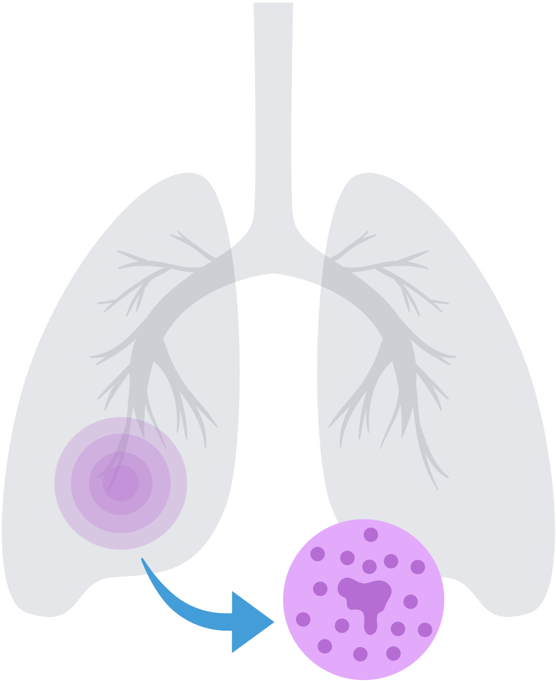


Classifying Lung Cancer

Michael Capparelli

Before We Start

I will be discussing lung cancer, some statistics of lung cancer, and I'll be showing you images of what cancer looks like in the lungs.

All images are CT scans.

Please feel free to tune out the presentation if this topic is personal to you.

Background

An overview of lung cancer

25% of cancer-related deaths

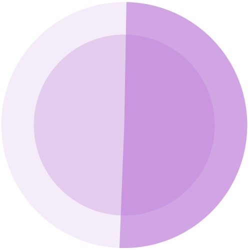
Third most common cancer behind skin and breast cancer.

More annual deaths than colon, breast, and prostate cancer-related deaths combined

New cases continue to decrease thanks to advances in early detection and treatment

2022 Infographics

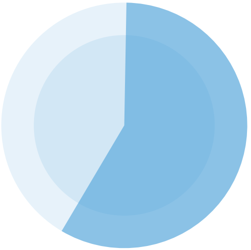
Estimates From	The American Cancer Society



