A semi-supervised solution for mesothelioma classification

Jiang Peichun	s330592
Khattar Reem	s321463
Tallone Samuele	s334046

Politecnico di Torino

Cancer overview

Malignant pleural mesothelioma (MPM) is a rare cancer originating from the mesothelial cells of the pleura. It is strongly associated with asbestos exposure

Epithelioid subtype:

- accounting for approximately 60–80% of cases
- composed of uniform epithelial-like cells and exhibits a tubulopapillary or solid growth pattern.
- better prognosis and higher treatment responsiveness

Sarcomatoid subtype:

- accounting for approximately 10–20% of cases
- more aggressive variant consisting of spindle-shaped cells
- associated with poor therapeutic response and a worse overall prognosis

Biphasic subtype:

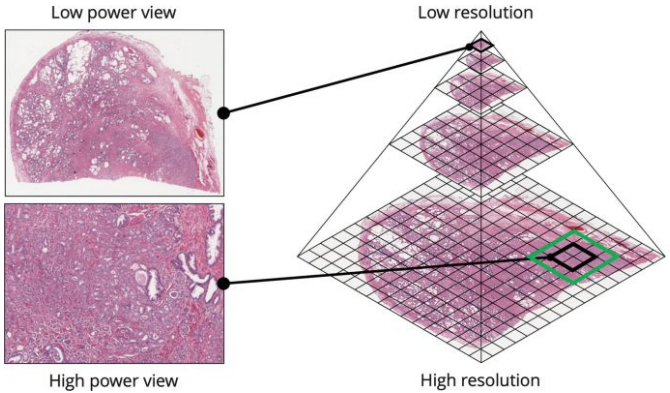
- contains both epithelioid and sarcomatoid components
- the clinical behavior lies between the two
- prognosis worsens as the sarcomatoid component increases

WSI and MIL paradigm

Whole Slide Images (WSI): pyramid structure

WSI stores the same image at multiple resolutions, creating a stack of layers. At the base, there is the full-resolution image with all its detail. As moving up the pyramid, each level contains a progressively lower-resolution version of the same image.

- require significant storage space
- can be efficiently managed and navigated



Multiple Instance Learning (MIL)

- weakly supervised solution
- treating each WSI as a ”bag” of many small image patches, or ”instances”.
- doesn’t require expensive pixel-level annotations
- learn from slide-level labels by identifying the most discriminative instances within the bag

CLAM and PINS paper

Clustering-constrained Attention Multiple Instance Learning (CLAM)

- uses an attention-based pooling function to **aggregate the patch level features** into a slide-level prediction.
- assign importance scores to each patch, allowing the model to **focus on tumor-rich regions** while ignoring background artifacts.
- introduces instance-level clustering, encourages the model to **learn distinct representations for different tissue classes**, improving both accuracy and interpretability.
- able to produce **heatmaps** that allow clinicians to visualize the relevant regions

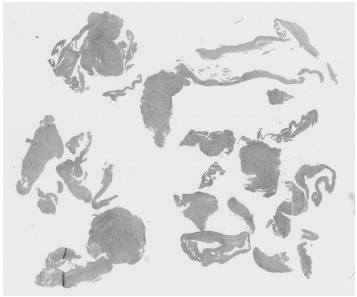
CLAM and PINS paper

Positive INstance Sampling (PINS)

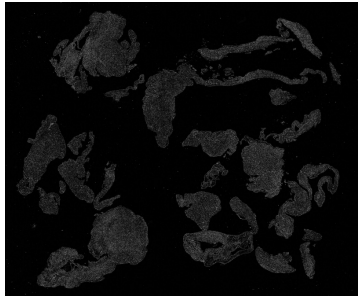
Eastwood et al. (2023) proposed an approach which suggests that manually engineered features, such as nuclear size, cell density, and spatial distribution, can provide a robust numerical signature for cancer subtyping. As It was implemented on a smaller version of the dataset using Tissue Micro-Array (TMA) cores, the novelty of our project:

- Implements a similar dual-path feature strategy on entire WSI.
- Extracting these handcrafted features via QuPath and concatenating them with ResNet50 deep features
- comparison with current state-of-the-art feature extractors

