Detection of Early Stage Lung Cancer from CT-Generated 3D Lung Volumes

Ankit Baghel and Lucas Ramos, {abaghel, ramosl}@stanford.edu

Goals

- Goal is to convert computerized tomography (CT) scans into 3-dimensional volumes that deep learning architectures could use to detect cancer.
- Secondary goal is to see if relevant features for cancer detection are maintained through downsizing of stored data.

Predicting

- For a baseline, we used a 2D ResNet and 2D DenseNet model using pre-trained weights on the ImageNet database.
- For a more advanced model, we implemented a Keras 3D UNET model that takes 3D volumes as input for 3D convolutions, downsamples them, and then upsamples them back to their initial size.
- Since UNET models are better optimized for segmentation tasks rather than binary classification tasks, we removed the upsampling steps for additional convolutional layers to deepen the model and learn additional features.

Data

- Our data comes from the Kaggle Data Science Bowl 2017 which contains lung CT scans of 2100 patients.
- Each patient was labeled either 0 for no cancer diagnosis within a year or 1 for cancer diagnosis within a year.
- Each patient folder contains a variable number of CT scan slices. Each slice is a 512 x 512 image provided in the DICOM format.
- Images can be represented as float arrays and contain one channel.

Features

- The CT scan slices for each patient were preprocessed by stacking them, converting their pixel values to Hounsfield Units (HU), thresholded to ignore non-lung tissues, and normalized and zero-centered.
- The resulting 3D volumes were then converted to NumPy arrays which were downsampled to be of size (64,64,64).
- These arrays are the raw input for the models used.

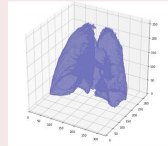

Figure 1: 64x64x500 lung volume after preprocessing

Models

- We first used a 2D ResNet and 2D DenseNet on the input by first squashing it along the height axis from 64 channels to 3 to conform to a pre-trained model.
- Fine-tuned existing weights on our set of data.
- ResNet was 152 layers and used ImageNet pre-trained weights (model not shown).
- DenseNet was 161 layers and also used ImageNet pre-trained weights (model not shown).

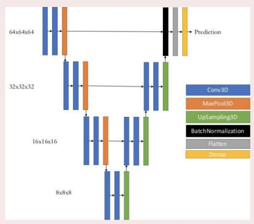

Figure 2: 3D UNet Model Implemented

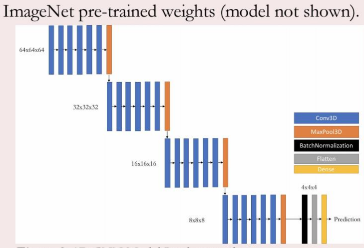

Figure 3: 3D CNN Model Implemented

Results

	Training Loss	Training Accuracy	Validation Loss	Validation Accuracy
ResNet-152	0.5926	0.7295	0.6239	0.7000
DenseNet-161	0.6007	0.7240	0.6206	0.7000
3D-UNet	23.636	0.7373	34.1301	0.6975
3D-CNN	1.4112	0.7373	0.1390	0.6975

Table 1: Model Results

Discussion

- All models performed roughly equivalently on the dataset (~70% accuracy)
- Later analysis of the 3D-CNN revealed that specificity was 1.0 and sensitivity was 0, which means that the classifier was not learning useful features and was instead taking advantage of the class bias of the data set.
- It is likely that downsizing all dimensions to 64 resulted in loss of detectable cancer features.
- To handle this issue, we re-ran the 3D-CNN on the dataset before it was downsized completely to see if any features could be learned.
- However, this was not sufficient to prevent overfitting to a single class (result not displayed).

Future Work

- As we had to squeeze our input data for the training to be more manageable, we could try training with the full 3D volumes so that no features are lost.
- Instead of a 2D ResNet/DenseNet, a 3D ResNet/DenseNet could be used instead as it maybe more accurate because the model would be more specific to our data.

References

1. "Lung Cancer." *American Cancer Society*, www.cancer.org/cancer/lung-cancer.html.
2. Kaggle, "Data Science Bowl 2017." https://www.kaggle.com/c/data-science-bowl-2017.
3. "Full Preprocessing Tutorial | Kaggle." *Countries of the World | Kaggle*, www.kaggle.com/gzuidhof/full-preprocessing-tutorial.
4. Ronneberger, Olaf. "Invited Talk: U-Net Convolutional Networks for Biomedical Image Segmentation." *Informatik Aktuell Bildverarbeitung Für Die Medizin 2017*, 2017, pp. 3–3., doi:10.1007/978-3-662-54345-0_3.