**BMJ Open Ophthalmology**

# Semantic segmentation of gonio-photographs via adaptive ROI localisation and uncertainty estimation

Andrea Peroni [1] Anna Paviotti,[2] Mauro Campigotto,[2] Luis Abegão Pinto [3] Carlo Alberto Cutolo [4] Jacintha Gong,[5] Sirjhun Patel,[5] Caroline Cobb,[5] Stewart Gillan,[5] Andrew Tatham [6] Emanuele Trucco [1]

¹VAMPIRE Project, Computing, School of Science and Engineering, University of Dundee, Dundee, UK
²Department of Research and Development, NIDEK Technologies Srl, Albignasego, Italy
³Department of Ophthalmology, Hospital de Santa Maria, Lisbon, Portugal
⁴Clinica Oculistica, Di.N.O.G.M.I., University of Genoa, Genoa, Italy
⁵Department of Ophthalmology, Ninewells Hospital, NHS Tayside, Dundee, UK
⁶Department of Ophthalmology, Princess Alexandra Eye Pavilion, NHS Lothian, Edinburgh, UK

**Correspondence to**
Mr Andrea Peroni; a.peroni@dundee.ac.uk

## ABSTRACT

**Objective** To develop and test a deep learning (DL) model for semantic segmentation of anatomical layers of the anterior chamber angle (ACA) in digital gonio-photographs.

**Methods and analysis** We used a pilot dataset of 274 ACA sector images, annotated by expert ophthalmologists to delineate five anatomical layers: iris root, ciliary body band, scleral spur, trabecular meshwork and cornea. Narrow depth-of-field and peripheral vignetting prevented clinicians from annotating part of each image with sufficient confidence, introducing a degree of subjectivity and features correlation in the ground truth. To overcome these limitations, we present a DL model, designed and trained to perform two tasks simultaneously: (1) maximise the segmentation accuracy within the annotated region of each frame and (2) identify a region of interest (ROI) based on local image informativeness. Moreover, our calibrated model provides results interpretability returning pixel-wise classification uncertainty through Monte Carlo dropout.

**Results** The model was trained and validated in a 5-fold cross-validation experiment on ~90% of available data, achieving ~91% average segmentation accuracy within the annotated part of each ground truth image of the hold-out test set. An appropriate ROI was successfully identified in all test frames. The uncertainty estimation module located correctly inaccuracies and errors of segmentation outputs.

**Conclusion** The proposed model improves the only previously published work on gonio-photographs segmentation and may be a valid support for the automatic processing of these images to evaluate local tissue morphology. Uncertainty estimation is expected to facilitate acceptance of this system in clinical settings.

## INTRODUCTION

Glaucoma is a group of neuropathies that irreversibly damage the optic nerve, possibly leading to blindness. Increased intraocular pressure, normally regulated in the anterior chamber angle (ACA) by the trabecular meshwork (TM), is widely recognised as a primary risk factor.[1] The underlying cause of a limited TM functionality can be used to categorise the disease and choose the best medical treatment, making the ACA assessment fundamental.[1]

## Key messages

**What is already known about this subject?**
► Conventional gonioscopy is an important eye examination with several known limitations (eg, long examination time, expertise required, etc). Digital devices have been recently developed to speed up the acquisition of data (images of the irido-corneal interface) and allow their storage. However, computer algorithms to support the analysis of this specific type of data are still very limited both in number and in capabilities. While algorithms for angle aperture classification have been proposed, we argue that a semantic segmentation approach could enable a more detailed description of the irido-corneal angle morphology and better support clinicians' work.

**What are the new findings?**
► With this study, we addressed all the limitations of the only available semantic segmentation algorithm for gonio-photographs (which is our previously published work). Our deep learning-based system allows to effectively segment anatomical layers in gonio-photographs, providing accurate and interpretable results. Our research fills a gap in current literature on clinical applications of computer algorithms. Segmentation systems designed for other purposes are not suitable for off-the-shelf-use in this case, since they do not deal properly with specific data characteristics.

**How might these results change the focus of research or clinical practice?**
► Semantic segmentation is a powerful image analysis technique, widely used for other types of medical data for research and clinical applications (eg, segmentation of retinal layers in optical coherence tomography B-scans). Among other advantages discussed in the manuscript, the automatic segmentation of anatomical layers in gonio-photographs may speed up digital data analysis, serve as processing backbone for the automatic measurement of relevant anatomical features (eg, width of synechiae, degrees of angle closure, etc) and be particularly useful in virtual clinics. Our pilot study reports solutions to deal with this kind of data and very promising results, to promote research in this important field.