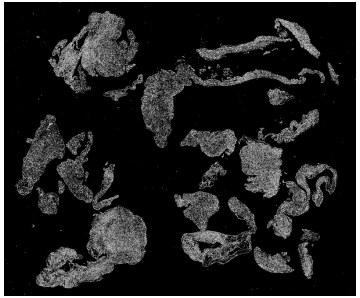
Methodology - Tissue segmentation



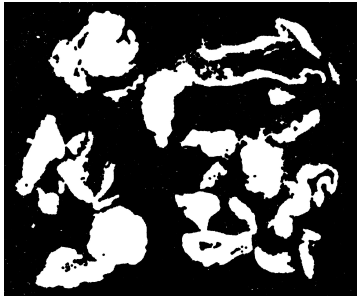
Grayscale



Laplacian



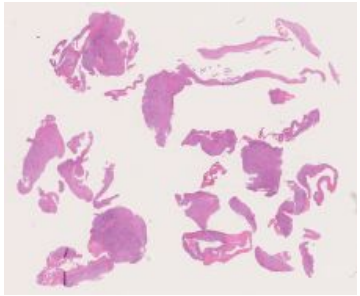
Otsu mask



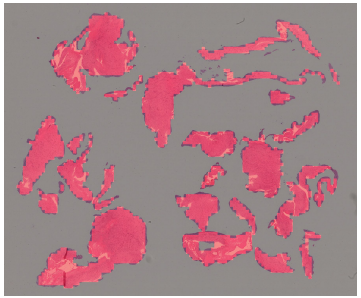
Morph closing

Final mask

Methodology - Patch Extraction



Thumbnail



Overlay

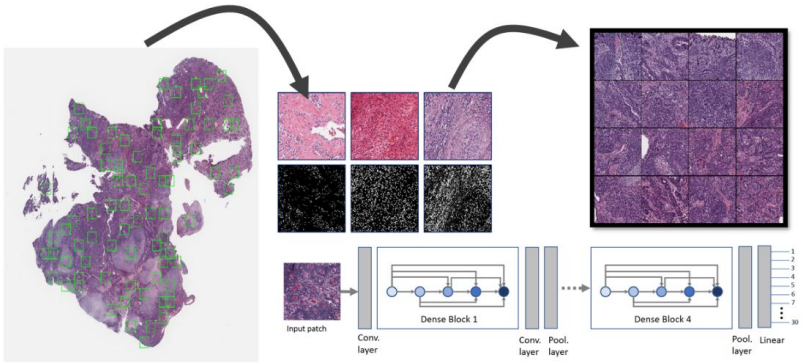
Patches extraction:
Patches were extracted at 512x512 pixels at level 0, meaning the level with the highest magnification.

Overlay creation:
For each slide an overlay was created so that we could visually check the sanity of the patch extraction process

Methodology - Feature Extraction

KimiaNet

- deep CNN based on DenseNet-121 architecture
- trained on histopathology WSIs
- produces discriminative deep features for medical image representation and analysis



Methodology - Feature Extraction

UNI

- uses self supervised learning
- built on Vision Transformer (ViT) architecture
- trained on large-scale pathology datasets
- learns tissue and cellular patterns automatically
- produces embeddings usable for downstream medical tasks

