PREDICTING THE PRIMARY TUMOR OF BRAIN METASTASES USING STIMULATED RAMAN HISTOLOGY

Mahsa Vali

Feb 2024

CMMC
CENTER FOR MOLECULARE MEDICINE COLOGNE

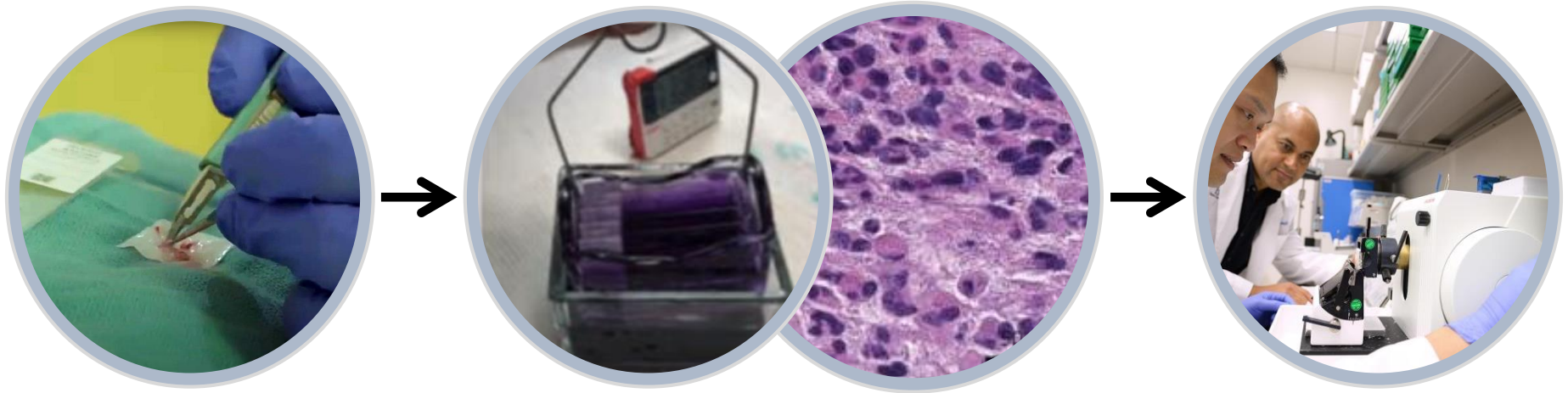
UNIKLINIK KÖLN

University of Cologne

Problem Statement

Predict the primary from brain metastases imaged using Raman spectroscopy

Traditional diagnosis workflow



tissue extraction → imaging (H&E staining) → interpretation

RESULTS IN UP TO 2 HOURS

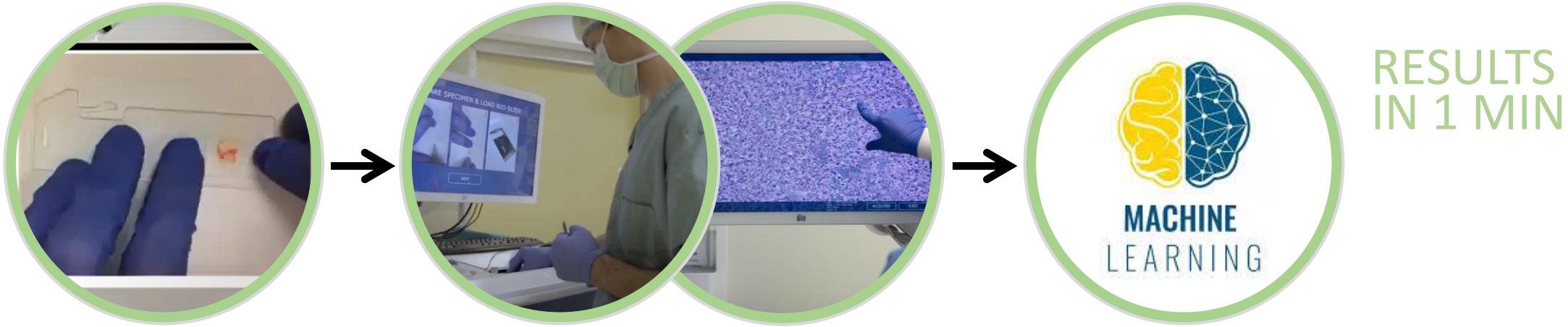