Current international guidelines recommend gonioscopy to assess the ACA.[2 3] However, conventional (manual) gonioscopy presents known limitations, for example, challenging acquisition of digital images and steep learning curve. To overcome these, new devices have been developed.[4 5] The acquisition of digital images of the ACA enables automated data analysis, for example, using deep learning (DL).

DL systems have pushed the state-of-the-art in image processing, with numerous applications in ophthalmology. Such systems can be divided into two main categories:

▶ *Classification algorithms*: assigning each input to one of a set of classes, for example, healthy/pathological sample discrimination in fundus images.[6]

▶ *Segmentation algorithms*: performing a pixel-wise classification of inputs, thus highlighting target structures like the vasculature in fundus images or retinal layers in optical coherence tomography (OCT) data.[7–9]

DL has been used for glaucoma-related tasks on OCT.[10] In particular, the detection of angle closure using anterior-segment (AS) OCT cross-sections of the ACA has proved effective.[11 12] However, AS-OCT does not allow direct inspection of the ACA surface, preventing the evaluation of important biomarkers, such as neovascularisations and TM pigmentation; the detection of anterior synechiae (SY) may also be problematic.

The literature on DL applications to gonio-photographs is much more limited. This imbalance and the complementarity of AS-OCT and gonio-photographs for ACA evaluation support the importance of research in automated gonio-photographs analysis.

The automatic detection of angle closure in gonio-photographs has been studied.[13] However, algorithms so far only provide a global characterisation of input images, possibly missing important local features and preventing a comprehensive analysis of layers' interfaces and their changes over time, for example, for follow-up.

We present a DL system for semantic segmentation of digital gonio-photographs to allow the assessment of local ACA morphology. Our algorithm provides an accurate segmentation of five anatomical layers: iris root (I), ciliary body band (CBB), scleral spur (SS), TM and cornea (C). We discuss the advantages of this model with respect to our previous work,[14] which is, to our best knowledge, the only full publication on automated ACA layers segmentation. In particular we address the main limitation in 'DL-based segmentation of gonioscopic images',[14] regarding results interpretability, in two ways: (1) implementing a region of interest (ROI) localisation module to refine segmentation results and (2) estimating output epistemic uncertainty through a Monte Carlo dropout approach.[15] Moreover, we refine the model architecture and fine-tune training hyperparameters to improve the overall segmentation performance.

Our work fills a gap in current research for clinical applications of DL (automated analysis of gonio-photographs). Our algorithm can be used as backbone processing for the measurement of structures of the ACA, for example, size of anterior SY, and their changes over time. Other possible use comprise, but are not limited to, the localisation of the trabecular meshwork, as a preprocessing step for pigmentation grading (eg, prior to laser trabeculoplasty), and the improvement of auto-alignment and auto-tracking systems for data acquisition in non-contact clinical examinations, that are gaining importance due to the COVID-19 pandemic.

It is also worth noting that segmentation algorithms widely employed in other clinical applications are not suitable for off-the-shelf use with our digital gonio-photographs. This is due to ground-truth limitations discussed in the Materials and methods section. Our model design has been specifically conceived to process ground-truth data to maximise output accuracy and interpretability.

This paper is structured as follows: Materials and methods section describes the dataset and the model architecture; Results section compares the performance of the proposed model with those of our previous work[14]; Discussion section summarises our findings, highlighting improvements and current limitations.

## MATERIALS AND METHODS
### Patients and public involvement
This research aims to develop artificial intelligence (AI) software for aiding digital gonioscopy. For this reason, the involvement of patients and public beyond the phase of data acquisition was neither appropriate nor possible.

### Data selection and characteristics
We modified the dataset of ACA sector images (960×1280 pixels, RGB) used in 'DL-based segmentation of gonioscopic images',[14] acquired with a NIDEK GS-1 device (NIDEK CO., LTD. Gamagori, Japan), removing 28 images due to non-compliance with requirements of the annotation protocol and adding 34 new gonio-photographs, for a total of 274 images from 214 examinations of 162 patients. New images were annotated by the same group of annotators and according to the same protocol as in 'DL-based segmentation of gonioscopic images'[14] to ensure consistency. All images were acquired with patients' agreement and following General Data Protection Regulation rules (including anonymisation at source) during routine clinical examinations. Minor acquisition conditions not compromising data consistency were left to the discretion of physicians.

