Machine Learning for Automated Detection of Transcellular Pores

Gabrielle Humphrey

Overby Laboratory, Department of Bioengineering
Imperial College London

Research supervised by Professor Darryl R. Overby,
with Jacques A. Bertrand

August 2026

IMPERIAL

LAIDLAW FOUNDATION

Abstract

Micron-scale pores in the endothelium of Schlemm's canal provide a route for aqueous humour to leave the eye, and reduced pore formation has been linked to glaucoma. Fluorescent permeability assays can now visualise these pores across large microscopy fields, but the analysis remains slow: a masked observer outlines every cell by hand and inspects it for pores, which takes days to weeks per cell strain. I developed a Python pipeline that segments individual Schlemm's canal endothelial cells, proposes candidate pore locations, and routes them to human confirmation. Automatic segmentation agreed closely with two independent manual segmentations across 605 matched cells, at a median intersection over union of 0.934 and 0.916. Candidate detection initially appeared to achieve perfect recall. Uniformly random in-cell points, however, recovered 95.6% of manual pore annotations at a comparable candidate density, which showed that raw recall was saturated. Re-analysis against this chance floor changed the conclusion twice over: adding the pre-stretch channel was not the improvement it appeared to be, and morphological background subtraction of the post-stretch image was. Scalable pore counting is feasible, but only when detection performance is measured against the probability of succeeding by chance.

1. Introduction

Every minute, aqueous humour is produced inside the eye and must leave at nearly the same rate. If drainage resistance rises, intraocular pressure can rise with it, and with it the risk of glaucomatous damage. Glaucoma is the leading cause of irreversible blindness worldwide and was projected to affect around 111 million people by 2040 [1]. The final continuous cellular barrier in the conventional outflow pathway is the inner-wall endothelium of Schlemm's canal [2, 3].

Aqueous humour is thought to cross this layer largely through micron-scale openings that pass either through individual endothelial cells or between neighbouring cells [4, 5]. These pores are small and transient, but their density appears to matter. After accounting for fixation conditions, Johnson and colleagues reported fewer than one fifth as many inner-wall pores in glaucomatous as in normal human eyes [6], and pore density correlates with outflow facility [7]. Later work

showed that glaucomatous Schlemm's canal cells are stiffer and less able to form pores under biomechanical loading, which links pore formation to a disease-relevant cellular phenotype [8, 9]. Pore formation is therefore a plausible therapeutic target, and counting pores is how anyone would test whether a candidate drug moves it.

Studying that phenotype requires a way to count pores repeatedly and at scale. Historically this has meant painstaking microscopy and manual classification. In stretch experiments, pore-like structures were re-imaged at high magnification and independently marked by multiple masked observers [10]. More recent fluorescent permeability assays make the problem more tractable experimentally: cells are grown on a substrate coated with biotinylated gelatin, and a fluorescent tracer binds the substrate only where it can pass through a cell, turning transcellular pores into bright spots that can be imaged over much larger fields [11]. A tracer applied before stretch labels the exposed substrate around each cell, so cells appear as dark silhouettes on a bright background. A second tracer of a different colour, applied after stretch, marks the pores. A third channel labels nuclei.

The downstream analysis still requires manual counting, and that creates an unusual bottleneck: the experiment can generate images faster than a person can reliably turn them into pore measurements. Pores in these images are dim, small and rare, and the same images contain bright staining artefacts at leaky intercellular junctions that are far brighter than any pore [12]. At present a masked observer segments each cell by hand, including hand-drawing separation lines between cells that touch, then inspects each cell for pores. That takes days to weeks per cell strain. It is also subjective, which makes counts from different observers difficult to compare.