operating room

outside the operating room

Problem Statement

Predict the primary from brain metastases imaged using Raman spectroscopy

Diagnosis workflow assisted by Deep Learning



RESULTS IN 1 MIN

tissue extraction

imaging (SRH)

DL-aided interpretation

operating room

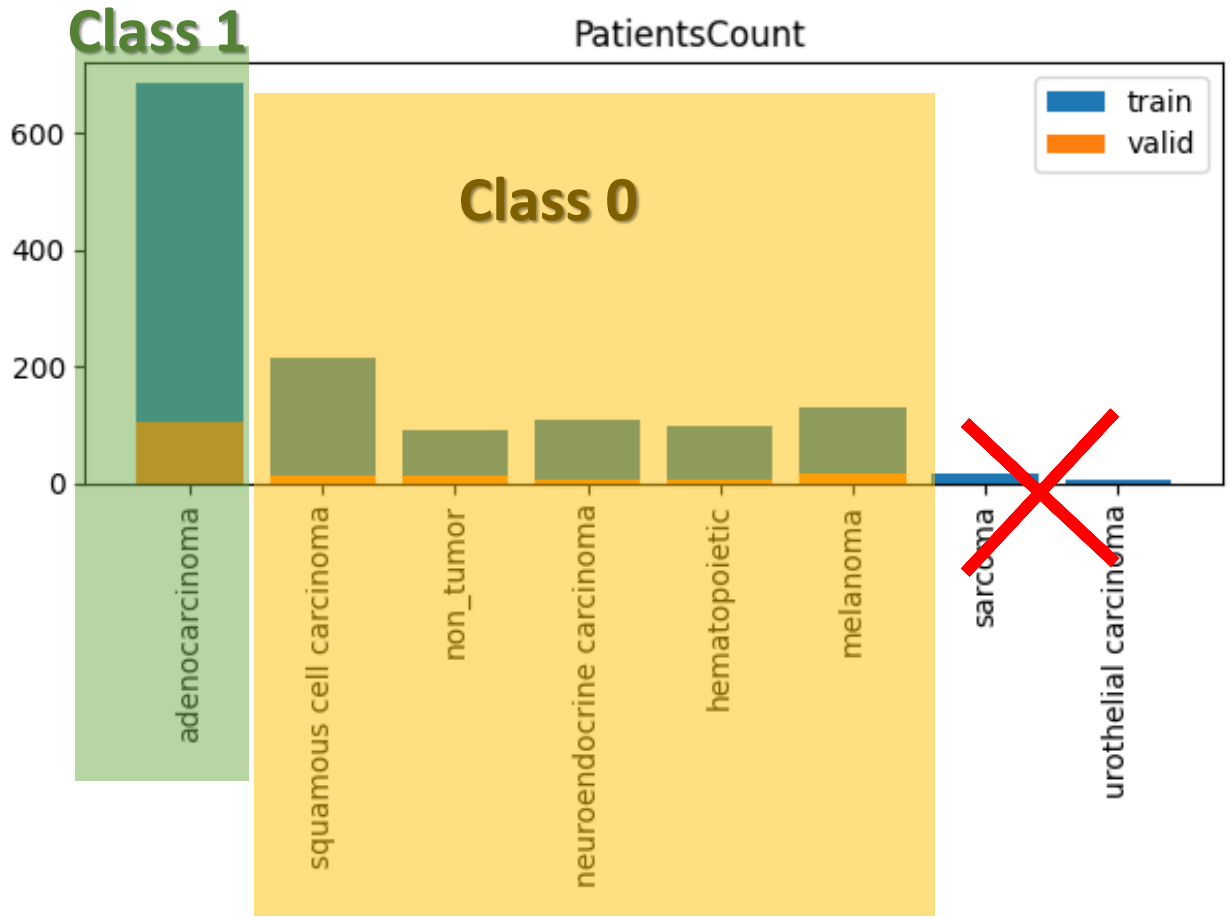
Dataset
Data analysis

The dataset:

collected from University Hospital Cologne
Includes ~296 patients
Recorded using Raman technology (SRH)
Stored in DICOM format

