

Lung Cancer Subclassification

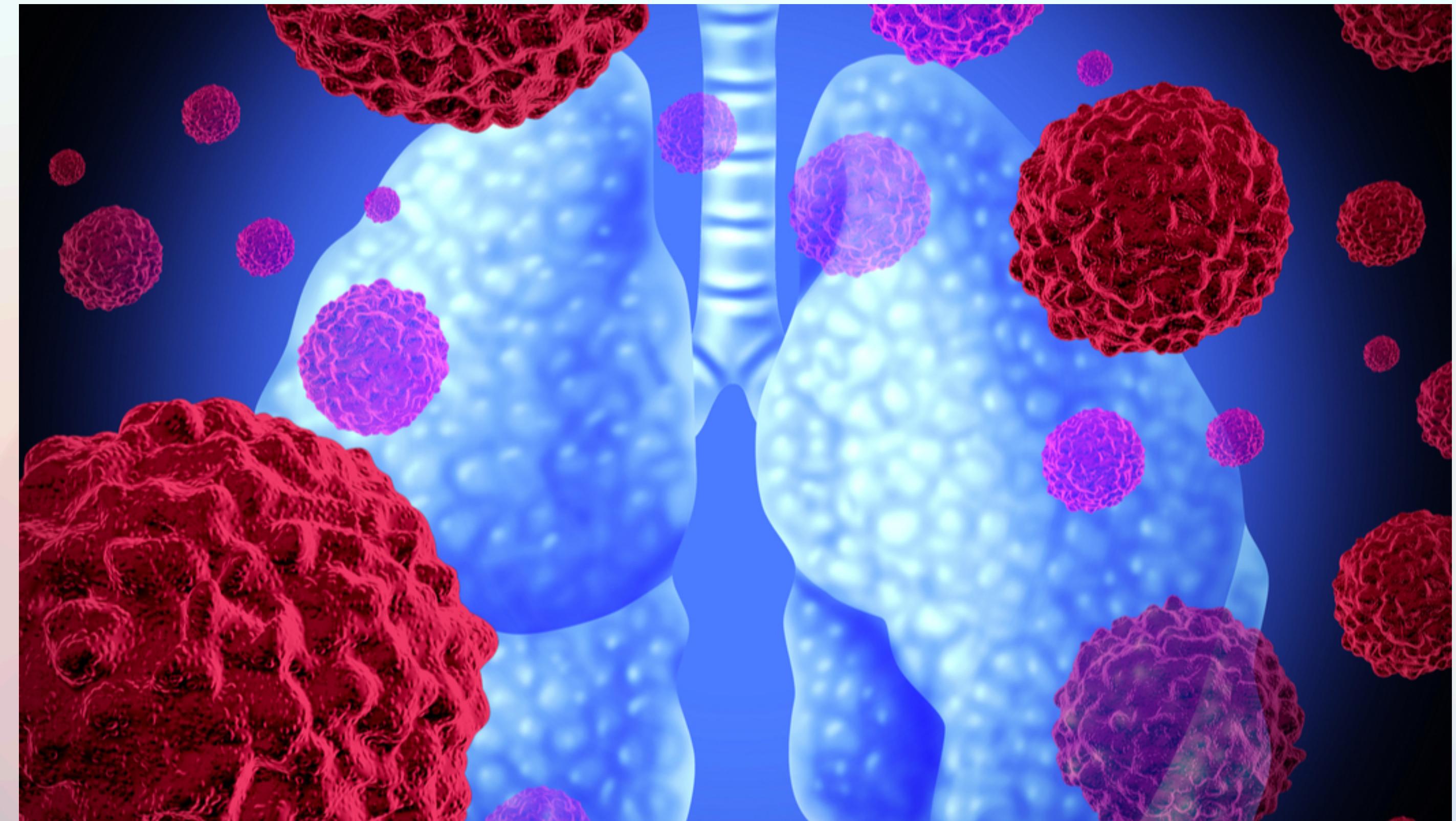
Using FastAI and Hugging Face

Contents

Context

Data

Recommendations



Context

Problem Statement

Given histopathological images of lung benign tissue, lung adenocarcinoma, and lung squamous cell carcinoma, can we build and deploy a computer vision model utilizing transfer learning that classifies new histopathological images of lung tissue more accurately than a domain expert?

