Healthcare Analytics 5 (2024) 100289

Contents lists available at ScienceDirect

Healthcare Analytics

journal homepage: www.elsevier.com/locate/health

A novel convolutional neural network for identification of retinal layers using sliced optical coherence tomography images

Akshat Tulsani [a,1], Jeh Patel [a], Preetham Kumar [a,*], Veena Mayya [a,1], Pavithra K.C. [a], Geetha M. [b], Sulatha V. Bhandary [c], Sameena Pathan [a]

[a] Department of Information and Communication Technology, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, 576104, India
[b] Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, 576104, India
[c] Department of Ophthalmology, Kasturba Medical College, Manipal, Manipal Academy of Higher Education, Manipal, 576104, India

ARTICLE INFO

Keywords:
Convolutional neural network
Retina
Image segmentation
Optical coherence tomography
Layer segmentation

ABSTRACT

Retinal imaging is crucial for observing the retina and accurately diagnosing pathological problems. Optical Coherence Tomography (OCT) has been a transformative breakthrough for developing high-resolution cross-sectional images. It is imperative to delineate the multiple layers of the retina for a proper diagnosis. A novel segmentation-based approach is introduced in this study to identify seven distinct layers of the retina using OCT images. The developed approach presents SliceOCTNet, a customized U-shaped Convolutional Neural Network (CNN) that introduces group normalization and intricate skip connections. Paired alongside a hybrid loss function, the SliceOCTNet outperformed most state-of-the-art approaches. The introduction of Group Normalization in SliceOCTNet stabilized the model and improved layer identification even when working with small datasets. The use of skip connections also contributed to an improvement in the spatial outlook of the model. Implementing a hybrid loss function addresses the class imbalance problem in the dataset. Duke University's spectral-domain optical coherence tomography (SD-OCT) B-scan dataset of healthy and Diabetic Macular Edema (DME) afflicted patients was utilized to train and evaluate the SliceOCTNet. The model accurately recognizes the seven layers of the retina. It can achieve a high dice coefficient value of 0.941 and refine the segmentation process to a higher level of precision.

1. Introduction

Optical Coherence Tomography (OCT) is a novel imaging technique that captures high-resolution features at a microscopic level using low-coherence light [1]. It is non-invasive in nature and poses little to no risk to the patient. The resulting scan is a two-dimensional cross-sectional image of the area of interest that is either color-coded or grayscale. At the time of conception, it was only able to capture transparent tissue accurately. However, as technology advanced, it provided clear real-time results on non-transparent tissue, vastly improving its applications and eventually becoming one of the most prominent techniques in the field of biomedical imaging.

The use of OCT for high-resolution imaging of the retina is well established. Fig. 1 illustrates a cross-sectional view of the retina [2]. It is commonly observed that diseases impacting the sub-retinal area manifest distinct indicators, for instance, fluid-filled cysts in diabetic macular edema (DME), depicted in Fig. 2, or hard exudates and degenerative debris in the case of DRUSEN. These indicators are discernible throughout the layers of the retina. Recognizing the distinct retinal layers facilitates the monitoring and tracking of various eye conditions such as DME [3–5], glaucoma [6], diabetic retinopathy (DR) [7], and age-related macular degeneration (AMD) [8]. Recently, the segmentation of retinal layers in the context of neurodegenerative disorders like Alzheimer's disease, Parkinson's disease, multiple sclerosis, and essential tremor has been explored and is a significant area of investigation in medical imaging and neurology [9]. Therefore, accurate delineation of the layers of the retina is a crucial task for the diagnosis and monitoring of retinal disorders. Manual interpretation by experts was previously the preferred method for identifying the boundaries of the layers in the retina. However, this approach had several drawbacks, including the fact that it was highly time-consuming and subject to the expertise and experience of the individual performing the delineation. Therefore, there is an urgent need for an automatic segmentation technique that can effectively and efficiently evaluate the progression of retinal disorders.

* Corresponding author.
E-mail address: preetham.kumar@manipal.edu (P. Kumar).
[1] These authors contributed equally to this work and designated as co-first authors.

https://doi.org/10.1016/j.health.2023.100289
Received 28 June 2023; Received in revised form 3 December 2023; Accepted 5 December 2023
Available online 7 December 2023

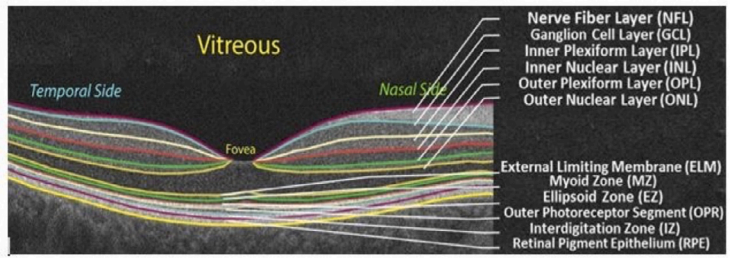

Fig. 1. Illustration of 12 layers of the retina.

