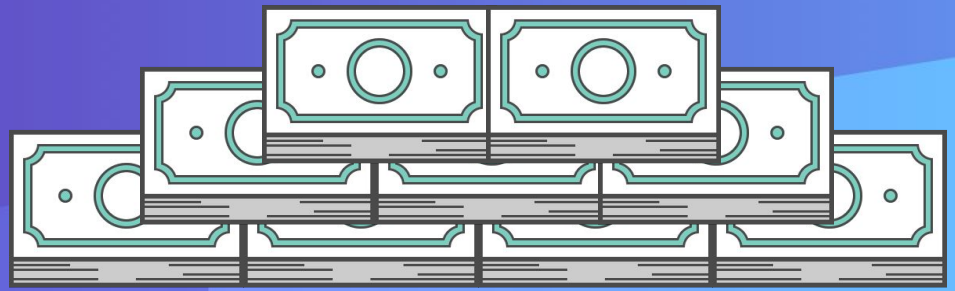
Fedorov A, et al. 2018. Standardized representation of the LIDC annotations using DICOM. PeerJ Preprints 6:e27378v1

Reina, G. A. et al. (2019).ADVERSE EFFECTS OF IMAGE TILING FOR AUTOMATIC DEEP LEARNING. Neuro-oncology

As well as



Lots… of



Lung Nodule Classification and Localization

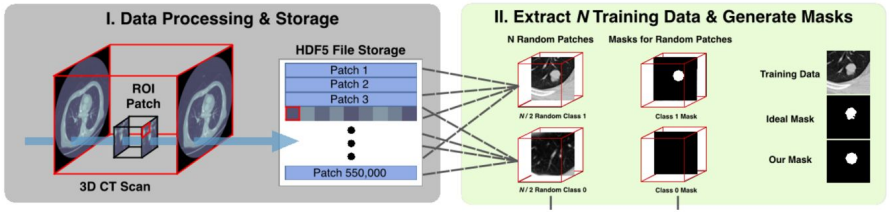
Suman Gunnala, Anil Luthra, Tony Reina, Kyle M Shannon

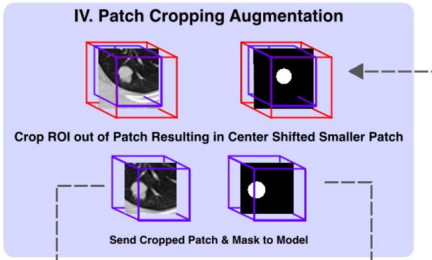
Advisors: Dr. Mehrdad Yazdani, Dr. Bradley Voytek, Dr. Gulaka

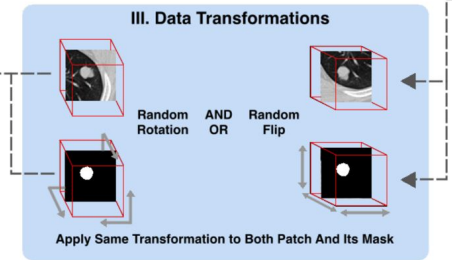
Abstract

Although chest CT screenings are highly effective at reducing lung cancer mortality, today in 2018, radiologists are overburdened with an ever-growing volume of scans to read. We offer a solution of software-assisted lung nodule detection to aid in faster, more reliable diagnosis.

We accomplish this by training a 3D convolutional neural network (3D CNN) that provides high sensitivity and low false positive predictions. Our models are trained using data from a publicly-available dataset called LUNA16. We created a scalable solution and have tested our model on predictions of whole chest CT scans. Our optimal model achieved a dice coefficient of .814.



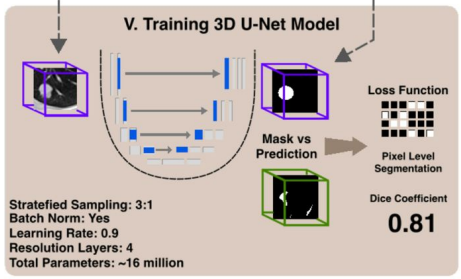




Methods

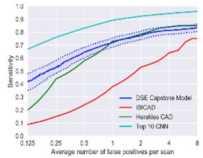
We created a scalable, reproducible processing pipeline that can be rapidly deployed. First, we wrote a patch extractor script that goes through each CT scan and saves regions of interest (ROI), along with associated metadata, into an HDF5 storage object. We faced several challenges in building this patch extractor, including edge clipping issues (where ROIs are too close to an edge); class overlap and centroid collisions (where part of the patch includes both class 0 & 1); and, significant class imbalance (700:1).

Due to the class imbalance, we used mini batch sampling, along with data augmentation techniques (flip, random off-center crops, and rotations) to generate a larger variance of training data. Although expert radiologists marked nodule centers and diameters, they did not segment out the nodules in 2D or 3D space. In order to train a 3D U-Net model for segmentation we generated our own 3D masks using the centroid and diameter measurements to create spherical masks. We trained 3D RESNET models for classification and 3D UNET models for localization tasks. The model architecture and code can be found on out github repository.

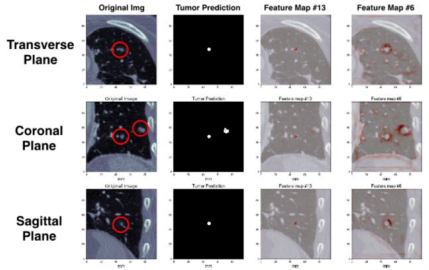


Results

Model scoring and validation are based on a free ROC script created by the LUNA16 challenge curators. We trained and tested our U-Net model with 10-fold cross validation. We observed improved average precision score per holdout data set as shown in the figure (below). The average precision score range was 0.563 to 0.756 for 64x64x64 patch. Our 3D segmentation model (right figure) has been shown to achieve good results with dice coefficient of 0.814.



Our models are predicting nodules of varying diameters, off centers, and multiple nodules within the same slice. Inferences are made on various test image sizes which achieves the true image size agnostic segmentation. The feature maps immediately prior to the final prediction indicate that nodule contour and texture are two features the model is using to make its predictions.