Task :

Multi-Class Classification
Initially decided to :
Adenocarcinoma VS others

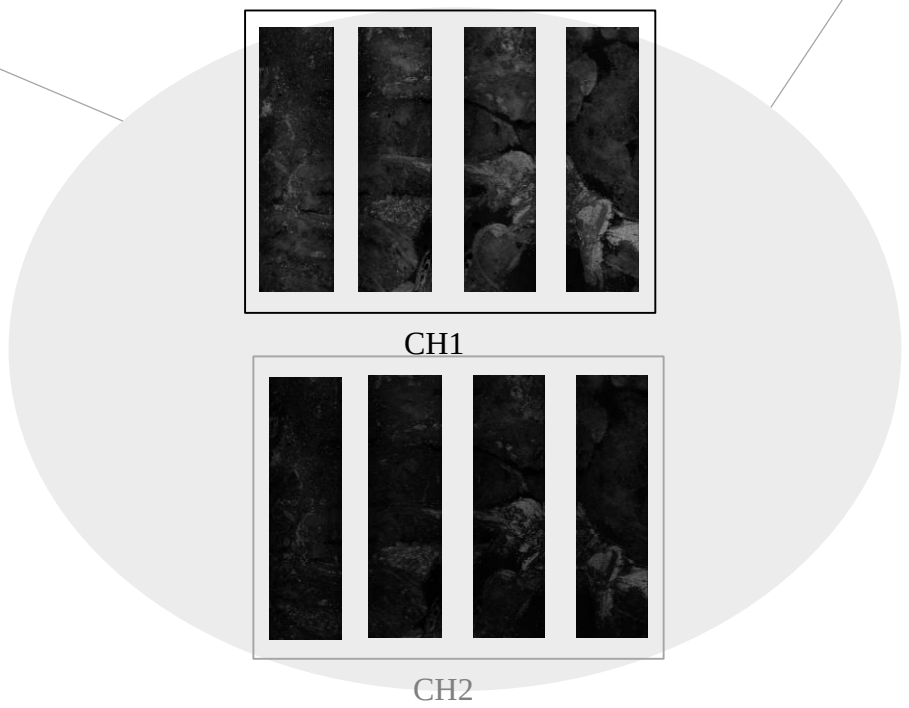


Dataset
Data Preprocessing

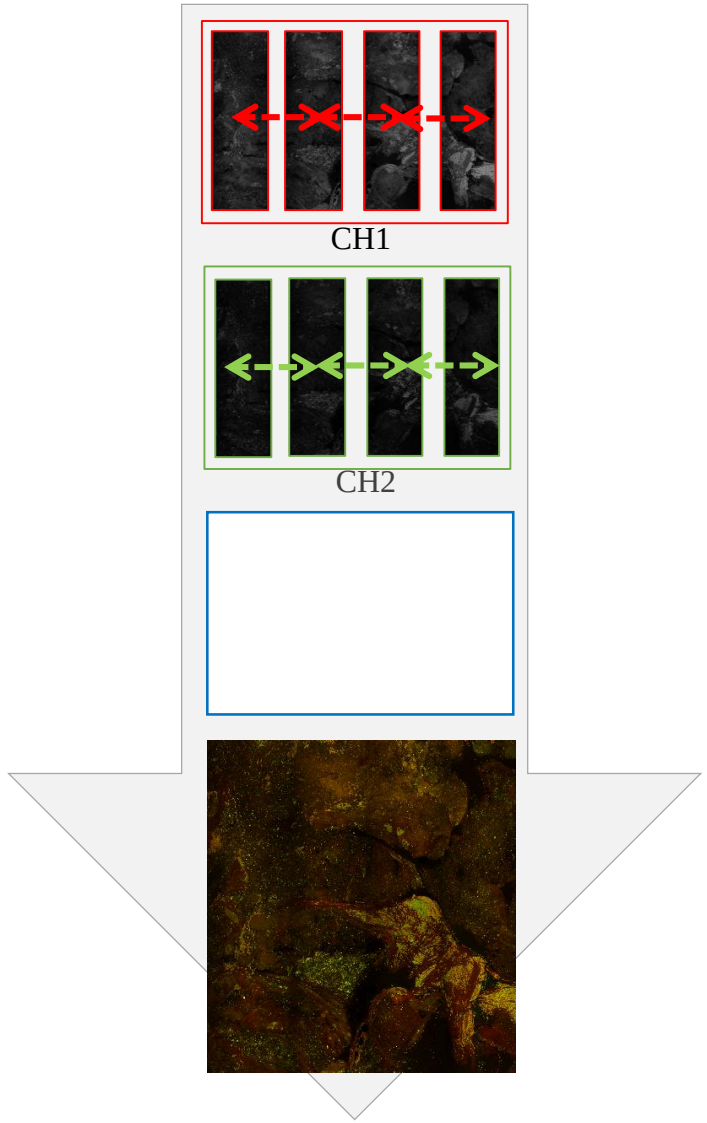
✓ Strips

Assigning 2 spectral strips to the R & G channels respectively, while the3rd input for B :

- Mean $[CH1\ , CH2]$