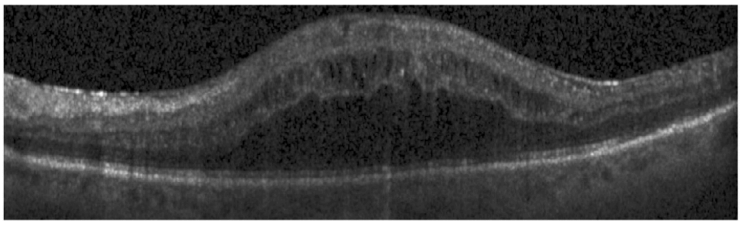

Fig. 2. A sample OCT B-scan from Duke University's OCT dataset depicting a fovea-centered cyst.

Automated techniques can significantly improve the speed and accuracy of the diagnostic process by providing a more objective and consistent method for identifying retinal boundaries. Additionally, by automating the process, these techniques can help reduce the dependence on manual interpretation by experts, thereby reducing the subjectivity and variability of the diagnosis as well as the time required for the diagnostic process. However, such methods are not infallible.

OCT scan results are in grayscale and typically contain a lot of noise. Furthermore, the layers of the retina may become deformed due to the presence of pathological biomarkers. Due to this, there is a need for a quality dataset containing a large number of images for accurate prediction and a long training time. The following are the key contributions of this work to advancing the state-of-the-art:

- A novel SliceOCTNet model developed from scratch proposed to delineate intraretinal layers in an OCT image. Inspired by CNN's ability to learn high-level semantic features and U-NET's ability to use features across contracting and expanding paths.
- The study explores the use of group normalization and skip connection which enhanced the ability of model to identify the intensities of various pixels in various layers, thereby improving the overall process of segmentation.
- The model proposed in the study improves the ability to distinguish between similar semantic features (as observed in retinal layers). The proposed work can be translated to solve similar problems observed in various biomedical segmentation challenges.
- Duke University's retinal scan dataset, consisting of healthy and DME patients, was used for training the model [10]. The proposed model achieved a dice score of 0.941, outperforming most state of art models.

The remainder of the paper is as follows: existing approaches to segmenting retinal OCT layers are discussed in Section 2. Section 3 details the proposed pipeline and the SliceOCTNet architecture, along with the implementation specifics. In Section 4, the experimental evaluation and benchmarking results are presented, followed by conclusions and a discussion of the scope of future research.

2. Related work

Traditionally, two approaches are used for segmentation tasks: mathematical model-based approaches like active contouring, graph theory, and Markov modeling, and deep learning approaches such as Convolutional Neural Network (CNN) and Fully Convolutional Networks (FCN). Approaches like Ronneberger's U-NET [11] or Long's FCN [12] frameworks serve as inspiration to many CNN-based approaches in the field.

During the last decade, the primary focus of the work involved combining mathematical modeling with machine learning models to enable automated retinal layer segmentation. Abràmoff et al. [13] extended the applications of graph search algorithms by assimilating varying feasibility constraints and applying them to SD-OCT scans. The work depicts the ability of such methods to segment intraretinal surfaces. A key problem faced in their work was reducing the memory requirements for processing large amounts of spectral OCT data. Li et al. [14] presented a fully automated approach for segmenting layers of the retina. They utilized Dijkstra's algorithm to identify the edges and a vessel detection algorithm that identifies blood vessels present in the SD-OCT scan. This makes the algorithm robust against distortion caused by their presence [15]. They propose the use of non-linear adaptive filters and reducing overall computational complexity to enhance results. Chen et al. [16] tackled the issue of identifying several macular pathologies. Retinal alignment is used to align the images slightly to better detect the varied appearance of the pathological markers present. The foveal slice used as input was manually selected by expert ophthalmologists, and an automated approach for fovea localization may be developed in the future. Kim et al. [10] presented a novel method that made use of histogram-of-gradient (HOG) descriptors and support vector machines (SVM) to classify every image as either a normal, AMD, or DME-affected OCT scan. This work highlights the possible use of techniques that do not require annotations of the layers of the retina, which can be potentially useful if there is a large alteration in the structure of the sub-retinal area. Michael et al. [17] were among the first to develop a model trained on a real-world dataset containing scans from patients severely affected by DME. They define each SD-OCT scan as a graph of interconnected nodes and traverse each boundary separately using minimum weighted paths to segment retinal layers. This study is constrained by a specific method of segmentation, as the classification method utilized failed to isolate layers on its own. Sivaswamy et al. [18] made significant efforts to achieve continuous segmentation of retinal boundaries. They tackle the problem of discontinuous edges caused by the presence of shadows in the OCT scans by performing edge detection and subsequent boundary-tracing. Due to the ground truth being sourced from multiple markers, they remark that there is ambiguity in their training process.

