



A Deep Learning Approach for Semantic Segmentation of Gonioscopic Images to Support Glaucoma Categorization

Andrea Peroni¹ , Carlo A. Cutolo² , Luis A. Pinto³ , Anna Paviotti⁴,
Mauro Campigotto⁴, Caroline Cobb⁵, Jacintha Gong⁵, Sirjhun Patel^{5,6},
Andrew Tatham^{7,8} , Stewart Gillan⁵ , and Emanuele Trucco¹ 

¹ VAMPIRE Project, Computing, School of Science and Engineering,
University of Dundee, Dundee, UK
{a.peroni,e.trucco}@dundee.ac.uk

² Clinica Oculistica, DiNOGMI, University of Genoa, Genoa, Italy

³ Department of Ophthalmology, Hospital Santa Maria, Lisbon, Portugal

⁴ NIDEK Technologies Srl., Albignasego, Italy

{annapaviotti,maurocampigotto}@nidektechnologies.it

⁵ Department of Ophthalmology, Ninewells Hospital, NHS Tayside, Dundee, UK

⁶ Department of Ophthalmology, University of Dundee, Dundee, UK

⁷ Princess Alexandra Eye Pavilion, NHS Lothian, Edinburgh, UK

⁸ Ophthalmology, University of Edinburgh, Edinburgh, UK

Abstract. We present a deep learning semantic segmentation algorithm for processing images acquired by a novel ophthalmic device, the NIDEK GS-1. The proposed model can sophisticate the current reference exam, called gonioscopy, for evaluating the risk of developing glaucoma, a severe eye pathology with a considerable worldwide impact in terms of costs and negative effects on affected people's quality of life, and for inferring its categorization. The target eye region of gonioscopy is the interface between the iris and the cornea, and the anatomical structures that are located there. Our approach exploits a dense U-net architecture and is the first automatic system segmenting irido-corneal interface images from the novel device. Results show promising performance, providing about 88% of mean pixel-wise classification accuracy in a 5-fold cross-validation experiment on a very limited size dataset of annotated images.

Keywords: Image segmentation · Deep learning · Gonioscopy

1 Introduction and Motivation

We present a deep-learning (DL) semantic segmentation algorithm for processing images acquired by a novel ophthalmic device, the NIDEK GS-1, in the context of computer-assisted risk evaluation and categorization of *glaucoma*. To our best knowledge, this is the first DL-based computational system to segment digital images of the interface between iris and cornea (irido-corneal angle).

© Springer Nature Switzerland AG 2020

B. W. Papież et al. (Eds.): MIUA 2020, CCIS 1248, pp. 373–386, 2020.

https://doi.org/10.1007/978-3-030-52791-4_29

Glaucoma is a leading cause of irreversible blindness and currently affects more than 70 million people worldwide [23]. The chances of developing this pathology may be affected by the failure of a specialized eye structure, the *trabecular meshwork* (TM), that bounds the periphery of the iris along its interface with the cornea in the eye anterior chamber (Fig. 1).

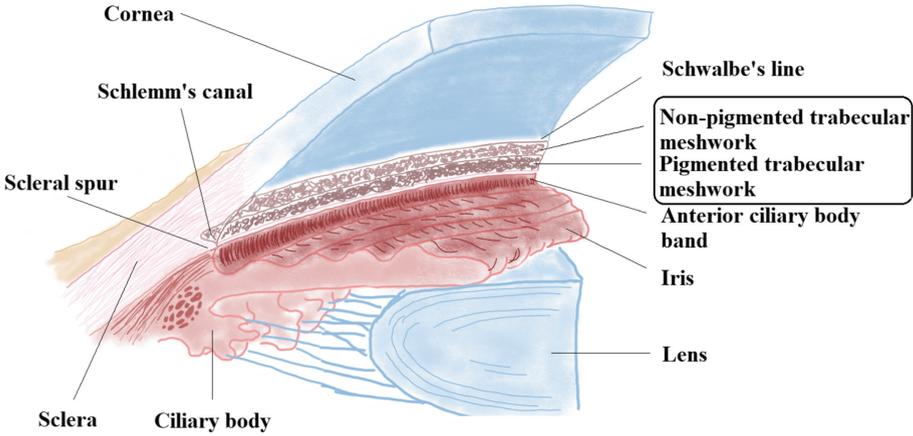


Fig. 1. A representation of a section of the 360° interface between iris and cornea.