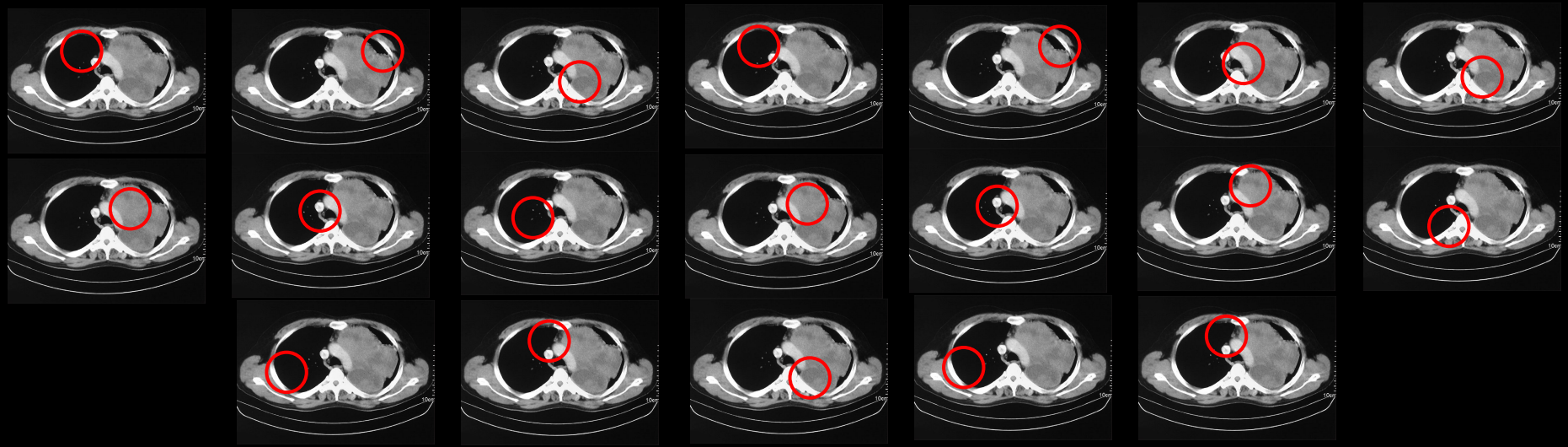
Code.
References.
More Info.
http://bit.do/github-dse

UC San Diego
JACOBS SCHOOL OF ENGINEERING

UCSDCSE
Computer Science and Engineering

UC San Diego
Cognitive Science

During this 45 Minute Talk



19 new cases & **10** deaths from lung cancer

Kyle Shannon

Cup of coffee?

kyle@kmshannon.com

linkedin.com/in/kmshannon/

Contributions & Interests

Improving access to healthcare in rural settings, (facilitated > 200k medical trips over 5 million miles last year in NY/CA)

Partnering with hospitals to reduce readmission rates

Building data & ML tools for clinicians

Starting center for the UC-Health weekend pick-up game basketball “league”

Backup Slides

References

[1] Aberle, D.R., Adams, A.M., Berg, C.D., Black, W.C., Clapp, J.D., Fagerstrom, R.M., Gareen, I.F., Gatsonis, C., Marcus, P.M., Sicks, J.D., 2011. Reduced lung-cancer mortality with low-dose computed tomographic screening. New England Journal of Medicine 365, 395–409.

[2] https://www.cancer.gov/types/lung/research/nlst-qa

[3] Depeursinge A, Foncubierta-Rodriguez A, Van De Ville D, Müller H. Three-dimensional solid texture analysis in biomedical imaging: review and opportunities. Med Image Anal 2014;18:176–196. doi: 10.1016/j.media.2013.10.005.

[4] Bayanati H, Thornhill RE, Souza CA, et al. Quantitative CT texture and shape analysis: can it differentiate benign and malignant mediastinal lymph nodes in patients with primary lung cancer? Eur Radiol 2015;25(2):480–487.

[5] S. C. Wong, A. Gatt, V. Stamatescu, and M. D. McDonnell, "Understanding data augmentation for classification: when to warp?" CoRR, vol. abs/1609.08764, 2016. [Online]. Available: http://arxiv.org/abs/1609.08764

[6] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: CVPR (2016)

[7]https://www.mathworks.com/content/mathworks/www/en/discovery/convolutional-neural-network/jcr:content/mainParsys/image_copy.adapt.full.high.jpg/1492406018870.jpg

[8] Long, J., Shelhamer, E., and Darrell, T. Fully convolutional networks for semantic segmentation. CoRR, abs/1411.4038, 2014.

[9] Springenberg, Jost Tobias, Dosovitskiy, Alexey, Brox, Thomas, and Riedmiller, Martin. Striving for simplicity: The all convolutional net.arXiv preprint arXiv:1412.6806 , 2014.

[10] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. InMIC-CAI, pages 234–241. Springer, 2015.

[11] Cicek, O., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. arXiv preprint arXiv:1606.06650 (2016)

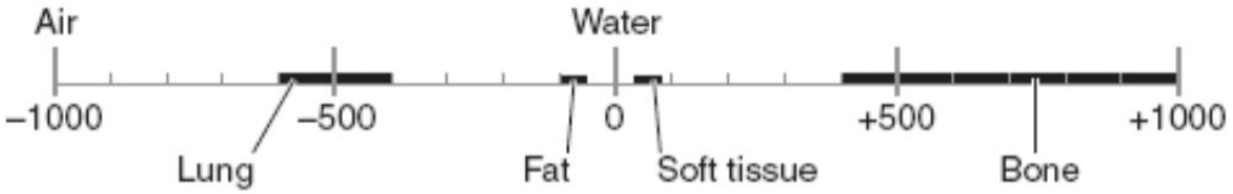
[12] H. R. Rotha, H. Odaa, X. Zhoub, N. Shimizua, Y. Yanga, Y. Hayashia, M. Odaa, M. Fujiwarac, K. Misawad, and K. Moria, "An application of cascaded 3D fully convolutional networks for medical image segmentation,"ArXiv e-prints, 2018.