Note that this work focuses on the development of an AI software tool; it is not an association study requiring cross-linked patient data, hence, none was sought nor used.

Table 1 shows the distribution of the features of interest in our dataset by two main traits. The first is the iris colour, light (blue or green) or dark (brown); the second is the predominant feature of the ACA sector, that can be only one of the following four: (1) the presence of anterior SY (local adhesions of the iris on the TM

**Table 1** Dataset features distribution (% rounded at first decimal)

|  | High pigm. | Slight pigm. | Synechia | Angle closure |
|---|---|---|---|---|
| Light iris | 43 (15.7%) | 45 (16.4%) | 7 (2.6%) | 1 (0.4%) |
| Dark iris | 60 (21.9%) | 51 (18.6%) | 43 (15.7%) | 24 (8.8%) |

Each image has been categorised by two main visual traits: the iris colour (rows) and an additional predominant feature of the sector (columns).

or C), (2) appositional closure of the angle (AC) (iris occluding the TM in at least half of the frame) or (3, 4) the TM pigmentation grade in all the other images (two cases: high, corresponding to Scheie's grades II, III and IV; slight, Scheie's grades None and I). Dark irises are predominant (65%), especially in the SY and AC subsets (86% and 96%). Moreover, structural changes in ACA layers represented by SY and AC images are considerably under-represented in the dataset (27.5%).

Images were annotated by four clinical experts by delineating target layers, following a protocol devised especially and using the VGG Image Annotator Tool (version 2.0.8),[16] after annotator training. Authors SP and JG were respectively, year-4 and year-7 specialty trainees with experience in gonioscopy, author CAC was glaucoma specialist with 5 years of clinical experience in glaucoma management and author LAP had more than 10 years of clinical experience in tertiary referral centres.

Our digital gonio-photographs show a narrow depth-of-field and vignetting, so that only part of them can be evaluated confidently, introducing a degree of subjectivity in the ground truth and features correlation between the annotated and the un-annotated (label NA) image regions (online supplemental figure 1). These limitations must be carefully addressed when designing and training a segmentation algorithm. In fact, semantic segmentation of ACA layers depends on anatomical features visible in the images, for example, layers' interfaces, while the (subjective) boundaries between the annotated and un-annotated regions of each image depend on the reduction of local informativeness due to de-focussing and vignetting.

In a previous study, we have reported a detailed analysis on the inter-annotator variability of anatomical layers delineations in digital gonio-photographs.[17] In essence, we observed lower average agreement (quantified using Dice scores) for CBB and SS (about 75% and 65%, respectively) than for I, TM and C (about 97%, 87% and 95%, respectively). This will be important when assessing system performance.

### Preprocessing and data augmentation

Following the approach in 'DL-based segmentation of gonioscopic images',[14] images of different ACA angular sectors were first rotated to a common orientation of layers (horizontal, iris at the bottom), resized to 240×320 pixels and normalised to (0, 1) range. The augmentation pipeline comprised both geometric and photometric transformations. All transformation parameters were randomly extracted from uniform distributions at each training epoch. Geometric transformations consisted of: translations along x-axis and y-axis (ranges (0, image_width/3) and (0, image_height/3), respectively); rotations (in degree, range (−30, 30)); shears along x-axis and y-axis (range (−10, 10)); and magnification (range (0.8, 1.2)). Photometric transformations consisted of: contrast variations (range (0.8, 1.2)); Gaussian noise injection (zero mean, SD range (0, 5)); uniform lightness variation (range (0.8, 1.2)); and non-uniform (sinusoidal) lightness variations (mixture of two sinusoids with amplitude, frequency and phase ranges derived experimentally).

### Network architecture and training

Figure 1 summarises our approach to provide an accurate segmentation map of ACA layers and deal with the limitations of our ground truth effectively. The data representation generated by the network encoder is processed in parallel by two independent network units: the *semantic decoder*, which returns the estimated class for each image pixel (the segmentation map), and the *ROI decoder*, which highlights an image ROI (the sharp and well-lit area). The two outputs are combined to provide a final segmentation map within the estimated image ROI.

The basic processing blocks are shown in figure 2A. The convolutional block (figure 2A1) (ConvB) is a sequence of dropout[18] (0.2 drop probability), 2D convolutional layer (3×3 kernel size, with Kaiming-He uniform initialisation,[19] zero-padding), instance normalisation[20] and Leaky-ReLU activation[21] (0.01 negative slope). Dense blocks (figure 2A2) (DenseB), inspired by,[22] are sequences of four ConvBs with two intermediate concatenations of output pairs. The encoder block (figure 2A3) (EncB) is a DenseB followed by an input-output concatenation. The decoder block (figure 2A4) (DecB) is a DenseB followed by a ConvB that reduces the number of feature maps.