The TM regulates the intra-ocular pressure by draining the aqueous humour. If its effectiveness diminishes the eye pressure may increase and, with it, the risk factors.

The current gold-standard examination, called *gonioscopy* [2], allows the clinician to inspect the structures located at the irido-corneal interface by manually holding a multi-faceted mirror in contact with the patient's eye. The visibility and conformation of these structures is of fundamental importance for assessing both risks and aetiology. The mirror is necessary as the interface lays well off the direct line of sight (Fig. 2(a)).

This system presents several limitations: it takes time and requires substantial expertise [11] with a long and steep learning curve, and the acquisition of images for follow-up (e.g. detection of changes over time, clinical trials) is problematic [4]. This paper focuses on images from a newly developed device, the NIDEK GS-1, capable of performing semi-automatic gonioscopy and of storing digital images of the irido-corneal interface, enabling a complete (360°) visualization (Fig. 2(b)).

Our contribution is the development of the first DL-based semantic segmentation algorithm to process this type of images. By providing immediate information about the location, extension and integrity of the interface structures, it can aid clinical decisions, making patient examination more effective, as well as more efficient.

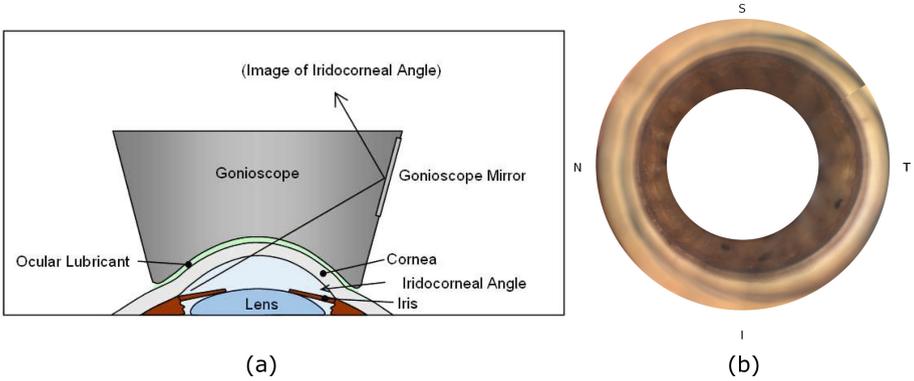


Fig. 2. (a) The working principle of gonioscopy: the mirrored prism allows to inspect the irido-corneal interface (image by Mick Lucas, distributed under CC BY-SA 3.0). (b) an example of 360° interface acquisition from NIDEK GS-1 device (the darker, inner region is the iris while the lighter one is the cornea; refer to Fig. 1 for a 3D representation of the same anatomical structures).

The paper is structured as follows: Sect. 2 discusses concisely some work relevant for our study; Sect. 3 our image dataset; Sect. 4 the annotations used and the related protocol, the data pre-processing, the network design and its training; Sect. 5 experiments and a brief discussion of the network results and performance; Sect. 6 the current strengths and limitations of the proposed model.

2 Related Work

DL techniques have been widely applied in ophthalmology as diagnostic aids for several pathologies, such as diabetic retinopathy [1] and macular degeneration [18].

Literature related to our project may be conveniently divided into two categories. The former comprises DL techniques aimed to segment structures that are geometrically similar to the irido-corneal interface layers. They were mostly developed to segment retinal layers in OCT images using convolutional (CNN) [7] and fully-convolutional (FCN) [3, 8, 16, 17, 19, 22] neural networks. The latter refers to DL algorithms designed for supporting glaucoma diagnosis with fundus images, mainly using CNN architectures [15, 20]. The main limitation of the existing deep-learning algorithms, none of which uses gonioscopic images, is the lack of results interpretability. Fundus images may be used to assist glaucoma diagnosis with good reported reliability, but gonioscopy remains necessary to characterise the aetiology of glaucoma.

In another paper [5], images of the irido-corneal interface acquired using a different device, called RetCam, were automatically graded by means of a non-DL algorithm. Despite the similarity of the images used, it does not perform

semantic segmentation of the interface structures, a very important source of information.

