Robust multi-view approaches for retinal layer segmentation in glaucoma patients via transfer learning

Mateo Gende[1,2]^, Joaquim de Moura[1,2]^, José Ignacio Fernández-Vigo[3]^, José María Martínez-de-la-Casa[3]^, Julián García-Feijóo[3]^, Jorge Novo[1,2]^, Marcos Ortega[1,2]^

[1]VARPA Group, A Coruña Biomedical Research Institute (INIBIC), University of A Coruña, A Coruña, Spain; [2]CITIC Research Centre, University of A Coruña, A Coruña, Spain; [3]Department of Ophthalmology, San Carlos Clinical Hospital, Madrid, Spain

Contributions: (I) Conception and design: All authors; (II) Administrative support: J de Moura, J Novo, M Ortega; (III) Provision of study materials or patients: J de Moura, J Novo, M Ortega, JI Fernández-Vigo, JM Martínez-de-la-Casa, J García-Feijóo; (IV) Collection and assembly of data: M Gende, JI Fernández-Vigo, JM Martínez-de-la-Casa, J García-Feijóo; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Joaquim de Moura. Universidade da Coruña, Campus de Elviña, s/n, 15071, A Coruña, Spain. Email: joaquim.demoura@udc.es.

Background: Glaucoma is the leading global cause of irreversible blindness. Glaucoma patients experience a progressive deterioration of the retinal nervous tissues that begins with a loss of peripheral vision. An early diagnosis is essential in order to prevent blindness. Ophthalmologists measure the deterioration caused by this disease by assessing the retinal layers in different regions of the eye, using different optical coherence tomography (OCT) scanning patterns to extract images, generating different views from multiple parts of the retina. These images are used to measure the thickness of the retinal layers in different regions.

Methods: We present two approaches for the multi-region segmentation of the retinal layers in OCT images of glaucoma patients. These approaches can extract the relevant anatomical structures for glaucoma assessment from three different OCT scan patterns: circumpapillary circle scans, macular cube scans and optic disc (OD) radial scans. By employing transfer learning to take advantage of the visual patterns present in a related domain, these approaches use state-of-the-art segmentation modules to achieve a robust, fully automatic segmentation of the retinal layers. The first approach exploits inter-view similarities by using a single module to segment all of the scan patterns, considering them as a single domain. The second approach uses view-specific modules for the segmentation of each scan pattern, automatically detecting the suitable module to analyse each image.

Results: The proposed approaches produced satisfactory results with the first approach achieving a dice coefficient of 0.85±0.06 and the second one 0.87±0.08 for all segmented layers. The first approach produced the best results for the radial scans. Concurrently, the view-specific second approach achieved the best results for the better represented circle and cube scan patterns.

Conclusions: To the extent of our knowledge, this is the first proposal in the literature for the multi-view segmentation of the retinal layers of glaucoma patients, demonstrating the applicability of machine learning-based systems for aiding in the diagnosis of this relevant pathology.

Keywords: Computer-aided diagnosis (CAD); optical coherence tomography (OCT); glaucoma; deep learning; segmentation

^ ORCID: Mateo Gende, 0000-0003-1686-7189; Joaquim de Moura, 0000-0002-2050-3786; José Ignacio Fernández-Vigo, 0000-0001-8745-3464; José María Martínez-de-la-Casa, 0000-0001-9441-0542; Julián García Feijóo, 0000-0002-7772-5718; Jorge Novo, 0000-0002-0125-3064; Marcos Ortega, 0000-0002-2798-0788.

Introduction

According to the 2019 World Report on Vision published by the World Health Organization (1), glaucoma is the second leading cause of blindness overall, and the first cause of irreversible vision loss. This pathology starts with a progressive loss of peripheral vision. This gradual process may start to affect the central visual field and cause severe vision impairment. Since glaucoma is difficult to detect or even asymptomatic in its early stages, it may not be noticed before it causes irreversible vision loss, leading to significant under-diagnosis and over-treatment of the disease (2).

Glaucoma is characterised by the deterioration of the nervous fibres responsible for transmitting neural impulses from the photosensitive cells in the eye to the visual cortex. This process is typically caused by an increase in intraocular pressure. This pressure deforms the optical nerve head (ONH), damaging the nervous fibres and leading to a loss of vision (3). Other forms of glaucoma include angle-closure glaucoma, which progresses much faster, or normal tension glaucoma, in which there is no appreciable increase in intraocular pressure. Due to the slow progression of this disease, coupled with the variety of forms it may present and the irreversibility of its effects over the visual field, an early and accurate assessment of the glaucoma is necessary in order to preserve patient vision.

In the context of healthcare, the domain of computer-aided diagnosis (CAD) systems is rapidly advancing, thanks to the development of new and improved machine learning algorithms coupled with more accessible high performance computing architectures. These systems provide assistance to healthcare workers in a variety of tasks by means of image classification, segmentation or the extraction of biological markers. In this sense, CAD systems based on deep learning are achieving remarkable progress thanks to the ability of convolutional neural networks to automatically extract and select visual features that are relevant to the diagnosis task (4). In this way, the systems can be developed directly by having convolutional models learn visual patterns from annotated data. Furthermore, the visual patterns learned by a model trained on a particular domain where data is readily available (source domain) can be used to develop a model intended to work in a similar domain where the annotated data is scarcer (target domain), a technique known as transfer learning (5,6). This data scarcity can be caused by several factors such as data sensitivity and privacy, difficulty or subjectivity in annotation and high costs associated to data acquisition. All of these are factors that particularly affect medical imaging. There are different approaches to transfer learning, such as selectively freezing some of the model layers and fine-tuning the remaining ones, or pre-training a model in the source domain and then re-training all of the layers using the pre-training as an initialisation, which can improve model generalisation (7).

Several computer-aided approaches exist for the assessment of glaucoma. These are mainly aimed at evaluating the damage to the neural fibres of the eye, by either assessing secondary signs such as the cupping of the ONH or by measuring changes and deviations to the retinal layer structure. Out of these, approaches related to measuring the cupping of the ONH in retinal fundus images are the most prevalent. By measuring the size of the optic disc (OD) and the optic cup in retinal fundus images, a cup-to-disc ratio can be extracted to measure the effects of the pressure on the ONH. In this line, many works in the literature have tackled the automatic segmentation of the OD and cup in this image modality by means of computer vision or machine learning [for reference (7-11)]. A recent work in retinal fundus images has also employed transfer learning for the segmentation of the OD, optic cup and fovea (11), highlighting the advantages of using transfer learning for related ophthalmological diagnosis tasks with results comparable with trained experts.