Women

119,000 new cases
61,360 deaths

51% mortality rate



Men

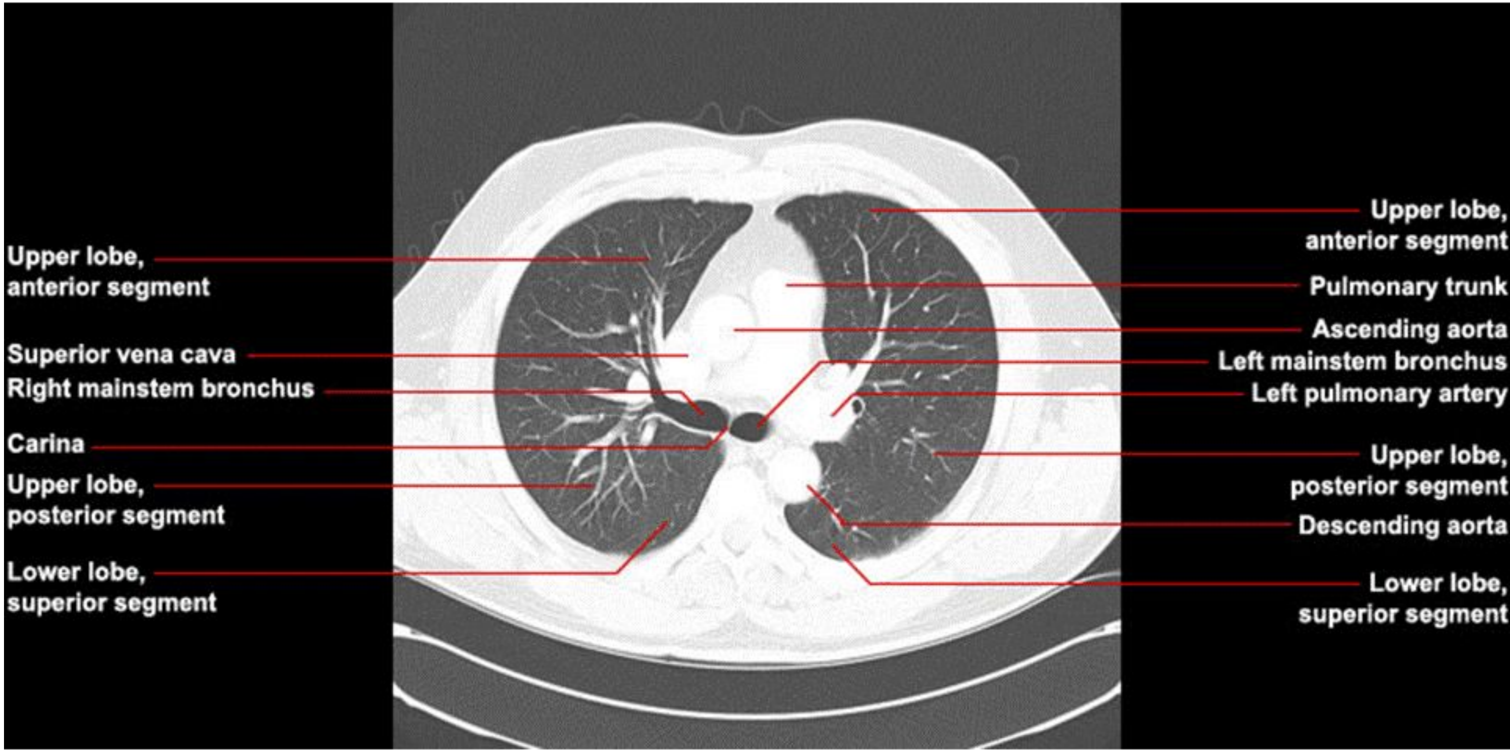
118,000 new cases
68,820 deaths

58% mortality rate

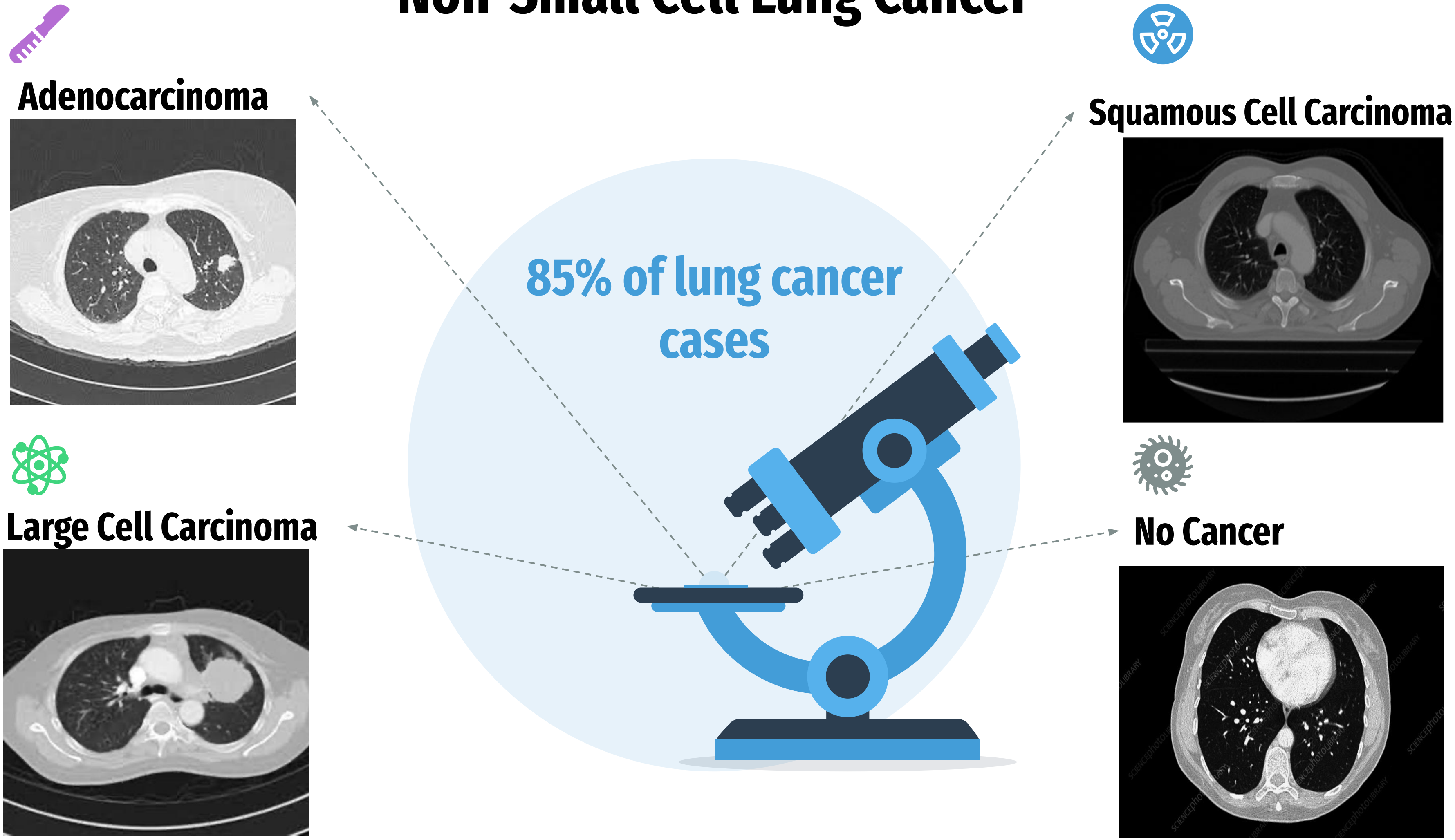
Understanding Lung Cancer

Cheat Sheet:

- Heart is the middle
- To the left and right of the heart are the lungs
- The white “branches” are your bronchioles

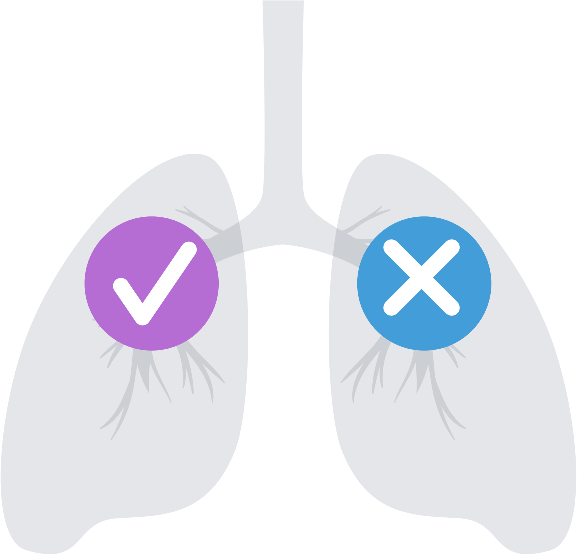


Non-Small Cell Lung Cancer



Why CT Scans?

CT Scan



X-ray

- Prior to the early 2000s, standard practice was to use an x-ray.
- Regular chest x-rays haven't shown to help most people live longer, and therefore aren't recommended for lung cancer screenings.
- CT scans have shown to find abnormal areas in the lungs that may be cancer.
- Research has shown that annual CT scans can save lives, especially in high-risk patients.