Figure 2B captures the three network components in detail. The encoder figure 2B1 is composed of two initial ConvBs with eight filters (drop out of the first ConvB is disabled), followed by three combinations of max pooling and EncB, each doubling the number of feature maps. A final DenseB generates the latent data representation. The semantic decoder figure 2B2 up-samples feature maps via max un-pooling[23] and concatenates them with those forwarded by the encoder. The resulting feature maps are processed through DecBs, each reducing by a factor 4 the number of feature maps. The processing ends
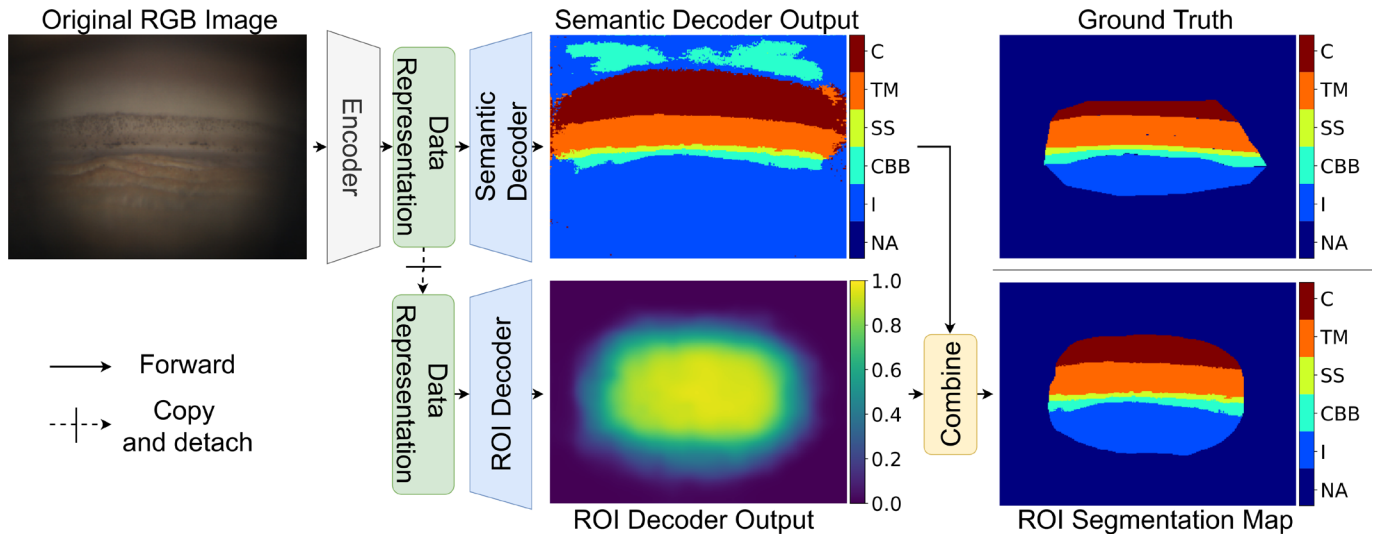
**Figure 1** Network architecture overview with example of intermediate and final outputs for a given input image. C, cornea; CBB, ciliary body band; I, iris root; NA, un-annotated region; ROI, region of interest; SS, scleral spur; TM, trabecular meshwork.

with a DenseB and a 1×1 2D convolution, followed by a six-class softmax activation (the NA class is considered for consistency but does not contribute to semantic decoder optimisation). We can interpret the sequence of encoder and semantic decoder as a segmentation U-Net[24] trained via weighted categorical cross-entropy with equal weights for all the anatomical layers considered and weight 0 for the region of the image not annotated by the experts.

