Cancer Classification

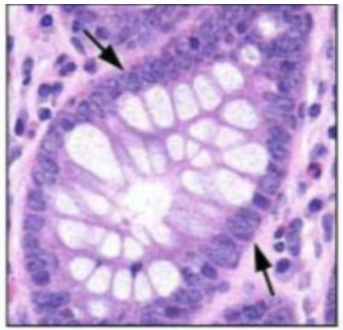
An investigation of CNN Architecture

Table of Contents

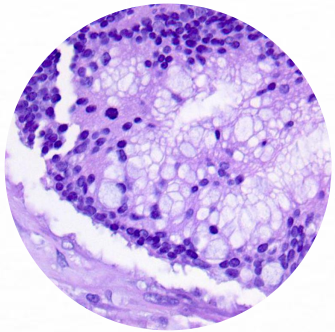
Introduction

Histopathology - the diagnosis and study of diseases of the tissue

Normal Colon



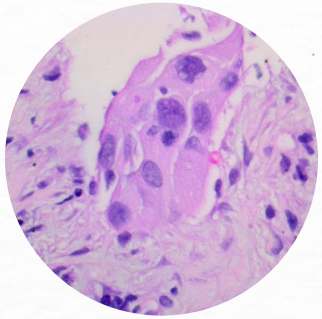
Colon

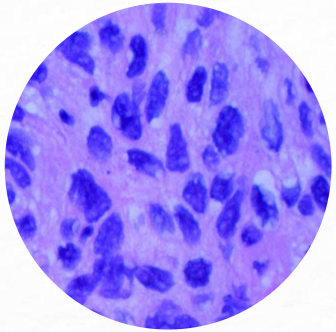


ACA

Lung

ACA





SCC

Question:

How can we most effectively identify cancer in microscopic slides of tissue and once the cancer is identified how can we classify the type of tissue involved?

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

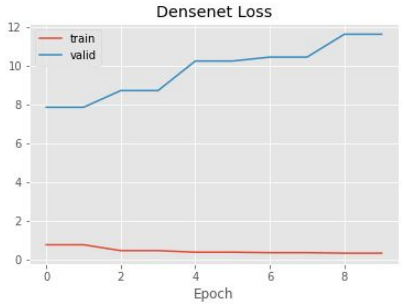
How can we improve on our baseline model?

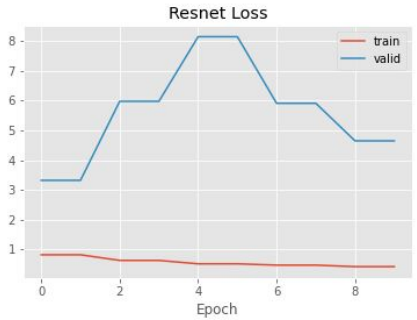
Apply Regularization

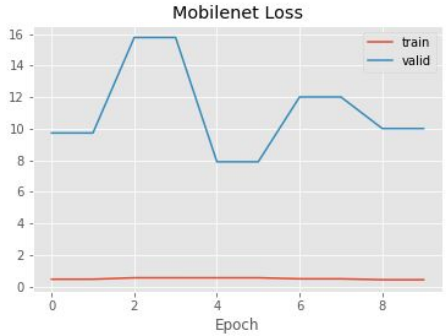
What techniques can we apply to generalize our model?