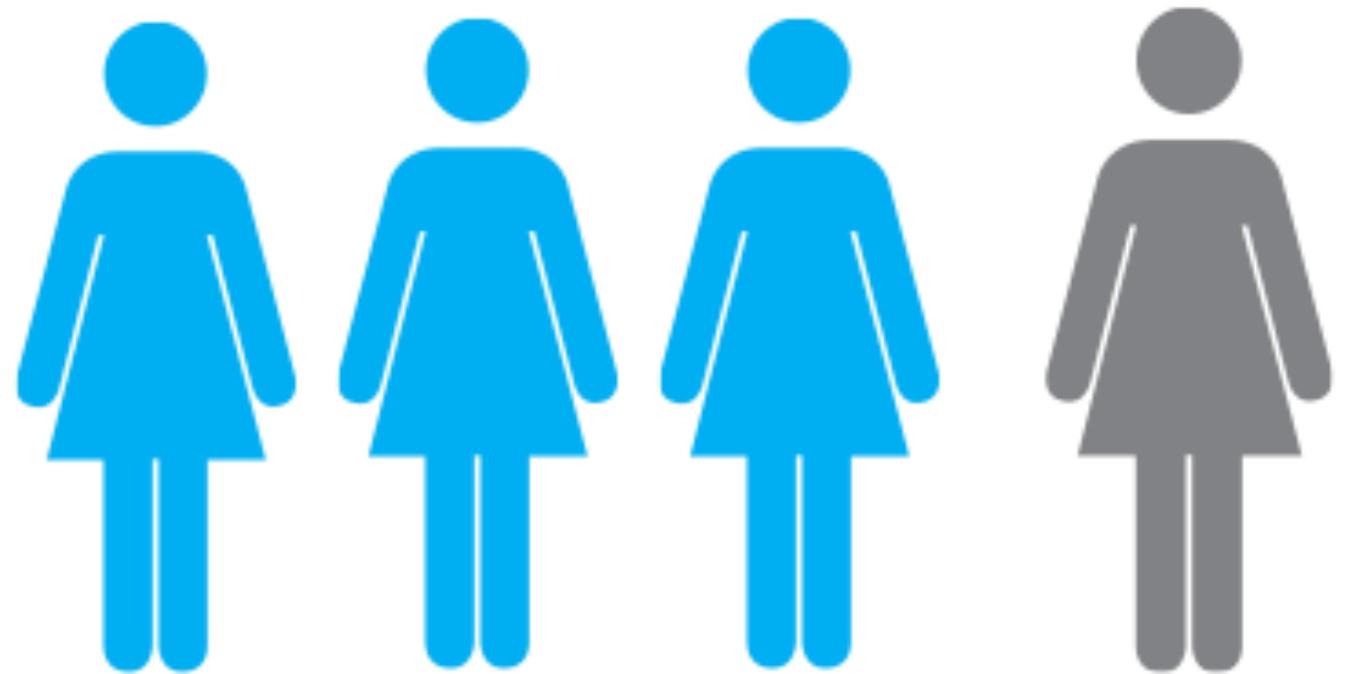


Lung Cancer Facts

- 237,000 Americans will be diagnosed this year.
- 130,000 will die annually.
- 382 killed each day. That's a Boeing 777 airplane.
- 22% survival rate: lowest among common cancers.
- Leading cause of cancer death, least amount of federal research funding.



LUNG CANCER kills almost **3 TIMES** as many men as prostate cancer. ([source](#))



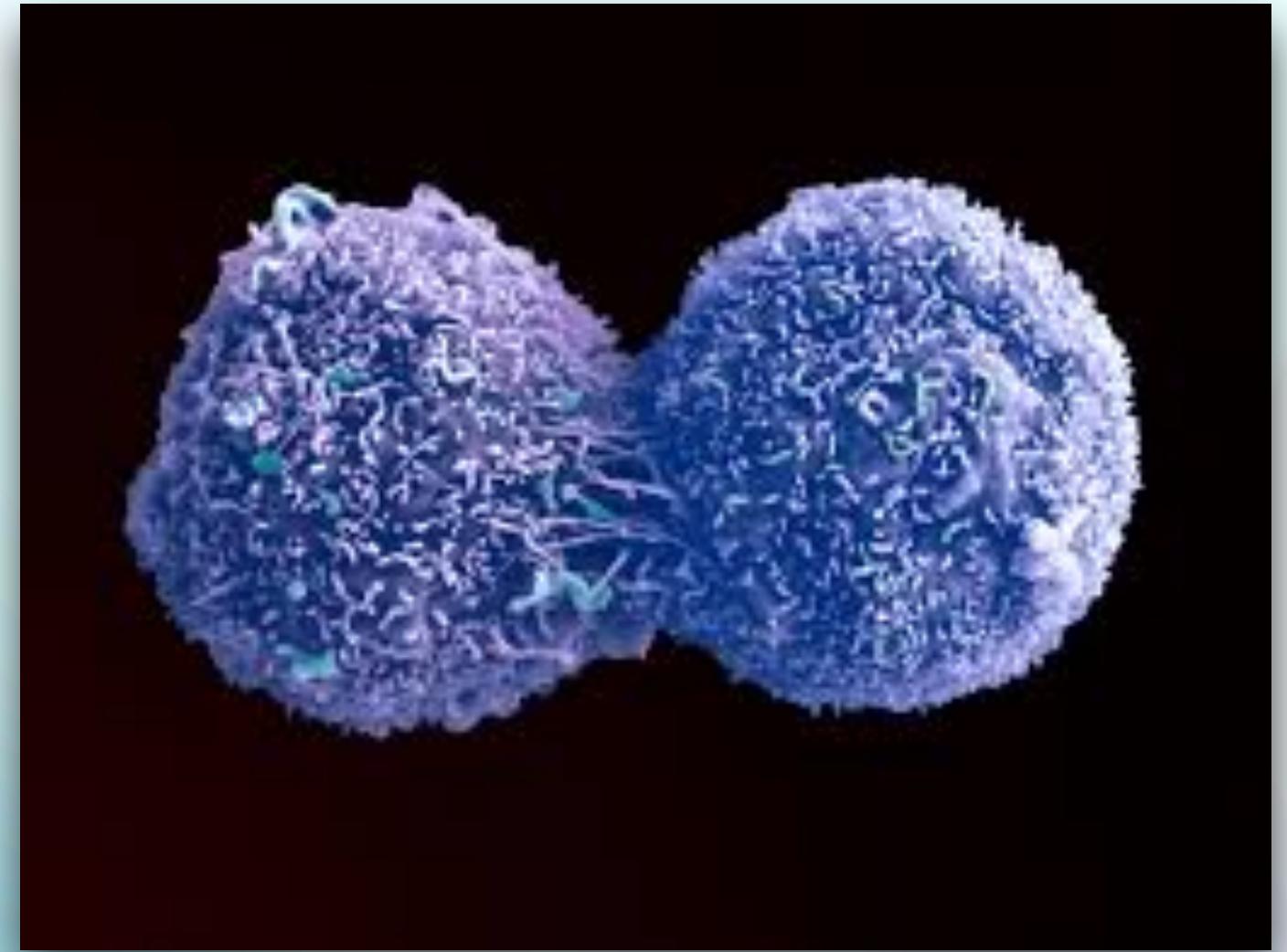
LUNG CANCER kills almost **3 times** as many women as breast cancer. ([source](#))

541,000

People in the U.S. today have been diagnosed with lung cancer.

