Application of a Deep Learning Classification Model in Anterior Segment Optical Coherence Tomography Images for Improving Limbal Stem Cell Deficiency Diagnosis

SARAH WU [1], Micah Vinet [2,3], William Speier [2,4], Sophie Deng [3]

[1] BIG Summer Program, Institute for Quantitative and Computational Biosciences, Department of Computational & Systems Biology, UCLA
[2] Department of Bioengineering, UCLA [3] Stein Eye Institute, UCLA [4] Department of Radiology, David Geffen School of Medicine, UCLA

Background & Motivation

- **Limbal Stem Cell Deficiency (LSCD)** is a degenerative eye disease caused by loss of corneal epithelial cells, impairing wound healing and vision
- **Anterior segment optical coherence tomography (AS-OCT)** measures corneal epithelial thickness, a key LSCD severity marker, but there are no standardized criteria for grading severity or distinguishing scarred from healthy tissue
- **Goal:** Build a deep learning framework to classify LSCD severities and scarred vs. healthy corneal tissue
- **Approach**: Use a modified InceptionV3 deep learning model on AS-OCT images and integrate radiomic features

Methods & Workflow

Data:

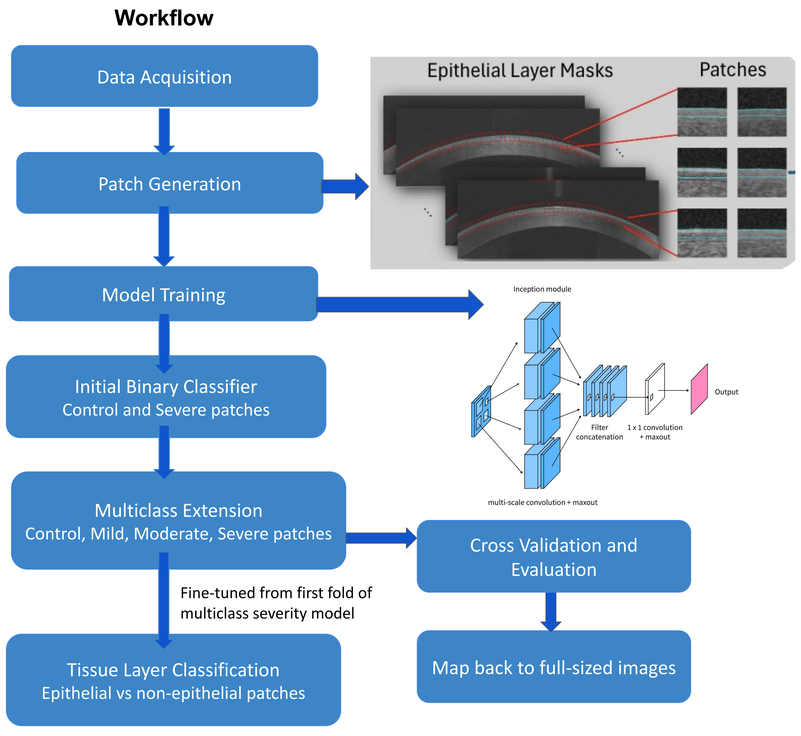
- AS-OCT scans from LSCD patients at the UCLA Stein Eye Institute
  - Severity classification: Control, Mild, Moderate, Severe (from clinic)
  - Tissue classification: Epithelial, Non-epithelial (based on clinical masks)

Patches

- Severity classification:
  - Extracted 100×100 px patches from full-sized AS-OCT images
  - Balanced dataset: 5,588 patches per class (control/mild/moderate/severe)
    - 5-fold cross validation
- Epithelial vs Non-epithelial classification
  - Extracted 75x75 px patches to increase data availability
  - Balanced dataset: 3,298 patches per label (epithelial/nonepithelial)
    - 70% training, 15% testing, 15% validation
  - Fine-tuned from multiclass severity model for binary tissue classification
- Patch-level predictions were mapped back to the corresponding full-sized image with final class assigned by majority voting over patch predictions

Model Development: The InceptionV3 model, pretrained on ImageNet, was integrated with radiomic features of entropy, contrast & homogeneity

Workflow



Results

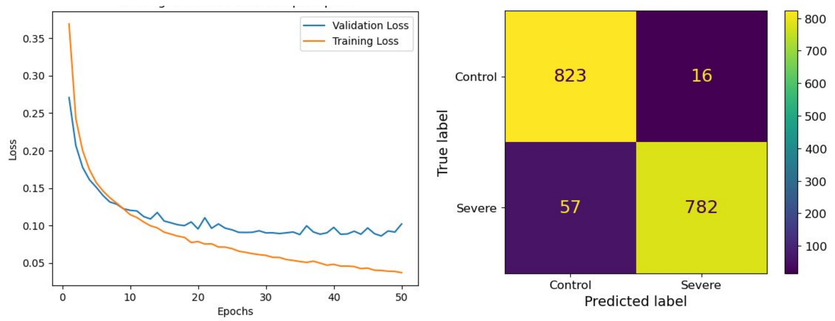


Figure 1. Binary Classification of **severe vs control** image patches over 50 epochs demonstrated converging loss curves and 97.99% precision, 93.21% recall, and a F1 score of 95.54% indicating strong precision–recall balance and effective generalization

