Final Project Presentation

Whole Slide Image Multi-Class Classification

人工智慧機器人碩士學位學程 NM6124012 葉庭均
數據所 RE6111058 黃子軒
數據所 RE6121029 張立勳

Content

- Goal
- Project Overview
- Method
- Model Workflow
- Data
- Results
- Conclusions

Goal

- Automate the diagnostic process, reducing the workload on pathologists and allowing them to focus on more complex cases.
- Reduce variability in diagnosis that might occur due to human factors, leading to more consistent and accurate results.
- Increase the accuracy of model for WSI classification
- Distinguish WSI between rare cancers successfully

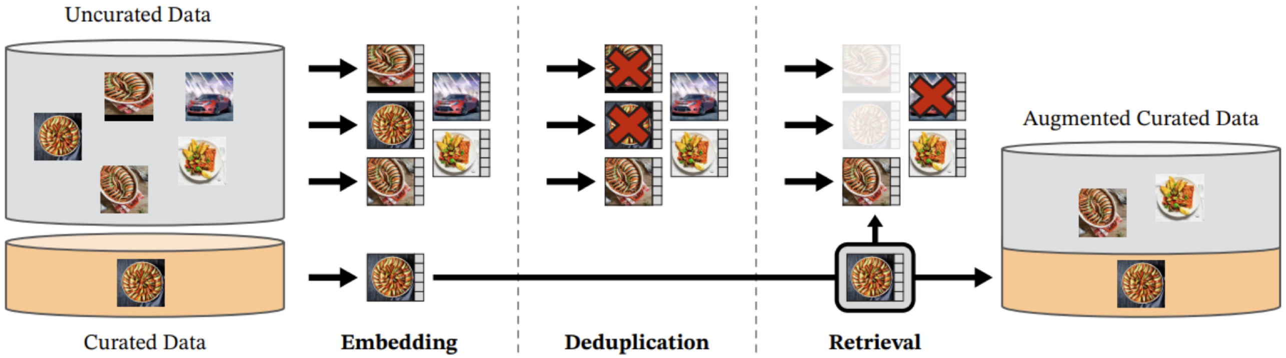
Project Overview

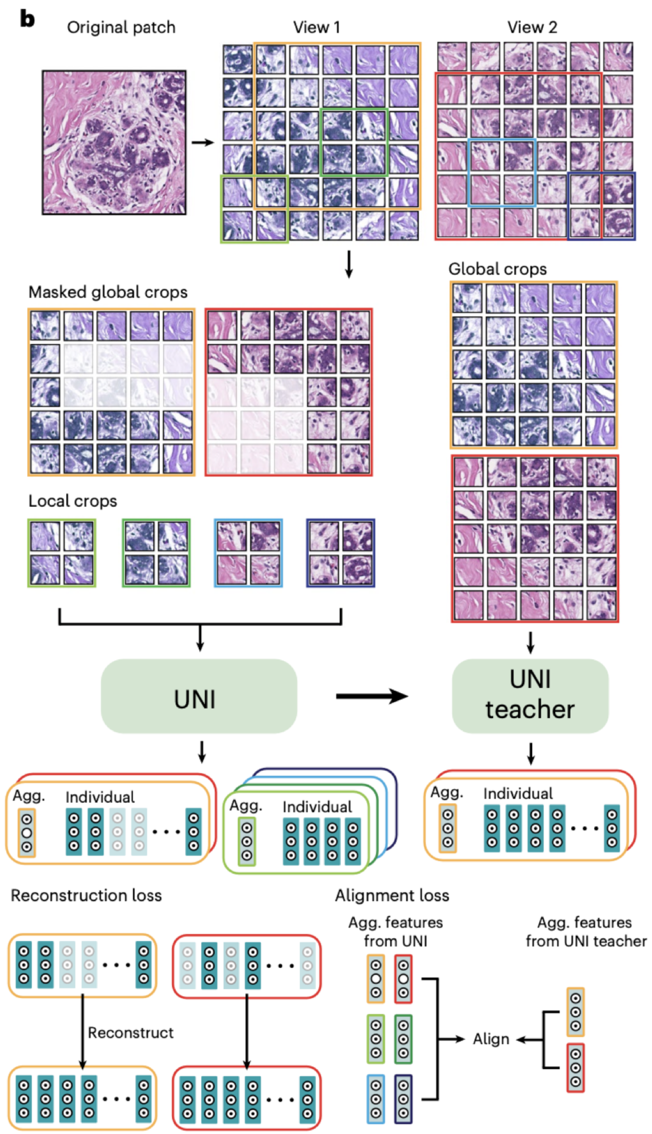
- Baseline model: UNI
- Use the concept of Naive Bayes
- Focus on task: multi-class classification