3 Materials

All the images selected to develop the proposed segmentation algorithm were acquired using the new NIDEK GS-1 device during clinical trials at partners' institutions. The whole irido-corneal interface circumference is captured by a sequence of sixteen 960×1280 pixels, RGB images, stored in *jpeg* format, each covering 22.5° .



Fig. 3. Three angle sectors with different orientations and specific visual features.

The entire dataset used in our experiments consists of 268 images. The images were accurately chosen to form a pilot-level set capturing the variability of the main visual features that characterize the anatomical region (Fig. 3). 76% of dataset images are of healthy cases, divided between dark (53%) and light (47%) irises and taking into account the variety of TM pigmentation grades (52% slightly pigmented and 48% highly pigmented).

Pathological images comprises: closed angles (9% of the whole dataset), where there is contact between the iris and pigmented TM; peripheral anterior synechie (12%), i.e. anomalous extensions of iris tissue that partially occludes the TM; images where implants are present (e.g. shunts or stents) (1.5%) and images acquired after surgical intervention (trabeculotomy) (1.5%). Pathological cases introduce high complexity to the segmentation problem, since anomalous tissue, as well as implants, modifies the expected structure of layers unpredictably.

Because of the novelty of the device, the availability of GS-1 exams was very limited at the time of writing this paper. For this reason, images were acquired from 138 different patients and multiple images from the same patient were selected considering criteria aimed to minimize data correlation. In particular, whenever possible, multiple images from the same patient were chosen from different eyes, and they had to be either characterized by a very different TM pigmentation (since the TM texture is not uniform along the irido-corneal interface circumference), by the presence of pathological tissue (i.e. synechiae) to ensure structural variation, or by a different grade of angle closure (which may also considerably vary within the same eye).

4 Methods

4.1 Annotation Process and Ground Truth

As mentioned above, NIDEK GS-1 is a recent device (launched in October 2018) and there are no available datasets of annotated images acquired by it. For this reason, it was essential to devise an annotation process to generate reliable ground truth. An annotation protocol was designed in close collaboration with the clinicians involved in this project (including authors Cutolo C. A., Pinto L. A., Gillan S., Tatham A., Cobb C., Gong J., Patel S.) and based on the experience of our research group in previous, similar projects [24]. The annotation tool chosen for this task was the VGG Image Annotator [6]. Ophthalmologists involved in annotating images were trained in the use of the annotation tool to generate consistent information.

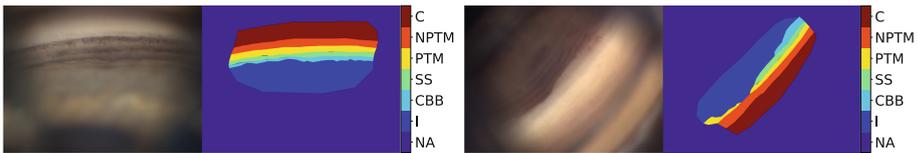


Fig. 4. Two sector images and their annotations.

Each image is paired with six binary maps, representing the regions covered by target layers. They are: the iris root (I), the ciliary body band (CBB), the scleral spur (SS), the pigmented trabecular meshwork (PTM), the non-pigmented trabecular meshwork (NPTM) and the cornea (C) (refer to Fig. 1 for a 3D visualization of layers location). The remaining area is the unannotated region (NA). Two annotation examples are reported in Fig. 4.

Considering the limited amount of training data and in order to ease the segmentation problem without compromising the outputs utility from a medical perspective, pigmented and non-pigmented trabecular meshwork annotations were merged to generate a single target region, the trabecular meshwork (TM).

It is worth specifying that the unannotated region of the images is not an actual segmentation target and is not used at training time. It is just the image portion that was impossible to classify clearly (e.g. because of blurring or vignetting).

The whole image dataset was divided into subsets and each was assigned to one of our collaborators to be annotated. Subsets were non-overlapping except for six common images, that were segmented by all five the ophthalmologists in order to evaluate their agreement. By visually comparing these data it is possible to qualitatively notice that some interpretation disagreement exists among the annotators, as exemplified by Fig. 5.

