Recurrent Models for Detecting Cancer in 3D CT Volumes

Luran He, Alexander Lin & Amanda Zhang

Harvard AC209b/CS109b

Objectives

- Create a model for lung nodule segmentation from CT scans.
- Create an RNN model for improving performance given a sequence of scans.

Dataset

Lung Image Data Consortium: LIDC

- Largest public database for lung CT scans.
- 1018 CT scans from 1010 patients.
- DICOM (512x512) files representing horizontal (axial) cross-sections of the chest cavity.
- XML document with metadata (annotations by radiologists).

Data Processing

- Downsized image dims from 512x512 to 256x256.
- To deal with class imbalance, filtered for and trained only on slices that contained nodules.
- Subset 'windows' of image that contain nodule.

Examples

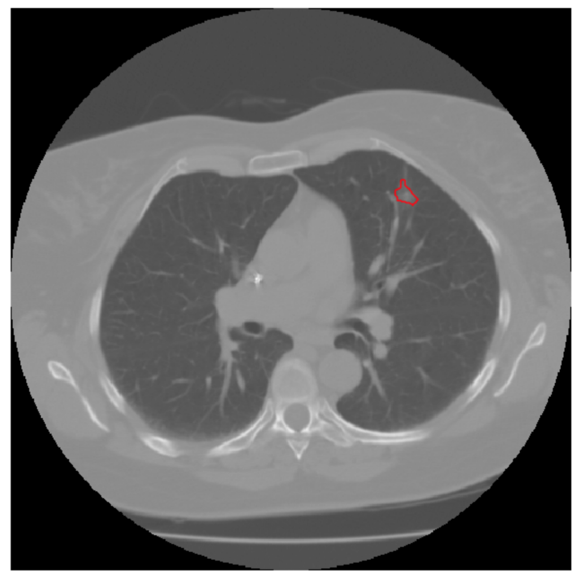


Figure 1: Bounding region of example nodule.

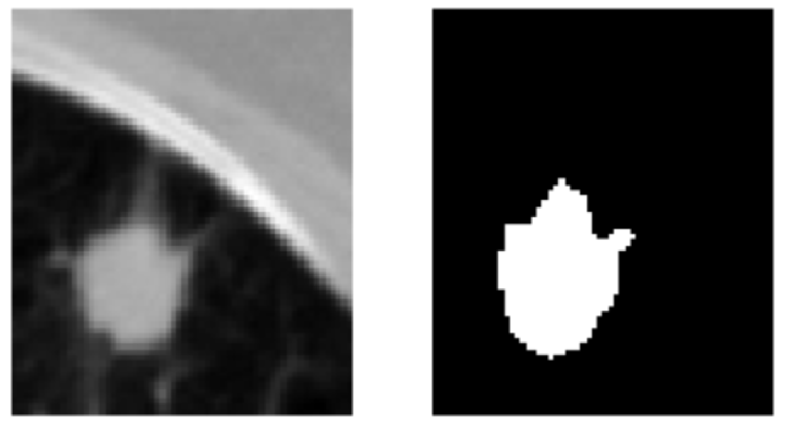
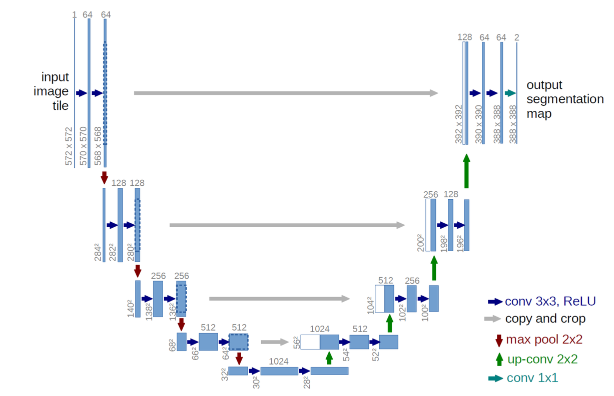


Figure 2: Crop of nodule and ground truth segmentation.

U-net [1]



- High performing architecture originally designed for cell segmentation.
- Layers that upsample and downsample symmetrically.
- Combines information from various levels of convolution to provide a localization for the output.
- Requires few training images to converge.