The ROI decoder figure 2B3 is a simple and convenient way to filter out the artefacts expected in the segmentation map periphery (since dark and blurred image areas are not informative enough to be correctly classified), overcoming the issue on results visualisation in 'DL-based segmentation of gonioscopic images'.[14] A detached copy of the data representation is processed through a DecB, a ConvB (dropout is deactivated) and a 2D convolution
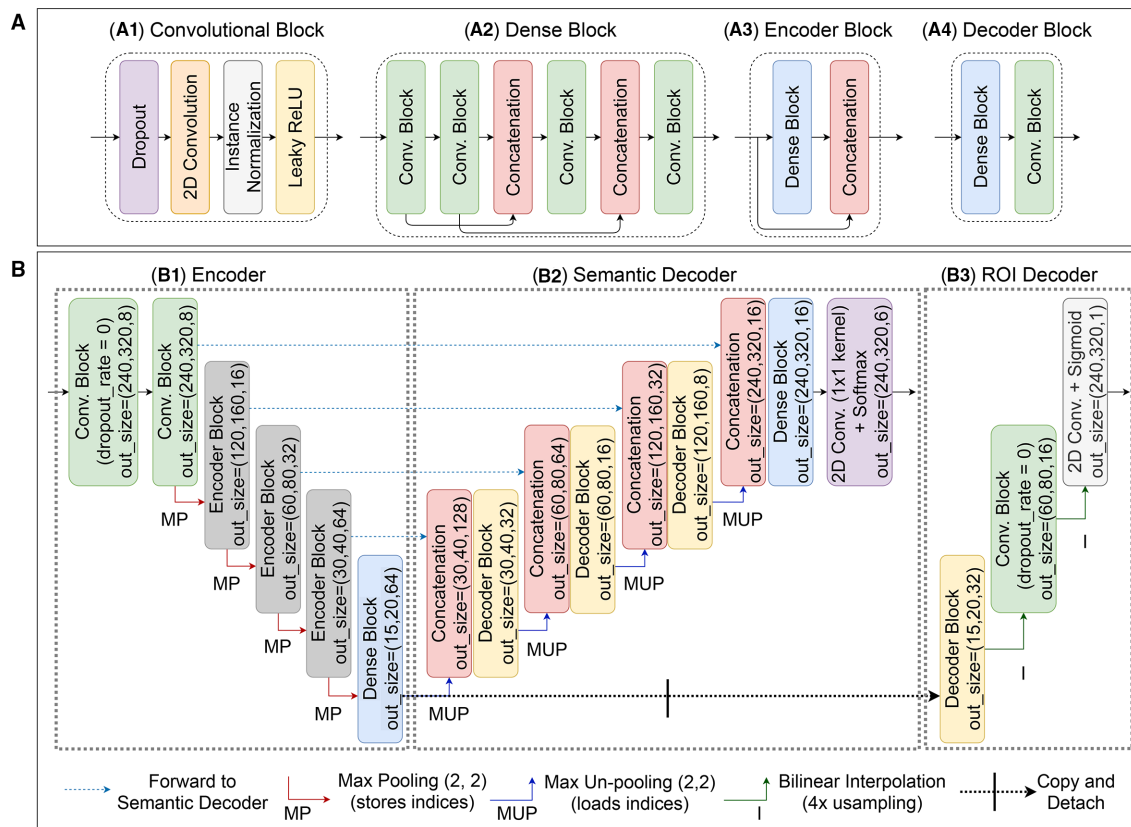


**Figure 2** (A) Basic processing blocks: convolutional (A1), dense (A2), encoder (A3) and decoder (A4) blocks; (B) detail of the proposed architecture: encoder (B1), semantic decoder (B2) and region of interest decoder (B3).

with a sigmoid activation. The intermediate feature maps are up-sampled by a factor 4 using bilinear interpolation.

To train the ROI decoder we first generated an ROI likelihood map from the original semantic annotations: all annotated pixels were first assigned value 1 (binarisation), gaps between layers (if any) were filled to obtain a dense region that is then smoothed to simulate a probabilistic distribution of annotator's evaluation confidence (online supplemental figure 2). ROI likelihood maps were used as reference when computing the mean squared error loss for the ROI decoder, independently from the semantic segmentation optimisation.

After ROI binarization (threshold 0.5 in our results), the two intermediate results are multiplied.

Semantic segmentation of ACA layers and ROI localisation are two un-correlated tasks that rely on different interpretations of the same image features. The former looks for patterns and textures that characterise the different anatomical layers; the latter evaluates sharpness and illumination variations across the frame. We verified experimentally that other systems, like attention mechanisms,[25] are not effective at addressing this problem since they do not deal with the two tasks independently.

Optimisation of model's weights was performed through stochastic gradient descent with Nesterov momentum (0.9), learning rate equal to 0.01 and 8 images per batch. The model was designed and trained using Python (3.7.9) and PyTorch (1.7.0).