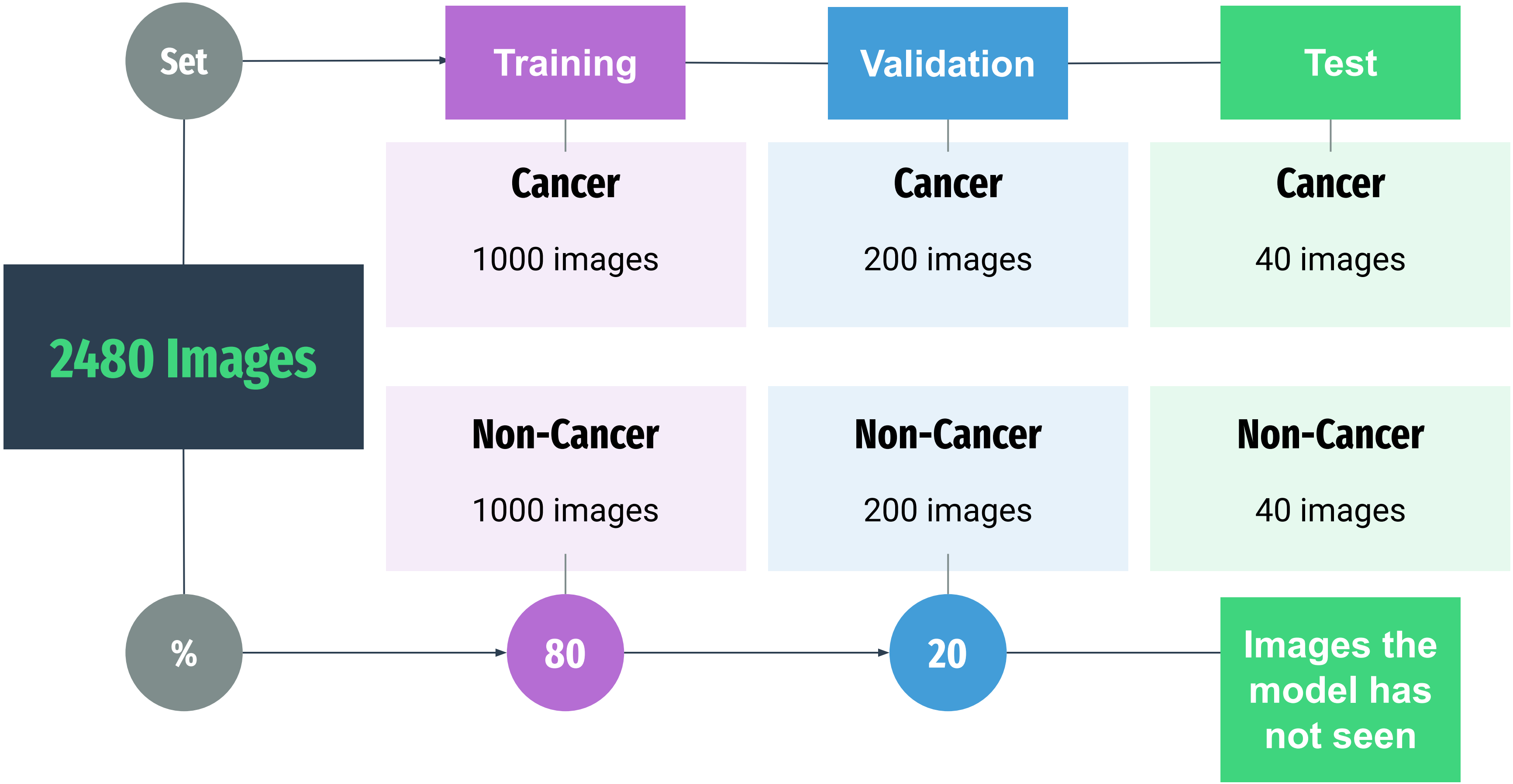
Problem Statement

- Using convolutional neural networks, classify chest computed tomography scans containing various forms of lung cancer with high accuracy and recall.

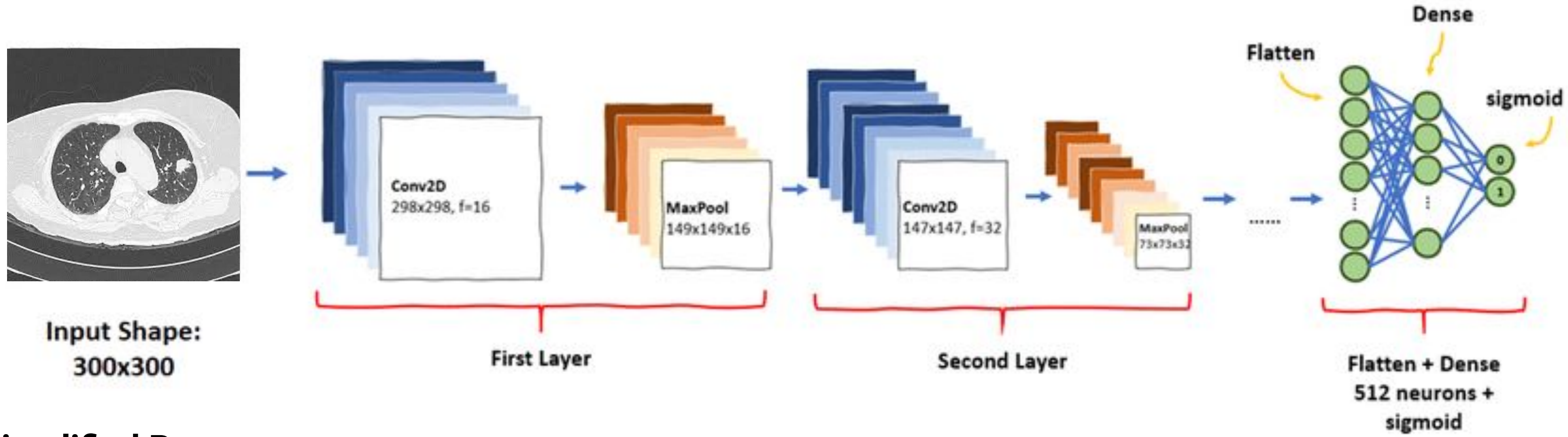
- This is a binary problem that uses deep-learning neural nets to determine whether a CT scan contains cancer or not.