Over the years, there has been a significant interest in developing U-Net based deep learning models to facilitate automated retinal layer segmentation. Conjeti et al. [19] designed a novel model for segmenting layers of the retina from OCT images, known as ReLayNet, which was based on U-Net. This model was tested on the publicly available Duke SD-OCT B-scan dataset. The loss function used is a weighted multi-class logistic function along with the dice loss. Along with retinal layers, it delivers remarkable results while identifying fluid-filled cystoids inside the surface of the retina, as commonly seen in patients suffering from DME. Schlemper et al. [20] developed attention U-Net, a variant of the original U-Net architecture. They utilized attention gates to improve the sensitivity of predictions and minimize manual supervision while decreasing computational overhead. Heisler et al. [21] focused on segregating OCT layers with fluid present in them by using a novel LF-U-Net model. LF-U-Net aims to leverage and combine the strengths of U-Net and FCN while making use of dilated convolutional layers to extract more features without inflating the number of parameters. Loncaric et al. [22] explored transfer learning to segment four distinct retinal layers. Their model was pretrained on a non-OCT dataset, known as ImageNet, and it modifies the existing U-Net architecture by introducing a ResNet-based encoder [23]. Through their work, they were able to demonstrate that a U-shaped architecture may benefit from the inclusion of an encoder with pre-trained weights. Sun et al. [24] created a novel network termed MPG-Net for segmenting OCT retinal layers. Their base U-shaped model was attached to two different modules: a feature refinement module and a prediction guidance module. The dataset used for the model was Duke's SD-OCT B-scan dataset, and they were able to segment seven distinctive layers. Jin et al. [25] designed a novel two-stage graph CNN known as MGU-Net that was able to segment seven layers of the retina and the optic disc. To deal with the fact that each retinal layer is different thickness, they added a multi-scale global reasoning module that divides the input paths into three so that features of retinal layers with different thicknesses can be encoded into global feature maps. These maps are then upsampled and joined together to make a global feature map. This model was evaluated on Duke University's DME dataset and achieved a dice score of 0.871. Wang et al. [26] designed a novel architecture known as SGNet. It utilizes both finer class labels, which are composed of the finer attributes present in the image, and an aggregated group of finer class labels, known as super-class labels. The first strategy focuses on predicting the super-class and, consequently, the finer class, whereas the second strategy directly operates on the finer class labels. The majority of existing U-Net based works enhance performance by introducing additional layers or employing multiple networks, leading to an increase in network parameters and subsequently extending training and inference times. There is scope for hyperparameter tuning, involving variations in loss functions and regularization techniques, without the necessity of increasing parameters.

Recently, Moradi et al. [27] combined a graph-cut algorithm with a cubic spline to autonomously label eleven retinal boundaries for analyzing AMD and monitoring its progression. Subsequently, the enhanced images were utilized in a deep ensemble approach, merging a bagged tree and end-to-end deep learning classifiers. The authors achieved 93% mean intersection over union (IOU) scores for layer segmentation task. Lu et al. [28] proposed a boundary-enhanced semi-supervised network (BE-SemiNet) aiming to enhance retinal layer segmentation by leveraging an auxiliary distance map regression task. This approach proved effective in scenarios with limited labeled data and abundant unlabeled data, achieving an average accuracy of 87.88%. Gende et al. [29] proposed dual methodologies for segmenting retinal layers in glaucomatous OCT images. The initial method leverages inter-view resemblances, employing a unified module to segment all scan patterns as a unified domain. Meanwhile, the second approach employs individual modules tailored to each scan pattern, automatically identifying the appropriate module for analyzing each image. They obtained an average accuracy of 87% for the segmentation of all layers. Although these significant accomplishments have been made, there is still potential for enhancing performance without making significant changes to the model parameters.

The literature discussed above highlights the challenges that arise due to the grayscale format of OCT images, the presence of discontinuous boundaries, and the nearly uniform shape of the layers. These aspects can pose difficulties for both CNN and mathematical models in terms of accuracy. Additionally, pathological retinal markers can introduce morphological changes that obscure layer boundaries. These changes must either be eliminated from the dataset or integrated into the model's design. This necessitates the development of a model capable of offering higher semantic perspectives, distinguishing similar layers in OCT images, and adeptly managing speckle noise within the images.