## Uncertainty estimation

We included dropout layers in our model and used Monte Carlo dropout[15] to generate multiple softmax activations for an input image at inference time. The predicted class for each pixel is the argmax of the average activations, and the activations' variance for the assigned class estimates the model (epistemic) uncertainty. If the pixel activations for a given final class are consistent across several output candidates, the variance is low; otherwise, the variance is high (high uncertainty).

Model calibration was verified on the hold-out test set after every cross-validation fold (see the next section for more details) by computing the expected calibration errors (ECE).[26] We obtained an average (across folds) ECE of ~0.01, suggesting that our model is well calibrated and that activation variance may be used to estimate epistemic uncertainty.

## RESULTS

Figure 3 compares an example of combined network output (edges of the segmentation map refined by the ROI, bottom right) with the ground truth delineation (top right). The segmentation is noticeably accurate and the ROI very similar to that highlighted by the annotator. The uncertainty (variance) map provides useful information about the model confidence in the results. In this case, the variance map only highlights layers interfaces, as expected even when the segmentation is accurate, since layers boundaries are often not very sharp features.
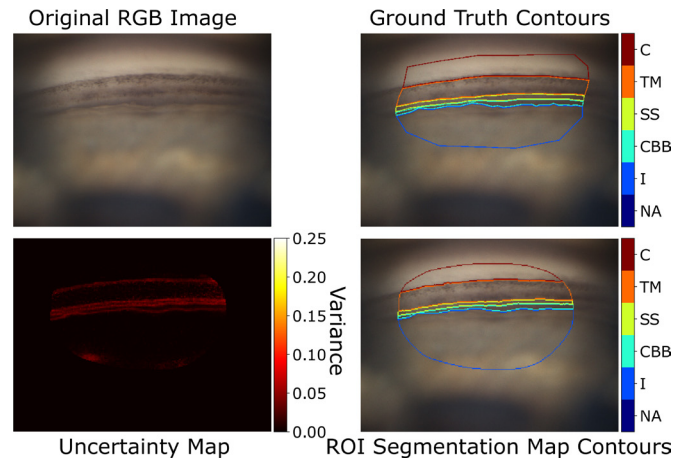


**Figure 3** Example of gonio-photograph (top left) and ground truth delineations of the visible layers (top right); boundaries of the segmentation map output by the semantic decoder and refined by the ROI (bottom right); uncertainty (variance) map (bottom left). ROI, region of interest.

We compared the performance of our previous[14] and current models on the new dataset. First, we split it into a test set (31 images from 25 examinations of 17 patients) and a training-validation set (243 images from 189 examinations of 145 patients), with similar distributions of the features described in the Materials and methods section. We then randomly split the training-validation set into five folds of 29 patients each, to cross-validate the two models, each fold consisting of a variable number of images (mean: 48.6, SD: 5.3 for the previous[14] and 5.7 for the new implementation). The algorithm in 'DL-based segmentation of gonioscopic images',[14] used as baseline, was trained as per the original approach: assigning loss class weights equal to the ratios between the average size of each layer and the average size of a reference layer (I), and storing best model weights after each epoch returning a lower validation loss. The new system was trained to maximise segmentation accuracy within the annotated images area, storing model weights after each epoch returning an increased validation accuracy. Both loss and accuracy only account for pixels within the annotated image regions, since segmentation performance on the un-annotated area is, by definition, not measurable.

Comparing average per-layer metrics on the hold-out test set (table 2) between the two models, our new one returned overall higher Dice scores, in particular for CBB and SS (+2.4% and +5.6% on average), two layers of fundamental clinical importance. To interpret these values correctly, one must consider that, importantly, the average inter-observer Dice scores for CBB and SS reported in our pilot inter-annotator variability study[17] were about 75% and 65% (the lowest), making our results consistent with the average agreement between experts. Our previous approach promoted sensitivity (less false negatives) for layers with lower inter-annotator agreement (CBB and SS) while the new one promotes precision (less false positives).

**Table 2** Performance comparison between the algorithm proposed in 'DL-based segmentation of gonioscopic images'[14] and ours

