



# Cancer Classification

An investigation of CNN Architecture

# Table of Contents

**01**

## **Introduction**

Domain  
background

**02**

## **Goals**

What do we want  
to learn from this  
project?

**03**

## **Model Selection**

What is the most  
effective pretrained  
model?

**04**

## **Model Architecture**

One model or two?

**05**

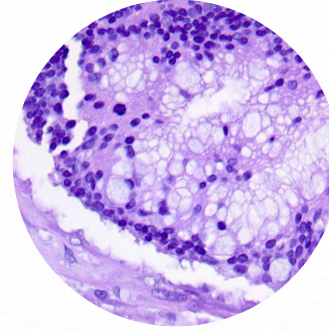
## **Fine Tuning**

Hyperparameters,  
regularization

# Introduction

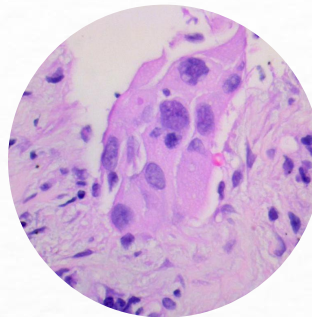
Histopathology - the diagnosis and study of diseases of the tissue

Colon

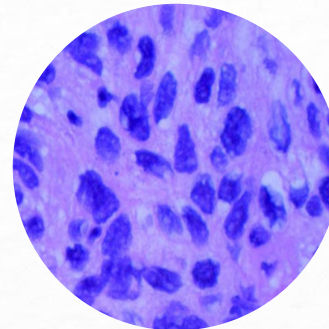


ACA

Lung



ACA



SCC



## Question:

How can we most effectively  
identify cancer in an x-ray of tissue?

# Goals



## Explore

What pre-trained model will give the best results?



## Refine

How can different approaches to the same problem yield different results?



## Fine Tune

How can we improve on our baseline model?



## Apply Regularization

What techniques can we apply to generalize our model?

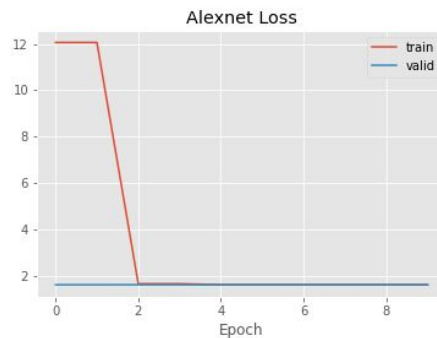
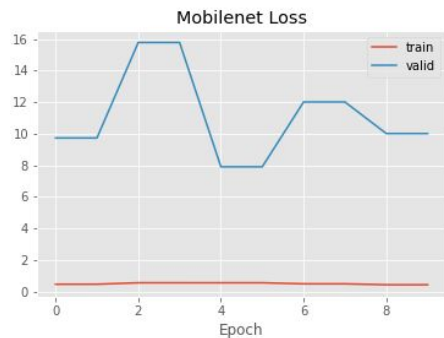
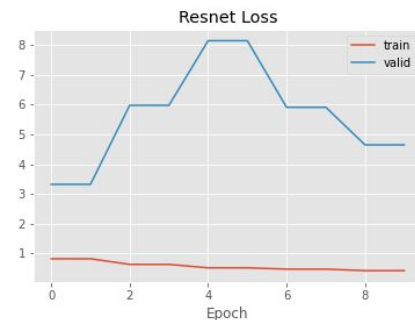
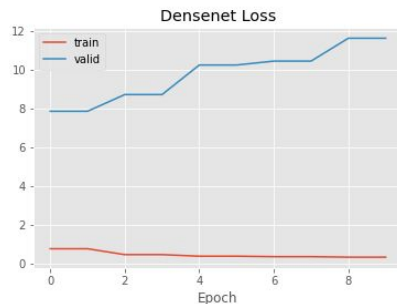
# Model Selection