- Zero
- Minus $[CH1\ , CH2]$



✓ Stitching



Dataset
Data Preprocessing: Patching

There are two assumptions:

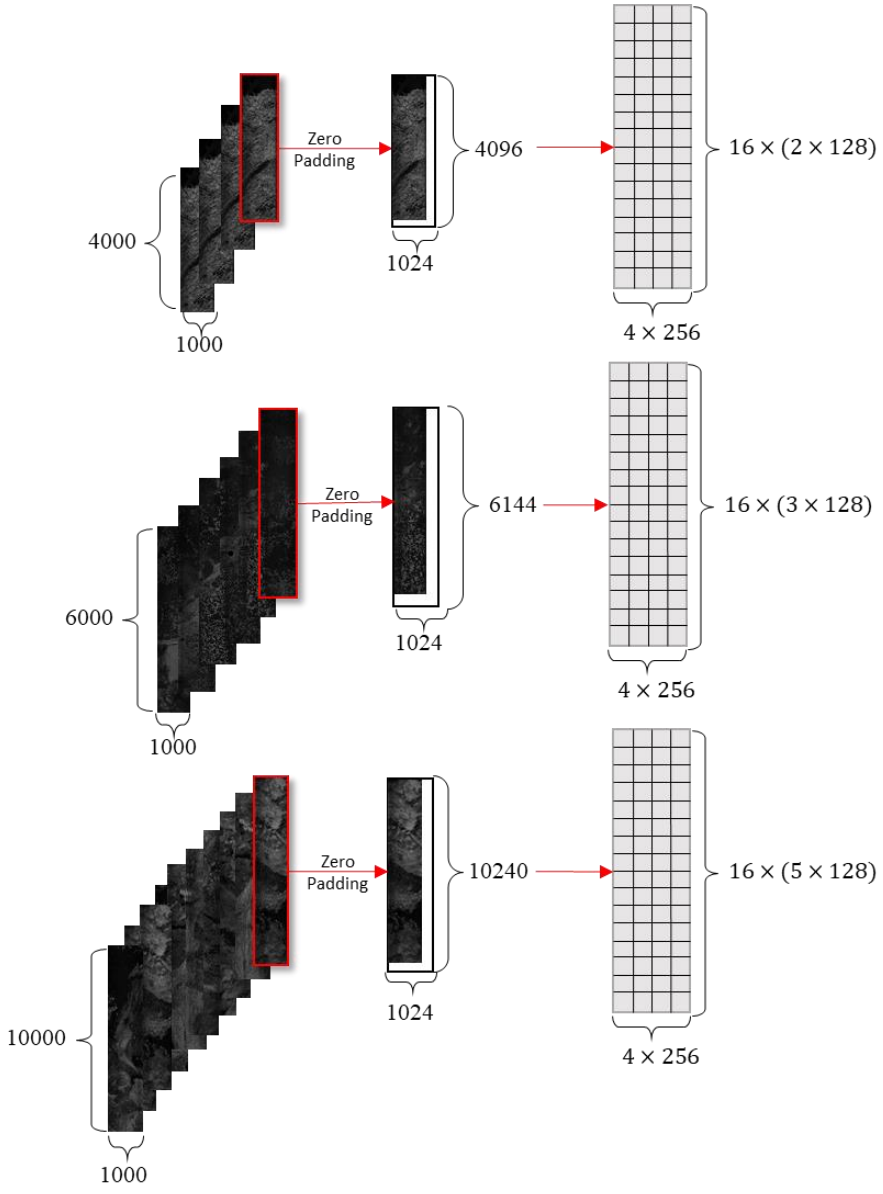
1. Extracting an equal number of patches from images with different dimensions.
2. Extracting patches with the same dimensions but varying numbers in each image.