A preliminary quantitative analysis on the inter-annotator agreement was performed. This analysis accounted for classification differences within the annotated region of images, since, as already stated, leaving a portion of an image

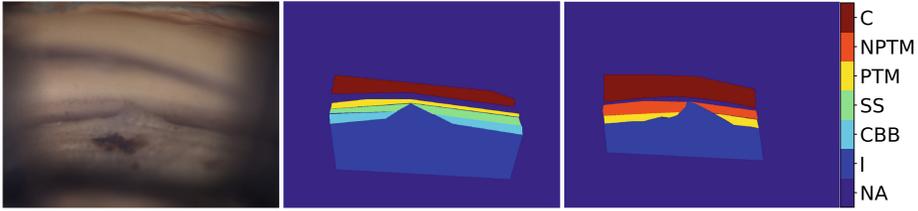


Fig. 5. Comparison of two different annotations of the same dataset image. Annotations were performed by ophthalmologists with experience in gonioscopy.

unannotated does not consist in a mistake itself. The agreement value among two annotators on a specific layer was calculated as the classification precision of the first with respect to the second (i.e. considering the second annotator the reference and calculating the $TP/(TP + FP)$ ratio where $TP =$ true positives and $FP =$ false positives). Results show that, while the inter-annotator agreement on larger and better identifiable regions (e.g. iris and cornea) is high ($>90\%$), it decreases when considering smaller regions. For example, the agreement between two annotators on the scleral spur, computed as the mean precision value over the six common images, may be $<50\%$, and may be even lower when comparing two annotations of a single image (e.g. 0% in Fig. 5).

The difficulty of classifying some of the irido-corneal interface structures is explained by the high intra-layer variability and inter-layer similarity of their features (e.g. texture and colour). Moreover boundaries among subsequent structures may not be sharp.

The design of a more accurate inter-annotator agreement assessment is a fundamental part of our present and future work.

4.2 Data Pre-processing

The data pre-processing consists in several phases aimed to favour the network at performing the segmentation. They are:

- *data re-sizing* to 240 by 320 pixels to limit memory usage. Whole images are input to the network;
- *data rotation* to a reference orientation, to relax the necessity of learning rotational invariance. This is done by considering the angles associated with acquisition sectors and rotating all the input images so that they show the iris at the bottom and the cornea at the top of the frame;
- *local contrast enhancement* through contrast limited adaptive histogram equalization [10]. Images are converted to HLS colour space and the L channel is equalized before converting them back to RGB;
- *data range scaling* to the $[0, 1]$ values range by simply dividing every pixel value by 255, which is the maximum theoretical value in RGB format;
- *data augmentation* through small rotations (0° to $\pm 30^\circ$), width and height shifts (0% to $\pm 30\%$ along each dimension), zoom (0% to $\pm 20\%$), shear (0°

to $\pm 20^\circ$ along each dimension) and horizontal flips. The augmentation is performed in order to simulate small misalignment of the device with the patient's eye.

Augmentation is performed online, generating new transformed images at each training epoch. Augmentation is not meant to generate a larger quantity of artificial data, indeed, the amount of augmented images at each training epoch is equal to that of input images, but to make the system insensitive to small geometrical input variations by introducing new, random modifications, every epoch. Input images and the corresponding annotations are transformed accordingly to maintain the correspondences between target annotations and anatomical structures.

4.3 Deep Learning Architecture

The architecture of the network developed is depicted in Fig. 6. It is an adaptation of two deep-learning models: the U-net, originally devised for semantic segmentation of medical images [21], and the Dense-Net, for image classification [12]. It may be also considered a simplified version of the architecture in [14]. Dense blocks made up of four convolutional layers are integrated in both the encoding and the decoding path of the U-net in order to increase the model abstraction capabilities, effectively re-use information and speed up the convergence. Long skip connections forward the feature maps of the encoder towards the corresponding resolution level of the decoder where they are concatenated with the upsampled maps before being processed. The number of filters each dense-block is made up of is reported in Fig. 6, inside the rectangles, while the dimension of the output feature matrices is written over each block that performs convolutions (single convolutional layers and dense blocks).

The structure of dense blocks, depicted in Fig. 7(a), consists of four convolutional blocks (Fig. 7(b)), that are simply a sequence of batch normalization [13], 2D convolution and ReLu activation. Each convolutional block forwards its outputs to all subsequent filters within the same dense block. The network terminates with a (1, 1) convolutional layer that performs a reduction along the third dimension in order to match the number of ground truth targets and, finally, a softmax activation.