Non-Small Cell Lung Cancer Types

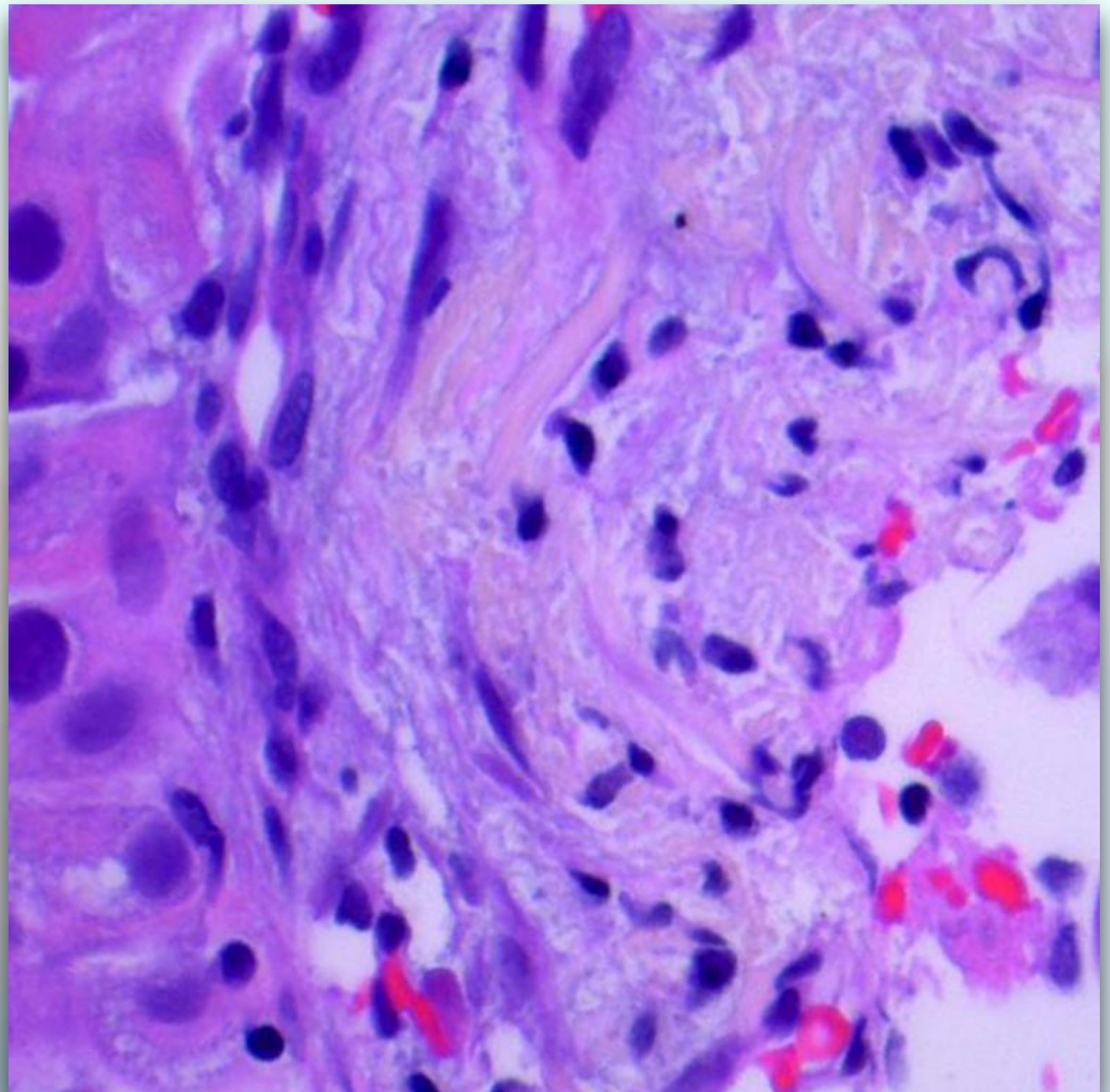
- Adenocarcinoma (40% of lung cancers)
- Squamous cell carcinoma (25% of lung cancers)
- Large cell carcinoma (10% of lung cancers)



Data

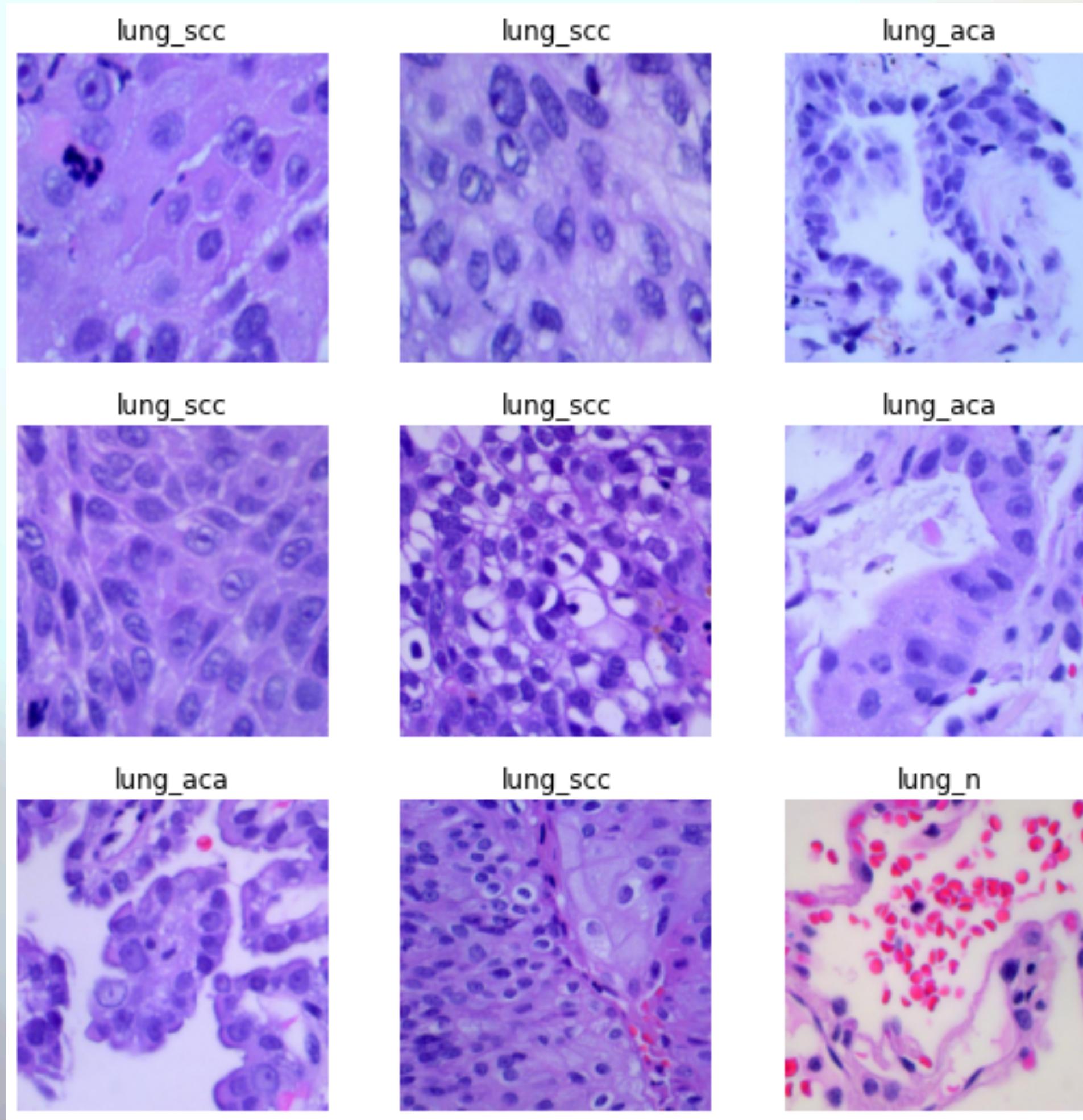
What are we working with?

