STAT390: CMIL project

All team meeting

22nd April 2025

Problem statement & Data

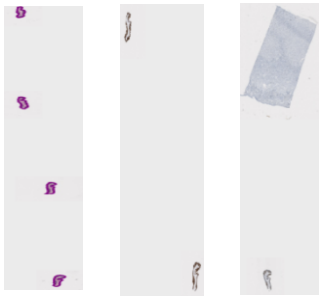
- The objective of the project is to develop a machine learning model to accurately classify ***Conjunctival melanocytic intraepithelial lesions (C-MIL)*** as per the ***WHO 2022 classification system***

- 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 107 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

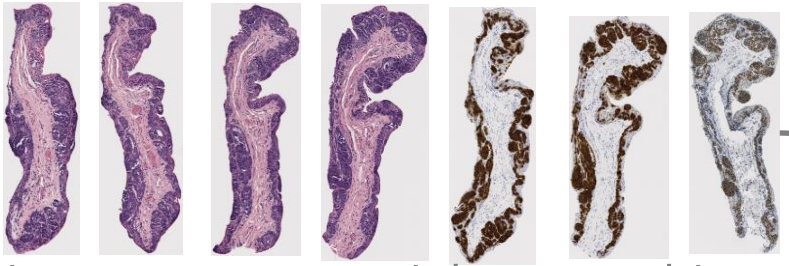
Slice extraction for annotation (example: case 1)



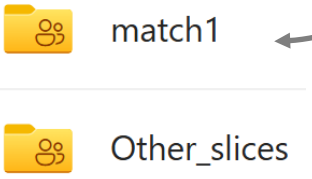
H&E Melan-A Sox-10

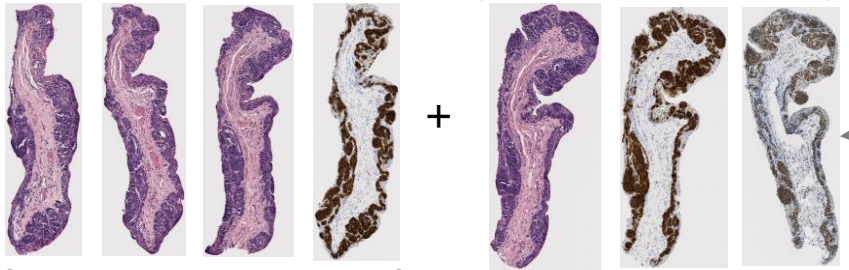
WSI (Whole slide images)

Extract tissue slices



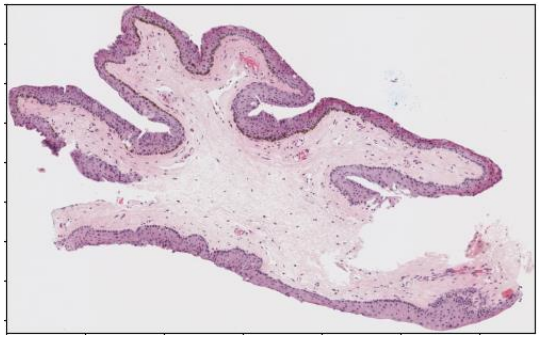
H&E Melan-A Sox-10



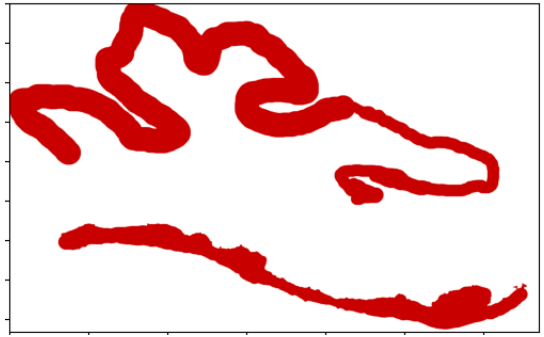


Unmatched slices Matched slices

Patching algorithm (example: case 85)