The number of downsamplings is related to the maximum useful perceptive field we expect to obtain as a consequence of the input dimensions. The model has 525.000 trainable parameters. The reduced size is mainly explained by the limited size of the training set. Experimental results show that increasing the number of trainable parameters leads to training set overfitting with our set-up.

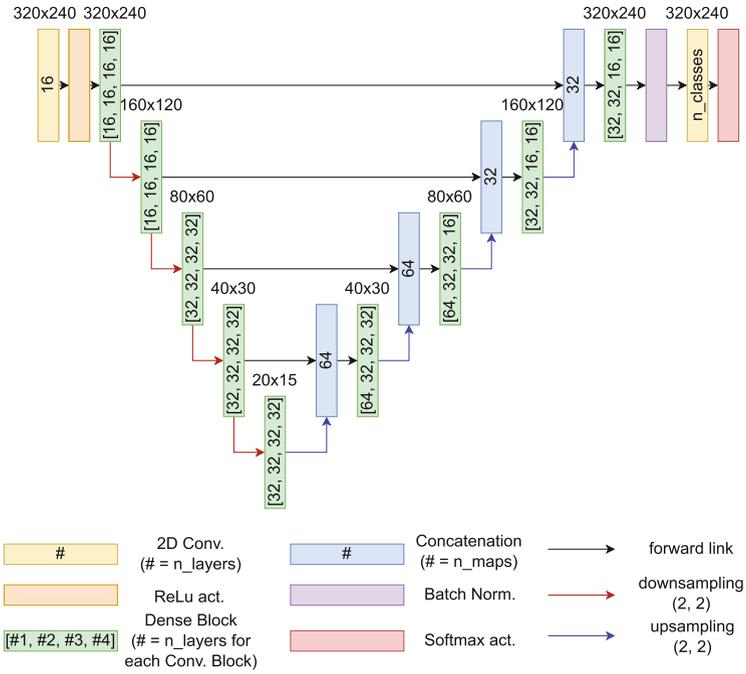


Fig. 6. The proposed network architecture.

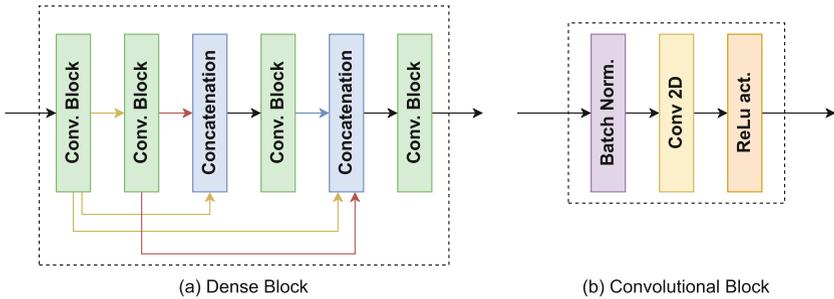


Fig. 7. The structure of dense and convolutional blocks.

4.4 Network Training

Our annotations comprise only the part of the target structures the clinicians were very confident about. Hence, textural information in the unannotated regions may be correlated with the target annotations and, for this reason, can not be used for training since it would decrease the overall network accuracy. This is clearly visible in Fig. 8.

Network training is, thus, performed minimizing a weighted categorical cross-entropy loss function. Each class weight is computed by considering the ratio

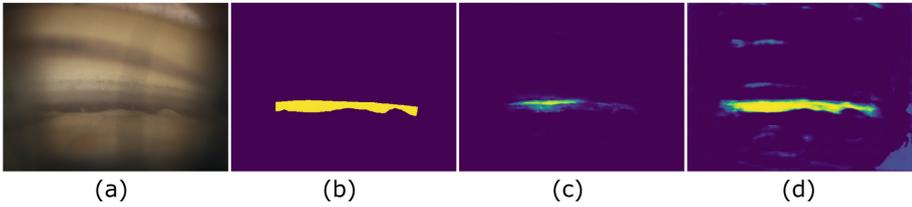


Fig. 8. The consequences of using the unannotated layer over training: (a) original image; (b) ciliary body band ground truth; (c) network output obtained including the unannotated layer during training; (d) network output obtained excluding the unannotated layer during training.

between the number of pixels of the training set that belong to that class and the number of pixels of a reference target, the iris, to account for class imbalance. Pixels belonging to the unannotated layer are given weight 0. The accuracy metric used to monitor the learning process was modified to consider only the region of the ground truth associated with the real target layers.