Method

UNI, a pretrained encoder. (ViT-L)

Pretrained using DINOv2 self-supervised learning approach on dataset Mass-100K.





Method

- Why
  - The overall accuracy of UNI decreases as the number of the labels increases
- How
  - Set the first label as our baseline label
  - Bulid UNI models to compare the baseline label with other labels
  - Calculate the odds of classes in each UNI model
  - Choose the class with the maximum odds

Method

Modified:

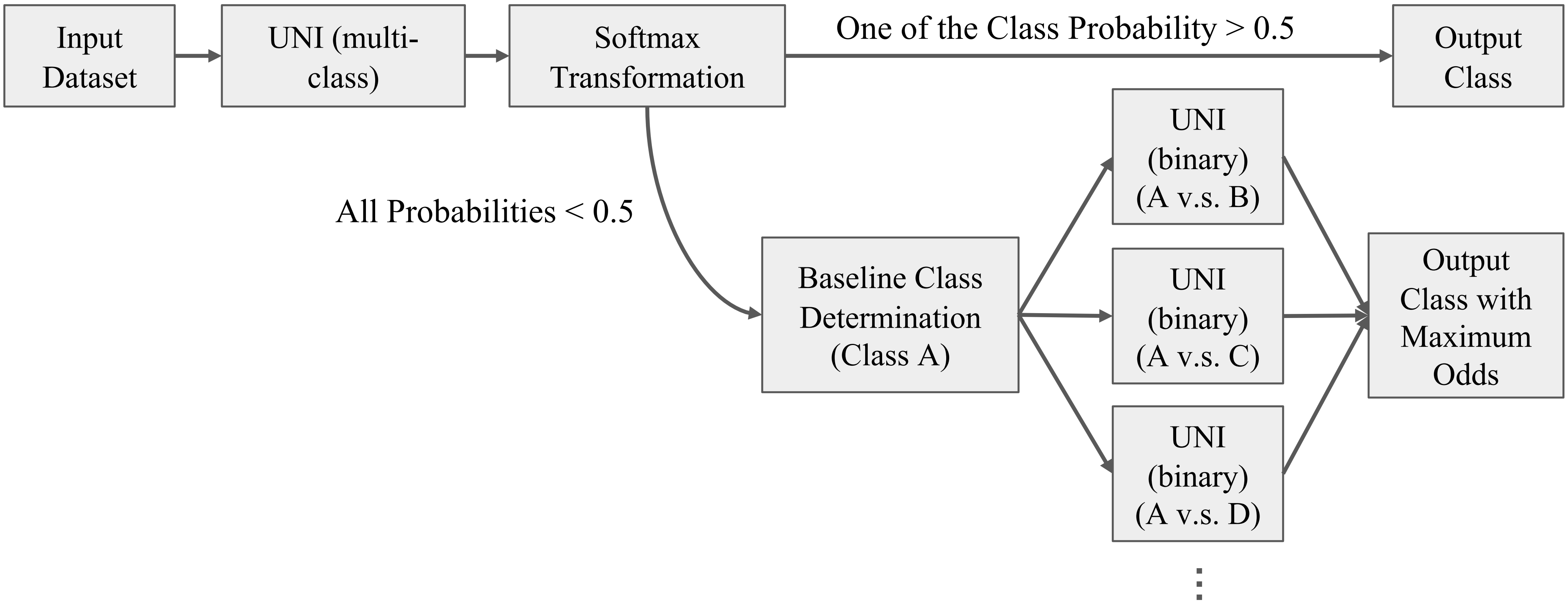
If all of the probabilities in one sample are not more than 0.5, which means there is no dominant class, this sample will be predicted additionally. We will select one class as baseline and train each of the other classes against the baseline class. That is, if we have n class in this case, we will have n-1 trained model.