- Lung Benign Tissue: 5,000 images
- Lung Adenocarcinoma Tissue: 5,000 images
- Lung Squamous Cell Carcinoma Tissue: 5,000 images
- HIPAA Compliant, Validated, De-identified

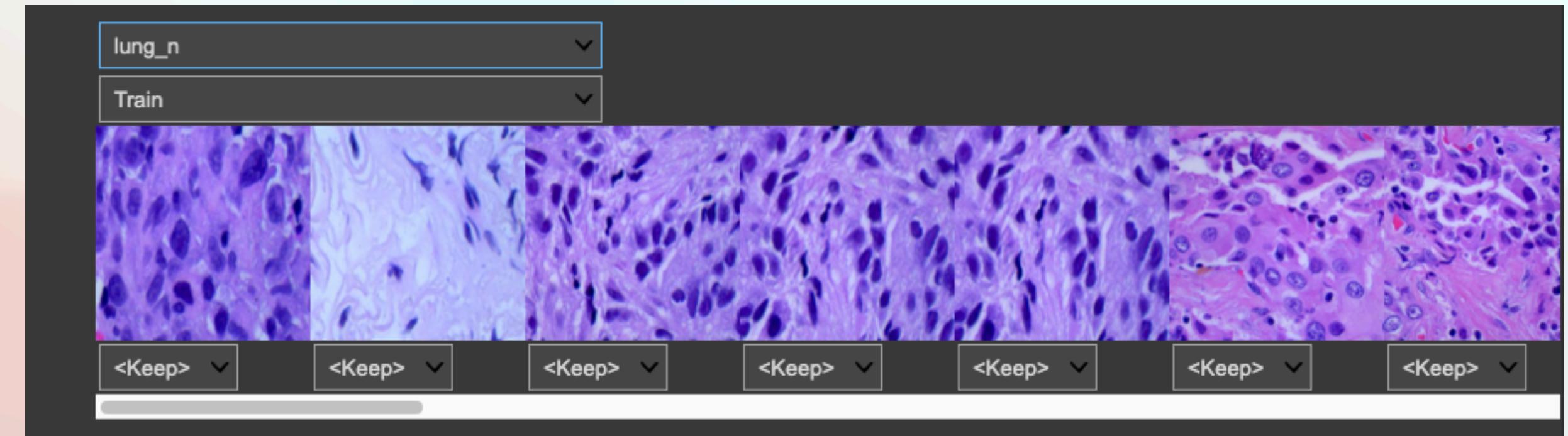


Exploring our Data

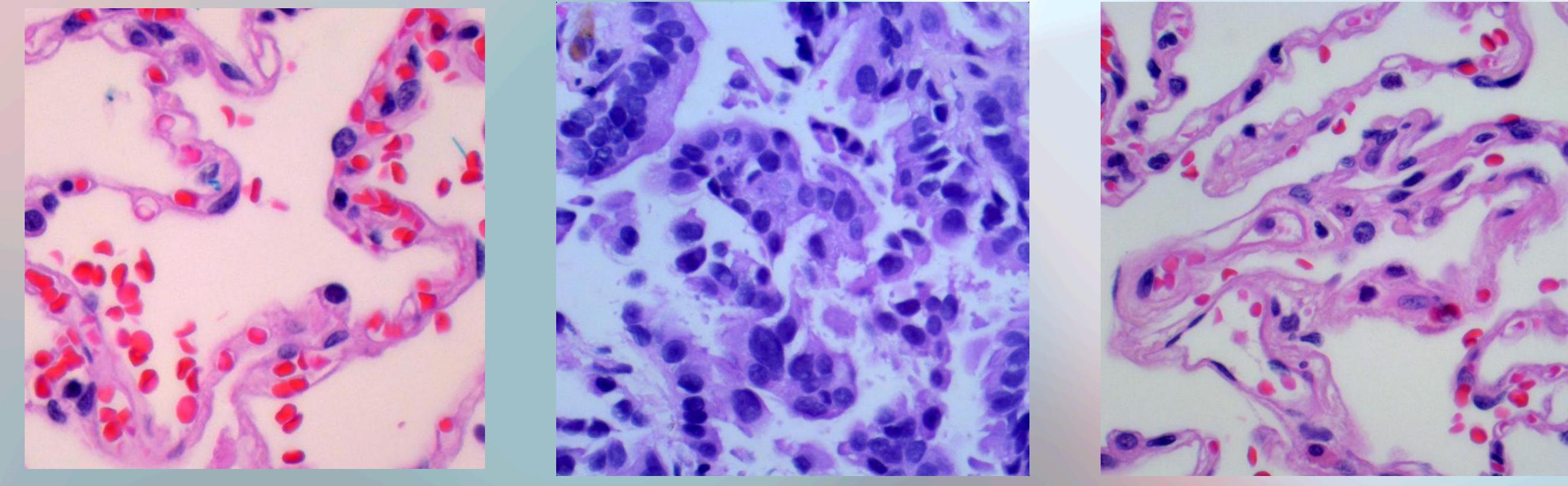
We can show a batch of data!