Methodology – Feature Extraction

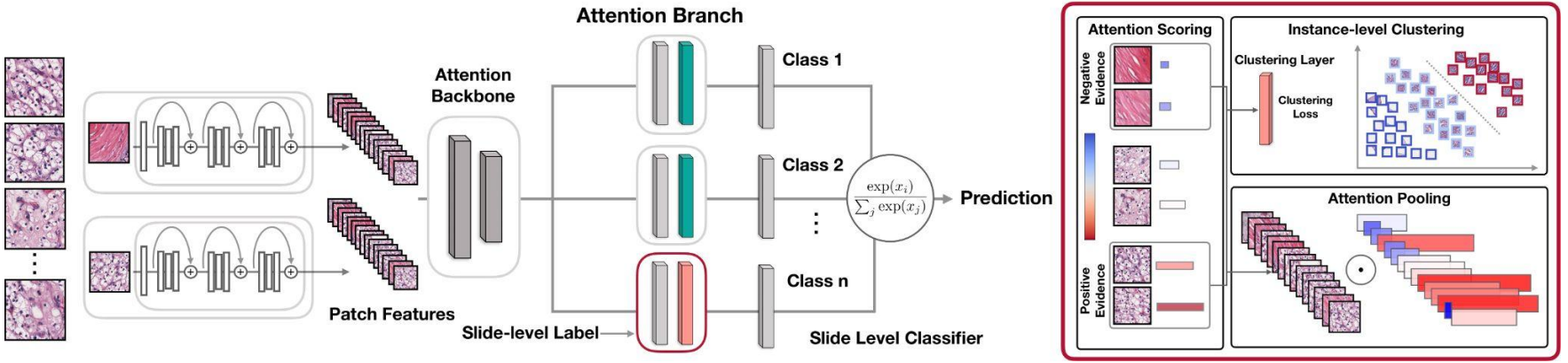
Manual Features

- texture, intensity, spatial, morphological features,... extracted with QuPath
- combined with ResNet50 deep learned features
- encode domain knowledge that may not be easily captured by neural networks
- enrich feature representation and improve model interpretability and performance

Methodology - Prediction Model

CLAM

- Clustering-Constrained Attention Multiple Instance Learning
- designed for WSI analysis using Multiple Instance Learning (MIL)



Training details & Dataset used

- **Learning Rate**
  10^(-5) UNI, 10^(-4) the rest;
- **Epochs**
  20;
- **Weight Decay**
  10^(-5);
- **4-fold cross validation**
- **model configuration**
  CLAM single branch small model configuration;

  8 sample patches for instance-level loss.

Subtype	Total Dataset	Used Dataset
Epithelioid	96	19
Biphasic	19	19
Sarcomatoid	5	4
Total	120	42

Restricted dataset because of class imbalance, poor performance on test runs and complexity of manually extracting features

Results

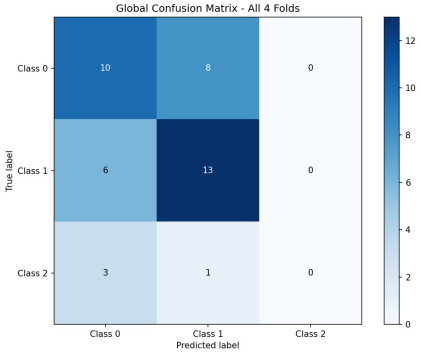
In this project we wanted to evaluate:

1) Contribution of manually engineered features

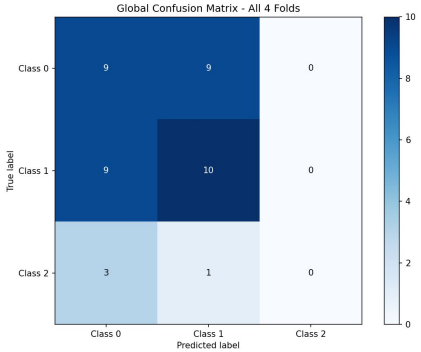
2) How manually engineered features compares with current state of the art base of foundation models:

Feature Extractor	F1-score	Accuracy	AUC
ResNet50	0.516	0.559	0.709
UNI	0.603	6340	0.713
KimiaNet	0.751	0.777	0.869
ResNet50 (M)	0.393	0.464	0.684

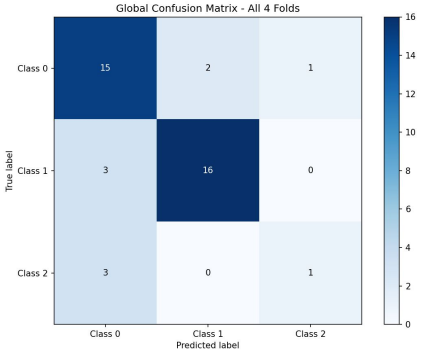
Average of 4 fold cross validation computed on the model that achieved the best validation accuracy for each fold