|  |  | Previous model[14] | | New model | |
|---|---|---|---|---|---|
|  |  | Mean (%) | SD (%) | Mean (%) | SD (%) |
| I | Prec. | **96.2** | **0.4** | 92.8 | 1.3 |
|  | Sens. | 91.0 | 2.2 | **97.4** | **0.8** |
|  | Dice | 93.6 | 1.0 | **94.8** | **0.4** |
| CBB | Prec. | 67.0 | 5.0 | **84.8** | **1.0** |
|  | Sens. | **75.2** | **4.2** | 64.0 | 4.3 |
|  | Dice | 70.4 | 2.0 | **72.8** | **2.9** |
| SS | Prec. | 53.2 | 4.2 | **68.4** | **1.7** |
|  | Sens. | **78.6** | **2.9** | 69.6 | 3.3 |
|  | Dice | 63.2 | 2.8 | **68.8** | **1.2** |
| TM | Prec. | 84.0 | 1.9 | **84.6** | **2.0** |
|  | Sens. | 85.0 | 3.0 | **91.2** | **1.2** |
|  | Dice | 84.4 | 1.9 | **87.6** | **0.5** |
| C | Prec. | 95.8 | 1.2 | **96.4** | **0.5** |
|  | Sens. | 89.2 | 3.2 | **92.2** | **1.2** |
|  | Dice | 92.0 | 1.5 | **94.2** | **0.4** |

Bold values indicate the best model in each of the comparisons (each row of the table)
C, cornea; CBB, ciliary body band; I, iris root; SS, scleral spur; TM, trabecular meshwork.

The pixel-wise accuracy within the annotated image region was much higher for the new model (91% vs 87%).

ROIs estimated by the model and those identified by the annotators cannot be compared quantitatively given the subjective estimation of their boundaries based on smooth illumination and focus transitions. To qualitatively validate the ROI decoder of our model we asked a clinician, blinded to our ground truth data, to verify that ROIs estimated by the model did not leave out any ACA feature of clinical interest. The result was that our model highlighted an appropriate ROI in every test image and in every cross-validation fold, suggesting a stable, reliable approach.

Uncertainty estimation highlighted local segmentation irregularities correctly.

## DISCUSSION

Gonioscopy is a clinical standard. New imaging devices allow to acquire digital gonio-photographs more consistently, enabling automatic analysis. The anatomy of ACA layers is clinically relevant as related to a high-prevalence disease (glaucoma), therefore automatic systems supporting diagnosis will be increasingly important. Systems for automated assessment of ACA conditions have been developed on AS-OCT data.[11 12] However, AS-OCT scans do not allow the evaluation of important features of the ACA, such as the TM pigmentation, and make the measurement of others, for example, anterior SY, problematic. For these reasons, AS-OCT will likely not replace gonioscopy soon, remaining a complementary source of information.

Automatic systems for detecting angle closure in digital gonio-photographs have been proposed[13] assigning only a single *state* to each image (a label that globally represent what the image shows) that may obscure relevant local conditions. In order to assess the morphology of ACA accurately and account for local variations of layers' interfaces, a segmentation approach is certainly more suitable. State-of-the art semantic segmentation networks cannot deal with limitations posed currently by the annotation of our gonio-photographic images. Our previous DL-based segmentation model[14] (the only one in the literature to our best knowledge) achieved a promising performance within the area annotated by experts, but failed to filter out the un-informative image periphery.

We presented here a solution to overcome previously reported limitations by adaptively identifying an ROI to refine segmentation maps and improve results readability. Moreover, our new calibrated model can support data analysis and interpretation further by highlighting uncertain segmentation areas. This is done estimating pixel-wise epistemic uncertainty as activations variance over multiple segmentation candidates. This will be further investigated in our future work.

The overall segmentation accuracy (~91%) is promising, although within a limited dataset. The layer-wise Dice scores on the test set correlate well with results from a pilot inter-annotator variability study.[17]

Finally, we discuss the current limitations of our research. We acknowledge the limited amount of data used. Although our dataset is representative for a variety of ACA features, larger datasets of annotated images are needed to train and evaluate comprehensively DL systems for gonio-photographs analysis in clinical practice.

The relative novelty and still limited use of digital devices for gonioscopy currently limits the availability of data for rarer conditions. In particular, more images representing complex layers morphologies, for example, SY, shall be collected and annotated in the future to further improve the generalisation capabilities of our models.

Our segmentation model fills a gap in current applications of DL for ophthalmology, enabling a more accurate data analysis than classification algorithms for angle closure and possibly providing a processing backbone for the measurements of clinical parameters, for evaluating changes of layers morphology over time and for developing auto-tracking systems for device alignment in non-contact clinical settings.