With the support of the Laidlaw Foundation, I spent the summer working on that bottleneck in the Overby Laboratory. As a first-year undergraduate Mathematics student, I had studied a range of subjects including pure and applied mathematics, statistics and computational methods, but before beginning this placement I had little prior knowledge of glaucoma, fluorescence microscopy or image analysis. The project began with a straightforward goal: automate the analysis with machine learning. In practice it became more fundamental. An earlier PhD project in the laboratory had already established the right shape for a solution, in which an interest point detector proposes candidate pores at high recall and a classifier ranks them so a human confirms a shortlist instead of searching whole images by brute force [12]. That design is sound and I kept

it. But before a classifier could be trusted, the cells had to be segmented correctly, candidate pores had to be localised, and the detector had to be evaluated in a way that did not reward it simply for proposing thousands of points.

The central finding is methodological, and it directly changes the biology that can be extracted from these images. A detector can report nearly perfect recall and still contain almost no useful information, if the candidate density is high enough that a pore is likely to sit near some candidate by chance. Once I introduced a matched random baseline, two conclusions I had already reached disappeared, and a different improvement emerged in their place.

2. Methods

2.1 Imaging and data

Images are three-channel fluorescence micrographs of cultured human Schlemm's canal endothelial cells: a pre-stretch tracer channel, a post-stretch tracer channel, and a nuclei channel. Two fields were available with independent manual cell segmentations, referred to here as Field 1 (4,980 by 5,010 px) and Field 2 (4,984 by 6,502 px). Both are 8-bit. Channel order is not consistent across acquisition sessions, so it is recorded per image rather than assumed. The pixel scale is 0.1625 μm/px, defined once in the codebase; a unit test fails if any module hardcodes a second copy of it.

2.2 Cell segmentation

Segmentation runs on the pre-stretch channel, where cells are dark on a bright substrate. The channel is smoothed with a Gaussian filter (sigma = 1.0) and thresholded by keeping pixels darker than half the Otsu value [13]. Otsu's method picks a brightness cut-off automatically by finding the value that best separates a two-peaked intensity histogram. Halving it is a deliberately conservative choice, made so that neighbouring cells do not bleed into one another. A binary dilation with a disk of radius 1 then closes hairline gaps.

Nucleus seeds come from the nuclei channel, smoothed at sigma = 2.0, thresholded by Otsu and labelled by connected components, giving one seed per nucleus. Cells are separated by a watershed transform seeded from those nuclei and constrained to the cell mask. A watershed

treats the image as a landscape and floods it outward from seed points, drawing a boundary wherever two floods meet. Here the landscape is the negative Euclidean distance transform, in which each pixel's height is its distance from the nearest background pixel. Cells touching the image border are dropped, because their area and pore counts are truncated.

2.3 Pore candidate detection and measurement

Candidate detection thresholds the post-stretch channel at a percentile of in-cell intensity, then applies local maximum finding restricted to the cell masks. Defaults are the 65th percentile and a minimum separation of 8 px between peaks. The separation is the real control on output volume: moving it from 4 to 40 px moves the candidate count from 48,648 to 2,436 on a single field, while the percentile is close to inert by comparison.

Pore size is measured by half-maximum thresholding, so that a bright pore and a faint pore of the same physical size measure the same rather than size tracking exposure. Local background is estimated from a ring rather than the whole window, because a pore large enough to fill the window would otherwise contaminate its own background estimate and be reported as no pore at all. Pores are assigned to a cell within a tolerance of 10 px (1.63 μm) of the cell boundary. That tolerance is not a workaround. Pores form where the cell is thinnest, which is at its edge, and on the annotated field 59 of the 74 manual pore marks sit outside the automatic mask by a median of 2 px.

2.4 Validating segmentation against manual tracing

Automatic and manual segmentations are paired through traced nucleus centroids. Each centroid falls inside exactly one cell in either mask, which pairs the two without a distance heuristic or a greedy overlap search. A nucleus is skipped, with the reason recorded, when it falls out of bounds, when it lands on background in either mask, or when the pairing would reuse a cell already matched.

Agreement is reported as intersection over union, or IoU: the area where the two outlines overlap, divided by the total area they cover between them. An IoU of 1.0 means the outlines are identical and 0 means they do not touch. It is computed over the union of the two bounding boxes rather than the full field. The point of anchoring on nuclei is that this matching is

threshold-free and therefore cannot be tuned to improve the number it reports, which is not true of distance-based or overlap-based matching. Manual segmentations and pore annotations were produced in Fiji [14].