Model Tuning Table

	Batch Norm	Resolution Layers	Learning rate	Weighted Loss Function	HU Norm.	Batch size	Training Dice Coefficient	Classification Accuracy
Model 1	No	3	0.001	No	No	1 to 1	0.64	N/A
Model 2	Yes	4	0.001	Yes	No	1 to 1	0.38	0.5
Model 3	Yes	4	0.001	Yes	Yes	1 to 1	0.43	0.5
Model 4	Yes	4	0.003	No	No	1 to 1	0.563	N/A
Model 5	Yes	4	0.006	No	No	1 to 1	0.781	N/A
Model 6	Yes	4	0.009	No	No	1 to 1	0.8155	N/A
Model 7	Yes	4	0.02	No	No	1 to 1	0.7711	N/A
Model 8	Yes	3	0.009	No	No	1 to 1	0.6274	N/A
Model 9	Yes	4	0.009	Yes	No	1 to 1	0.556	0.5
Model 10	Yes	4	0.02	Yes	No	1 to 1	0.554	0.5
Model 11	Yes	4	0.009	Yes	Yes	1 to 1	0.4247	0.5
Model 12	Yes	4	0.02	Yes	Yes	1 to 1	0.5752	0.5
Model 13	Yes	4	0.009	No	No	3 to 1	0.8029	N/A
Model 14	Yes	4	0.02	No	No	3 to 1	0.48	N/A
Model 15	Yes	4	0.009	No	Yes	1 to 1	0.6813	N/A
Model 16	Yes	4	0.02	No	Yes	1 to 1	0.7932	N/A

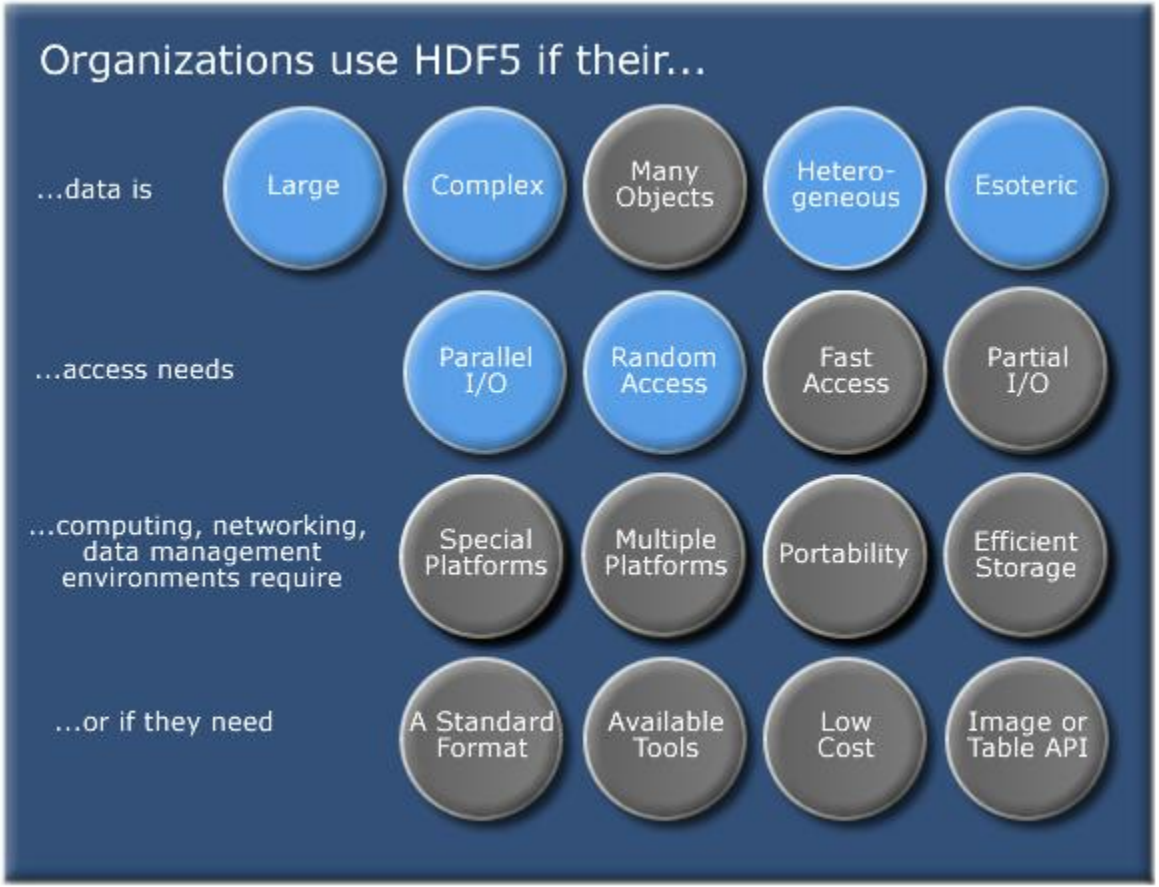
Table 1: Summary of results from different models on holdout subset 5 dataset

Hounsfield Units For Biological Material



Bone	+400 → +1000
Soft tissue	+40 → +80
Water	0
Fat	–60 → –100
Lung	–400 → –600
Air	–1000