Other works focus on the analysis of the retinal layers in optical coherence tomography (OCT) images. As opposed to retinal fundus images, which only show a view of the surface of the tissue, OCT can produce volumes of data which visualise the histological structure of the retina in cross-sectional images. This allows direct measurements of the thickness of the layers at different points in the retina. In order to better assess the glaucoma-related damage that a patient may present, ophthalmologists often use different patterns, generating multiple views to scan different regions of the eye. The most prevalent of these views are circumpapillary circle scans around the ONH, macular cube scans around the macula and radial scans centred on the OD.
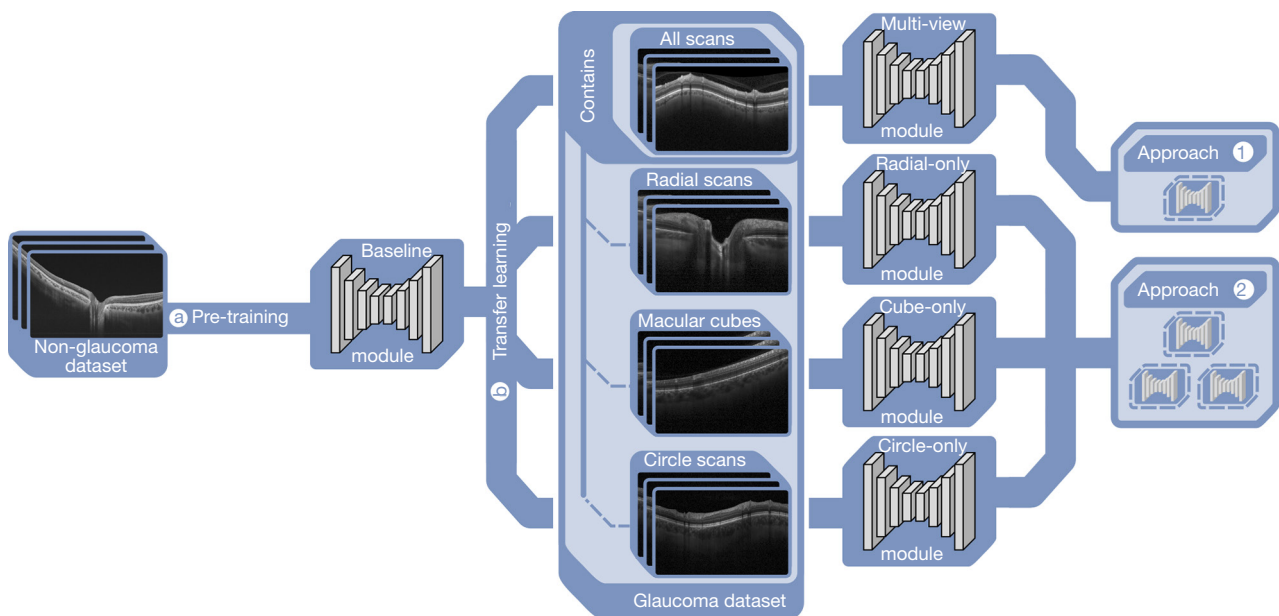
Figure 1 Summary of the transfer learning methodology employed to train the different segmentation modules. A baseline module is pre-trained on a non-glaucomatous dataset. This model is used for weight initialisation of all of the modules employed in the two proposed approaches. These are trained in a multi-view or view-specific set of images.

Circumpapillary scans allow an assessment of nervous degeneration close to the optic nerve in all directions. Macular scans can provide a visualisation of the area responsible for fine-detail vision, with the biggest impact on patient quality of life. Finally, OD radial scans allow the analysis of ONH deformations and degeneration related to glaucoma. These OCT views allow a more comprehensive assessment of the disease than retinal fundus images, providing insight into how the different anatomical structures of the eye may be affected by glaucoma.

Out of all the retinal layers, the inner-most layers of the retina: the retinal nerve fibre layer (RNFL), ganglion cell layer (GCL) and inner plexiform layer (IPL) show the best discriminative power for glaucoma detection (12). Among these, the RNFL around the circumpapillary region has shown great potential for discrimination (13). The automatic detection and segmentation of these layers can be approached with different classical digital image processing techniques, such as the use of graph search algorithms (14-16), active contour models (17-19), level sets (20), geodesic distance (21), or fuzzy histogram hyperbolisation (22). Other, more recent works have made use of machine learning algorithms for the retinal layer segmentation, more specifically convolutional neural networks architectures. However, these approaches are focused on the segmentation of the retinal layers

of patients suffering from peripapillary atrophy and cataract (23), multiple sclerosis (24), age-related macular degeneration (25,26), central serous retinopathy (24,27), or diabetic retinopathy (24,26,28). Fewer works have approached the problem of segmenting the layers of the retina in glaucoma patients, attending to the disease-specific degeneration that affects them. Even then, these are limited to a single scan pattern, namely the circumpapillary scan (28-31) or the ONH (32), while foregoing the analysis of the other relevant views for the diagnosis of glaucoma. In this sense, to the best of our knowledge, the segmentation of the relevant retinal structures for the diagnosis of glaucoma in all of the OCT views used for its assessment remains to be addressed.

In this work, we present two approaches for the fully automatic, multi-view segmentation of the retinal layers in OCT images of glaucoma patients. By making use of transfer learning (*Figure 1*), these approaches are able to make the most of the available data in order to provide a robust segmentation of the anatomical regions of relevance for glaucoma diagnosis: the RNFL, retinal layer with the highest reported discriminative power for glaucoma diagnosis; the inner retina, which comprises the ganglion cell bodies and the neural connections of the photosensitive cells; and the outer retina, containing the bodies of the

photoreceptor cells as well as the remaining retinal layers. This proposal allows the segmentation of these layers from the three most prevalent scans for glaucoma diagnosis: circular peripapillary scans showing the thickness of retinal layers around the periphery of the OD, and allowing an evaluation of early signs of the disease; parallel macular cube scans, which allow an assessment of severe cases of glaucoma with deterioration that may affect the central visual field; and radial scans extracted from the ONH, showing the morphology of the OD and deformations close to the area most affected by glaucoma. Furthermore, two different approaches are proposed for the multi-region segmentation of the layers. The first of these approaches consists in a common module for the segmentation of the layers in all three considered views, taking advantage of visual patterns that may be shared between the segmented structures in all three views. The second one takes a more discrete approach by using a classification network to automatically discern between the different scan types and use a view-specific module to analyse images from each of the scan types. These two approaches have different advantages in terms of data availability and performance for each of the considered scan types. To the extent of our knowledge, this work is the first in the literature to propose the multi-view segmentation of retinal layers for patients of glaucoma. By providing a robust and fully automatic segmentation of the retinal layers in the views most commonly used for glaucoma assessment, the clinical diagnosis process for this disease can be greatly simplified and performed in an objective and less time-consuming manner, helping to provide an early diagnosis and preserve patient vision and quality of life.