2.5 Evaluating detection against a null model

Candidate detection is deliberately high recall, which means recall on its own can be bought by proposing more candidates. Detector recall was therefore evaluated against a null model: points scattered uniformly at random inside the same cell masks, scored by exactly the same matching procedure, at matched candidate density. I refer to the recall those random points achieve as the chance floor. The match radius is 15 px (2.4 μm), the value used for region-of-interest labelling in the laboratory previously. Cell masks cover 21.1% of the field. Margin over the floor, rather than recall itself, is reported throughout.

Scoring functions were compared at matched candidate budgets rather than at matched parameters, so that no variant could win by simply producing more candidates. Two robustness checks were applied to anything that appeared to help: a sweep of its free parameter, to distinguish a plateau from one lucky value, and a four-quadrant spatial split with budgets scaled by mask area, to check the effect is not driven by one region of one image.

2.6 Software, reproducibility and blinding

The pipeline installs as a Python package with four graphical entry points and six batch subcommands. Segmentation methods are registered behind a documented extension point, so a new method, including a learned one, can be dropped in and benchmarked against the same reference. The validated method is frozen and pinned by a regression test, so the quoted agreement figure cannot drift silently. Of 581 tests, 571 run without any private data, which means the code is testable by someone who does not hold the images. Every number in this paper is re-runnable from a single command line, which matters because a disputed figure can then be recomputed rather than defended from memory.

All analysis was performed blind to experimental condition. No comparison between normal and glaucomatous cell strains is reported here. That is a deliberate choice rather than an omission,

and it is why this paper reports a validated method and descriptive measurements rather than a group difference.

3. Results

3.1 Segmentation became reliable only after the algorithm changed

The first obstacle was not pore classification. It was defining the cell at all. Heavy smoothing followed by an intensity-based watershed produced masks that looked plausible at a glance but merged closely apposed cells and erased thin processes. The intensity-based variant scores only 0.62 against manual tracing, because the smoothing needed to suppress noise also destroys the gradients the flood relies on. Once segmentation was changed to light smoothing followed by a distance-transform watershed, which separates touching spindle-shaped cells by their shape instead, the automatic masks agreed closely with independent manual tracing. That is the automated equivalent of the separation lines a human draws by hand.

Across 247 matched cells in Field 1, mean IoU was 0.911 and median IoU was 0.934. Automatic and manual cell area were strongly correlated ($r = 0.984$) with a mean bias of only +5.6 μm^2. In Field 2, 358 cells were matched, and the mean was lower at 0.817 while the median stayed high at 0.916. Across both fields, 605 cells were matched one to one.

Table 1. Automatic segmentation agrees with two independent manual tracings at a median IoU above 0.91. Agreement between automatic and manual cell segmentation across both annotated fields.

Metric	Field 1	Field 2
Matched cells	247	358
Mean IoU	0.911	0.817
Median IoU	0.934	0.916
IoU > 0.5	98.4%	91.3%
IoU > 0.7	96.4%	78.8%
Area correlation, r	0.984	0.856

Metric	Field 1	Field 2
Area bias	+5.6 µm²	+1.1 µm²

The gap between mean and median on Field 2 matters, and reporting only one of them would misdescribe the result. A mean of 0.817 against 0.911 reads as a large drop in quality. Medians of 0.916 against 0.934 are nearly identical. The difference is a tail of badly matched cells rather than a uniform reduction in accuracy, and the matching diagnostics support that reading directly: Field 2's nucleus anchors were derived from its own nuclei channel rather than traced by hand, and 61 of its nuclei landed on manual-mask background against 6 on Field 1.

3.2 The first detector comparison was underpowered

The original detector used the post-stretch channel alone, even though an earlier classifier analysis assigned more feature importance to the pre-stretch channel (0.60 against 0.37). That suggested an obvious question: was candidate generation ignoring its best signal? I compared six scoring images at matched candidate budgets, namely post, pre, post plus pre, post minus pre, and the pixelwise minimum and maximum.