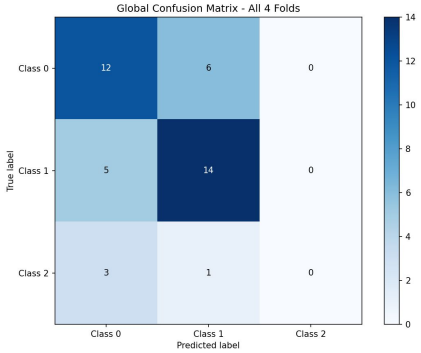
a) ResNet 50



b) ResNet 50 (M)



c) KimiaNet



d) UNI

a) Resnet 50:

Struggles to classify the data correctly, but is to be expected.

b) Resnet 50 (M):

Worst performances than simply using ResNet 50 without manual features.

c) KimiaNet:

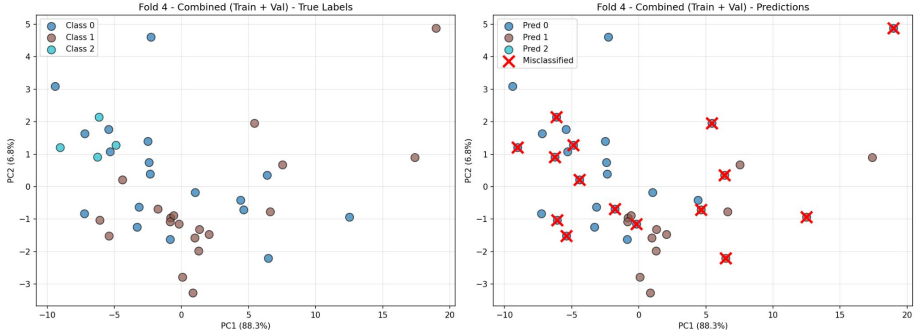
The best performing it is evident that it is able to classify the data the best, it is the only one that manages to correctly classify a sarcomatoid sample.

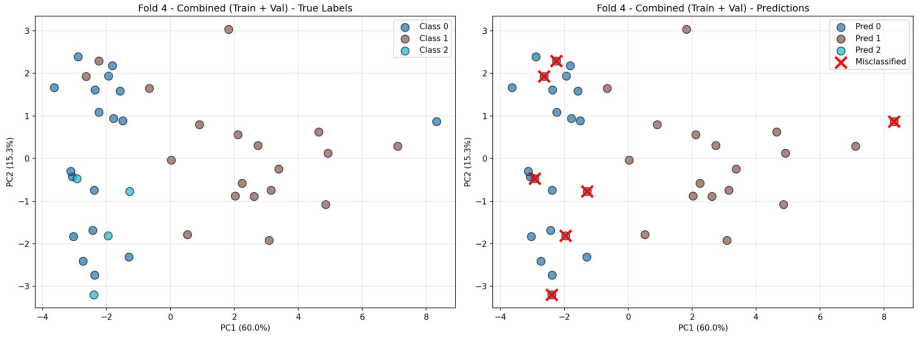
d) UNI:

Not the best feature extractor despite being trained on the largest dataset.

