Exploring 3D Convolutions for Lung Cancer Detection

Shubhang Desai

shubhang@stanford.edu

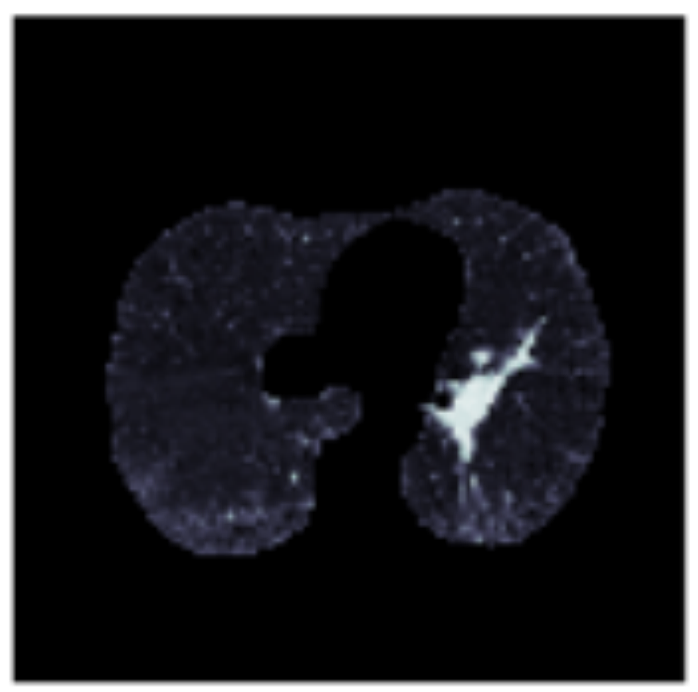
Motivation

It's often difficult for radiologists to find nodules in CT scans of lungs, and then distinguish cancerous nodules from non-cancerous nodules.

We wanted to build a system which could take a CT scan as input and predict if it contains cancer

Data

LUNA16 Dataset [2]**:** lung CT scans and locations of nodules in scans.



Kaggle CT Data [1]: lung CT scans and binary labels of presence of cancer.

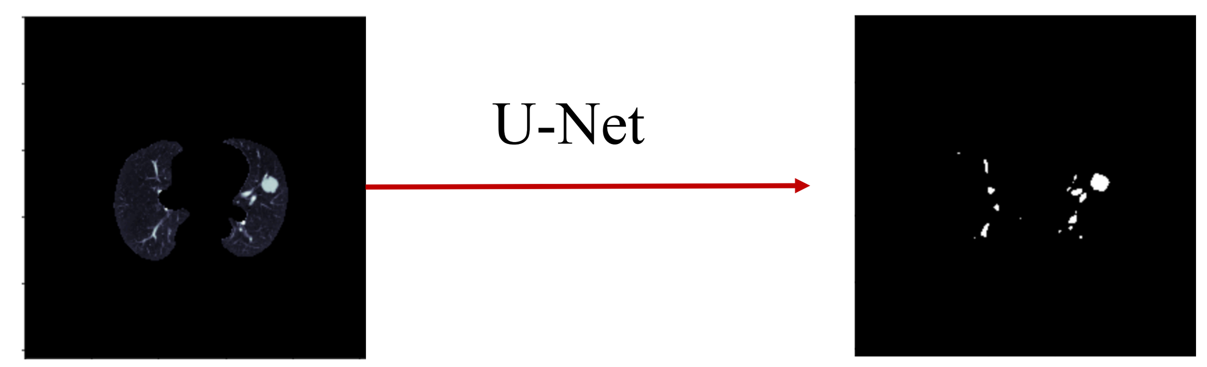
Approaches

Find nodule candidates by training segmentation on LUNA16 set, and use candidates to classify cancer. [1]

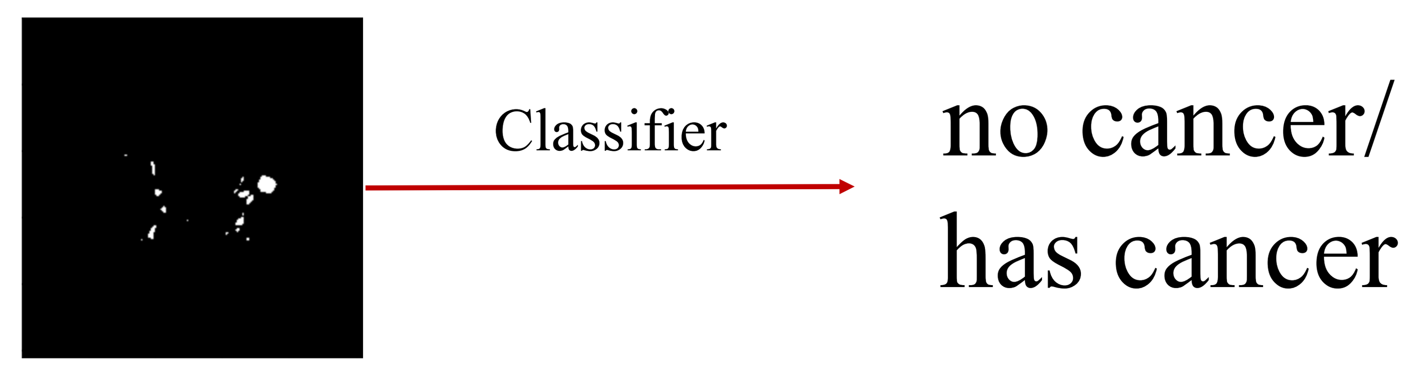
One classifier that we use is a 3D-ConvNet, which uses convolutions in x-, y-, and z-axis to transform input.

Model

Step 1: Use U-Net (popular biomedical segmentation network) to turn scan into predictions for where there may be nodules:



Step 2: Use variety of models (Naïve Bayes, MLP, ResNet) to transform top candidates into prediction of cancer:



Results

We trained our algorithm pipeline on a subset of the full training set, and tested on a subset of the dev and test sets.

Below are the results for Naïve Bayes, MLP, and 3D-ResNet:

Model	Dev Acc	Test Acc
Naïve Bayes	68.97%	50.00%
MLP	75.86%	50.00%
3D-ResNet	79.31%	62.50%

Analysis

Our deep architecture does much better than naïve methods on both dev and test sets, and may even pave way for SOTA results.

All of our methods, however, do much better on dev than on test—perhaps some hyperparameter tuning is necessary.

As a proof-of-concept, we are looking quite good!

Future Work

Train models on **full** datasets to get much better predictions and more meaningful accuracies.

Experiment with more architectures (such as DenseNet or others).

Tune hyperparameters (positive class weight, learning rate, etc.).

References

[1] Albert Chon, Niranjan Balachandar, and Peter Lu. Deep convolutional neural networks for lung cancer detection. 2017.
[2] Booz Allen Hamilton and Kaggle. Data sci- ence bowl 2017, 2017. data retrieved from LUNA16 site, https://www.kaggle.com/c/ data-science-bowl-2017/data.
[3] Colin Jacobs, Arnaud Arindra, Adiyoso Setio, Alberto Traverso, and Bram van Ginneken. Lung nodule analysis 2016, 2016.