STAT390: CMIL project

All team meeting

22nd April 2025

Problem statement & Data

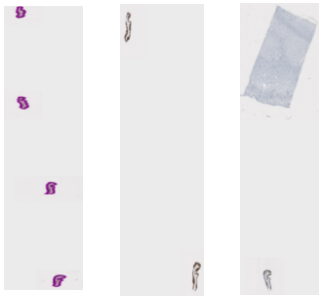
- The objective of the project is to develop a machine learning model to accurately classify ***Conjunctival melanocytic intraepithelial lesions (C-MIL)*** as per the ***WHO 2022 classification system***

- The motivation behind the project is to provide a ***consistent, and accurate grading*** of C-MIL so that the most appropriate management plan for the patient can be developed

- We have CMIL data for 105 cases from 97 patients from the following 3 ocular oncology/pathology centers:
  - Liverpool University Hospitals NHS Foundation Trust (Liverpool; cases from 2018 to 2021),
  - Royal Hallamshire Hospital (Sheffield; from 2011 to 2021), and
  - Rigshospitalet (Copenhagen; from 1996 to 2021)

Source: https://www.sciencedirect.com/science/article/pii/S0023683723002246

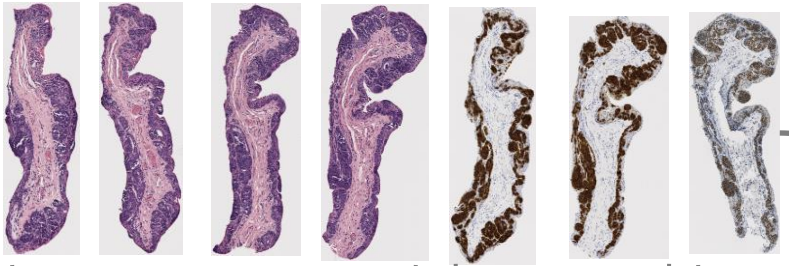
Slice extraction for annotation (example: case 1)



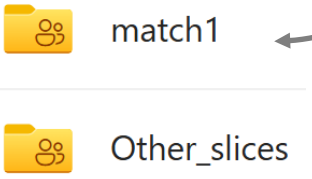
H&E Melan-A Sox-10

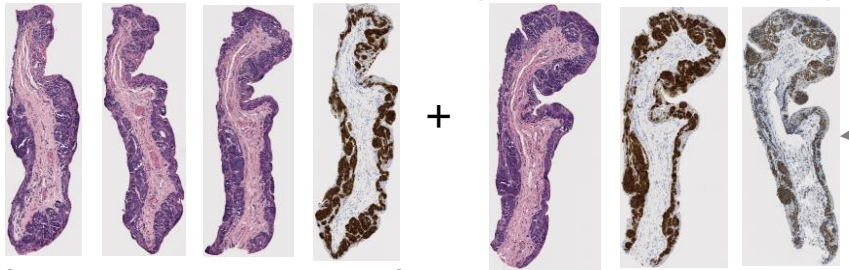
WSI (Whole slide images)

Extract tissue slices



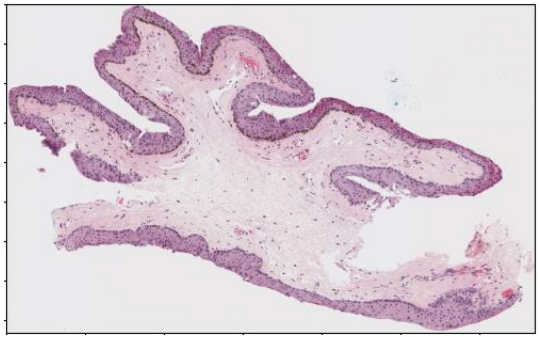
H&E Melan-A Sox-10



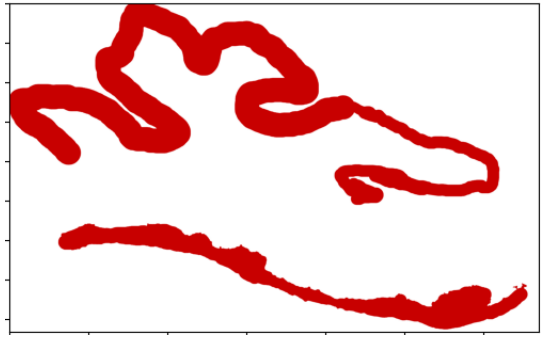


Unmatched slices Matched slices

Patching algorithm (example: case 85)