The first result appeared clean. None of the alternatives beat post alone. The problem was that the experiment was answering the wrong question. Detection had been restricted to pixels inside the segmented masks, but only 15 of the 74 manual pore clicks fell strictly inside those masks, because pores form at the cell margin and the mask stops at the boundary. With n = 15, one pore changed recall by 6.7 percentage points, so differences between the leading methods amounted to a single annotated event. The apparent negative result did not carry enough information to support any conclusion.

3.3 Fixing the power reversed the conclusion, and a null model reversed it again

The project already used a 10 px tolerance for assigning edge-adjacent pores to cells. Applying the same tolerance to the evaluation brought 70 of the 74 manual pore locations into scope. Repeating the comparison gave the opposite answer. The post minus pre difference image reached full recall with 14,467 candidates, while post alone required 24,899. Read literally, that was a 42% reduction in reviewer workload.

An underpowered test does not fail quietly. It returns a confident, specific and wrong answer.

Before treating the new result as real, I asked a simpler question: how often would a random point be close enough to a pore to count as a hit? The answer changed the project. Cell masks cover 21.1% of the field, and the 15 px match radius is large relative to the density of proposed points. At roughly 25,000 random in-mask candidates, 95.6% of manual pores were recovered by chance, and at 40,000 candidates random points reached full recall (Figure 1). At the detector's high-density operating point, perfect recall was almost guaranteed.

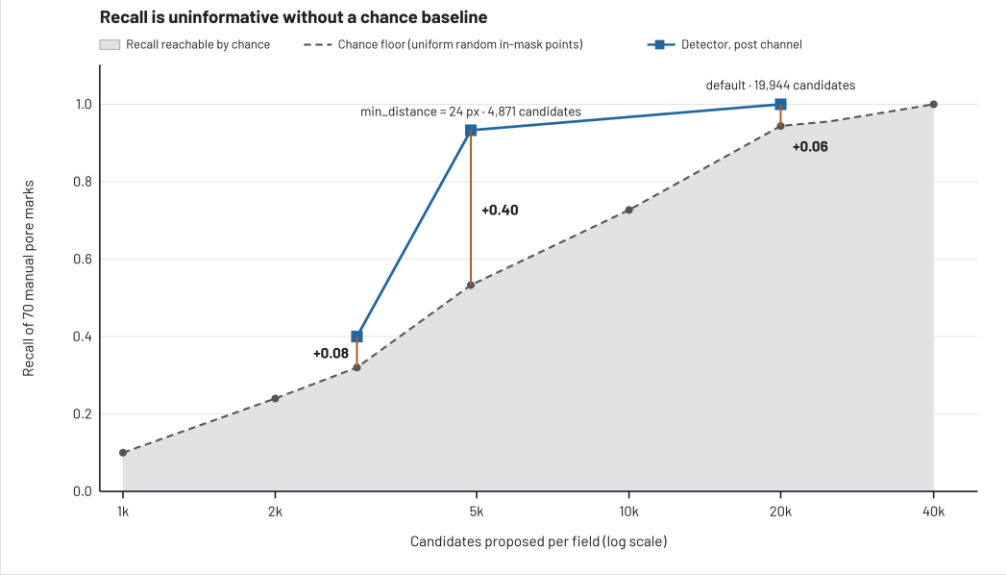

Figure 1. Detector recall converges onto the chance floor well before the default operating point is reached. *Recall against 70 manual pore marks, plotted against the number of candidates proposed per field. The shaded region is what uniformly random points inside the same cell masks achieve at the same candidate density. Vertical bars mark the detector's margin over the floor at three candidate budgets. At the default operating point the detector recovers every manual pore mark, but random points recover 0.944 of them, so the margin over chance is about three pores.*

Recomputing performance as margin over chance changed the picture. The apparent post minus pre advantage shrank to about +14.6 percentage points above the matched random baseline at 90% recall. It was also spatially unstable. A four-quadrant split gave gains of +31%, minus 39%, +50% and +15%. Because one quadrant reversed outright, I treated the effect as suggestive rather than established and discarded the 42% claim.