Model Features

Feature Map Outputs from U-net

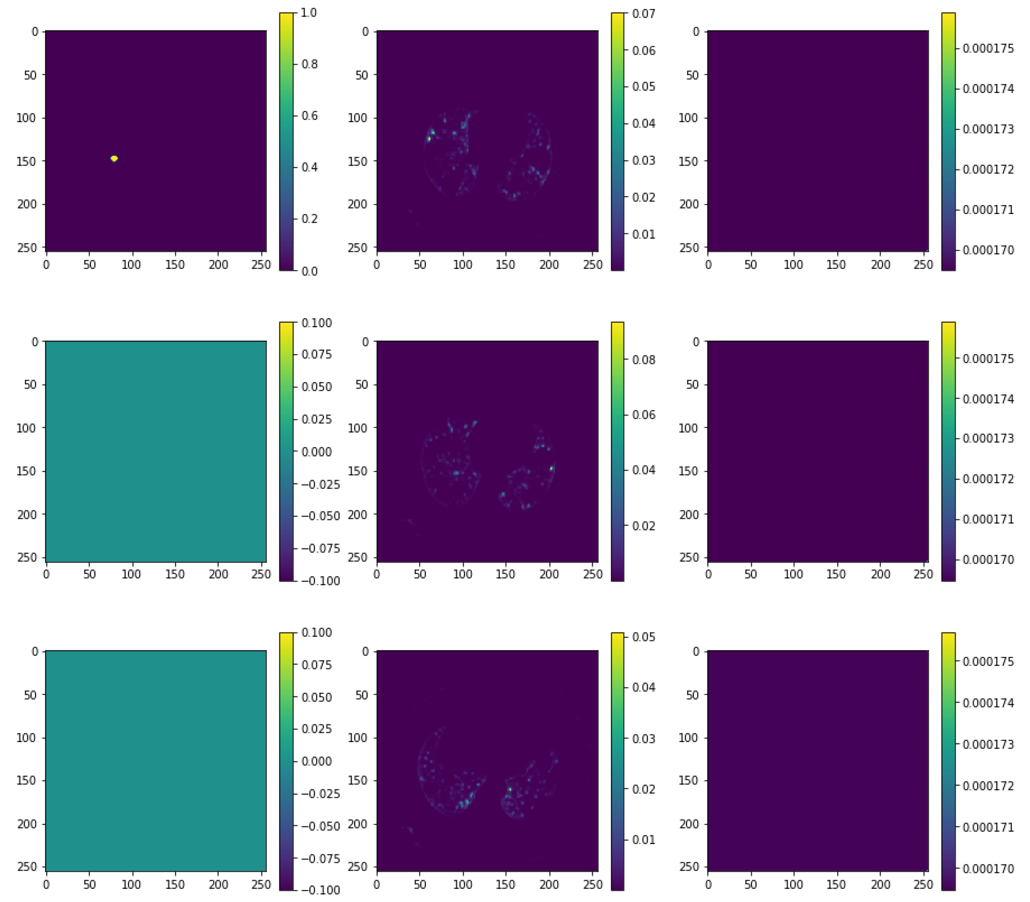
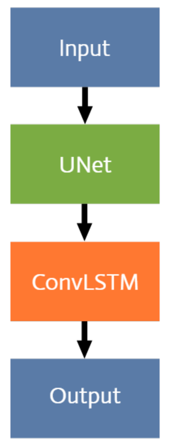


Figure 3: Sample Feature Map

The first column represents the ground truth; the second column is the feature output of the U-Net, and the last column is the output of the LSTM. The U-Net appears to be performing well segmenting the edges of the body in the CT scan while the LSTM seems to flatten those features to near-uniform value.

U-net + RNN



- Fed feature maps outputted by the U-Net model into a Convolution LSTM layer (input-state, and state-state transitions contain convolutional structures) [2].
- Predictions in each layer influenced by the predictions from past and future layers.

Model Evaluation

Evaluation metrics

For a set of random patients, we have ROC curves for our UNet model and two different UNet + RNN models. Most of the time, the RNNs fails to contribute positively to the U-Net model.