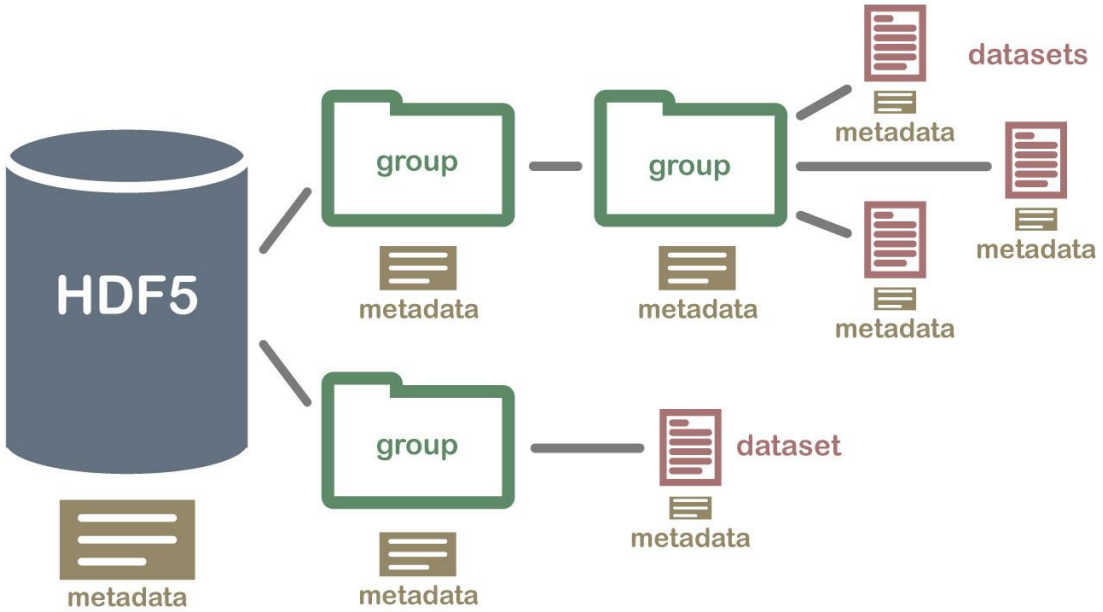
HDF5 - Why And When To Use



HDF5 - It’s FOSS

Group: A folder like element within an HDF5 file that might contain other groups OR datasets within it.

Dataset: The actual data contained within the HDF5 file. Datasets are often (but don't have to be) stored within groups in the file.



UNET Model Approaches

Model A

Generic U-Net Model

Model B

With Batch Normalization and Localization + Classification

Model C

With Batch Normalization, Hounsfield Normalization and Localization + Classification

Dice Metric - Think F1 Score

$$Dice\ Coefficient = \frac{2 * TP}{FN + (2 * TP) + FP}$$

$$Jaccard\ Index = \frac{TP}{TP + FN + FP}$$

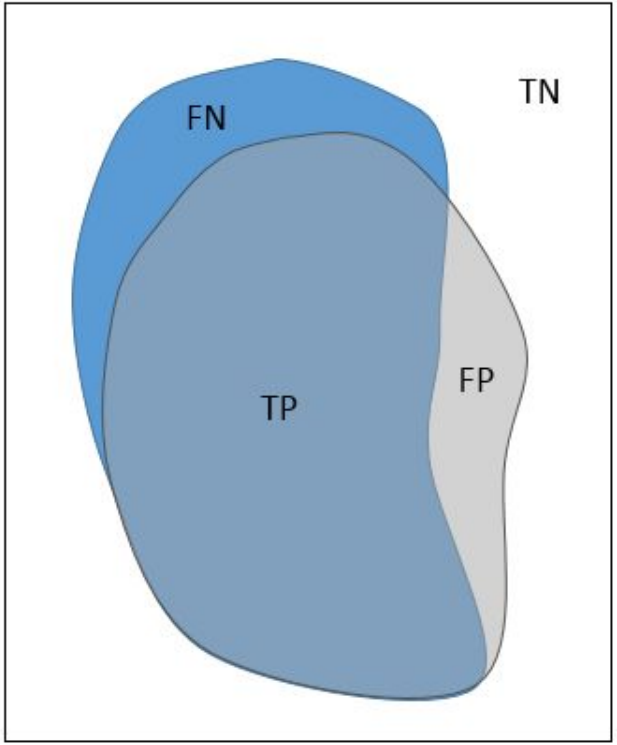
$$Sensitivity = \frac{TP}{TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