This manuscript is structured as follows: In the Methods section we provide a description of the dataset that was employed, as well as a detailed explanation of the proposed multi-region segmentation approaches. The Results section presents the experimental results that were obtained, with the Discussion section offering a discussion of said results. Finally, in the Conclusions section we present the conclusions that were reached and describe possible future lines of work.

Methods

In order to facilitate the replicability of this proposal, in this section we explain the dataset employed in the development of this work (Dataset subsection), its annotation (Annotation subsection), as well as the segmentation methodology that

was followed (Methodology subsection). The technical details of model training and transfer learning can be found in the Training details section.

Dataset

Two datasets were used in this work, the first dataset (22) was used for training the baseline modules employed in the transfer learning process. This public dataset contains 732 images from 61 subjects. They were acquired using a Topcon DRI OCT-1 Atlantis in an original resolution of 1,024×992. It contains cases of high myopia, peripapillary atrophy and cataract. The second collected dataset was used to train and validate the proposed approaches, as well as to compare to the baseline in a glaucomatous domain. The collected glaucomatous dataset consists of a total of 811 OCT images. These were acquired with a Heidelberg SPECTRALIS® optical imaging platform. The images were inspected to ensure that the retinal layers were visible and any patient displaying signs of any non-glaucoma-related lesions were not considered for inclusion. The final set of images was collected from the eyes of 11 different glaucoma patients. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

The distribution of the images in terms of the scan modality is as follows:

❖ Circle class: 139 images. These are circumpapillary circular scans taken around the OD at different distances, with a resolution of 768×496 pixels. This view is most commonly used to estimate the thickness of the retinal layers at different sectors around the OD. This allows the assessment of early glaucoma-related damage.

❖ Cube class: 540 images. Images from the cube class are parallel scans of the macular region. These images have a resolution of 512×496 pixels. The macular cube scan pattern allows the measurement of the thickness of retinal layers close to the area responsible for high-acuity vision. Glaucoma-related damage to the light-sensitive and ganglion cells of this area can result in a deterioration of the central visual field and a considerable loss of vision.

❖ Radial class: 132 images. Images belonging to this class are taken radially centred on the OD. These have a resolution of 768×496 pixels. The radial scan is used to assess the morphology of the OD, as well as to measure surrounding retinal layer thickness. This allows an evaluation of changes to the shape

2850

Gende et al. Multi-view approaches for layer segmentation in glaucoma

Table 1 Glaucoma dataset details displaying the number of images contained of each class per patient

Patient	Circle	Cube	Radial
Patient 1	6	94	0
Patient 2	15	117	20
Patient 3	8	50	0
Patient 4	24	0	0
Patient 5	9	38	0
Patient 6	12	83	0
Patient 7	19	0	0
Patient 8	14	0	0
Patient 9	11	0	57
Patient 10	12	0	32
Patient 11	9	158	23
Total per scan	139	540	132
Total		811	

of the OD caused by an increase in ocular pressure or other glaucoma-related causes.

The distribution of the images per patient is displayed in *Table 1*.

These images were annotated indicating the RNFL, the inner retina and the outer retina, which are detailed in the next section.

Annotation

The collected glaucoma dataset was annotated by an expert in ophthalmological imaging using the computer vision annotation tool (CVAT) (publicly available online: https:github.com/opencv/cvat). The annotation process was carried out in a sequential manner so that each image in a volume or patient was annotated in a consecutive manner, ensuring that the annotations were coherent among parallel or consecutive scans. The layers were annotated using the polygonal drawing tool, reusing the keypoint nodes between layers to ensure that no gaps were left. The images were annotated with the following layers of relevance for glaucoma diagnosis:

❖ RNFL: the RNFL layer is formed by the fibres of the retinal nerve cells which allow the connection of the photoreceptor cells to the ONH. It can be found between the inner limiting membrane which

separates the vitreous from the retina and the GCL. Damage of this layer is heavily associated with loss of vision caused by glaucoma, and has been reported to have great discriminatory power for glaucoma diagnosis (13).

❖ Inner retina: this region can be found between the RNFL and the outer nuclear layer, comprising the GCL, the IPL and the inner nuclear layer as well as the outer plexiform layer. This region contains the ganglion cells as well as the neural connections of the photoreceptor cells. Although not as heavily associated to glaucoma as the RNFL, the inner retina contains the parts responsible of transmitting the impulses detected by the photoreceptor cells to the RNFL, and a glaucoma-related deterioration of this layer can have a significant impact on vision.

❖ Outer retina: this region is situated between the outer plexiform layer and the choroid, formed by the outer nuclear layer, the photoreceptor cells, and the retinal pigment epithelium. This layer contains the bodies of the photosensitive cells responsible for vision. Glaucoma-related damage or deterioration of these cells can have a direct impact patient sight.

❖ OD: the OD only appears in the radial scan images. This region is formed by the bundles of nerve fibres in the ONH. Although not a retinal layer per se, deformations in the OD can be indicative of glaucoma-related deterioration and can translate into damage to other related retinal layers. Furthermore, the OD is a blind spot in the eye, with no photosensitive cells or retinal layers situated beneath.

An illustration displaying the general appearance of the three views with their corresponding segmentation annotations can be found in *Figure 2*.

Methodology

As mentioned in the Introduction, two different approaches for the multi-view segmentation of the retinal layers were explored. These share the same internal segmentation architecture, with each of them having a different focus.