ROC curves for random sample of patients

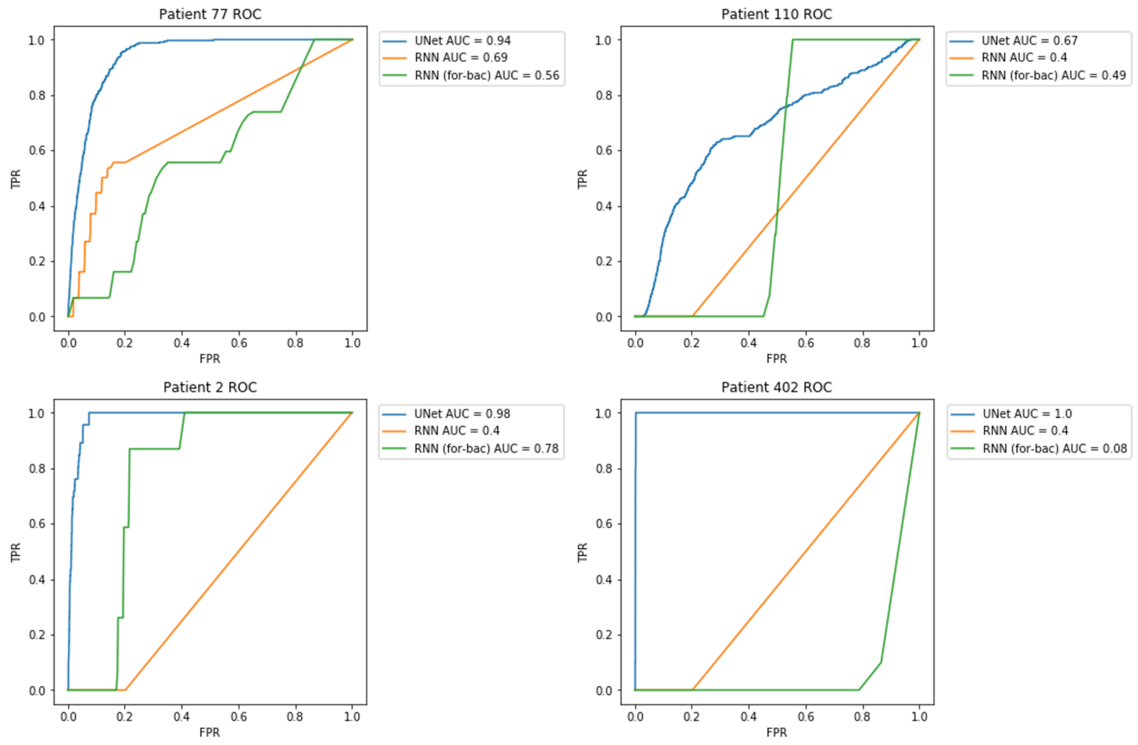
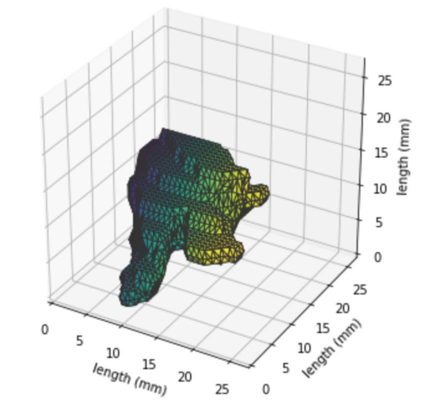


Figure 4: ROC Curves for 4 patients



Conclusion

Unfortunately, our models seem to suggest that LSTM doesn't improve upon the performance otherwise given by the U-Net. This wasn't entirely expected since we hoped the sequential nature of the problem would fit and RNN model quite nicely. However looking at the outputs of both the U-Net and the LSTM model, it seems like the LSTM is learning to output uniform values. This may be due to the uniformity of the ground truth itself since the majority of pixels will be non-tumorous in most scans.

Future Directions

Further ideas we would have liked to pursue include use training the U-net and LSTM layers in one model instead of taking the output of U-net as input to the LSTM. Other ideas include improving training by data augmentation, potentially using Generative Adversarial Networks.

References

[1] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[2] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *CoRR*, abs/1506.04214, 2015.

Acknowledgements

Special thanks to our TFs Eric, Eleni for helping us with infrastructure, and to the CS109b team!