TP - true positive
TN - true negative
FP - false positive
FN - false negative

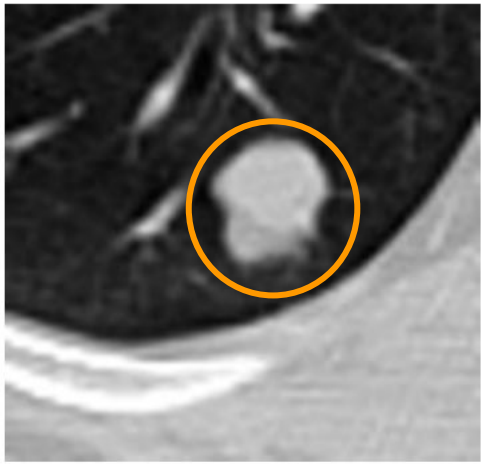
Manual Segmentation

Automated Segmentation



Class 1 vs Class 0

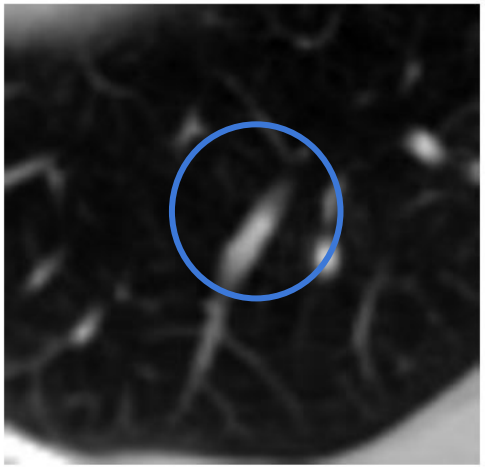
1,186



Class 1

Marked and Scored
by 4 Radiologists

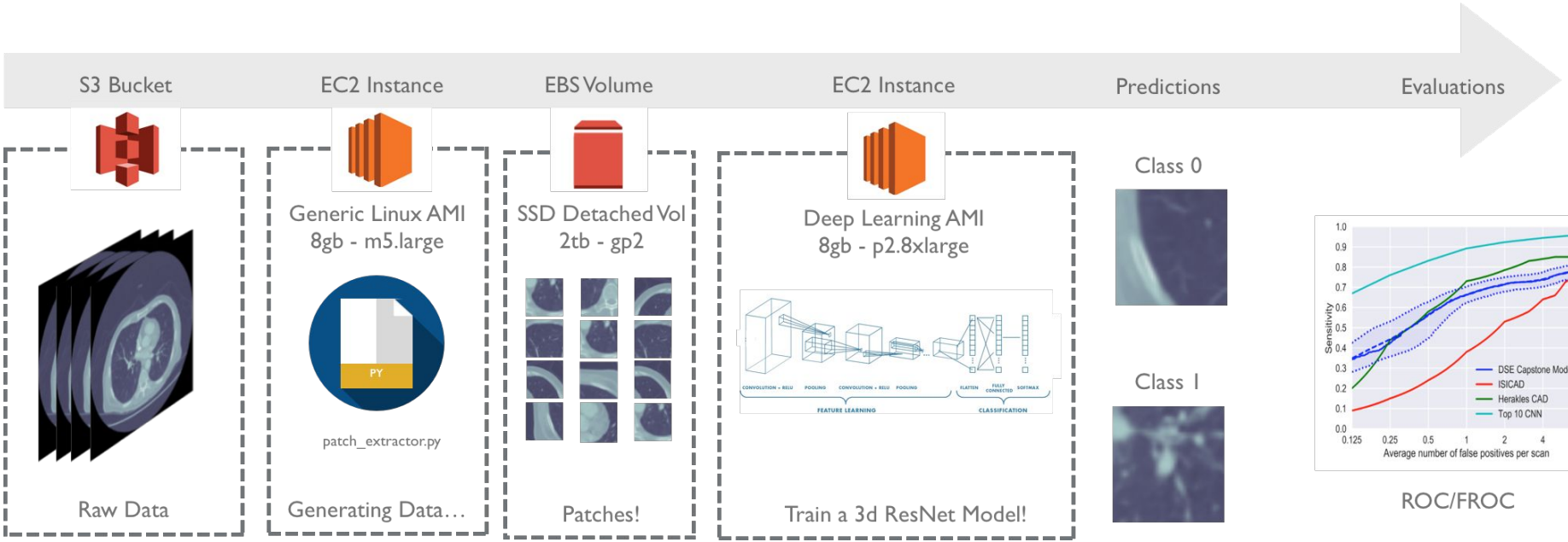
551,065



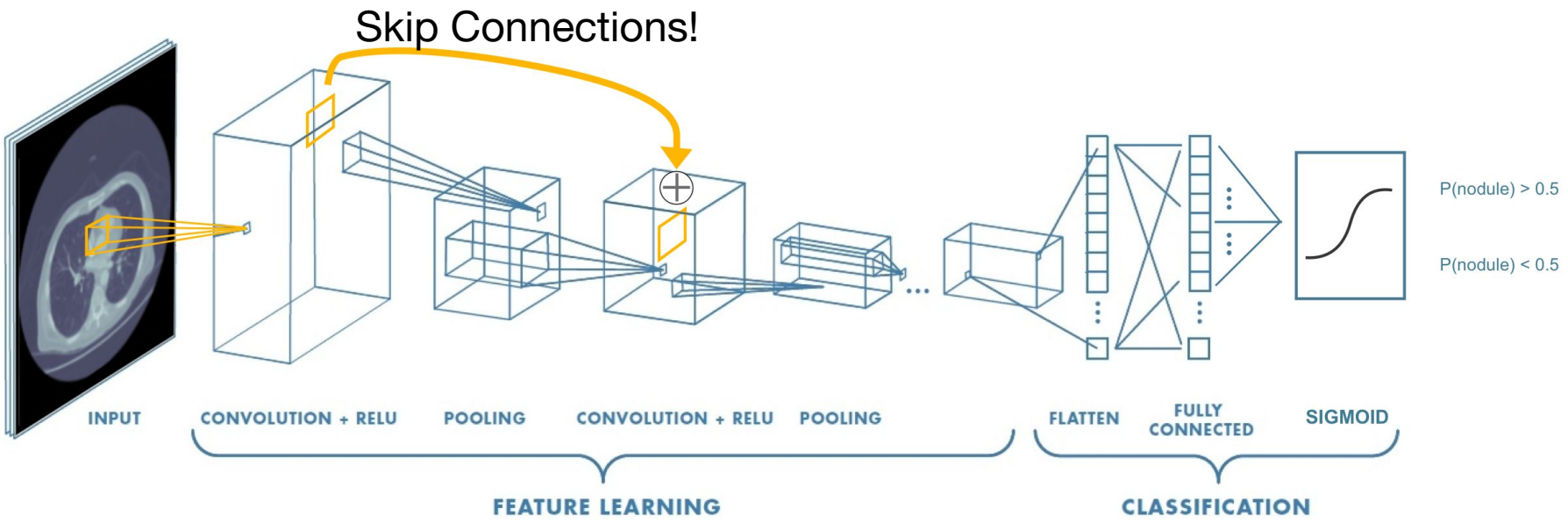
Class 0

Marked by CAD
Algorithm

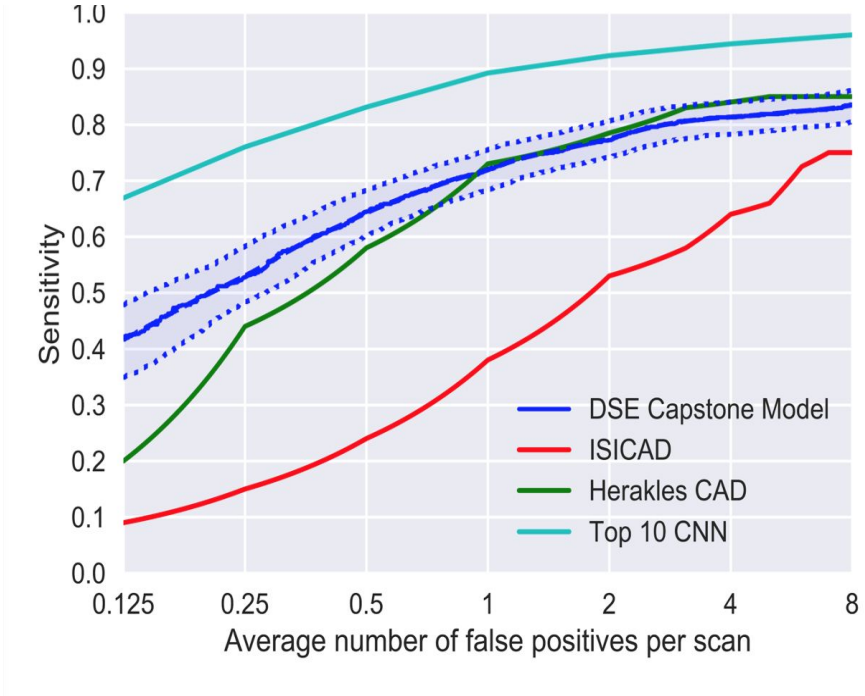
End-to-End Classification Pipeline - RESNET