Retinal layer segmentation
The segmentation of the retinal layers was modelled as a two-stage method based on the MGU-Net architecture (22). This architecture has demonstrated its ability to outperform
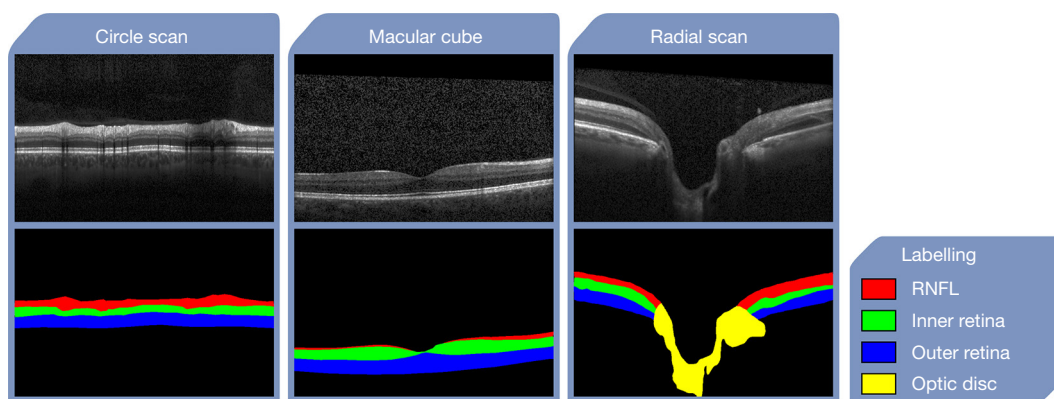
Figure 2 Examples of images belonging to the three views with their corresponding segmentation annotations. RNFL, retinal nerve fibre layer.
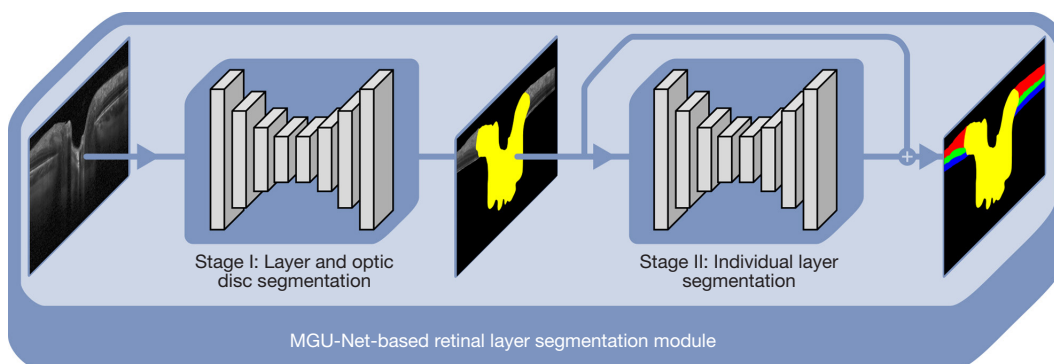

Figure 3 Summary of an MGU-Net based retinal layer segmentation model.

other similar architectures such as U-net (33), ReLayNet (27) or DRUNET (32) in terms of their ability to segment retinal layers, demonstrating its suitability for this task. The first stage consists in a preliminary segmentation of the OD and the retinal layers as a whole. The aim of this stage is the removal of the background (formed by the vitreous humour and the choroid) as well as the OD from the image that is analysed by the second stage of the architecture. The second stage performs the actual segmentation of each of the considered anatomical structures from the retina, using the pre-segmented image from the first stage as input. A summary of this architecture can be found in *Figure 3*.

Internally, each stage is performed by a MGU-Net. Each of these is in turn formed by an encoder and decoder parts with a multi-scale global reasoning (34) module in-between. This module uses multiple effective receptive fields in order to lessen the influence of the differences in thickness of the various segmented anatomical structures. After the two

stages are complete, the final segmentation is reconstructed using the output from both of the stages.

The MGU-Net architecture is used as the basis of all the segmentation modules employed in this work, while two different approaches were explored in order to study how to better adapt to the multiple views employed for the clinical assessment of glaucoma. These two approaches focus on different aspects of the adaptation to the particular visual features of each OCT view. The aim is to determine whether sharing said common features can improve the robustness of the segmentation, or whether the use of view-specific modules can better exploit the anatomic particularities of each specific view. These approaches are detailed hereunder.

First approach: single multi-view module

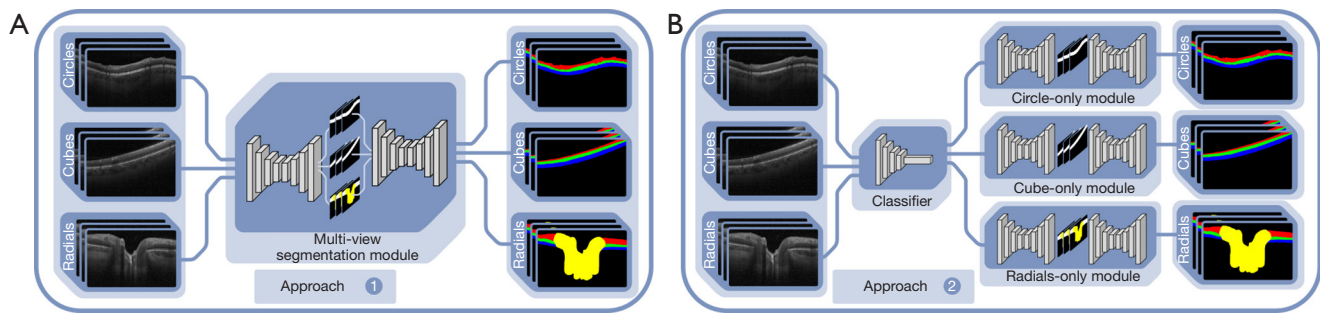The first of the proposed approaches takes on an extensive focus, using a single MGU-net based segmentation

Figure 4 The two multi-view segmentation approaches proposed in this work. (A) Approach 1, use of a single segmentation module for images belonging to the three considered views. (B) Approach 2, use of a classifier to select the appropriate architecture to segment each of the considered views separately.

module for the three considered views. This module is trained to segment images from each of these regions as if they belonged to a singular domain (*Figure 4A*). The motivation behind this approach is to explore whether the visual similarities in the appearance of the retinal layers between views can be useful to the segmentation. This way, the module can be trained to segment the anatomical structures in a view and take advantage of that training in order to segment the same structures in the similar domain of a different scan pattern. However, the use of a single architecture for the segmentation of the three considered views forces the model to adapt to all three different scan patterns at once. While visually similar, the anatomical structure of the segmented layers varies greatly between different views.

Second approach: view-specific modules

The second approach is complementary to the first, taking on a discrete focus, discriminating between the different scan patterns and using a view-specific segmentation module for each of them. The process of segmenting an OCT image through this approach can be understood as a two-stage procedure: first, a classifier is used to determine the correct segmentation module to use. Next, the image is segmented using the selected segmentation module (*Figure 4B*). Rather than forcing the modules to generalise for all of the scanning presets that are considered, this approach enables each of them to be fine-tuned to the appearance of the anatomical regions displayed in each view. This allows a finer adjustment of each module to their corresponding domain, without having to consider the visual features that each layer presents in any of the other scanning patterns and freeing each module to better fit the distinctive

characteristic of each anatomical structure in the different parts of the eye.