The detector's own performance also looked much less impressive. At the default it produces 19,944 candidates, recovers all 70 marks, and beats chance by 0.056. Raising the minimum peak separation to 24 px produces 4,871 candidates, recovers 0.933, and beats chance by 0.400. The default is not the best operating point. It is the point at which the metric stops being able to say anything.

3.4 The real improvement was morphological background subtraction

The weak post minus pre gain still needed an explanation. My working hypothesis was that the pre-stretch channel was not contributing pore information at all. Subtracting it from post removes whatever the two channels share, which is mostly uneven illumination and autofluorescence, so the difference image may simply be a crude background estimate. If that were true, a purpose-built background correction on post alone should do better.

That prediction held. At a budget of 2,000 candidates the uncorrected post detector recovered 0.235 of manual pores against a random baseline of about 0.26, placing it slightly below chance. Post minus pre reached 0.382, a margin of about +0.12. A white tophat on post with a 12 px structuring element reached 0.634, a margin of +0.37. A white tophat is the morphological version of background subtraction: the image minus its own opening, an operation that erases features larger than the filter and leaves behind exactly the small bright spots a pore looks like. A Gaussian high-pass control produced only a small gain, which shows the effect was not simply more high-pass filtering. The local morphological background estimate is what mattered.

Table 2. Detection at a matched budget of 2,000 candidates.

Scoring image	Recall	Margin over chance
post (baseline)	0.235	−0.02
post minus pre	0.382	+0.12
white tophat, r = 12 px	0.634	+0.37
Gaussian high-pass, sigma = 12	0.339	+0.08

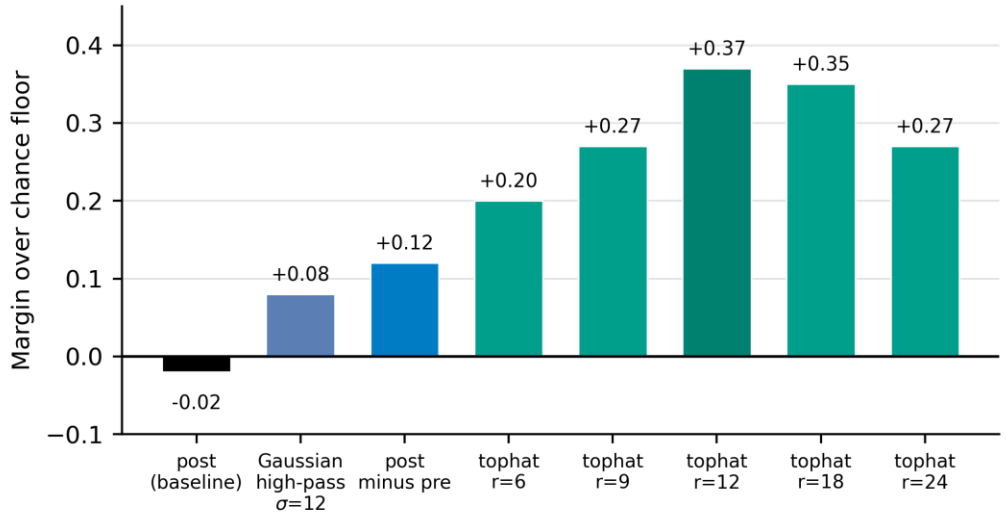

Figure 2. Morphological background subtraction gains far more over chance than any combination of the two fluorescence channels. *Margin over the chance floor at a matched budget of 2,000 candidates, by scoring function. The untouched post channel sits at or slightly below chance. The Gaussian high-pass control gains little, which locates the effect in the morphological operation rather than in high-pass filtering generally.*

The improvement also survived the checks that had eliminated the earlier result. Every tophat radius from 6 to 24 px beat the baseline at 2,000 candidates, with a broad maximum around 12 to 18 px rather than a single lucky value. In the same four-quadrant split the tophat gain stayed positive in every region, at +0.53, +0.52, +0.27 and +0.34. The method is implemented as an option but left disabled by default, because one pore-annotated field is not enough evidence to lock a new operating point into a pipeline other people will quote.