Results - Confusion matrix

PCA visualization

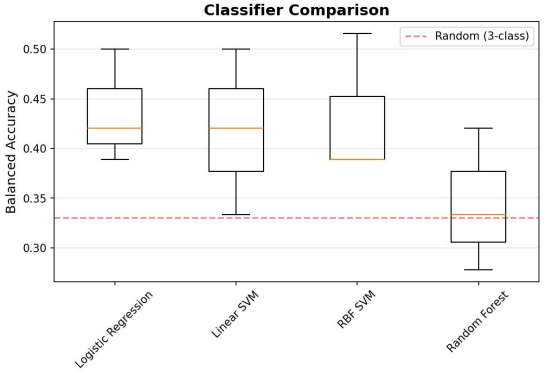


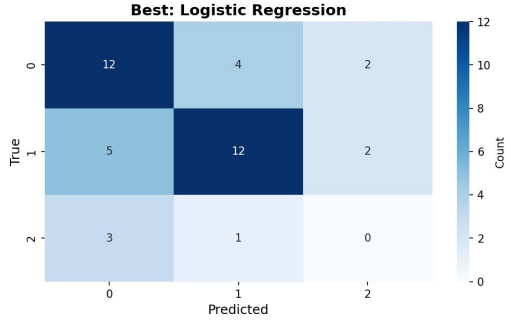


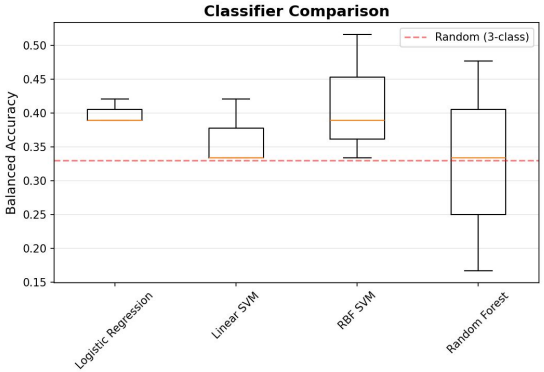
Resnet 50 (M):
From the PCA visualization it is possible to see that internally the features are not well separated leading to confusion when making the prediction.

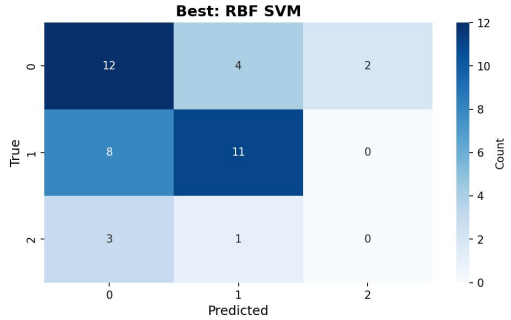
KimiaNet:
On the other end it performs much better since we are able to clearly separate the two main classes, it still struggles with sarcomatoid but that is understandable given the number of WSIs.

Shallow classifiers benchmark



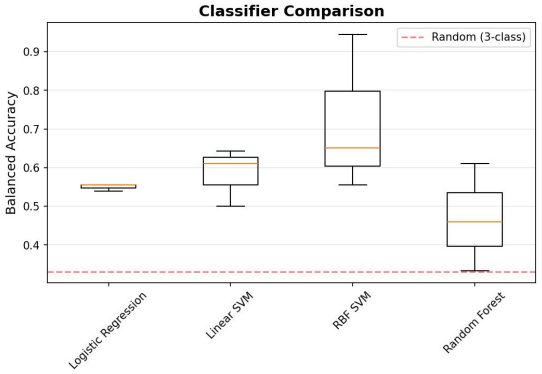


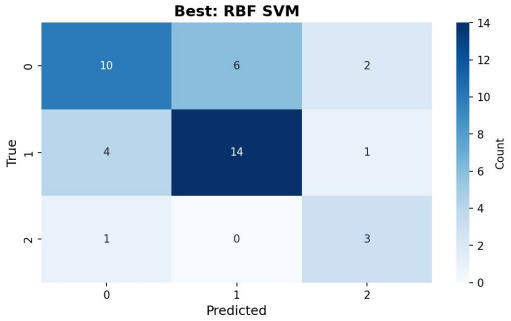




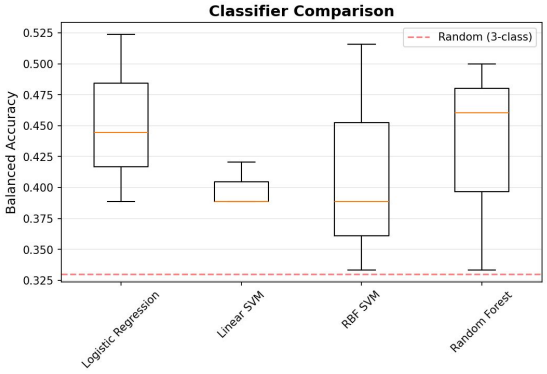
Resnet 50:
Okay performance across different shallow classifiers, this is what we expect from a feature extractor that has not been trained on histopathological data.