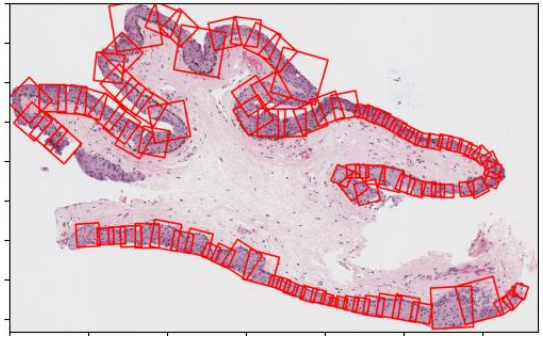


1. Tissue slice

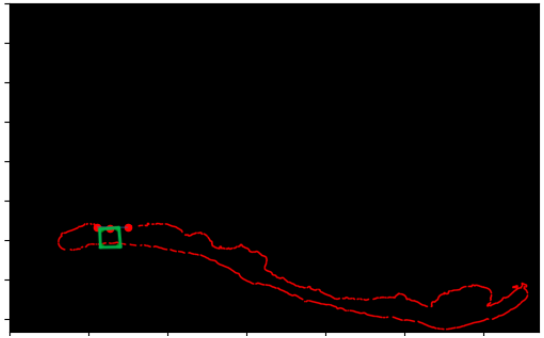


2. Epithelium mask

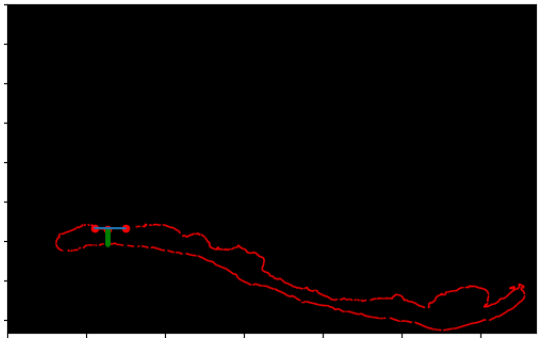
3. Epithelium contour



6. Result



5. Patch



4. Point on contour with tangent & normal

Preliminary modeling result

- We have developed a preliminary modeling framework to train and test the model

- To evaluate our modeling framework, we performed a random train-test split at the patch level, since we anticipate high classification accuracy at this level. This serves as a preliminary test, prior to evaluating performance at the case level, which is our intended deployment level

- As expected, we got a high prediction accuracy of 90.5% when we trained our model on 600 patches of similar size, and tested it on 400 patches