Training details

First, an MGU-Net module was trained on the training partition of its original dataset, which did not contain glaucoma patients (22). This model was used to set a baseline and to perform transfer learning on to develop the segmentation modules used in the two proposed approaches. The model was trained using the conditions described in the original publication, using an Adam (35) optimiser with a learning rate of 0.001, dynamically reduced by an order of magnitude every 20 epochs, with $\beta_1=0.9$ and $\beta_2=1\times10^{-4}$, for a maximum of 50 epochs. The loss that was used for training was a sum of Cross-Entropy and Dice loss, doubling the weight of the second stage loss.

To train and validate the segmentation modules, the collected glaucoma dataset was partitioned following a leave-one-out cross-validation at the patient level. In each partition, the images belonging to one patient were set aside as the test set, and the images from the remaining patients were further partitioned, using 80% for training and 20% for validation. Following this leave-one-out cross-validation, we ensure that each patient is equally represented in testing for all of the segmentation modules, and that no module can see images of the test patient in its corresponding training and validation sets.

Transfer learning was performed by initialising each model to the baseline weights that were pre-trained on the non-glaucomatous dataset (22). Then, the initialised models were allowed to re-train in their corresponding training sets while taking a snapshot of the weights at each
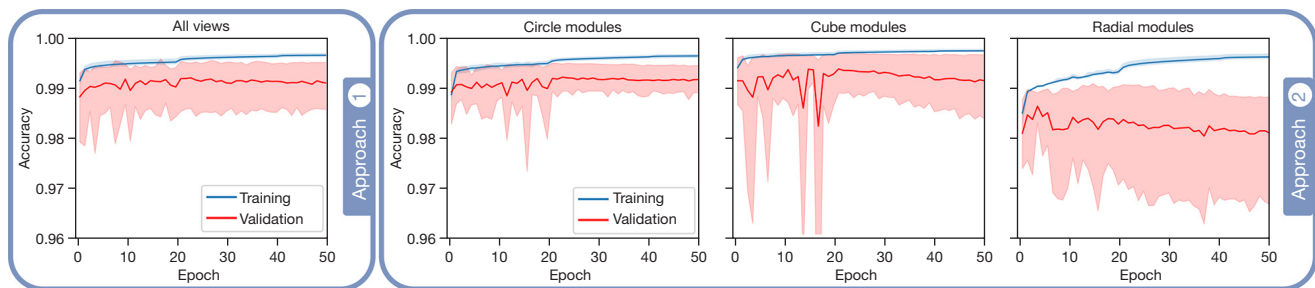

Figure 5 Average training and validation accuracy evolution for each of the trained modules of the first approach and the scan pattern-specific modules for the second approach.

epoch in such a manner that the models can be allowed to adapt to the specific domain of the glaucomatous dataset. Once this training was over, the checkpoint with the lowest validation loss was selected for testing. This way, for the first approach, all of the images in each of the sets were gathered, and transfer learning was performed on a module for each of the 11 partitions.

For the second approach, transfer learning was performed on each scan pattern-specific module separately using only the images belonging to the corresponding scanned region. Additionally, for this second approach, a series of classification models were trained and validated separately. In order to evaluate how different models performed on this task, three different architectures were selected for this: a DenseNet-121 (36), a ResNet-18 and a ResNet-34 (37). These architectures have seen widespread use in similar medical imaging classification and segmentation tasks with remarkable results (38-41).

Regarding the details of the transfer learning process, Adam (35) was used for model optimisation, with identical parameters to the baseline training. All of the segmentation models were trained using a sum of cross-entropy and dice loss, doubling the weight of the second stage loss of each segmentation module. These models used a mini-batch size of 1 image and were allowed to train for a maximum of 50 epochs. The classification models were trained on the collected dataset in a 5-fold cross-validation, using Cross-Entropy loss, with 16 images being used per batch. These models were allowed to train for a maximum of 100 epochs. Additionally, data augmentation was employed in the form of horizontal flipping during the training.

Results

In this section we describe the results that were obtained by performing transfer learning on the modules that constitute each of the proposed approaches. The metrics used for this evaluation were accuracy (Eq. [1]), precision (Eq. [2]), recall (Eq. [3]), and the dice coefficient and F1 score (Eq. [4]).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad [1]$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad [2]$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad [3]$$

$$\text{Dice} = \text{F1} = \frac{2 \times TP}{2 \times TP+FP+FN} \quad [4]$$

The training and validation accuracy evolution during the transfer learning process for the modules of both of the proposed segmentation approaches can be found in *Figure 5*. After the transfer learning process, the models were evaluated on their respective test sets and compared with the test results produced by the baseline which trained on a non-glaucomatous dataset (*Table 2*). A breakdown of the dice score achieved by each of the proposed approaches for each scan pattern can be found in *Table 3*. The metrics for all layers were calculated as the micro-average of all of the segmented layers. Complementarily, *Figure 6* displays some examples of the segmentation results produced by the baseline and the two proposed approaches for each of the scan patterns. Regarding the classification stage of the second approach, the results that were obtained are displayed in *Table 4*. In light of the relative similarity in terms of results, the ResNet-18 architecture was selected to be used as the classifier for the second approach due to its simplicity in terms of trainable parameters and efficiency, a desirable trait for the deployment of CAD models in clinical practice.

Discussion

The results that were obtained highlight the substantial

Table 2 Precision and recall achieved by each of the proposed approaches, compared with the baseline non-glaucomatous training

Location	Baseline			Approach 1			Approach 2		
	Precision	Recall	Dice	Precision	Recall	Dice	Precision	Recall	Dice
RNFL	0.67±0.12	0.67±0.14	0.67±0.13	0.73±0.13	0.90±0.06	0.79±0.08	0.87±0.09	0.82±0.14	0.84±0.10
Inner retina	0.70±0.13	0.75±0.15	0.72±0.13	0.85±0.08	0.82±0.10	0.83±0.08	0.84±0.14	0.88±0.07	0.85±0.11
Outer retina	0.89±0.08	0.60±0.15	0.71±0.13	0.92±0.08	0.89±0.07	0.91±0.04	0.92±0.07	0.90±0.06	0.91±0.05
All layers	0.76±0.10	0.67±0.15	0.71±0.13	0.85±0.07	0.87±0.06	0.85±0.06	0.88±0.09	0.87±0.07	0.87±0.08

Values are presented as mean ± standard deviation. Approach 1: single multi-view module; Approach 2: view-specific modules. RNFL, retinal nerve fibre layer.