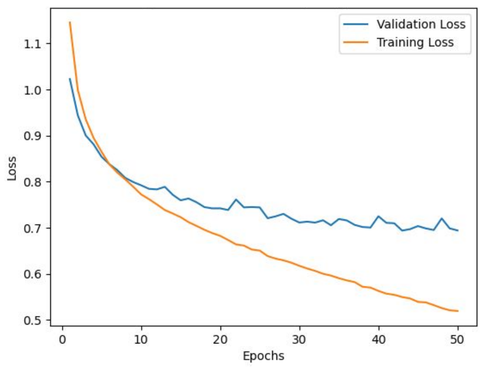


Figure 2: 5-fold cross validation for **multiclass severity classification** showed overfitting tendency at 50 epochs where training loss decreased steadily while validation loss plateaued around ~0.75

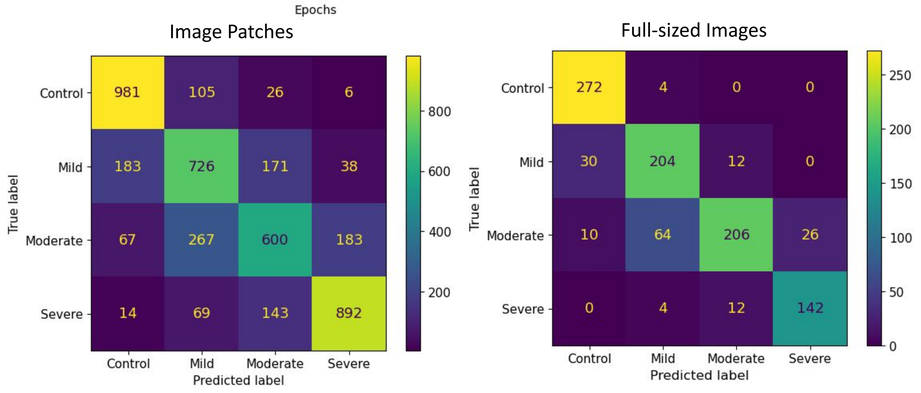


Figure 3: Fold 1 of **multiclass classification** (control vs mild vs moderate vs severe) demonstrated 71.55% accuracy on image patches and 83.57% accuracy on full sized images

- Aggregating patch predictions to assign full-image labels improved classification accuracy
- The model distinguished between control vs severe full-sized images with 100% accuracy and had the most difficulty distinguishing between mild and moderate images

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Accuracy on image patches	71.55%	71.42%	72.51%	72.19%	71.41%
Accuracy of full sized images	83.57%	82.93%	81.95%	81.58%	81.17%

Figure 4. 5-fold cross validation on the entire dataset demonstrated consistently high multiclass severity patch accuracy (~71-73%) and full image accuracy (~81-84%) across all folds

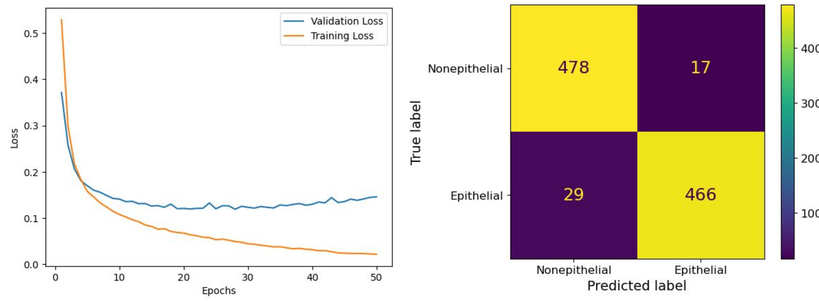


Figure 5. Training and validation loss curves for **epithelial vs. non-epithelial classification** stabilized at ~0.02 and ~0.15, respectively. The fine-tuned model achieved 96.48% precision, 94.14% recall and a 95.3% F1 score on the test set.

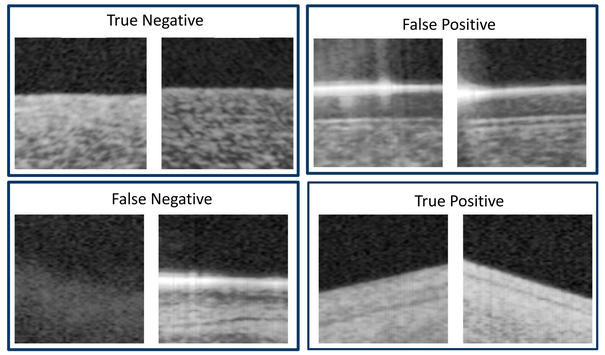


Figure 6: Epithelial layer patches- true negatives (top left), false negatives (bottom left), false positives (top right), true positives (bottom right)

Discussion

- Radiomics feature integration better captured subtle textural differences in AS-OCT images and improved robustness of model
- Smaller 75×75 patches increase data availability and maintain high accuracy, making them well-suited for limited datasets such as scar vs. healthy classification
- In full-image reconstruction, tied patch-level predictions defaulted to the less severe class. Future work could explore alternative tie-breaking strategies, such as confidence-weighted voting
- 5-fold cross-validation yielded stable severity classification accuracy across folds, indicating good model generalizability within the dataset
- High accuracy of epithelial vs. non-epithelial classification shows transfer learning from related AS-OCT tasks is effective and will be applied to scar vs. healthy classification

Conclusions

- Deep learning and radiomics enables high-accuracy LSCD severity grading and tissue classification from AS-OCT images
- This project has potential to minimize diagnostic variability and increase efficiency in LSCD evaluation, reducing workload for opthamologists
- Next steps: Generate and label 75x75 px patches for healthy vs scar classification which will be trained from epithelial/non-epithelial model

Acknowledgements

A huge thank you to Micah Vinet for his unwavering support and mentorship. Thank you to William Speier for his input and guidance and the clinicians at Stein Eye Institute for their data. Finally, thank you to the B.I.G. summer research program for this invaluable research opportunity