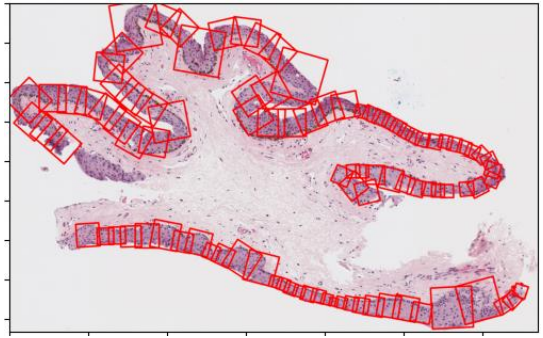


1. Tissue slice

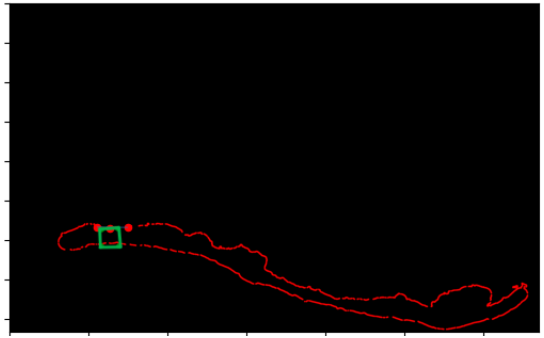


2. Epithelium mask

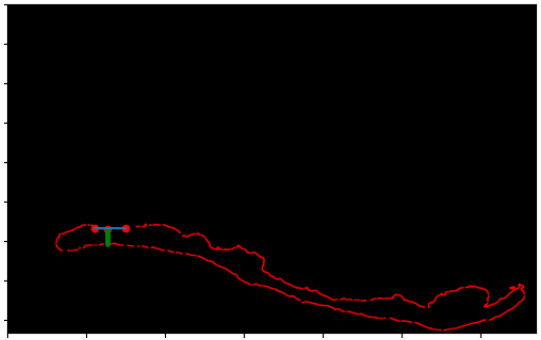
3. Epithelium contour



6. Result



5. Patch



4. Point on contour with tangent & normal

ML model

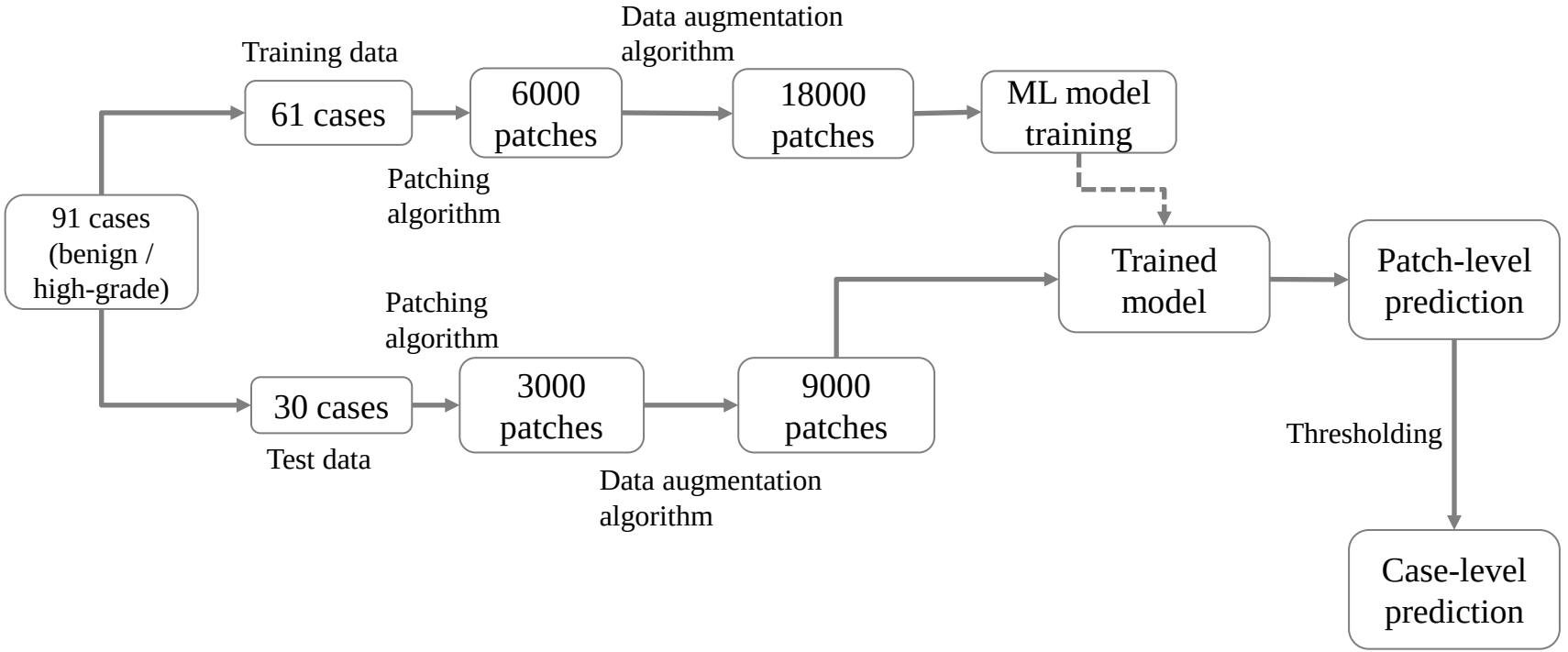
- We are currently trying out models to distinguish between benign and high-grade lesions
- The dataset consists of 91 cases:
  - 28 benign
  - 63 high-grade
- We assume each lesion in a case is either entirely benign or high-grade
- Approximately 10,000 image patches have been extracted from the 91 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



Reference

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