We can remove or relabel images if needed



Or we can examine image data individually

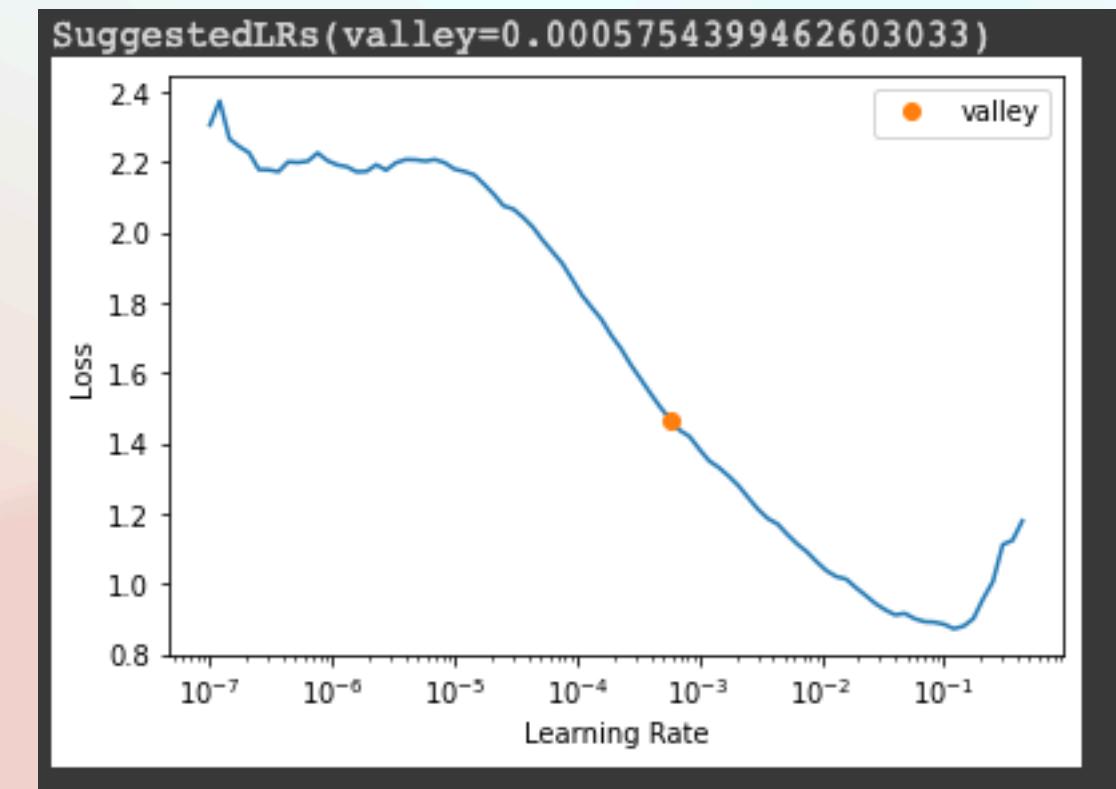


Chosen Model?

Resnet50

- Utilizes transfer learning
- Best accuracy: 0.999
- Perfect recognition of benign tissue
- More accurate than a domain expert/pathologist.

1. We visualized our learning rate

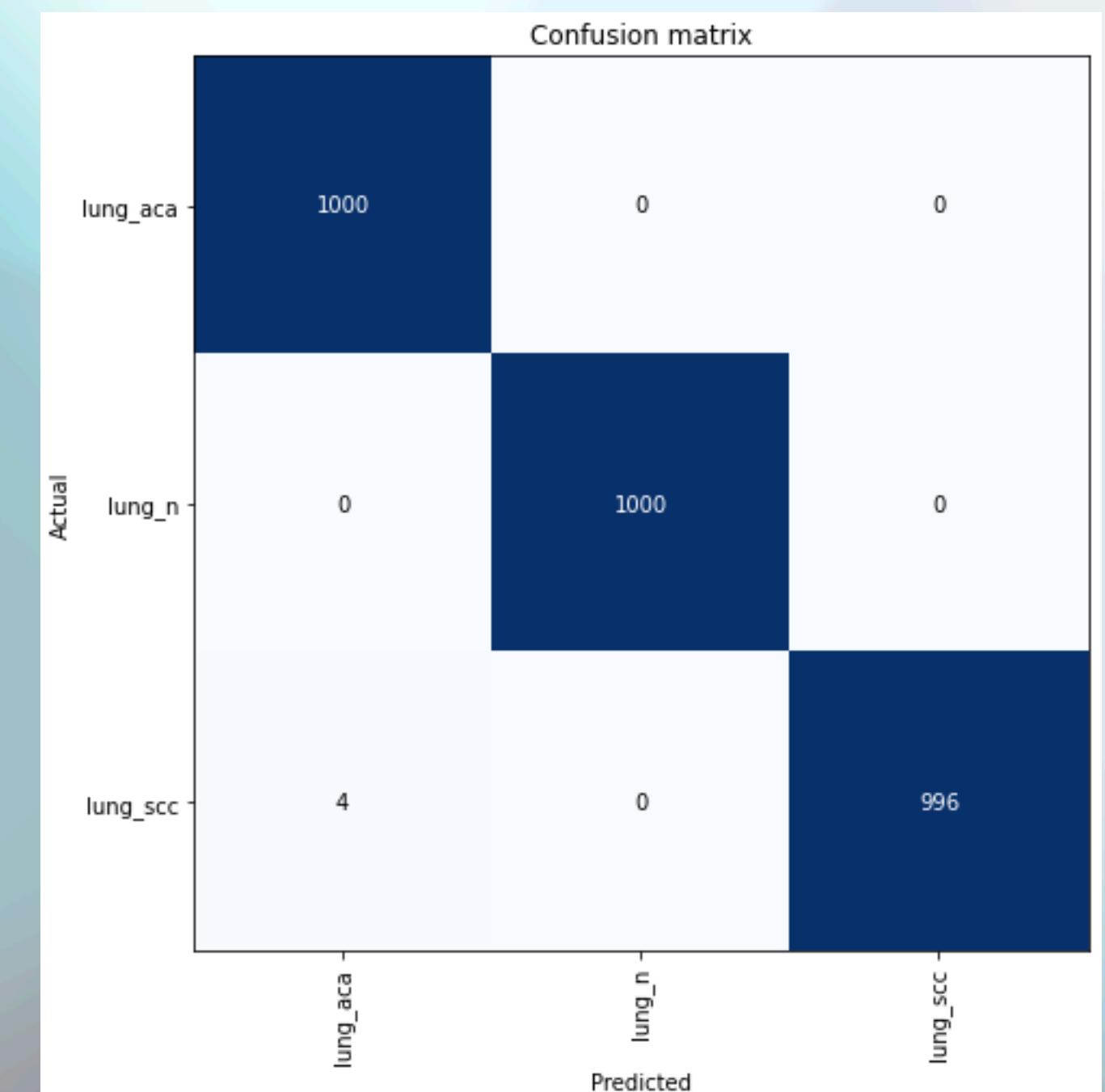
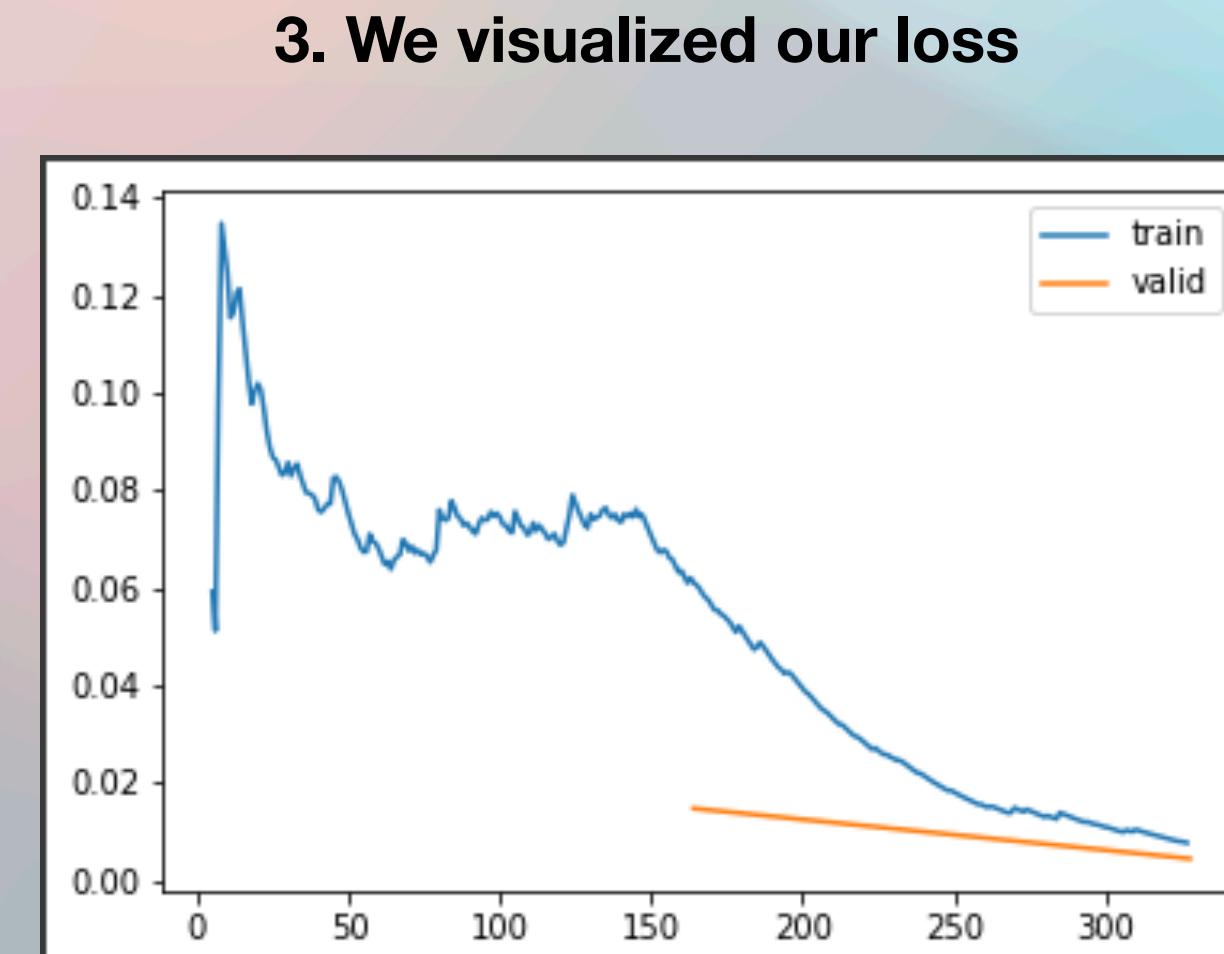