**Patient consent for publication** Not applicable.

**Ethics approval** This is a retrospective study performed on data acquired during the clinical evaluation of the NIDEK GS-1 imaging device. All the patients signed a consent form before being examined and data have been managed according to GDPR rules. Gonio-photographs were shared and used for our research with permission from the clinical institutions that acquired the data. The acquisition of data was carried out after ethics approval was obtained in each of the centres.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** No data are available. Not applicable.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**ORCID iDs**
Andrea Peroni http://orcid.org/0000-0002-3230-5269
Luis Abegão Pinto http://orcid.org/0000-0002-9960-7579
Carlo Alberto Cutolo http://orcid.org/0000-0002-2433-0704
Andrew Tatham http://orcid.org/0000-0003-0372-3100
Emanuele Trucco http://orcid.org/0000-0002-5055-0794

## REFERENCES

1 Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma: a review. *JAMA* 2014;311:1901–11.
2 European Glaucoma Society Terminology and Guidelines for Glaucoma, 4th Edition - Part 1Supported by the Egs Foundation. *Br J Ophthalmol* 2017;101:1–72.
3 Gedde SJ, Chen PP, Muir KW, *et al*. Primary angle-closure disease preferred practice Pattern®. *Ophthalmology* 2021;128:P30–70.
4 Shi Y, Yang X, Marion KM, *et al*. Novel and semiautomated 360-Degree gonioscopic anterior chamber angle imaging in under 60 seconds. *Ophthalmol Glaucoma* 2019;2:215–23.
5 Perera SA, Baskaran M, Friedman DS, *et al*. Use of EyeCam for imaging the anterior chamber angle. *Invest Ophthalmol Vis Sci* 2010;51:2993–7.
6 Christopher M, Belghith A, Bowd C, *et al*. Performance of deep learning architectures and transfer learning for detecting glaucomatous optic neuropathy in fundus photographs. *Sci Rep* 2018;8:1–13.
7 Mookiah MRK, Hogg S, MacGillivray TJ, *et al*. A review of machine learning methods for retinal blood vessel segmentation and artery/vein classification. *Med Image Anal* 2021;68:101905.
8 Roy AG, Conjeti S, Karri SPK, *et al*. ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. *Biomed Opt Express* 2017;8:3627–42.
9 Wei H, Peng P. The segmentation of retinal layer and fluid in SD-OCT images using Mutex dice loss based fully Convolutional networks. *IEEE Access* 2020;8:60929–39.
10 Ran AR, Tham CC, Chan PP, *et al*. Deep learning in glaucoma with optical coherence tomography: a review. *Eye* 2021;35:188–201.
11 Fu H, Xu Y, Lin S, *et al*. Angle-Closure detection in anterior segment Oct based on multilevel deep network. *IEEE Trans Cybern* 2020;50:3358–66.
12 Porporato N, Tun TA, Baskaran M, *et al*. Towards 'automated gonioscopy': a deep learning algorithm for 360° angle assessment by swept-source optical coherence tomography. *British Journal of Ophthalmology* 2021;188:bjophthalmol-2020-318275.
13 Chiang M, Guth D, Pardeshi AA, *et al*. Glaucoma Expert-Level detection of angle closure in Goniophotographs with Convolutional neural networks: the Chinese American eye study. *Am J Ophthalmol* 2021;226:100–7.
14 A deep learning approach for semantic segmentation of gonioscopic images to support glaucoma categorization. *Annual Conference on Medical Image Understanding and Analysis*, 2020.
15 PMLR. Dropout as a Bayesian approximation: representing model uncertainty in deep learning. *International conference on machine learning*, 2016.
16 The via annotation software for images, audio and video. *Proceedings of the 27th ACM international conference on multimedia*, 2019.
17 Peroni A, Paviotti A, Campigotto M, *et al*. On clinical agreement on the visibility and extent of anatomical layers in digital Gonio Photographs. *Transl Vis Sci Technol* 2021;10:1.
18 Srivastava N, Hinton G, Krizhevsky A. Dropout: a simple way to prevent neural networks from Overfitting. *J Mach Learn Res* 2014;15:1929–58.
19 Delving deep into Rectifiers: Surpassing Human-Level performance on Imagenet classification. *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*, 2015.
20 Ulyanov D, Vedaldi A, Lempitsky V. Instance normalization: the missing ingredient for fast Stylization. *arXiv preprint* 2016.
21 Rectifier nonlinearities improve neural network acoustic models. *Citeseer* 2013.
22 Densely connected Convolutional networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017.
23 IEEE. Learning deconvolution network for semantic segmentation. *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015.
24 Springer. U-Net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*, 2015.
25 Oktay O, Schlemper J, Folgoc LL. Attention U-Net: learning where to look for the pancreas. *arXiv preprint* 2018.
26 Obtaining well calibrated probabilities using Bayesian binning. *Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.