https://www.kaggle.com/datasets/maedemaftouni/large-covid19-ct-slice-dataset
https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images
https://www.kaggle.com/datasets/adityamahimkar/iqothnccd-lung-cancer-dataset
https://www.cancerimagingarchive.net/

Data Oversight



Convolutional Neural Network



Simplified Process:

- Feed the network images of equal input shape
- Conv2D layer will “slide” over the image, multiplying and summing the “slide” into a single output pixel.
- The MaxPool Layer selects brighter pixels from the image. Where Conv2d multiplies and sums, MaxPool simply selects the largest number aka the brightest pixel.
- The Sigmoid layer is the output which classifies the image as either a 0 or 1. In this case, cancer or non-cancer.

Modeling Results

	Model	Accuracy	Cancer Recall	Non-Cancer Recall
	ResNet50	99%	97%	100%
	VGG16	99%	97%	100%
	MobileNetV2	94%	88%	100%
	Custom	95%	97%	93%

ResNet50 Architecture

CT scans being the grayscale images they are, are prone to overfitting.

The architecture of ResNet50 along with the regularization methods I implement counteract overfitting adequately.

- **CONV2D**
- **Global Average Pooling**
- **Batch Normalization**
- Dropout Layer (40%)
- Dense Layer (64 Neurons)
- Batch Normalization
- Dropout Layer(40%)
- Sigmoid Output

Real World Use

- The purpose of this model is to demonstrate the precision and efficiency that deep learning brings to the healthcare field.

- You should never base a diagnosis on what a model says. The real-world purpose is for a physician to use a model such as this as a reference. By eliminating cases where both a doctor and the model is confident the patient does not have cancer, they can focus on cases they believe do have cancer.

- Doing so could save lives and using CT scans has already proven to be the case!

Conclusions

- The original problem statement asks if it is possible to accurately and precisely classify whether a CT scan displays cancer or not.

- After constructing many models, I can conclude that it is possible. The ResNet50 model answers this question best, bolstering an accuracy of .99 on the entirety of the testing data as well as a precision of 1.0 and .98 for cancer and non-cancer images respectively.

- A focus was to minimize false negatives, and the ResNet50 model only misclassified one image. This one image was recurring through all models which leads me to believe there may be an issue with how the image is pre-processed.

Recommendations

- From here, I would mask and segment the lungs. Doing so would allow for a better idea of where cancer could be in the lungs. I attempted to do this however due to time constraints I was unable to complete this task.

- High frequency in file repetition caused a lot of time to be spent looking over the images. To keep the train-validation split at an acceptable rate I ended up having only 80 images to test my model on. I would gather more data and test how my model performs on larger datasets.

- Additionally to segmenting the lungs, training the model to localize where it believes cancer is would be ideal as well.

References

https://www.kaggle.com/datasets/maedemaftouni/large-covid19-ct-slice-dataset
https://www.kaggle.com/datasets/mohamedhanyyy/chest-ctscan-images
https://www.kaggle.com/datasets/adityamahimkar/iqothnccd-lung-cancer-dataset
https://www.cancerimagingarchive.net/
https://www.cancer.org/cancer/lung-cancer/about/key-statistics.html
https://www.cancer.gov/types/common-cancers#:~:text=The%20most%20common%20type%20of,are%20combined%20for%20the%20list.

Thank You

Instructions for use

In order to use this template, you must credit Slidesgo in your final presentation.

You are allowed to:

- Modify this template.
- Use it for both personal and commercial projects.

You are not allowed to:

- Sublicense, sell or rent any of Slidesgo Content (or a modified version of Slidesgo Content).
- Distribute Slidesgo Content unless it has been expressly authorized by Slidesgo.
- Include Slidesgo Content in an online or offline database or file.
- Offer Slidesgo templates (or modified versions of Slidesgo templates) for download.
- Acquire the copyright of Slidesgo Content.

For more information about editing slides, please read our FAQs or visit Slidesgo School:
https://slidesgo.com/faqs and https://slidesgo.com/slidesgo-school