2. We trained our model

epoch	train_loss	valid_loss	accuracy	time
0	0.145546	0.065961	0.977000	15:47

epoch	train_loss	valid_loss	accuracy	time
0	0.062059	0.014836	0.993000	03:31
1	0.007783	0.004542	0.998667	03:29

4. We view our predictions

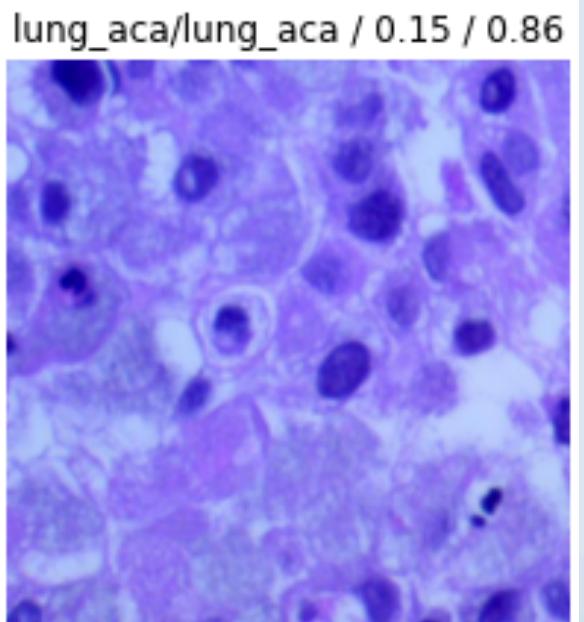
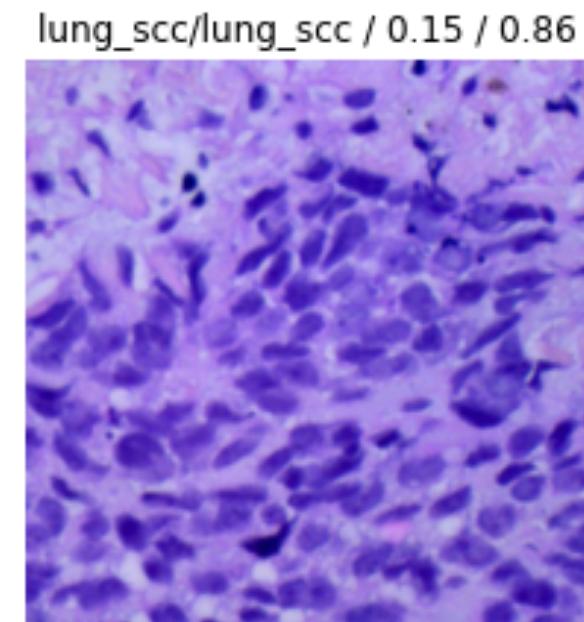
