Machine Learning for Automated Detection of Transcellular Pores

Gabrielle Humphrey · supervised by Professor Darryl R. Overby with Jacques A. Bertrand

IMPERIAL

LAIDLAW FOUNDATION

Why Schlemm's canal pores matter

Aqueous humour drains from the eye by crossing the inner wall of Schlemm's canal, and it does so through micron-scale transcellular pores. Fewer pores means higher outflow resistance and higher intraocular pressure, the one modifiable risk factor in glaucoma. Pore density is reduced in glaucomatous eyes, so counting pores reliably is a prerequisite for studying how the disease and its treatments act on the outflow pathway.

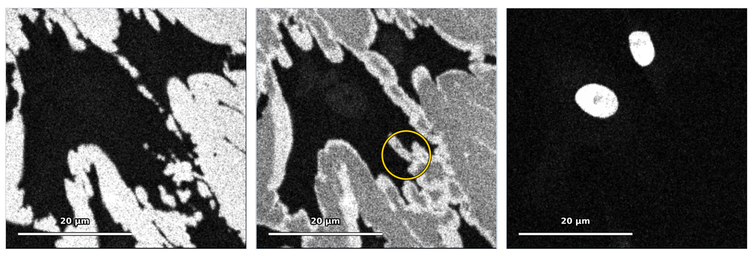

Fig 1. The same field in all three channels. A tracer applied before stretch labels the exposed substrate, so cells read as dark silhouettes. A second tracer applied after stretch turns each transcellular pore into a bright spot. One is circled.

The bottleneck

Pores are currently found by hand in Fiji/ImageJ: an observer outlines every cell in the field, then searches each one for pores. A single field is 25 megapixels and holds several hundred cells, so the work is slow and hard to scale.

The aim of this project was to reduce that field to a set of candidate pores a human can verify, keeping the observer's judgement where it is needed and removing the part that is only visual search.

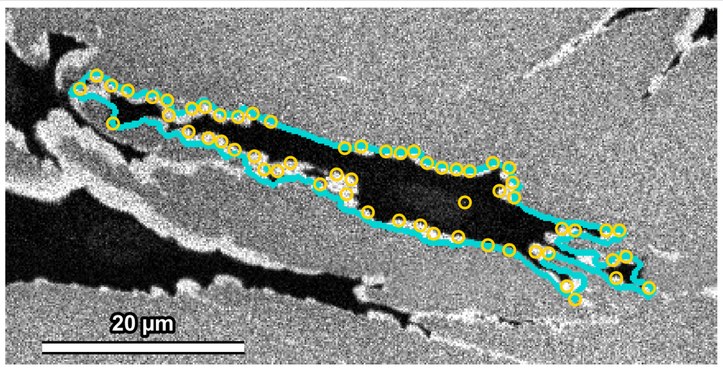

Fig 2. Candidate points proposed inside one segmented cell. Every circle is a spot the observer has to judge, and at the default minimum separation of 8 px this cell alone carries 60.

Pipeline overview

Cells are segmented from the pre-stretch channel with a nucleus-seeded distance-transform watershed. Candidates are then detected as local intensity maxima in the post-stretch channel, restricted to the segmented masks with a 10 px edge tolerance, and presented for human review. Confirmed pores feed the measurements of density, size and position.

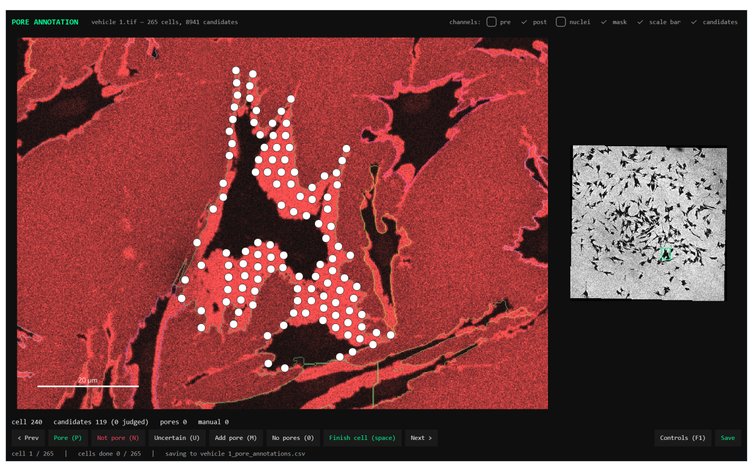

Fig 3. The review interface. Candidates are presented cell by cell to confirm or reject. Reviewed candidates are also training data for a classifier that would rank future candidates automatically, which remains future work.