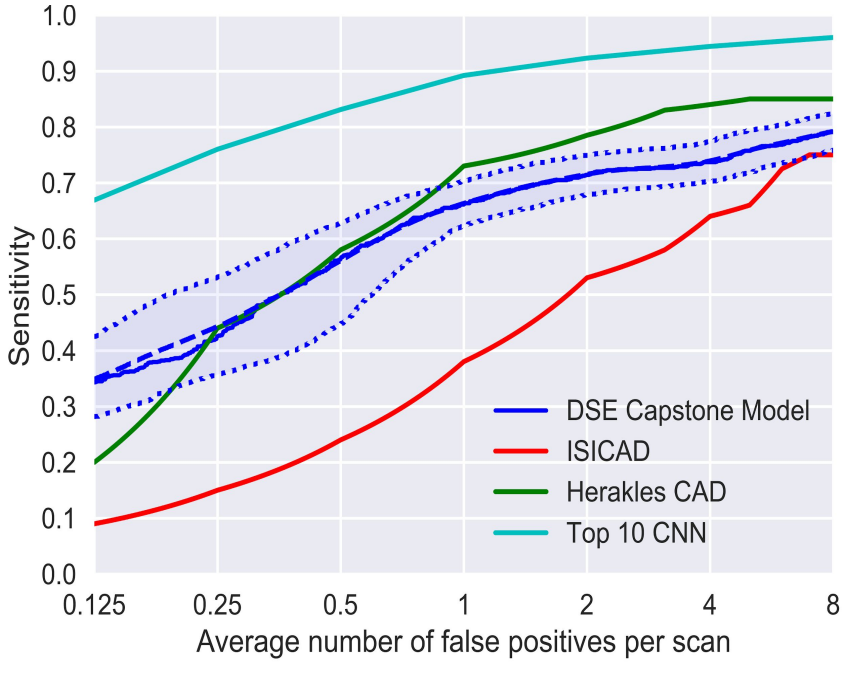


RESNET 3D

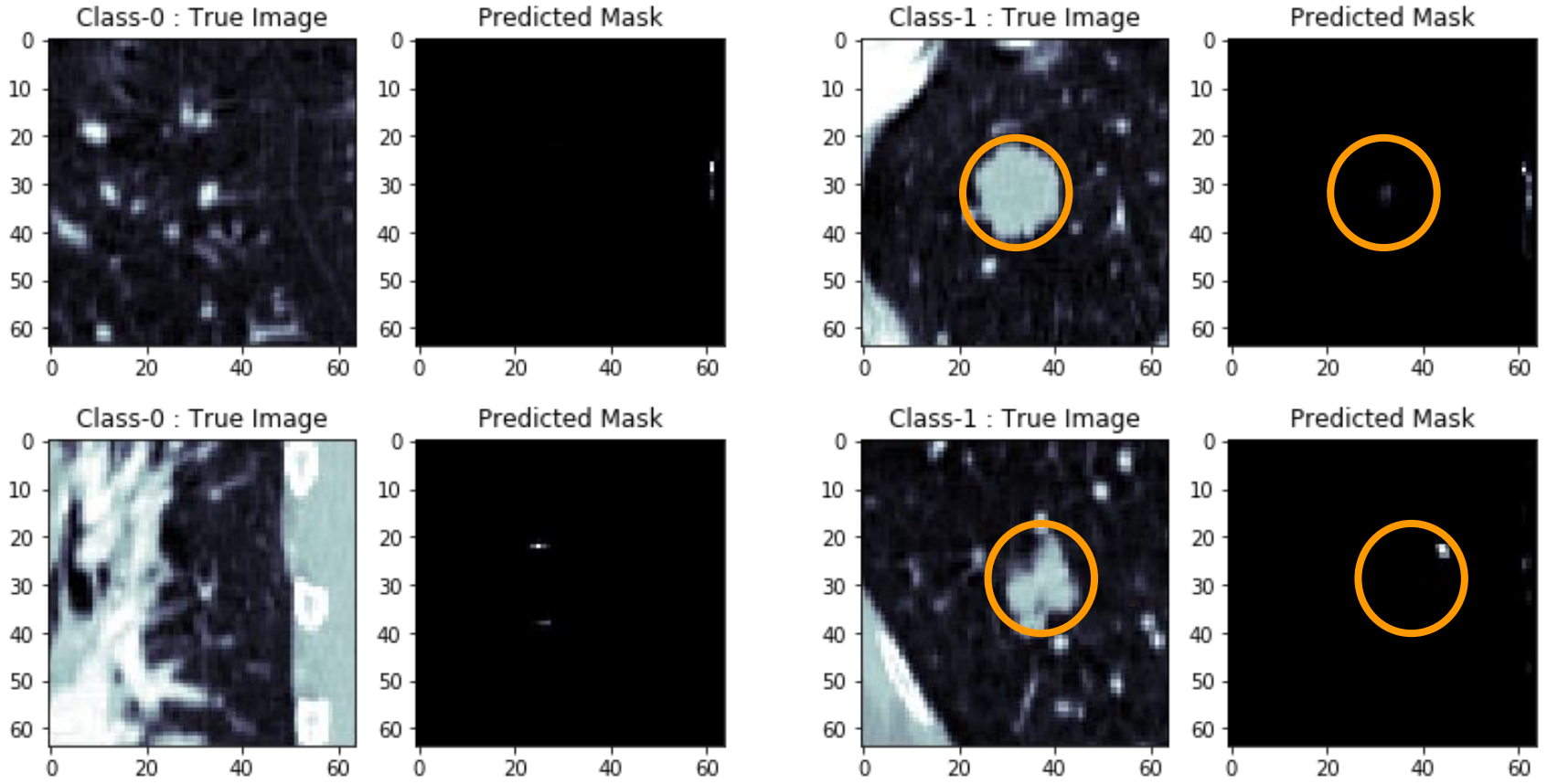

Skip Connections!
INPUT
CONVOLUTION + RELU
POOLING
CONVOLUTION + RELU
POOLING
FLATTEN
FULLY CONNECTED
SIGMOID
P(nodule) > 0.5
P(nodule) < 0.5
FEATURE LEARNING
CLASSIFICATION