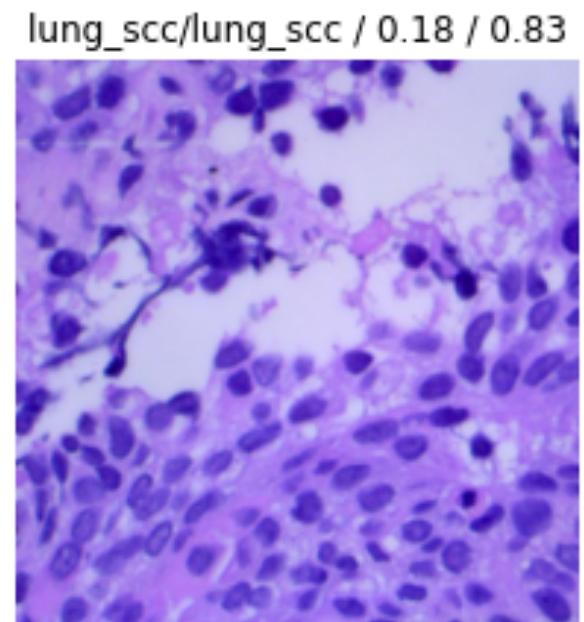
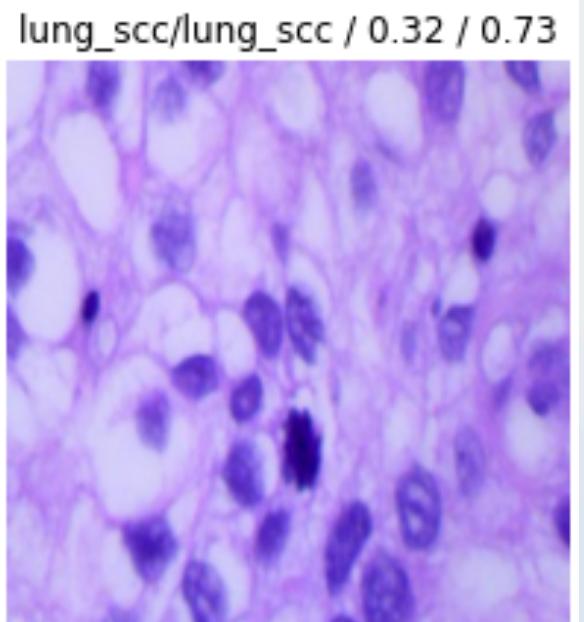
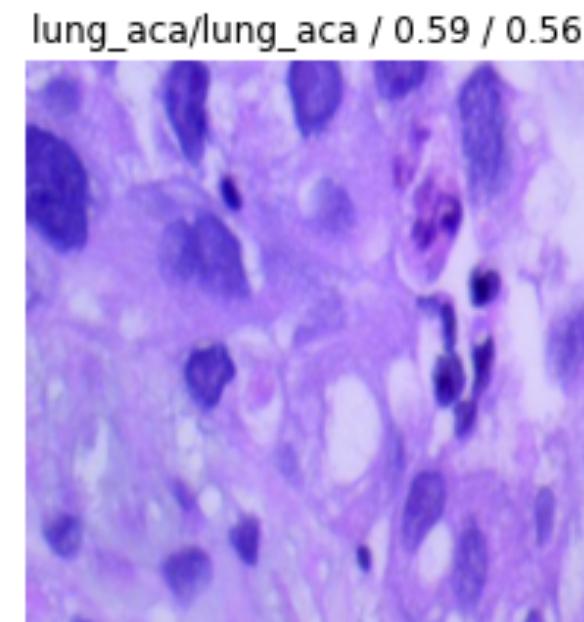
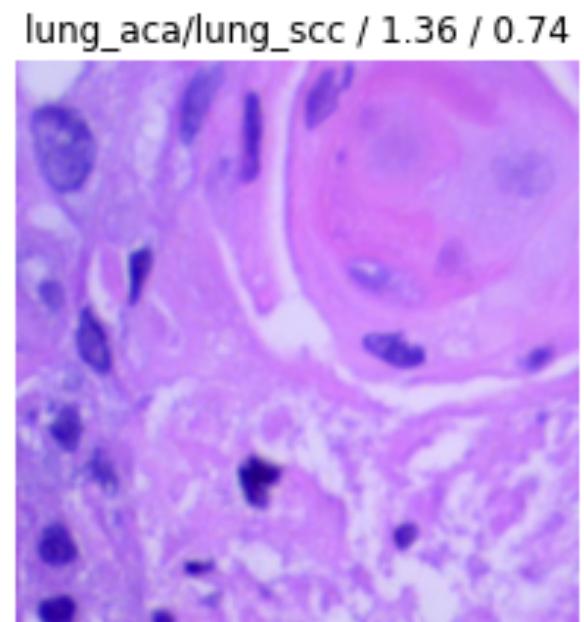
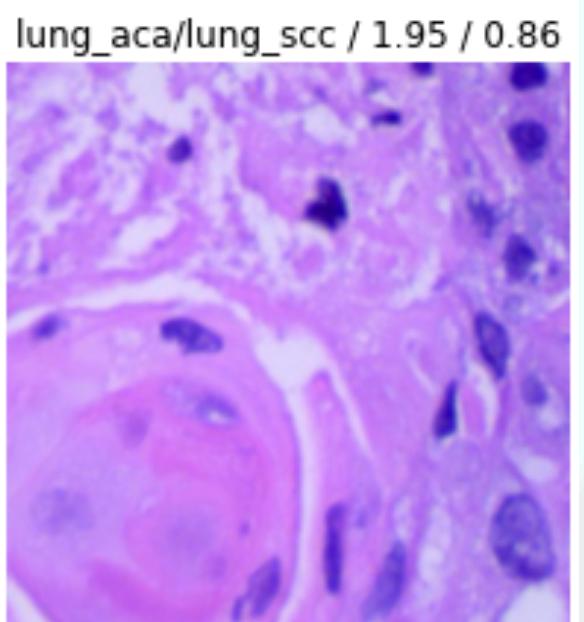
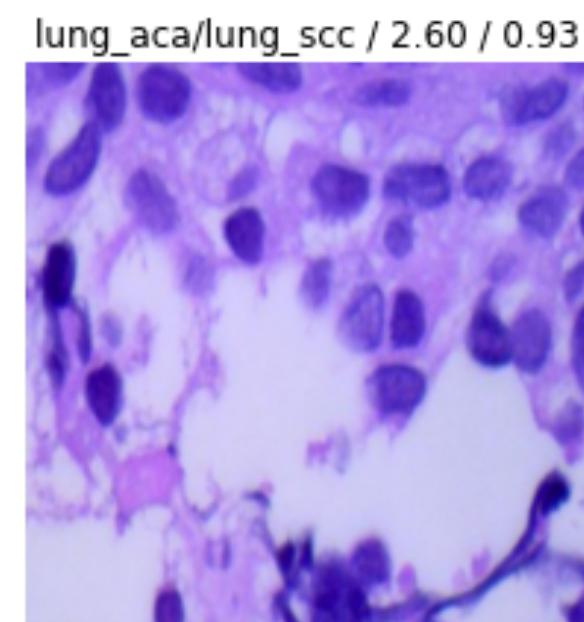
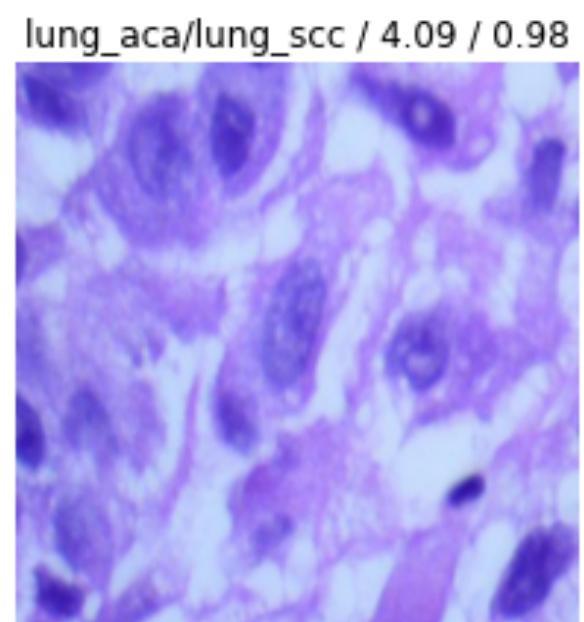