ML model

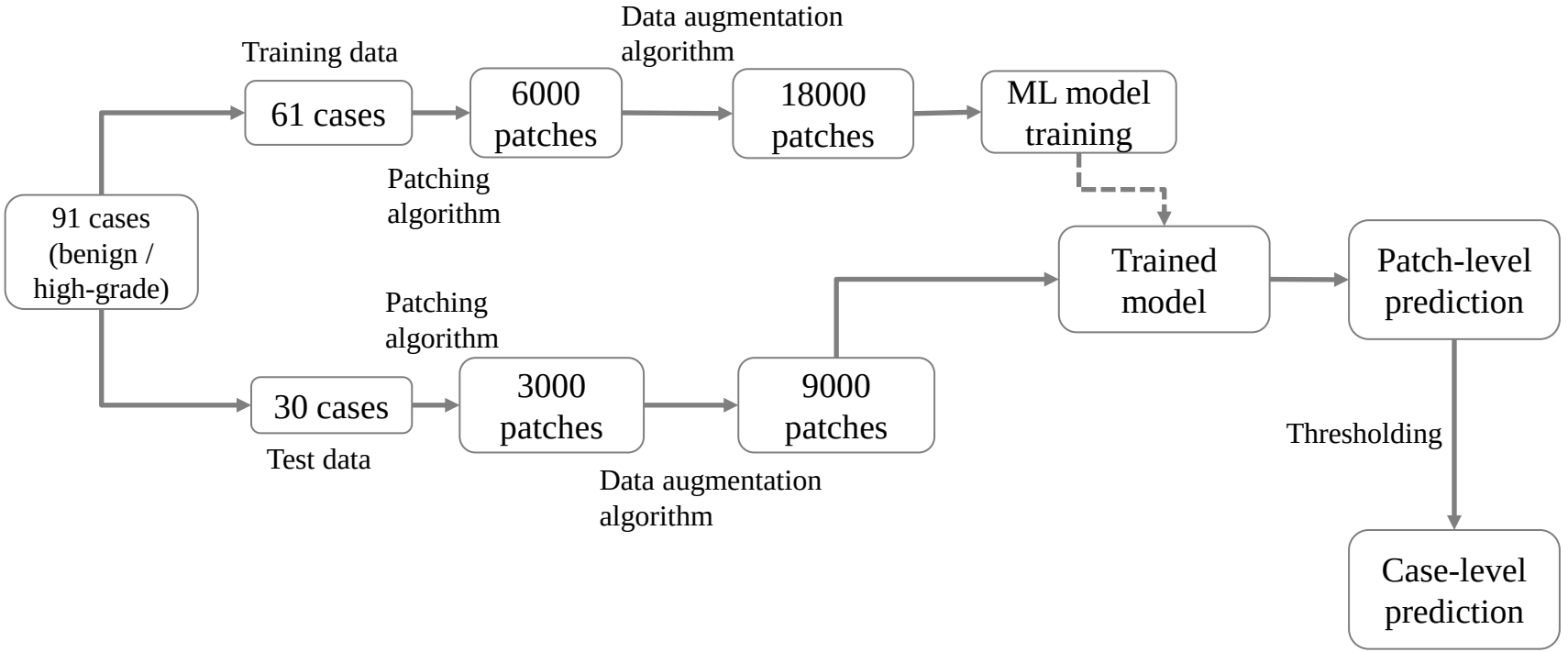
- We are currently trying out models to distinguish between benign and high-grade lesions
- The dataset consists of 91 cases:
  - 28 benign
  - 63 high-grade
- We assume each lesion in a case is either entirely benign or high-grade
- Approximately 10,000 image patches have been extracted from the 91 cases.
- Data augmentation techniques (flipping, rotation, noise addition) will be applied until no further performance benefit is observed.

ML model

- We are testing several commonly used models for histopathology images:
  - VGG12
  - ResNet50
  - EfficientNetB3

- Model tuning is not our current priority, as models will need to be re-tuned once low-grade annotated patches are available.
- The focus at this stage is to enhance the patch extraction algorithm to obtain better quality patches.

ML model - workflow



Current challenges and next steps

- A couple of current challenges that we face in training the model are:
  - Training with patches of different sizes
  - Combining information from all the three stains

- We have some ideas in the next couple of slides to address these challenges

Next steps: Patches of different sizes

- There are three ways of training the model with patches of different sizes:
  - **Naïve method:** Resize the patches to the same size before plugging them in the model
  - **Adaptive pooling:** Select important pixels from different parts of an image, using windows that adjust in size depending on image size
  - **Advanced methods:** (e.g., FCN – Unet architecture): Reduce the number of pixels, while maintaining the patterns in the image
- Currently, our model is using the naïve way of resizing the patches, so that all of them are of the size 256 x 256 before being used as input to the model
- We will use one of the advanced method to ensure that the shape and relative size of cells doesn't change while compressing images
- Also, we intend to discard patches below a certain size threshold

Next steps: Combining information from all stains

- There are two ways of combining information from all stains:
  - Train a single model with patches from all stains
  - Train a separate model for each stain and ensemble the models
- We will test both these approaches and their variations to identify the best approach

Reference

- *https://www.sciencedirect.com/science/article/pii/S0023683723002246*