Cell segmentation

- Pre-stretch channel smoothed with a Gaussian filter; half-Otsu threshold gives the cell mask.
- Nuclei provide one seed per cell.
- Watershed run on the distance transform, so touching spindle-shaped cells separate by shape.
- Edge-touching cells removed, because their area and pore counts are truncated.

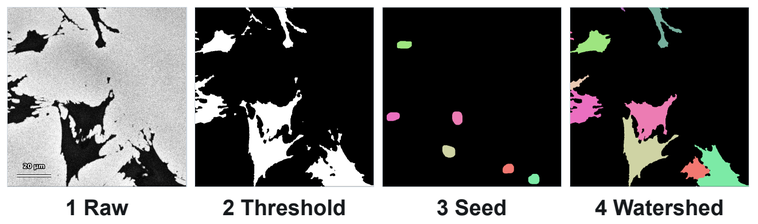

Fig 4. One field of view carried through the four stages of segmentation.

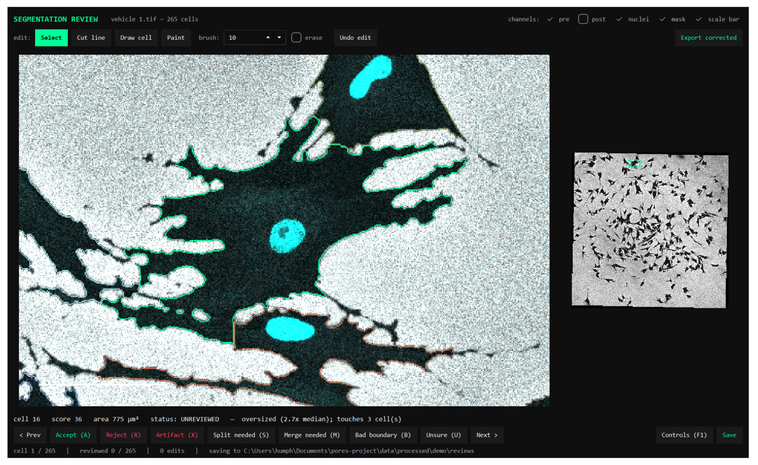

Fig 5. Segmentation output: automatic cell boundaries and nucleus seeds over the pre-stretch image.

Validation against manual tracing

Segmentation was checked against two independent hand tracings covering 605 cells in two fields. Automatic and manual outlines are paired through traced nucleus centroids, so the matching has no tunable parameter and cannot be adjusted to flatter the result. Median overlap was 0.934 in the first field and 0.916 in the second.

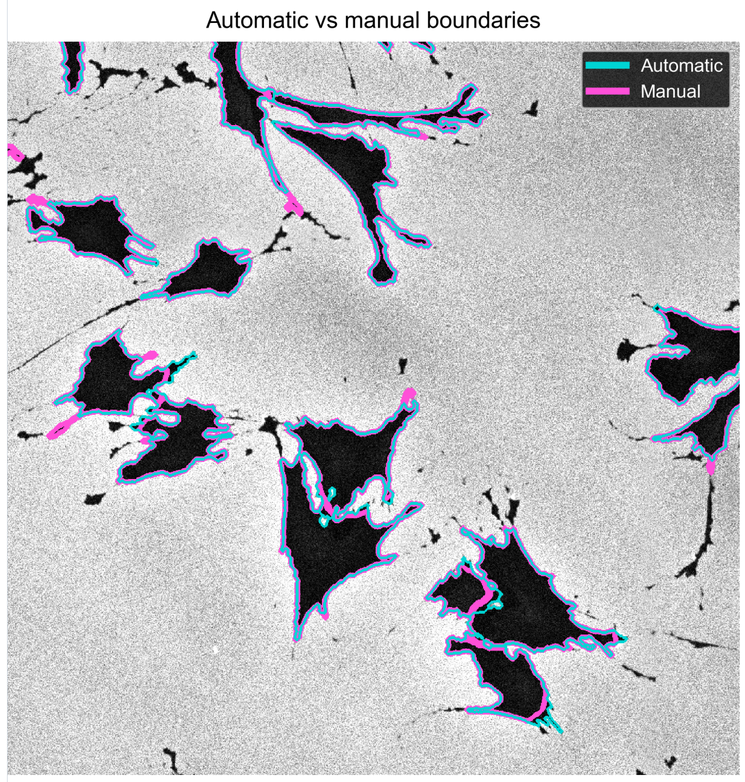

Fig 6. Automatic (cyan) and manual (magenta) boundaries across a cluster of touching cells, the case the method was built to handle. Outlines are paired through nucleus centroids, so the match has no tunable parameter.