The formula to determine the output class can be written as:

$$\hat{Y} = argmax \frac{P_i}{P_{1i}}$$

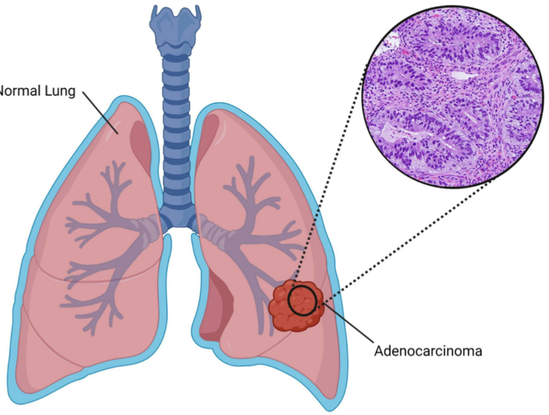
Pi is the probability of class i and P1i is the probability of the baseline class in the specific trained binary classification model.

Model Workflow

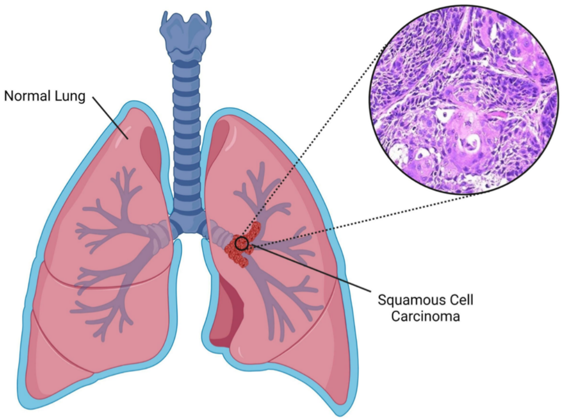


Data

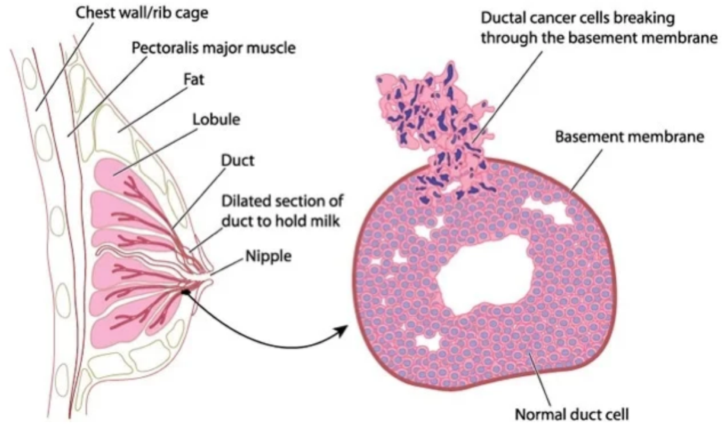
TCGA-LUAD



TCGA-LUSC



TCGA-BRCA



	LUAD	LUSC	BRCA
Train	70	63	70
Test	30	27	30

Results

UNI Classification

	Acc	Balanced Acc	Kappa	Weighted F1
Linear Probe (original)	0.862	0.822	0.772	0.851
Linear Probe (modified)	0.931	0.933	0.897	0.931

Conclusions

- In the original model, the UNI model gets a robust performance in binary classification.
- However, UNI model in multi-class classification has poor performance.
- In the modified module, we take the advantage of UNI model in binary classification and get better performance.
- The more classes we add, the more time it spends.
- The trade-off between computational time and the general applicability of the model is an important topic for discussion.
- The model becomes more robust and discriminative.

Thanks for Your Listening