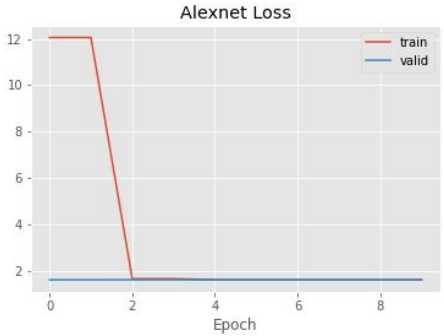
Model Selection

Mobilenet — Densenet — Alexnet — Resnet

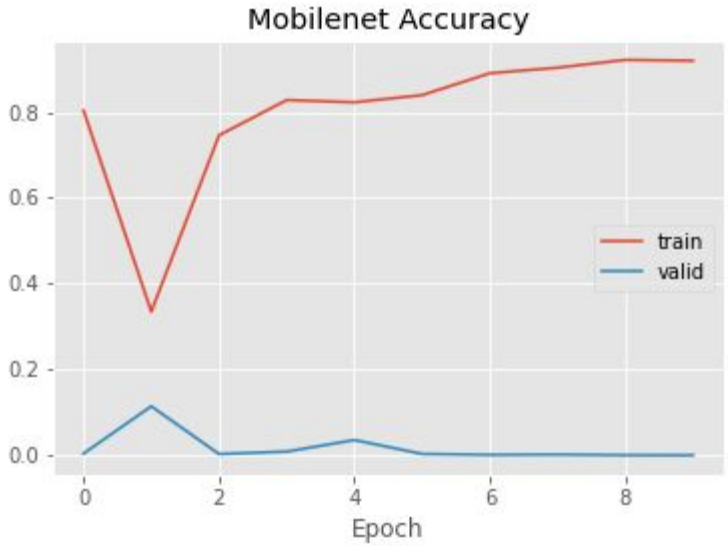


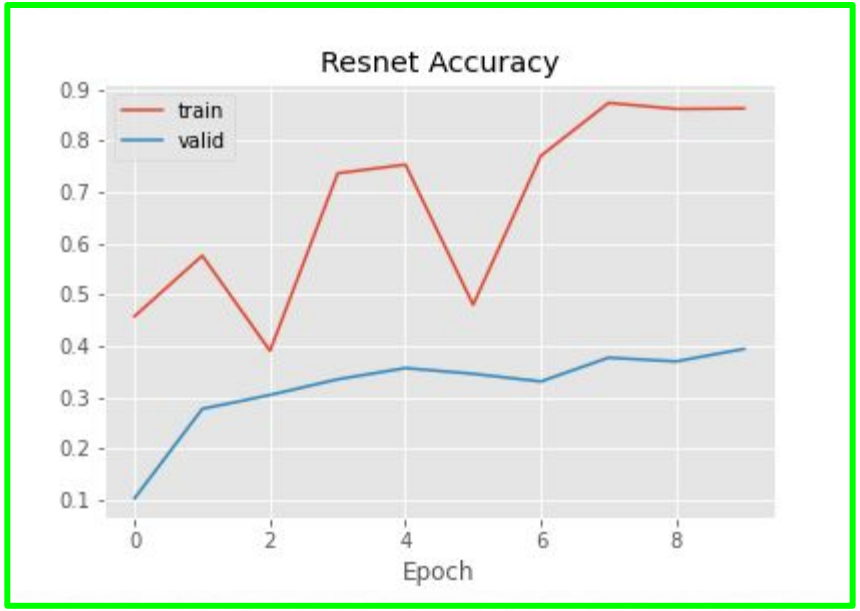






Model Selection (Cont)





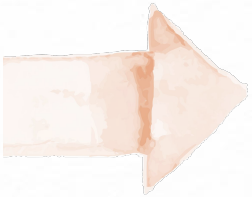
Model Architecture

Goal

Accurately identify cancer OR malignancy

Strategy 1:

Does cancer exist in the image?



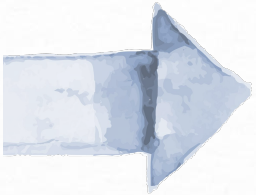
Model:

Single binary classification model

Strategy 2:

Does cancer exist in the image of:

a. Colon
b. Lung

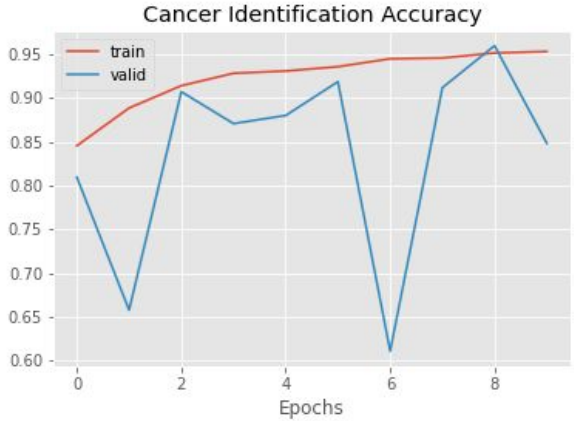


Models:

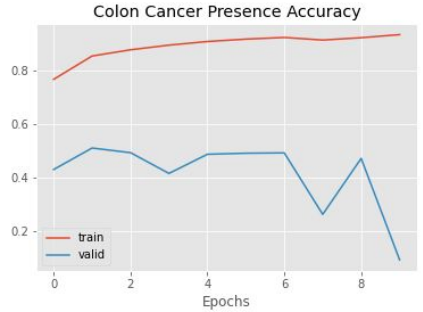
TWO binary classification models

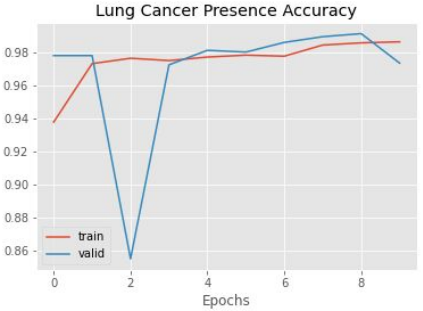
Model Architecture Results

Strategy #1



Strategy #2





Model Architecture

Goal

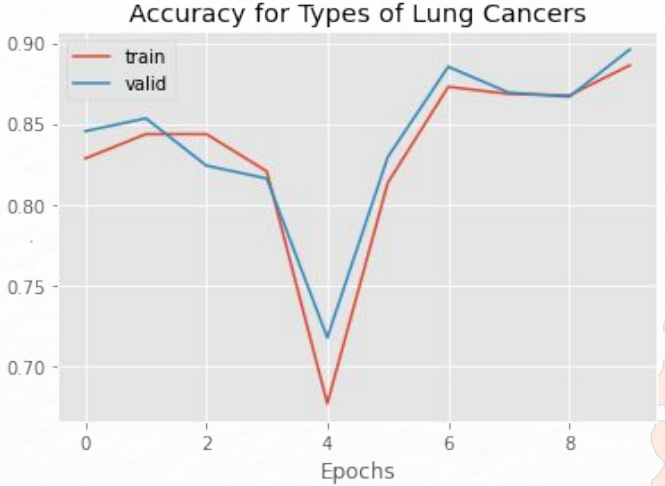
Accurately identify Types of Lung cancer

Strategy 3:

Types of Lung cancer, Adenocarcinoma or Squamous cell carcinoma

Model:

Single binary classification model



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

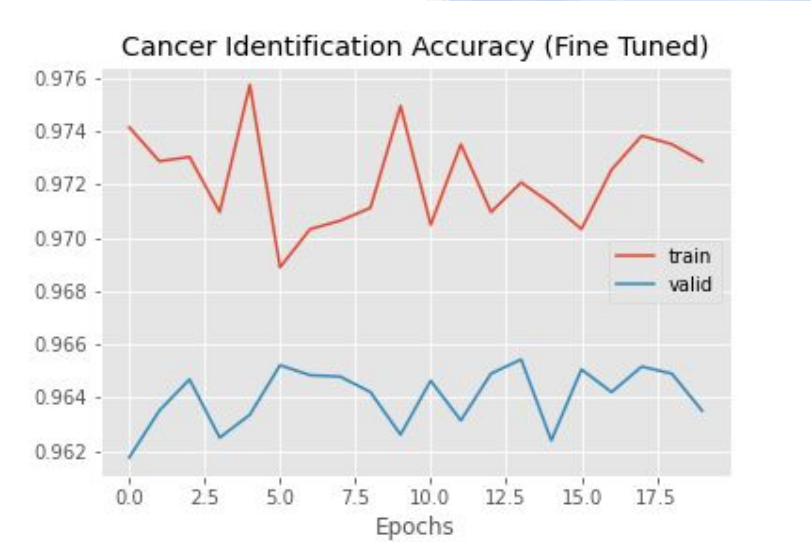
Weight Decay

Apply L2 regularization

Early Stopping

Stop training when metric decreases

Results


Cancer Identification Accuracy (Fine Tuned)
0.976
0.974
0.972
0.970
0.968
0.966
0.964
0.962
train
valid
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
Epochs

96.5%