Layers' weights are initialized using the He normal distribution [9], while biases are all initialized to the constant 0.1. The model backpropagates the error through stochastic gradient descent (SGD) with 0.01 learning rate and 0.9 momentum. It is trained for 200 epochs, after each of those its metrics are computed on the validation set. Reaching minimum validation loss value prompts the saving of network weights.

Optimization hyper-parameters were chosen experimentally, in order to avoid training overfitting (e.g. due to higher learning rate values) and local minima of the loss function (e.g. due to lower learning rate values).

5 Results

We measured the network performance through a 5-fold cross-validation experiment. The original dataset was split into a test set, made up of 43 images that were selected according to the features distribution reported in Sect. 3, and a main set of 225 images to be split into training and validation subsets during the cross-validation process. It means that in each cross-validation fold, the model was trained on 180 images, validated on 45 to save the best performing set of weights, and, finally, tested on 43.

The metric we chose to evaluate the model performance is the accuracy, i.e. the ratio between the number of correctly predicted pixels within the annotated portion of the images and the total number of annotated pixels.

Figure 9 compares weighted loss and accuracy values during the five cross-validation folds, on both the training and the validation set. The accuracy plots suggest that the network can generalize the information learnt from the training set when processing the validation set.

A further evaluation of network generalization capabilities was performed on the above mentioned test set. The validation and test mean accuracy values

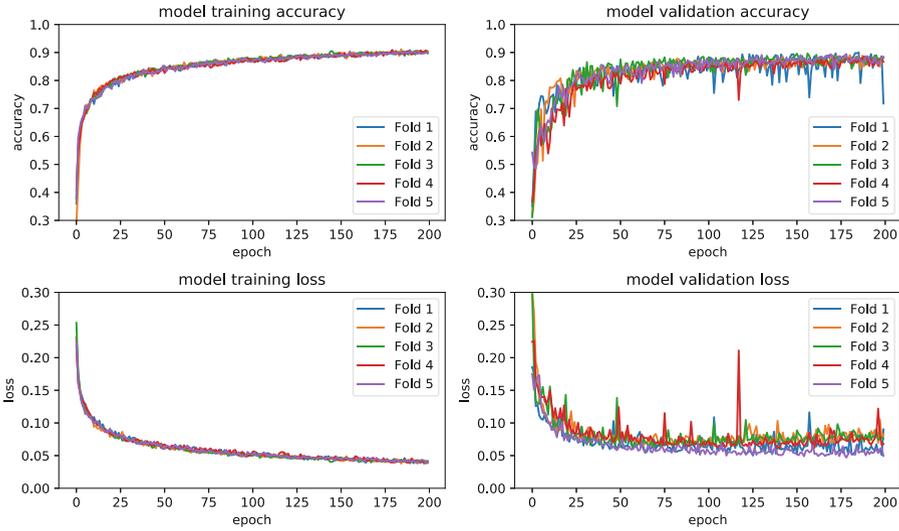


Fig. 9. Weighted cross-entropy loss and accuracy metrics of the proposed network on both the training and the validation sets.

were, respectively, 0.87 and 0.88, with standard deviations equal to 0.03 and 0.01. These values suggest that the network can generalize the information learnt and, even if the model weights are saved as soon as the minimum validation loss is met, the network does not overfit the validation set.

Figure 10 and 11 show two test images together with their ground-truth and network predictions.

Figure 10 reports an healthy sector of the irido-corneal interface. The corresponding segmentation of the regions annotated is very good. Some false positives are present outside the area of main interest. Since the training is performed using only the annotated part of the image and since the unannotated region is

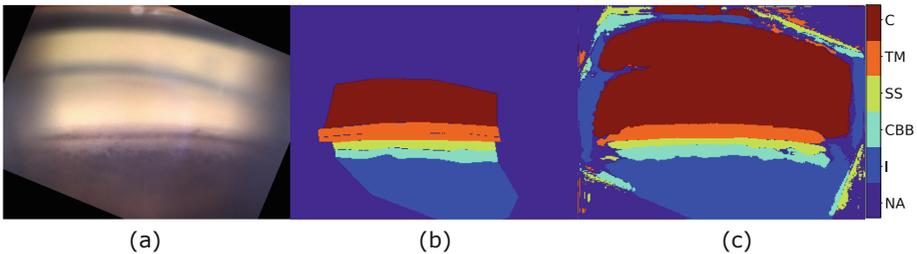


Fig. 10. (a) Healthy sector; (b) ground truth; (c) network prediction after 0.5 thresholding.

not an actual target class, currently we can not expect the model to correctly classify the image periphery.