**Mobilenet**

**Densenet**

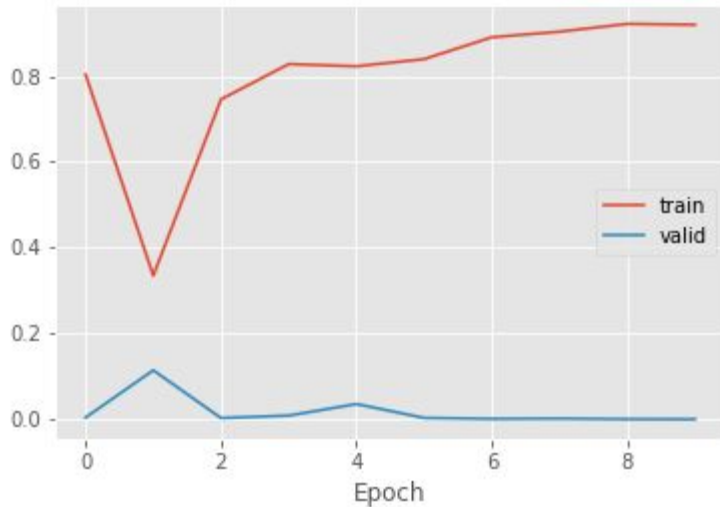
**Alexnet**

**Resnet**

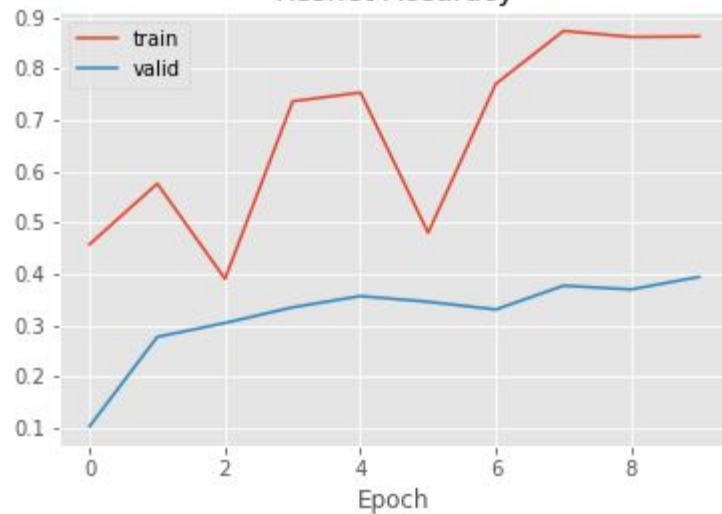


# Model Selection (Cont)

Mobilenet Accuracy



Resnet Accuracy





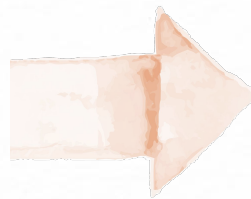
# Model Architecture

## Goal

Accurately  
identify cancer  
OR malignancy

## Strategy 1:

Does cancer exist  
in the image?



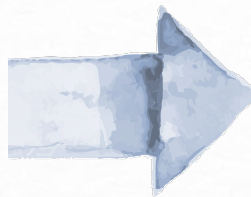
## Model:

Single binary  
classification model

## Strategy 2:

Does cancer exist  
in the image of:

- a. Colon
- b. Lung



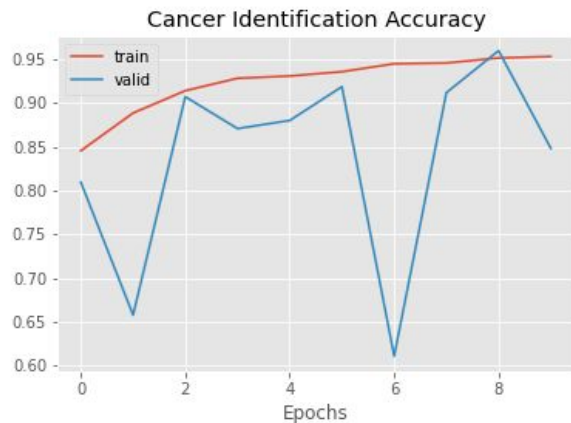
## Models:

TWO binary  
classification  
models

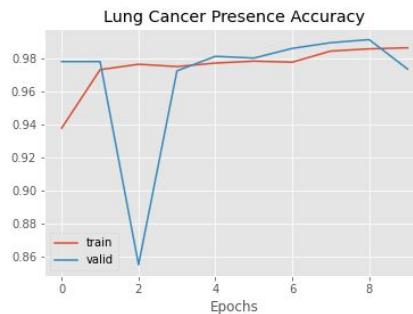
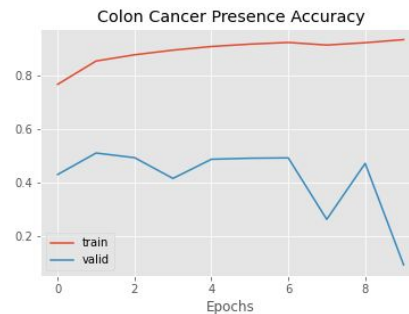


# Model Architecture Results

## Strategy #1



## Strategy #2



# Fine-Tuning

## Adjust Learning Rate

- Triangular Rate
- Layer Scaling

## Unfreeze Gradients

Gradual training of resnet layers

## Data Augmentation

Make more data

- physical transformations
- brightness
- fuzziness

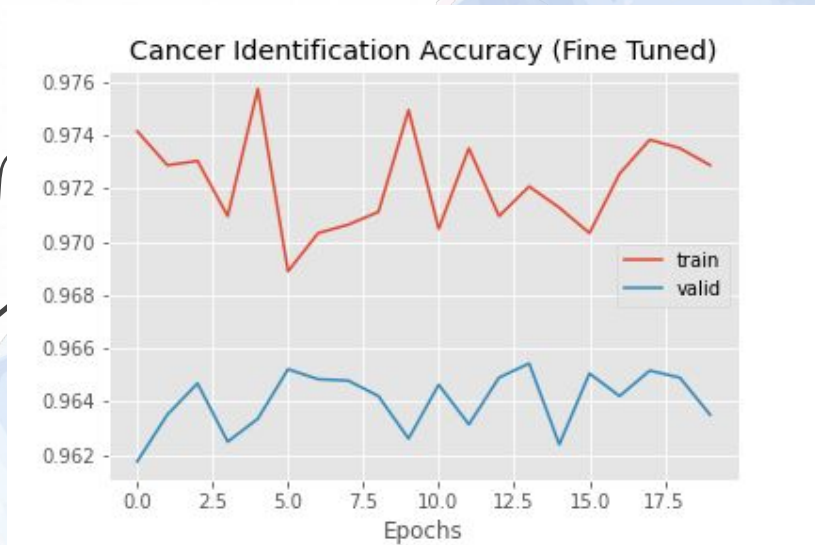
## Weight Decay

Apply L2 regularization

## Early Stopping

Stop training when metric decreases

# Results



96.5%