FROC Assessment Results Comparison 64 Slices (1 vs 4 Epochs)

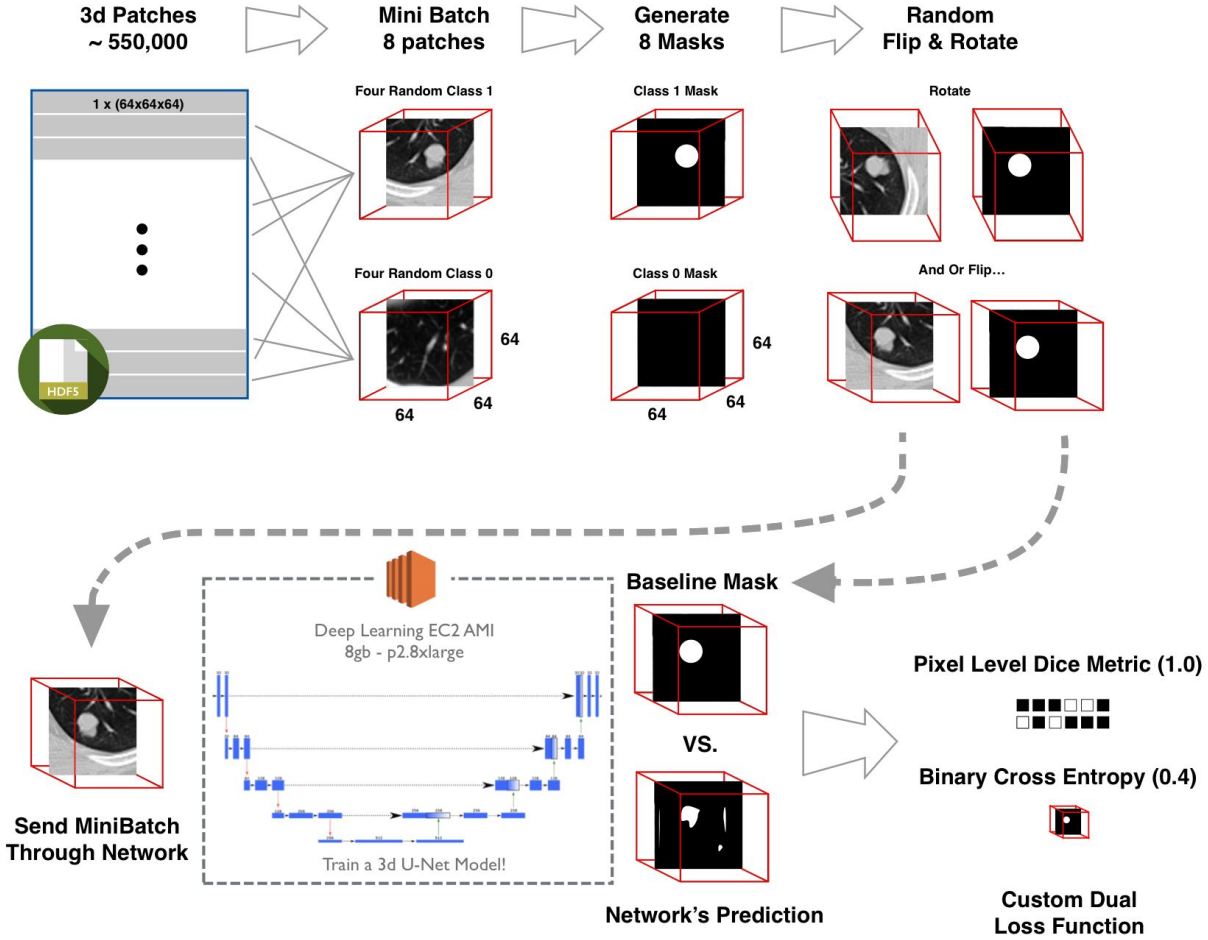




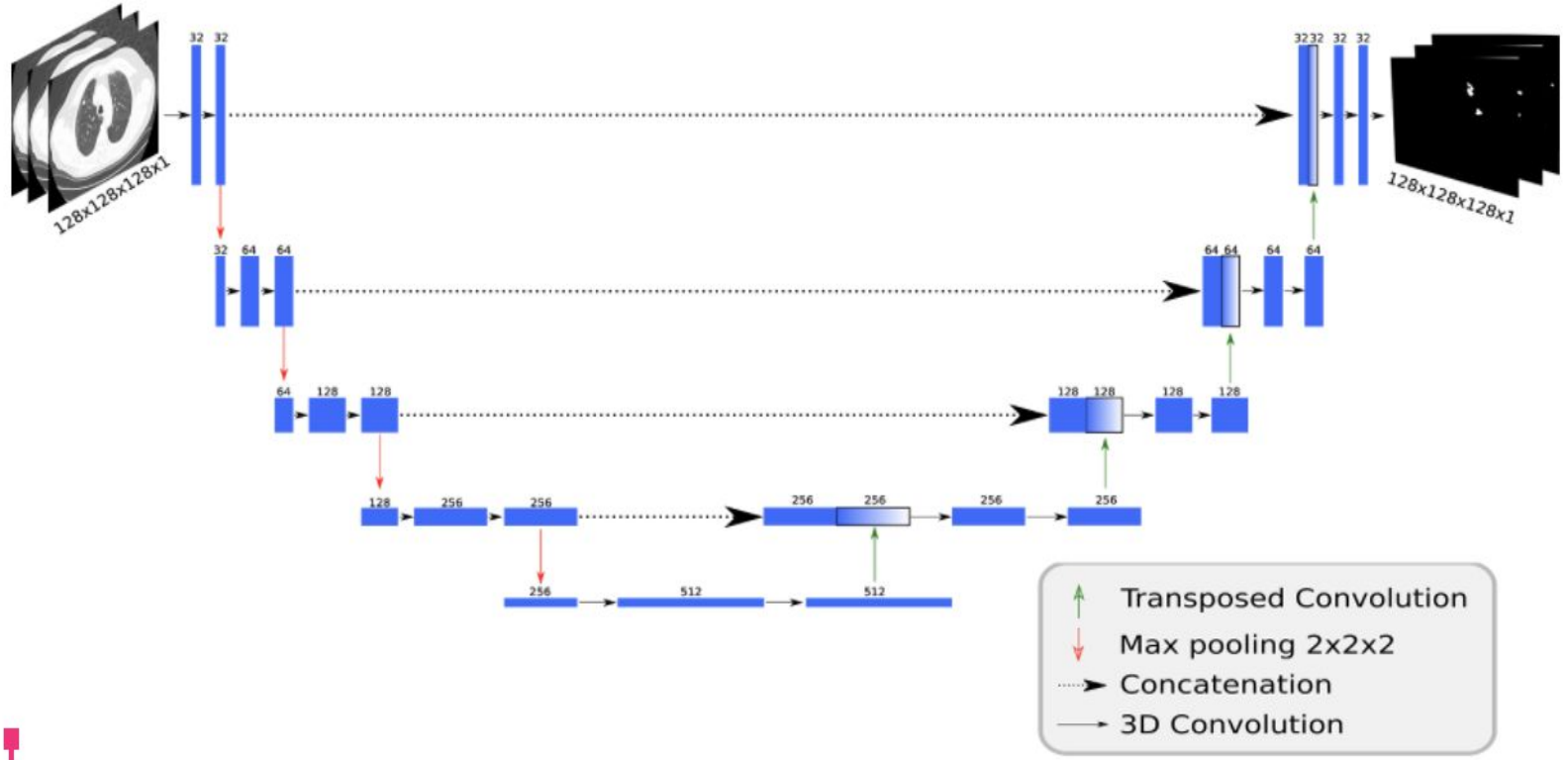
Initial Trial Run UNet Results