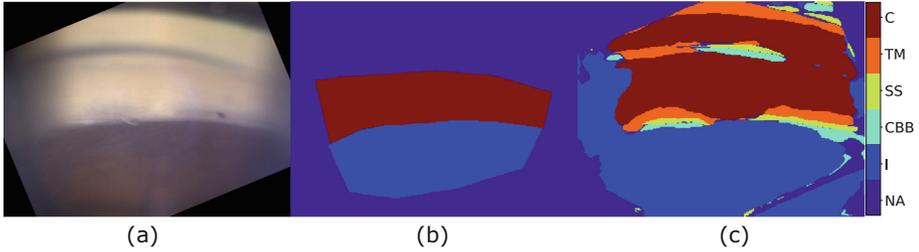


Fig. 11. (a) Unhealthy sector; (b) ground truth; (c) network prediction after 0.5 thresholding.

Figure 11 shows a pathological case where the cornea is directly in contact with the iris. The segmentation fails because of the incorrect classification of some shadows as targets. This is likely due to the limited amount of pathological images in the training dataset, as reported in Sect. 3.

Table 1. Class-specific precision and sensitivity on the test set.

Class	Precision	Sensitivity
Iris	99%	88%
Ciliary body band	52%	77%
Scleral spur	53%	83%
Trabecular meshwork	85%	85%
Cornea	95%	94%

Network precision and sensitivity mean values at classifying pixels into the five target classes were calculated over the test set during the cross-validation experiment and are reported in Table 1.

Precision and sensitivity values are high for three of the five target classes (iris, trabecular meshwork and cornea), while the precision values for the ciliary body band and the scleral spur are considerably lower, but in line with the inter-annotator agreement when labelling pixels that belong to these layers, as discussed in Sect. 4.

6 Discussion

Assessing the integrity of the structures visible in the irido-corneal interface via gonioscopy is very important both for evaluating risks of developing glaucoma

and for understanding its aetiology, but the current manual method has several limitations.

In this paper, we have presented the first DL semantic segmentation model developed to process gonioscopic images from a digital, semi-automatic instrument launched very recently, the NIDEK GS-1. Our aim is to assist the risk evaluation and the categorization of this pathology.

Our results seem very promising, especially considering the pilot-level size of our current image set and the considerable variation in the annotations from different experts. The system provides a meaningful computationally augmented visualization of the irido-corneal interface.

Our approach is currently limited by the modest size of the dataset, the limited representativeness of pathological data and the intrinsic difficulty, in some cases, of annotating the target structures consistently, an important observation that needs careful consideration in the assessment of the system performance. The current training strategy, which does not consider the unannotated area, considerably improves the classification accuracy within the central and more informative region of images but makes impossible to manage false positives in their periphery. Finding a suitable solution for overcoming this side effect is an important part of our future work.

Acknowledgement. We thank the CVIP/VAMPIRE research team of the University of Dundee, Dundee (UK) and the VAMPIRE research team of the University of Edinburgh, Edinburgh (UK) for useful discussions and suggestions.

Disclosure

This work is fully funded by a Ph.D. studentship from NIDEK Technologies Srl.

References

1. Abràmoff, M.D., et al.: Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Invest. Ophthalmol. Vis. Sci.* **57**(13), 5200 (2016). <https://doi.org/10.1167/iovs.16-19964>
2. Alward, W.: *Color Atlas of Gonioscopy*. American Academy of Ophthalmology, San Francisco (2008)
3. Ben-Cohen, A., et al.: Retinal layers segmentation using fully convolutional network in OCT images (2017)
4. Chandra, A., Gupta, A., Gupta, V., Sihota, R., Azad, R., Chandra, P.: Comparative evaluation of RetCam vs. gonioscopy images in congenital glaucoma. *Indian J. Ophthalmol.* **62**(2), 163 (2014). <https://doi.org/10.4103/0301-4738.116487>
5. Cheng, J., et al.: Closed angle glaucoma detection in RetCam images. In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology. IEEE, August 2010. <https://doi.org/10.1109/iembs.2010.5627290>
6. Dutta, A., Zisserman, A.: The VIA annotation software for images, audio and video. In: Proceedings of the 27th ACM International Conference on Multimedia, MM 2019. ACM, New York (2019). <https://doi.org/10.1145/3343031.3350535>