3.5 The pre-stretch channel held useful information, but in the wrong place

A second contradiction helped clarify what the pre-stretch channel is actually good for. Features measured at the human pore click separated confirmed pores from the detector's own false positives well: the ratio of pre-channel core intensity to outer-ring intensity gave an area under the ROC curve of 0.864, where 0.5 is chance. Yet the same feature barely helped when candidates were ranked at the detector's own local maximum. Reaching half the pores took 6,837 reviews, against roughly 12,450 for a random order and the roughly 3,250 an AUC of 0.864 implies.

The reason was localisation. The detector peak sat a median 5.7 px from the human click, while the feature's core zone extends only 4 px from its centre. Measuring at the detector peak therefore read part of the surrounding cell rather than the pore. The same feature fell from 0.864 at the click to 0.644 at the peak. Re-centring within a 7 px disc recovered it to 0.778. As a control, the equivalent post feature sits near chance at both positions (0.501 and 0.559), which is what makes the drop in pre attributable to position rather than to measurement noise.

Implementing the re-centring exposed two further defects that the reasoning had not. A first run reported pores moving a median of 7.1 px within a 7 px radius, which is impossible for a disc, and revealed a square search window. Constraining it properly surfaced a second problem: on a flat patch every pixel ties, and the search was breaking ties by scan order instead of staying put. Correcting both improved the result rather than reducing it, from 0.729 to 0.778, because constraining the search makes it land more truly on the pore. The pre channel is useful for describing a correctly localised pore. It is not a better first-pass detection image.

3.6 Cell contraction between stretch states

The pipeline also measures a quantity the laboratory currently records by hand as retracted area. Comparing each cell's footprint in the pre-stretch and post-stretch channels, matched through its nucleus, gives a median contraction of 42.0% (interquartile range 11.5 percentage points, n = 228 nucleus-paired cells). Sweeping the segmentation threshold from 0.40 to 0.60 of Otsu moves that estimate by 1.8 percentage points against a 42% effect, so it is not an artefact of the threshold parameter.

It may still carry a systematic bias, and I would not present it as a settled biological number. The threshold sweep moves both channels together, so it rules out the parameter but not the weaker contrast of the post channel. Cell-to-background separation is 164 grey levels in pre (25.8 against 189.4) but only 62 in post (91.1 against 152.9), which makes the post mask the less reliable of the two.

3.7 What the pipeline costs a reviewer

The metric that decides whether this tool gets used is not recall but candidates per real pore. At the default operating point a reviewer judges 285 spots per pore found. Raising the minimum

peak separation to 24 px brings that to 83 while retaining 0.933 recall. Combining the tophat with a separation of 40 px brings it to 32. Against whole-image visual search of a 25-megapixel field all three are an improvement, and the difference between 285 and 32 is the difference between a review session someone will finish and one they will not.

4. Discussion

The most important result of this project is that a high recall value is not, by itself, evidence that a rare-object detector is useful. In this dataset the candidate generator could report perfect recall while random points recovered almost the same pores. The failure came from geometry, not from a coding bug. Cell masks occupy a finite area, each manual pore is credited if any candidate lands within a non-trivial radius, and thousands of candidates are proposed. Under those conditions a fraction of hits are geometric coincidences, and that fraction grows with the candidate count until the metric saturates.

This matters beyond bookkeeping, because attributing an effect to the wrong cause misdirects the next piece of work. Had the 42% been accepted at face value, the obvious next step would have been more work on combining channels. The measurement that survived scrutiny points somewhere else entirely, at background subtraction, which is several times larger and needs only the channel already in use. The robustness checks are what separated the two: the same radius sweep and spatial split that confirmed the tophat had already eliminated post minus pre.