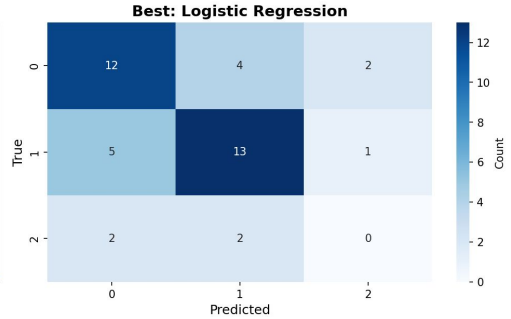
Resnet 50 (M):
The manually extracted features do not help in the classification process, for Random Forest the accuracy is on part with a random prediction.





UNI:
The best performing on shallow models, indicating that CLAM was the problem. Probably it struggles to separate the richer features with this few images.

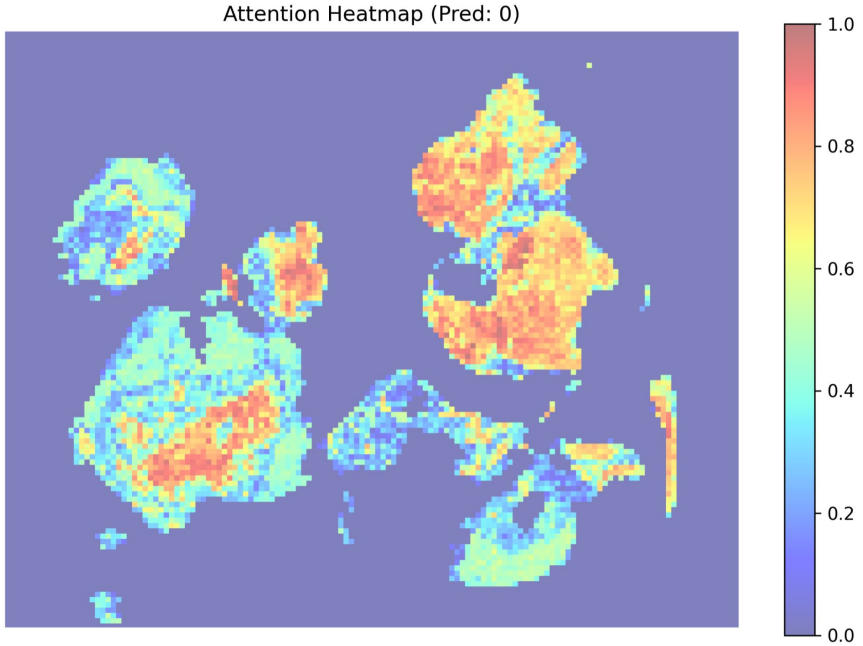




KimiaNet:
It performs okay, and it is better then ResNet 50, but worst than UNI, meaning that the features extracted by last are more informative overall.

Shallow classifiers benchmark

Advantages of CLAM model in real life diagnostics



Heatmap generation

Because of the attention mechanism it is possible to generate a heatmap of the attention weights.

In this way in a real world scenario the pathologist could analyze what areas are important for the classification.

Conclusions

Contribution of manually engineered features:

- The features extracted were not informative
- Lead to lower performances

Causes:

1) TMA cores are smaller and easier to work with, requiring less computational power than standard WSI.
2) Less manual features extracted in respect to the PINS paper
3) Extremely similar pathologies, difficult to separate with traditional methods

How manually engineered features compares with current state of the art base of foundation models:

- Manually extracting features was computationally expensive
- Less straightforward than using a base of foundation model
- Lower performances
- Challenging to understand what features are useful and what are not

Future works

Two branches

Move away from manual feature extraction:

- Improve on current state of the art foundation models that can also be adapted to more task
- Improve current data availability with public datasets

Focus on manual feature extraction:

- Understand and make publicly available what features work best with each pathology
- Introduce libraries to make the feature extraction process more simple
- Developing more robust methods to aggregate features extracted manually and extracted using DL.

Thank you for your attention

Politecnico di Torino