Patches from Strips :

$$1000 \times 4000 \rightarrow 1024 \times 4096 \Rightarrow 64 \times 256 \times 256$$
$$1000 \times 6000 \rightarrow 1024 \times 6144 \Rightarrow 64 \times 256 \times 384$$
$$1000 \times 10000 \rightarrow 1024 \times 10240 \Rightarrow 64 \times 256 \times 640$$

Patches from Stitch :

$$4000 \times 4000 \rightarrow 3584 \times 4096 \Rightarrow 256 \times 224 \times 256$$
$$6000 \times 6000 \rightarrow 5632 \times 6144 \Rightarrow 256 \times 352 \times 384$$
$$10000 \times 10000 \rightarrow 9216 \times 10240 \Rightarrow 256 \times 576 \times 640$$



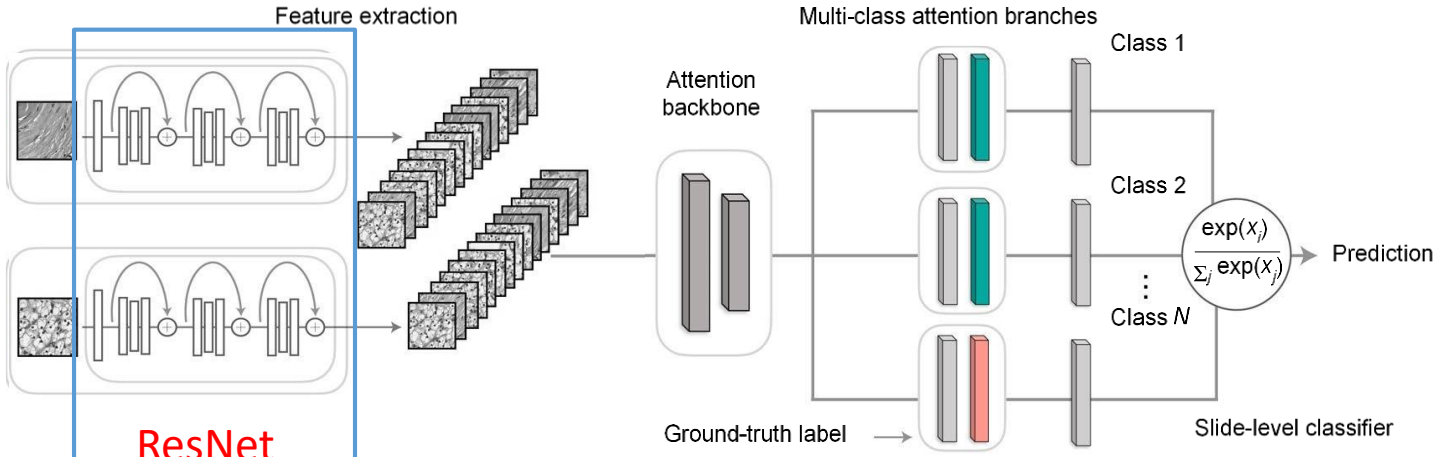
Model

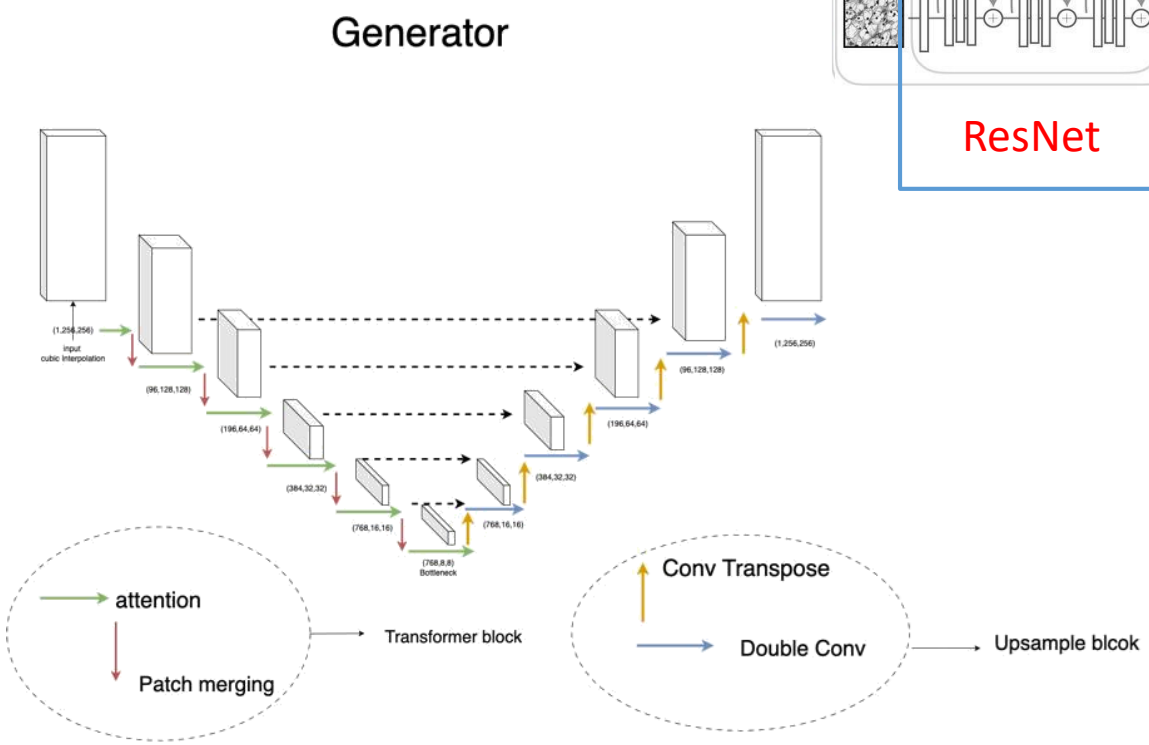
the CLAM conceptual framework and modified

Feature extractor :