7. Fang, L., Cunefare, D., Wang, C., Guymer, R.H., Li, S., Farsiu, S.: Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search. *Biomed. Opt. Express* **8**(5), 2732 (2017). <https://doi.org/10.1364/boe.8.002732>
8. Fauw, J.D., et al.: Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat. Med.* **24**(9), 1342–1350 (2018). <https://doi.org/10.1038/s41591-018-0107-6>
9. He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In: *The IEEE International Conference on Computer Vision (ICCV)*, December 2015
10. Heckbert, P.: *Graphics Gems IV*. AP Professional, Boston (1994)
11. Hertzog, L.H., Albrecht, K.G., LaBree, L., Lee, P.P.: Glaucoma care and conformance with preferred practice patterns. *Ophthalmology* **103**(7), 1009–1013 (1996). [https://doi.org/10.1016/s0161-6420\(96\)30573-3](https://doi.org/10.1016/s0161-6420(96)30573-3)
12. Huang, G., Liu, Z., van der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE (2017). <https://doi.org/10.1109/cvpr.2017.243>
13. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift. In: *Proceedings of the 32nd International Conference on International Conference on Machine Learning, ICML 2015*, vol. 37, pp. 448–456 (2015). [JMLR.org](http://jmlr.org)
14. Jegou, S., Drozdal, M., Vazquez, D., Romero, A., Bengio, Y.: The one hundred layers tiramisu: fully convolutional DenseNets for semantic segmentation. In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July 2017
15. Li, Z., He, Y., Keel, S., Meng, W., Chang, R.T., He, M.: Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. *Ophthalmology* **125**(8), 1199–1206 (2018). <https://doi.org/10.1016/j.ophtha.2018.01.023>
16. Liu, X., et al.: Semi-supervised automatic segmentation of layer and fluid region in retinal optical coherence tomography images using adversarial learning. *IEEE Access* **7**, 3046–3061 (2019). <https://doi.org/10.1109/access.2018.2889321>
17. Orlando, J.I., et al.: U2-Net: a Bayesian u-net model with epistemic uncertainty feedback for photoreceptor layer segmentation in pathological OCT scans. In: *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE, April 2019. <https://doi.org/10.1109/isbi.2019.8759581>
18. Pead, E., et al.: Automated detection of age-related macular degeneration in color fundus photography: a systematic review. *Surv. Ophthalmol.* **64**(4), 498–511 (2019). <https://doi.org/10.1016/j.survophthal.2019.02.003>
19. Pekala, M., Joshi, N., Liu, T.A., Bressler, N., DeBuc, D.C., Burlina, P.: Deep learning based retinal OCT segmentation. *Comput. Biol. Med.* **114**, 103445 (2019). <https://doi.org/10.1016/j.compbiomed.2019.103445>
20. Raghavendra, U., Fujita, H., Bhandary, S.V., Gudigar, A., Tan, J.H., Acharya, U.R.: Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images. *Inf. Sci.* **441**, 41–49 (2018). <https://doi.org/10.1016/j.ins.2018.01.051>
21. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *MICCAI 2015*. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28

22. Roy, A.G., et al.: ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. *Biomed. Opt. Express* **8**(8), 3627–3642 (2017). <https://doi.org/10.1364/BOE.8.003627>. <http://www.osapublishing.org/boe/abstract.cfm?URI=boe-8-8-3627>
23. Tham, Y.C., Li, X., Wong, T.Y., Quigley, H.A., Aung, T., Cheng, C.Y.: Global prevalence of glaucoma and projections of glaucoma burden through 2040. *Ophthalmology* **121**(11), 2081–2090 (2014). <https://doi.org/10.1016/j.ophtha.2014.05.013>
24. Trucco, E., et al.: Validating retinal fundus image analysis algorithms: issues and a proposal. *Invest. Ophthalmol. Vis. Sci.* **54**(5), 3546 (2013). <https://doi.org/10.1167/iovs.12-10347>