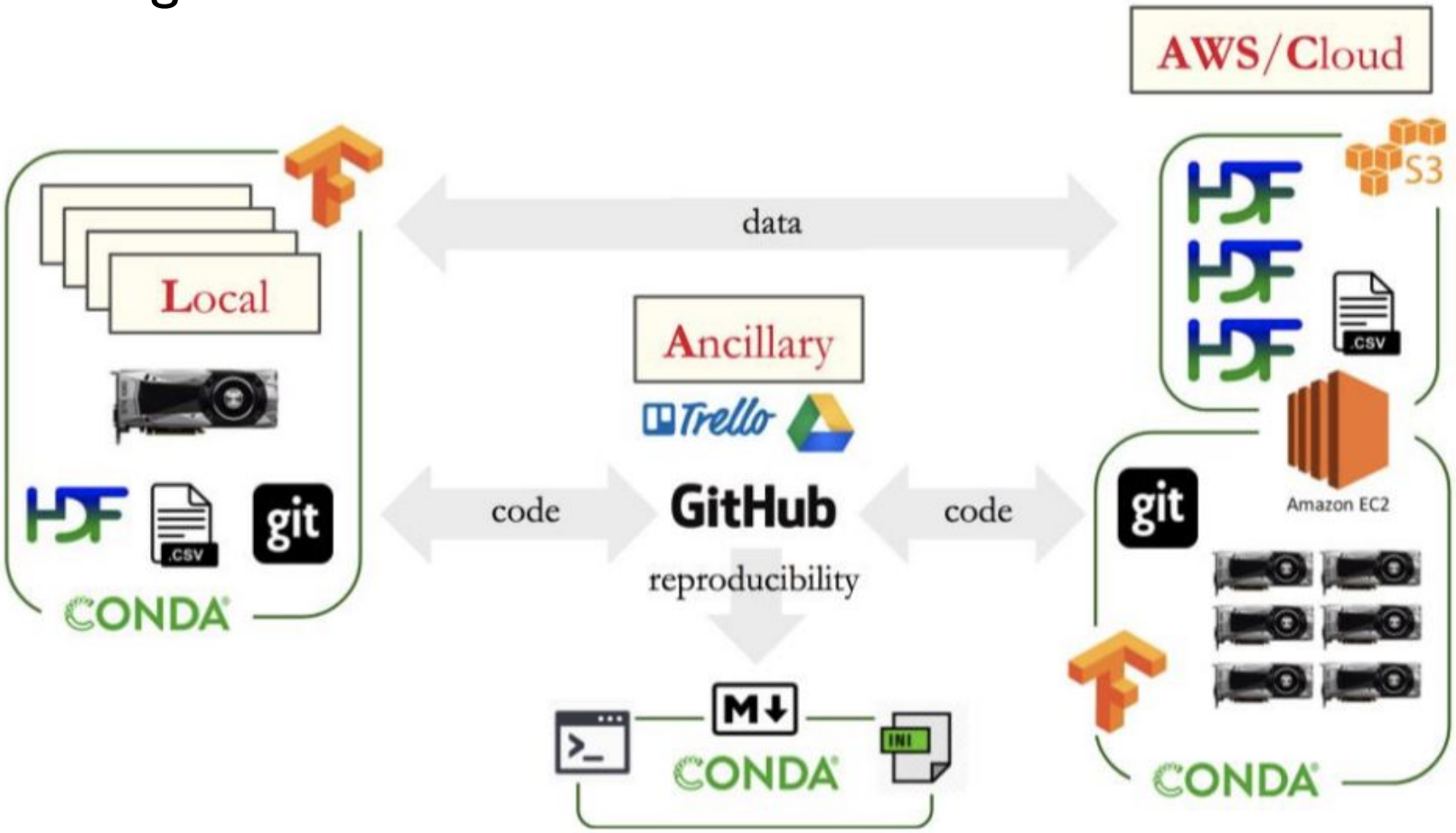
Track 2 Localization (Segmentation) + Classification Pipeline



UNet Model Overview



Project Organization



Patch Extractor FlowChart