Regnet, VGG, MobileNet

Auto Encoder from acratch





Result

CLAM

Image	Method	Augmentation	Normalization	Accuracy	precision	recall	f1-score	Support
Strips	Original CLAM	-	Min-max scaler	0.68	0.71	0.50	0.59	L0 = 54
					0.67	0.83	0.74	L1 = 66
	Reducing CLAM dimension	yes	quantile-normalize: 0.9999	0.71	0.66	0.74	0.70	L0 = 54
					0.76	0.68	0.72	L1 = 66
	Modified RegNet	yes	quantile-normalize: 0.9999	0.71	0.67	0.70	0.68	L0 = 54
					0.75	0.71	0.73	L1 = 66
Stitching	Original CLAM	-	quantile-normalize: 0.9999	0.62	1.00	0.18	0.31	L0 = 11
					0.59	1.00	0.74	L1 = 13
	CLAM & Modified RegNet	yes	quantile-normalize: 0.9999	0.67	0.80	0.36	0.50	L0 = 11
					0.63	0.92	0.75	L1 = 13
	Reducing CLAM dimension & Modified RegNet	-	quantile-normalize: 0.9999	0.67	1.00	0.27	0.43	L0 = 11
					0.62	1.00	0.76	L1 = 13
	CLAM & AutoEncoder	-	quantile-normalize: 0.9999	0.68	0.71	0.45	0.56	L0 = 11
					0.65	0.85	0.73	L1 = 13

Discussion

next steps for the project

Project error/problem:

Overfitting → $\begin{cases} Imbalance\ and\ low\ data \\ One\ tumor\ vs\ others \end{cases}$

Binary Classification:

Goal: Classify data into two distinct classes.

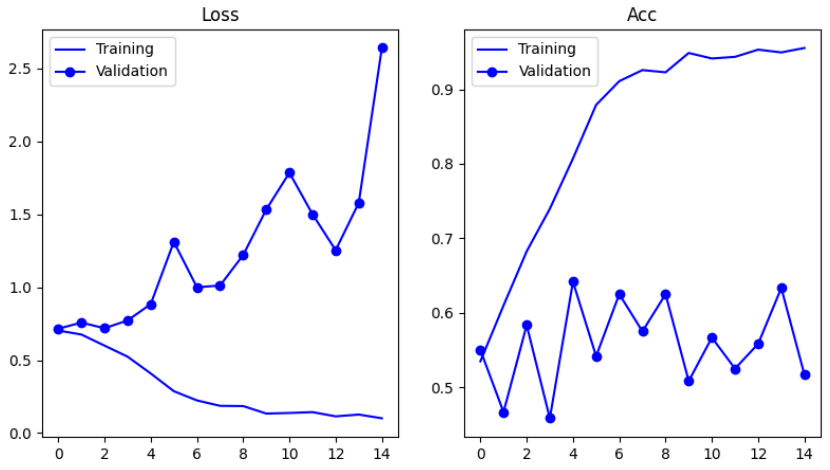
Focus: Distinguishing between the two classes.

One Class vs. Other Classes Classification:

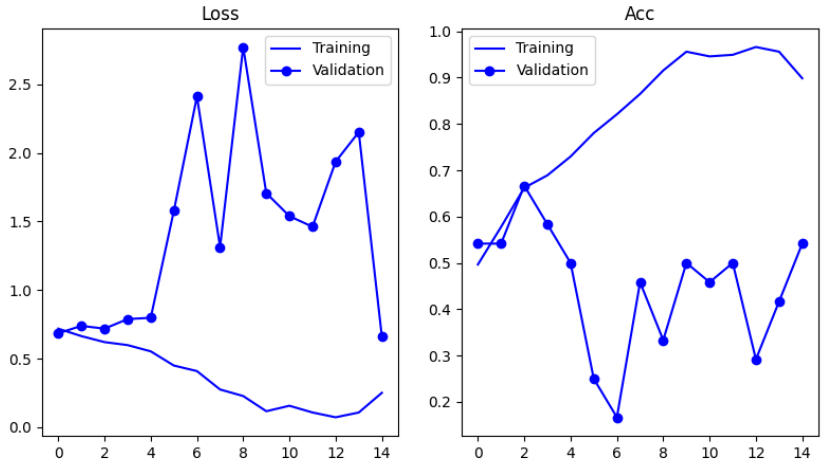
Goal: Distinguish one specific class from all other classes combined.

Focus: Identifying a specific class of interest.

As a result: More dataset is needed



Strips



Stitching


Thank you
for your attention

AutoEncoder