Table 3 Detailed breakdown of the dice coefficient achieved by the proposed approaches for each of the scan patterns

Location	Approach 1			Approach 2		
	Circle	Cube	Radial	Circle	Cube	Radial
RNFL	0.86±0.05	0.76±0.13	0.76±0.06	0.89±0.05	0.83±0.10	0.73±0.05
Inner retina	0.87±0.05	0.83±0.12	0.73±0.05	0.90±0.04	0.91±0.04	0.71±0.09
Outer retina	0.90±0.04	0.93±0.05	0.86±0.05	0.93±0.04	0.93±0.04	0.77±0.07
All layers	0.88±0.04	0.86±0.08	0.78±0.05	0.91±0.04	0.91±0.04	0.74±0.07

Values are presented as mean ± standard deviation. Approach 1: single multi-view module; Approach 2: view-specific modules. RNFL, retinal nerve fibre layer.
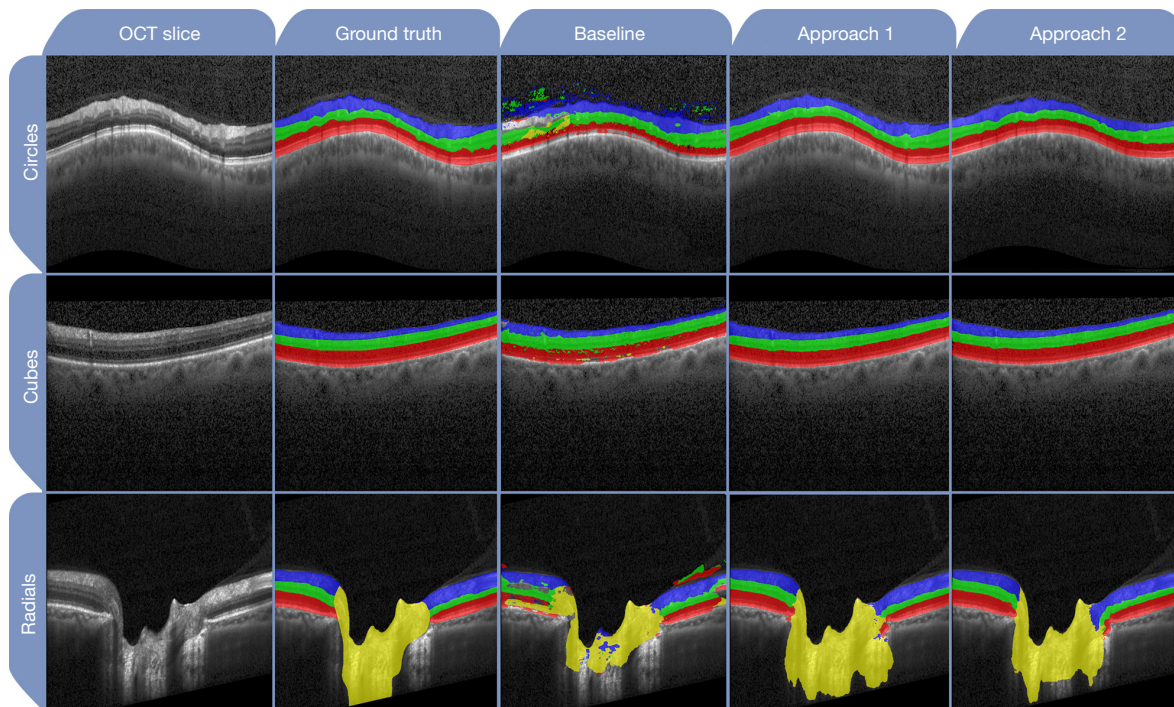

Figure 6 Example segmentation results of the different scan patterns. Approach 1: single multi-view module; Approach 2: view-specific modules. OCT, optical coherence tomography.

Table 4 Precision, recall and F1-score achieved by the classifiers used for the second approach

Variable	Precision	Recall	F1-score
DenseNet-121	0.98±0.02	0.98±0.02	0.98±0.02
ResNet-18	0.99±0.01	0.99±0.01	0.99±0.01
ResNet-34	0.99±0.01	0.99±0.01	0.99±0.01

Values are presented as mean ± standard deviation.

differences that may be found between images acquired with the same modality but extracted from different locations in the eyes, as well as the effect that these differences can have on the performance of the segmentation. These results indicate that while a model may be trained for retinal layer segmentation in OCT images extracted from non-glaucomatous patients, this training may not be sufficient to adjust the model to perform the same task in patients suffering from glaucoma, with the consequent decrease in performance. Visually, this can be seen in *Figure 6*, where the baseline fails to provide consistent layers, instead segmenting part of the speckle noise as retinal tissue while the proposed approaches provide an accurate segmentation of the retinal layers, with some variability being found on the OD due to the inherent uncertainty of its boundaries. When comparing the results of the baseline model, we can see that the application of transfer learning using glaucomatous patients translates into an increase in precision from 0.76 to 0.85 and in Recall from 0.67 to 0.87 for all layers. This supposes an overall increase in dice coefficient from a baseline 0.71 to 0.85 for the first approach which also processes images from the three scan patterns jointly.

When comparing the two proposed approaches among themselves, we can see that the first approach achieves a dice coefficient of up to 0.85±0.06 for all layers, while the second approach reaches up to 0.87±0.08 overall. However, this increase in performance does not translate equally to all of the scan patterns. From the results in *Table 2*, we can see that the view-specific segmentation outperforms the single-model approach for all the layers in the circle and cube classes, but not the radial class. Although the view-specific segmentation modules seem to adapt better to these first two classes than using a single module for all views, the benefit from learning visual patterns from the other views seems to benefit the radial class. This class is the least represented in this dataset, with only 132 images. Furthermore, this scan pattern shows the greatest variability

of the three, with the morphology of the OD varying greatly between patients, which can aggravate this data scarcity. This lack of representation for the radial class can be addressed in different forms. Aside from incrementing the overall amount of images belonging to this under-represented class, generative adversarial models could be used to generate more images that display the visual features that are not present in the original images. Furthermore, more specific approaches could be taken to this approach targeting the specific morphology of the OD, using the information contained in the A-scans to better delineate this region.

In order to validate the use of transfer learning that was used to leverage the data that was available, an experiment was performed in which the models were trained without an initialisation to the pre-trained weights of the baseline. A comparison between these models and the ones that were pre-trained can be found in *Table 5*. These results indicate that transfer learning can be especially useful for the least represented classes, when compared with the models that were not initialised to the baseline. These results show that after a pre-training on a source domain such as peripapillary radial scans belonging to patients suffering from other diseases, transfer learning can be used to increase the performance in the similar target domain of OCT images acquired with other scanning patterns such as circular or macular cube of patients suffering from glaucoma.