Interpreting our Model

What does it see?

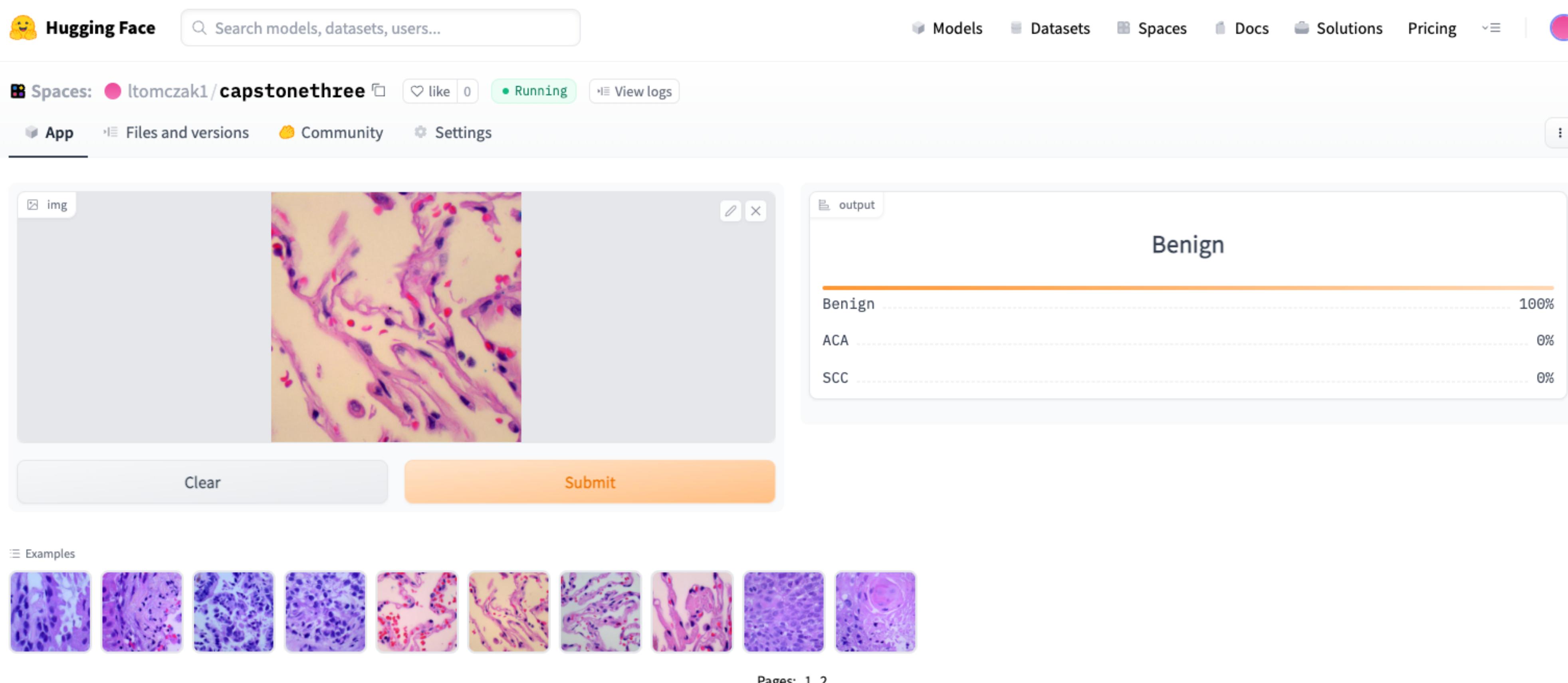
- Overly confident predicting lung adenocarcinoma.
- All 4 misclassifications were actually lung squamous cell carcinoma.
- High loss due to high confidence in incorrect prediction.
- Not so confident on some SCC predictions.
- Very certain with benign tissue.

Prediction/Actual/Loss/Probability



App in action!

- Hugging Face and Gradio make the process easy.
- We can test new histopathological images by dragging and dropping



Recommendations

What can we do about it?

- Take biopsies of suspected cancerous and precancerous lung tissue.
- Standardize histopathological images of biopsies.
- Pass images to our state-of-the-art image classifier for a diagnosis.



Benefits

Why change things?

- More accurate than a pathologist.
- Faster turnaround time than traditional analysis.
- Ability to create targeted treatment plan, faster.
- Reduced patient and provider costs.
- More image data on lung cancer.



**STOP
LUNG
CANCER**

Reorder: NHEP-19570 www.ComplianceSigns.com