The segmentation result is the more conventional contribution and it is solid. Median agreement above 0.91 against two independent hand segmentations, on 605 cells, through a matching procedure with no tunable parameter, is enough to use the masks for measurement. Cellpose and StarDist are the obvious modern alternatives [15, 16], and a generalist deep learning model is the natural first thing to reach for. In these images, though, cells are dark on a bright substrate, which is the inverse of what the pretrained models expect. The classical pipeline agrees with a human at a median IoU of 0.93, runs in seconds on a CPU with no training data, and stays interpretable when it fails. The method registry exists so that a learned method can be added later and compared on the same benchmark rather than argued about.

The human-in-the-loop design earns its place for a related reason. Pores here are dim, featureless and outnumbered by brighter artefacts, and candidates outnumber real pores by roughly 54 to 1. A classifier that labels everything negative in that setting is 98% accurate and useless, which is why average precision rather than accuracy is the right summary. Full automation would mean trusting a model on exactly the cases a trained observer finds hard. A ranked queue keeps the observer's judgement where it is needed and removes the part that needs none, which is looking at empty substrate.

5. Limitations

The strongest claims here are deliberately conservative, because the pore ground truth is small: 74 manual annotations from one microscopy field, of which 70 entered the tolerance-expanded evaluation. That is enough to expose metric saturation and to compare broad operating behaviours. It is not enough to optimise a detector or to claim the 12 px tophat radius is universally best, which is why background subtraction is implemented but not enabled by default.

Manual annotation is also not a perfect gold standard. The 15 px matching radius absorbs some localisation uncertainty, but that same tolerance is what makes chance matches possible at high candidate density. Future validation should report both localisation error and chance-adjusted detection performance, ideally with annotations from more than one reviewer.

The classifier is demonstrated rather than validated. Its training labels treat any unmarked candidate as a negative, which assumes the manual marking was exhaustive. With roughly 25,000 candidates against 74 marks, that is a strong assumption.

Two pieces of ground truth also disagree with each other. Growing the mask captures more pores but worsens agreement with the manual cell boundaries, and an adaptive segmentation variant does not beat 0.911: agreement falls monotonically as the mask grows. That negative result is worth recording, because the next person will otherwise try the same thing. The current global threshold may also miss extremely faint cell processes, and an adaptive method should be reconsidered if those processes prove important for pore assignment.

Segmentation was validated on two fields rather than the full image collection, from one laboratory and one imaging setup. Per-cell figures are kept descriptive throughout, because cells within a single image are not independent observations and attaching an inferential claim to them would be pseudoreplication.

Finally, this paper does not compare normal and glaucomatous cells. The analysis was kept blind to condition while the measurement pipeline was being built and validated. The next stage is to freeze the segmentation and detection rules, collect substantially more reviewed pore labels, train the classifier for the highly imbalanced pore-versus-non-pore problem, and then quantify pore density, size and spatial distribution. Any group-level count analysis should treat cell area as an exposure and account for clustering by image and donor rather than treating thousands of cells as independent observations.

6. Conclusion

This project began as an attempt to replace manual pore counting with a machine learning model. The more useful outcome was a validated way to decide when automation is actually working. The pipeline segments Schlemm's canal endothelial cells with close agreement to manual tracing, and it reduces the search for pores from whole-image visual inspection to a structured review queue of 32 to 285 candidates per real pore.

Its strongest contribution, though, is a control. Candidate recall must be reported against a matched chance baseline. In this dataset that single comparison overturned an apparently perfect detector, eliminated a 42% workload-reduction claim, and redirected the method toward morphological background subtraction. It costs one extra run of the same scoring code against uniformly random points. For future high-throughput studies of Schlemm's canal pores, the path to trustworthy automation is not to remove the human as quickly as possible. It is to make every automated step measurable, falsifiable, and cheap enough to verify.

Acknowledgements

I am grateful to Professor Darryl R. Overby and to Jacques A. Bertrand in the Department of Bioengineering at Imperial College London for their supervision and feedback, and for providing

the microscopy data and the manual segmentations and pore annotations that made validation possible. Every validation number in this paper is measured against Jacques's annotations. I also thank the members of the Overby Laboratory for their guidance throughout the project. This research was supported by the Laidlaw Scholars Leadership and Research Programme.