Metric	Field 1	Field 2
Matched cells	247	358
Median IoU	**0.934**	**0.916**
Mean IoU	0.911	0.817
Area correlation, r	0.984	0.856

Agreement remained strong at the median, indicating that lower mean performance in the second field reflected a tail of difficult cells rather than a broad failure of the method.

100% recall was almost meaningless

At its default operating point the detector proposed 19,944 candidates per field and recovered all 70 manual pore marks, a recall of 100%. Scattering the same number of points uniformly at random inside the same cell masks recovers 94.4% of them. The detector's real advantage over chance was about three pores.

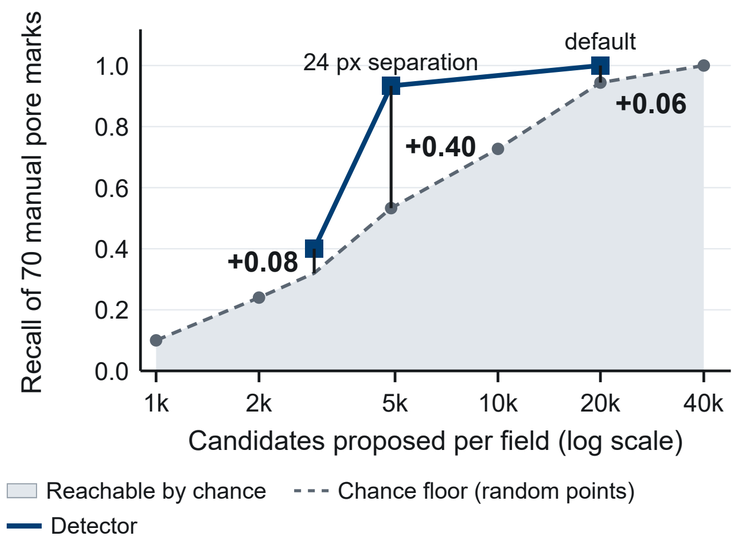

Fig 7. Pores are rare and candidate density is high, so recall saturates. Evaluating against a null model of random in-mask points was essential. At a budget of 2,000 candidates the shipped detector sat **below** chance.

The real improvement: background subtraction

Feature importance suggested the pre-stretch channel carried information the detector was ignoring, so subtracting it from the post-stretch channel looked promising, and at first it appeared to help. Re-tested against the chance floor at a matched budget of 2,000 candidates, that gain proved fragile: it did not survive testing in separate quadrants of the field.

The improvement that did survive was morphological. A white top-hat (the image minus its own opening) removes everything larger than the filter and leaves the small bright spots a pore looks like. The limitation was background subtraction, not channel choice.

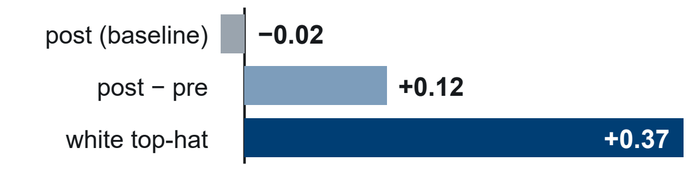

Fig 8. Margin over the chance floor at a matched budget of 2,000 candidates. The top-hat gain stayed positive at every radius from 6 to 24 px and in all four quadrants of the field.

Conclusions and next steps

Segmentation. Automatic masks agreed strongly with manual tracing across 605 cells.

Detection. Raw recall was misleading, because random candidates achieved similarly high recall.

Improvement. Morphological background subtraction produced the most robust gain over chance.

Foundation. The framework establishes a validated foundation for scalable pore analysis.

Next: expand the pore ground truth, train the candidate classifier on it, then quantify pore density, size and distribution. No normal-versus-glaucomatous comparison is reported here: analysis was blind to condition throughout.

REFERENCES

Johnson et al. *IOVS* 2002;43:2950. Braakman et al. *Exp Eye Res* 2014;127:224. Overby et al. *PNAS* 2014;111:13876. Rodrigues, PhD thesis, Imperial College London, 2022. Otsu, *IEEE Trans SMC* 1979;9:62. Schindelin et al. *Nat Methods* 2012;9:676.

ACKNOWLEDGEMENTS

Professor Darryl R. Overby and Jacques A. Bertrand provided the microscopy data and the manual segmentations and pore annotations against which every validation number here is measured. Supported by the Laidlaw Scholars Leadership and Research Programme.