Overall, both approaches provide satisfactory results for the segmentation of the layers in the three scan patterns, with the purpose-specific approach providing better results in general, and the multi-view approach showing better performance for the under-represented class that displays the greatest variability in the dataset. Due to the absence of other publicly available glaucoma-related OCT retinal segmentation datasets, it was not possible to compare this methodology with other existing methods using the same test data. *Table 6* displays a comparison with other works aimed at the segmentation of retinal layers in glaucomatous images, focused on a single OCT view. While this comparison is not done under equal criterions, the results achieved by the proposed approaches are in line with the ones achieved by other works in the literature, while able to simultaneously segment all the views relevant for clinical assessment of glaucoma.

Conclusions

Glaucoma is the global leading cause of irreversible vision

Table 5 Dice coefficient comparison for the use of transfer learning for all layers in the two proposed approaches

Variable	Approach 1	Approach 2		
	All views	Circle	Cube	Radial
No pre-training	0.87±0.04	0.91±0.05	0.91±0.04	0.63±0.16
With pretraining	0.85±0.06	0.91±0.04	0.91±0.04	0.74±0.07

Values are presented as mean ± standard deviation. Approach 1: single multi-view module; Approach 2: view-specific modules.

Table 6 Reported dice coefficient for the segmentation of the RNFL of other proposals compared to the results achieved in this work

Variable	All views	Radial	Cube	Circle
DRUNET (32)	–	0.92±0.03	–	–
CCU (15)	–	–	0.89	–
H-DLpNet (31)	–	–	–	0.90
Berenguer-Vidal *et al.* (30)	–	–	–	0.93
Baseline	0.67±0.13	0.59±0.04	0.77±0.07	0.68±0.13
Approach 1	0.79±0.08	0.76±0.06	0.76±0.13	0.86±0.05
Approach 2	0.84±0.10	0.73±0.05	0.83±0.10	0.89±0.05

Only the baseline, Approach 1 and Approach 2 share the same test dataset. Values are presented as mean ± standard deviation, when available. Approach 1: single multi-view module; Approach 2: view-specific modules. RNFL, retinal nerve fibre layer.

loss, requiring an early diagnosis in order to preserve sight. This diagnosis is usually done by detecting damage to the retinal layers in different parts of the eye, using different OCT scan patterns for this task. We present two different approaches for the multi-region segmentation of the retinal layers in OCT images of glaucoma. These approaches use state of the art segmentation modules in order to achieve a robust and accurate segmentation of three different anatomical structures of the retina of relevance for glaucoma assessment: the RNFL which has the highest discriminative power for glaucoma diagnosis, and also the most affected layer in this disease; the inner retina, containing the ganglion cells and the structure necessary for connecting the photosensitive cells to the RNFL; and the outer retina, comprising the photosensitive cells which enable vision. By making use of transfer learning from a similar non-glaucomatous domain, these approaches enable the fully automatic analysis of images from the three most commonly employed scan patterns. These are extracted from a different part of the eye, allowing a comprehensive assessment of the whole retinal anatomy.

The results that were obtained demonstrate an increase in performance thanks to the application of transfer learning to take advantage of the visual features that may

be learned from a similar source domain when applied to the target domain of glaucomatous OCT images from various anatomical regions. Furthermore, the two proposed approaches explore the different benefits that may be gained from sharing the visual features learned from different views or the use of view-specific modules fine-tuned to the segmentation in each scan pattern. The results show that the first general approach may benefit the less represented classes which show greater variability, allowing the visual features learned from other regions to supplement the lack of available visual data. Meanwhile, the view-specific approach can provide a more accurate segmentation of the retinal layers in the well-represented views. Overall, the proposed approaches provide a robust and accurate segmentation of the retinal layers in OCT images of glaucomatous eyes. To the best of our knowledge, this proposal is the first in the literature to address the segmentation of retinal layers in multiple anatomical regions and using different OCT scan patterns. By providing a robust and objective segmentation in the views that are more commonly used in clinical practice, the diagnosis process of glaucoma can be improved and simplified, helping to preserve patient vision and quality of life.

As future work, we plan to supplement the collected

dataset with additional images in order to better represent the more variable classes. Moreover, the effects of selectively freezing some of the retinal layers before fine-tuning the rest of the network merits further study. Furthermore, these approaches could be extended to the analysis of other relevant anatomical structures or pathological signs of additional relevant ocular diseases in domains where more data is available and allowing a fairer comparison from which more conclusions can be drawn regarding its suitability for the multi-view analysis in other similar domains.

Acknowledgments

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-22-959/coif). JIFV reports that he receives honoraria from Bayer, Roche, Abbvie, Brill Pharma and Novartis. JMMdlC reports that he receives honoraria from Santen and Abbvie. JGF reports that he receives honoraria from Santen, Abbvie and Glaukos. MG, JdM, JN and MO report that their institution receives funding from Consellería de Cultura, Educación, Formación Profesional e Universidades, Xunta de Galicia and the European Regional Development Fund. All authors report that their institution receives funding from Ministerio de Ciencia e Innovación y Universidades, Government of Spain. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. World Health Organization. World Report on Vision. World Health Organization; 2019. Available online: https://www.who.int/publications/i/item/9789241516570 (accessed on 01/04/2022).
2. Vaahtoranta-Lehtonen H, Tuulonen A, Aronen P, Sintonen H, Suoranta L, Kovanen N, Linna M, Läärä E, Malmivaara A. Cost effectiveness and cost utility of an organized screening programme for glaucoma. Acta Ophthalmol Scand 2007;85:508-18.
3. Casson RJ, Chidlow G, Wood JP, Crowston JG, Goldberg I. Definition of glaucoma: clinical and experimental concepts. Clin Exp Ophthalmol 2012;40:341-9.
4. Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. IEEE Trans Med Imaging 2016;35:1285-98.
5. Pan SJ, Yang Q. A Survey on Transfer Learning. IEEE Trans Knowl Data Eng. 2010;22:1345-59.
6. Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? Adv Neural Inf Process Syst. 2014; 27. Available online: https://proceedings.neurips.cc/paper/2014/file/375c71349b295fbe2dcdca9206f20a06-Paper.pdf
7. Fu H, Cheng J, Xu Y, Wong DWK, Liu J, Cao X. Joint Optic Disc and Cup Segmentation Based on Multi-Label Deep Network and Polar Transformation. IEEE Trans Med Imaging 2018;37:1597-605.
8. Cheng J, Liu J, Xu Y, Yin F, Wong DW, Tan NM, Tao D, Cheng CY, Aung T, Wong TY. Superpixel classification

based optic disc and optic cup segmentation for glaucoma screening. IEEE Trans Med Imaging 2013;32:1019-32.

9. Hervella ÁS, Rouco J, Novo J, Ortega M. End-to-end multi-task learning for simultaneous optic disc and cup segmentation and glaucoma classification in eye fundus images. Appl Soft Comput 2022;116:108347.