Data availability

The microscopy images analysed in this project are part of ongoing, unpublished work in the Overby Laboratory and are not publicly distributed. Analysis code, parameter definitions, and derived validation outputs can be shared subject to laboratory approval.

References

1. Tham YC, Li X, Wong TY, Quigley HA, Aung T, Cheng CY. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology*. 2014;121(11):2081–2090.
2. Bill A, Svedbergh B. Scanning electron microscopic studies of the trabecular meshwork and the canal of Schlemm: an attempt to localize the main resistance to outflow of aqueous humor in man. *Acta Ophthalmologica*. 1972;50(3):295–320.
3. Braakman ST, Moore JE Jr, Ethier CR, Overby DR. Transport across Schlemm's canal endothelium and the blood-aqueous barrier. *Experimental Eye Research*. 2016;146:17–21. doi:10.1016/j.exer.2015.11.026.
4. Ethier CR, Coloma FM, Sit AJ, Johnson M. Two pore types in the inner-wall endothelium of Schlemm's canal. *Investigative Ophthalmology & Visual Science*. 1998;39(11):2041–2048.
5. Braakman ST, Pedrigi RM, Read AT, et al. Biomechanical strain as a trigger for pore formation in Schlemm's canal endothelial cells. *Experimental Eye Research*. 2014;127:224–235. doi:10.1016/j.exer.2014.08.003.
6. Johnson M, Chan D, Read AT, Christensen C, Sit A, Ethier CR. The pore density in the inner wall endothelium of Schlemm's canal of glaucomatous eyes. *Investigative Ophthalmology & Visual Science*. 2002;43(9):2950–2955.

7. Allingham RR, de Kater AW, Ethier CR, Anderson PJ, Hertzmark E, Epstein DL. The relationship between pore density and outflow facility in human eyes. *Investigative Ophthalmology & Visual Science*. 1992;33(5):1661–1669.

8. Overby DR, Zhou EH, Vargas-Pinto R, et al. Altered mechanobiology of Schlemm's canal endothelial cells in glaucoma. *Proceedings of the National Academy of Sciences of the United States of America*. 2014;111(38):13876–13881. doi:10.1073/pnas.1410602111.

9. Siadat SM, Li H, Millette BA, et al. Endothelial cell stiffness and type drive the formation of biomechanically induced transcellular pores. *Cell Reports*. 2026;45(1):116668. doi:10.1016/j.celrep.2025.116668.

10. Braakman ST, Read AT, Chan DW, Ethier CR, Overby DR. Colocalization of outflow segmentation and pores along the inner wall of Schlemm's canal. *Experimental Eye Research*. 2015;130:87–96.

11. Li H, Wong C, Siadat SM, Perkumas KM, Bertrand JA, Overby DR, Sulchek T, Stamer WD, Ethier CR. Transient receptor potential vanilloid 4 modulates substrate stiffness mechanosensing and transcellular pore formation in human Schlemm's canal cells. *Acta Biomaterialia*. 2025;206:110–124. doi:10.1016/j.actbio.2025.09.039.

12. Rodrigues JR. *Computer vision tools to automate the study of Schlemm's canal inner wall cell biomechanics and pore formation*. PhD thesis, Department of Bioengineering, Imperial College London; 2022.

13. Otsu N. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*. 1979;9(1):62–66. doi:10.1109/TSMC.1979.4310076.

14. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for biological-image analysis. *Nature Methods*. 2012;9:676–682. doi:10.1038/nmeth.2019.

15. Stringer C, Wang T, Michaelos M, Pachitariu M. Cellpose: a generalist algorithm for cellular segmentation. *Nature Methods*. 2021;18:100–106. doi:10.1038/s41592-020-01018-x.

16. Schmidt U, Weigert M, Broaddus C, Myers G. Cell detection with star-convex polygons. In: *Medical Image Computing and Computer Assisted Intervention (MICCAI 2018)*. 2018:265–273.