10. Jiang Y, Duan L, Cheng J, Gu Z, Xia H, Fu H, Li C, Liu J. JointRCNN: A Region-Based Convolutional Neural Network for Optic Disc and Cup Segmentation. IEEE Trans Biomed Eng 2020;67:335-43.

11. Pascal L, Perdomo OJ, Bost X, Huet B, Otálora S, Zuluaga MA. Multi-task deep learning for glaucoma detection from color fundus images. Sci Rep 2022;12:12361.

12. Cifuentes-Canorea P, Ruiz-Medrano J, Gutierrez-Bonet R, Peña-Garcia P, Saenz-Frances F, Garcia-Feijoo J, Martinez-de-la-Casa JM. Analysis of inner and outer retinal layers using spectral domain optical coherence tomography automated segmentation software in ocular hypertensive and glaucoma patients. PLoS One 2018;13:e0196112.

13. Tan O, Li G, Lu AT, Varma R, Huang D; Advanced Imaging for Glaucoma Study Group. Mapping of macular substructures with optical coherence tomography for glaucoma diagnosis. Ophthalmology 2008;115:949-56.

14. Garvin MK, Abràmoff MD, Wu X, Russell SR, Burns TL, Sonka M. Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomography images. IEEE Trans Med Imaging 2009;28:1436-47.

15. Chiu SJ, Li XT, Nicholas P, Toth CA, Izatt JA, Farsiu S. Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation. Opt Express 2010;18:19413-28.

16. Kafieh R, Rabbani H, Abramoff MD, Sonka M. Intra-retinal layer segmentation of 3D optical coherence tomography using coarse grained diffusion map. Med Image Anal 2013;17:907-28.

17. Yazdanpanah A, Hamarneh G, Smith BR, Sarunic MV. Segmentation of intra-retinal layers from optical coherence tomography images using an active contour approach. IEEE Trans Med Imaging 2011;30:484-96.

18. Rossant F, Bloch I, Ghorbel I, Paques M. Parallel Double Snakes. Application to the segmentation of retinal layers in 2D-OCT for pathological subjects. Pattern Recognit 2015;48:3857-70.

19. González-López A, de Moura J, Novo J, Ortega M, Penedo MG. Robust segmentation of retinal layers in optical coherence tomography images based on a multistage active contour model. Heliyon 2019;5:e01271.

20. Novosel J, Thepass G, Lemij HG, de Boer JF, Vermeer KA, van Vliet LJ. Loosely coupled level sets for simultaneous 3D retinal layer segmentation in optical coherence tomography. Med Image Anal 2015;26:146-58.

21. Duan J, Tench C, Gottlob I, Proudlock F, Bai L. Automated segmentation of retinal layers from optical coherence tomography images using geodesic distance. Pattern Recognit 2017;72:158-75.

22. Dodo BI, Li Y, Eltayef K, Liu X. Automatic Annotation of Retinal Layers in Optical Coherence Tomography Images. J Med Syst 2019;43:336.

23. Li J, Jin P, Zhu J, Zou H, Xu X, Tang M, Zhou M, Gan Y, He J, Ling Y, Su Y. Multi-scale GCN-assisted two-stage network for joint segmentation of retinal layers and discs in peripapillary OCT images. Biomed Opt Express 2021;12:2204-20.

24. Liu X, Cao J, Wang S, Zhang Y, Wang M. Confidence-Guided Topology-Preserving Layer Segmentation for Optical Coherence Tomography Images With Focus-Column Module. IEEE Trans Instrum Meas 2021;70:1-12.

25. Fang L, Cunefare D, Wang C, Guymer RH, 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 2017;8:2732-44.

26. Li Q, Li S, He Z, Guan H, Chen R, Xu Y, Wang T, Qi S, Mei J, Wang W. DeepRetina: Layer Segmentation of Retina in OCT Images Using Deep Learning. Transl Vis Sci Technol 2020;9:61.

27. Xiang D, Chen G, Shi F, Zhu W, Liu Q, Yuan S, Chen X. Automatic Retinal Layer Segmentation of OCT Images With Central Serous Retinopathy. IEEE J Biomed Health Inform 2019;23:283-95.

28. Roy AG, Conjeti S, Karri SPK, Sheet D, Katouzian A, Wachinger C, Navab N. ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. Biomed Opt Express 2017;8:3627-42.

29. Wang J, Wang Z, Li F, Qu G, Qiao Y, Lv H, Zhang X. Joint retina segmentation and classification for early glaucoma diagnosis. Biomed Opt Express 2019;10:2639-56.

30. Berenguer-Vidal R, Verdú-Monedero R, Morales-Sánchez J, Sellés-Navarro I, Del Amor R, García G, Naranjo V. Automatic Segmentation of the Retinal Nerve Fiber Layer by Means of Mathematical Morphology and Deformable Models in 2D Optical Coherence Tomography Imaging. Sensors (Basel) 2021.

31. García G, Del Amor R, Colomer A, Verdú-Monedero R,

Morales-Sánchez J, Naranjo V. Circumpapillary OCT-focused hybrid learning for glaucoma grading using tailored prototypical neural networks. Artif Intell Med 2021;118:102132.

32. Devalla SK, Renukanand PK, Sreedhar BK, Subramanian G, Zhang L, Perera S, Mari JM, Chin KS, Tun TA, Strouthidis NG, Aung T, Thiéry AH, Girard MJA. DRUNET: a dilated-residual U-Net deep learning network to segment optic nerve head tissues in optical coherence tomography images. Biomed Opt Express 2018;9:3244-65.

33. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015) 2015:234-41.

34. Chen Y, Rohrbach M, Yan Z, Shuicheng Y, Feng J, Kalantidis Y. Graph-based global reasoning networks. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2019:433-42.

35. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. In: 3rd International Conference on Learning Representations (ICLR 2015). 2015. Available online: http://arxiv.org/abs/1412.6980

36. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ.

Densely Connected Convolutional Networks. In: 2017 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2017:2261-9.

37. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. 2016 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2016:770-8.

38. Moon WK, Lee YW, Ke HH, Lee SH, Huang CS, Chang RF. Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks. Comput Methods Programs Biomed 2020;190:105361.

39. Khened M, Kollerathu VA, Krishnamurthi G. Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. Med Image Anal 2019;51:21-45.

40. Khan S, Islam N, Jan Z, Din IU, Rodrigues JJPC. A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. Pattern Recognit Lett 2019;125:1-6.

41. Wu N, Phang J, Park J, Shen Y, Huang Z, Zorin M, et al. Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening. IEEE Trans Med Imaging 2020;39:1184-94.