Apart from the challenges mentioned above, the class imbalance between the layer and background can negatively impact the training cycle of the CNN model. Additionally, a small dataset and a deep network can lead to overfitting, diminishing the generalizability of the automated model. Thus, the following study introduces SliceOCTNet, a custom model focused on improving the segmentation of retinal OCT images and isolating the ILM, NFL-IPL, INL, OPL, ONL-ISM, ISE, and OS-RPE layers of the retina.

3. Methodology

To harness the benefits offered by group normalization, a fusion of loss functions, and skip connections, SliceOCTNet is proposed in this study. SliceOCTNet is built upon the foundational U-Net architecture. A comprehensive description of the dataset and methodology is provided in the following sections.

3.1. Dataset

The dataset used in this work is Duke University's SD-OCT B-scan dataset [17]. It contains 110 retinal scans taken from a total of ten patients and annotated by a group of two expert ophthalmologists. Annotation includes the following areas: inner nuclear layer (INL), outer plexiform layer (OPL), outer nuclear layer to inner segment myeloid (ONL-ISM), inner segment ellipsoid (ISE), a and pigment epithelium (OS-RPE). The dataset is a mix of scans of healthy and DME-affected retinas and contains images with dimensions of 496×768.

3.2. Preprocessing

The most common distortions in an OCT scan may be present due to noise or the quality of the image. For the model to learn well, it is important to get images that are as smooth as possible. Hence, it is imperative to denoise the entire dataset. For this task, a non-local means algorithm as given by (Eq. (1)) is used to preserve features. It is analogous to a means filter, where an average of the values of pixels around the point of interest is assigned to it. This is compared to other local patches, and the average of the patches closest to the point of interest is assigned as the value. A noise standard deviation is defined to subtract the noise variance when the localization of patches towards the point of interest takes place. Additionally, a constant is used to control the decay of weights, which may depend on the distance between the patches.

$$u(p) = \frac{1}{C(p)} \int u(q) f(d(p,q)) dq \tag{1}$$

Where $d(p,q)$ is the Euclidean distance between a patch of pixels p and q and f is the decreasing function applied to that term. $C(p)$ is the normalizing factor defined in the work done by Buades et al. [30].

Class imbalance is a very common problem in semantic segmentation and refers to the skewing of the distribution of the labeled classes present in an image. It commonly occurs when meaningful classes

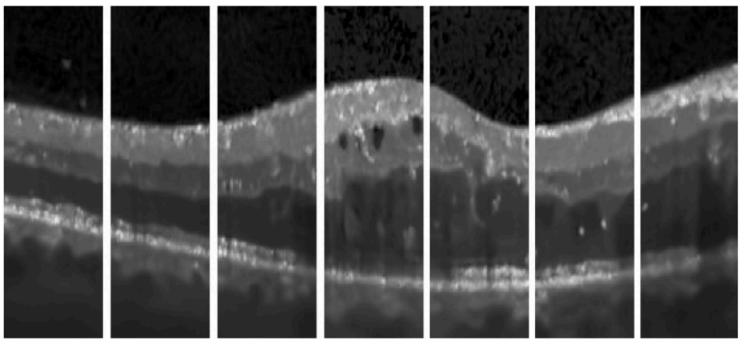

Fig. 3. A sample input OCT image sliced into 7 segments.

have a comparatively diminished distribution as opposed to unwanted features. In an OCT scan, the retinal layers occupy less than half of the area, and the rest is unwanted background. Being grayscale images, this background cannot be eliminated/ignored directly based on pixel value. The model would ideally have to understand that a specific region is background and not part of any layer. Here if the complete grayscale image is passed to the training module, the attention of the model would be swayed towards the background. To tackle this, image-slicing techniques were utilized [31,32]. To illustrate the process used, Fig. 3 was taken from the input dataset. For the proposed model, every image is sliced into 7 segments of size 216 × 64 to populate our dataset and increase the number of training images fed into the model to 770 images. This diminishes the extent of background in an image as compared to a complete image in the dataset by a significant margin and allows the model to focus on the desired area. Using the non-local means algorithm on the input dataset and extracting valuable information allows us to significantly improve the quality of the processed image dataset, making it fit for training [33].

3.3. SliceOCTNet architecture

SliceOCTNet is inspired by the U-Net model. Our study customizes and redevelops the U-shaped architecture from scratch, keeping in mind the likeness of OCT images and the requirement of the same task. SliceOCTNet is has a distinctive U-shape (refer Fig. 4). The left-hand side is known as the contracting path which has an FCN-like structure, whereas the right-hand side is termed the expansive path which is responsible for up-sampling the images.

In the contracting path, a typical layer features a series of convolutional operations that increase the depth of the image by increasing the number of channels present. Fig. 5 suggests the typical structure of a layer of the proposed SliceOCTNet model. It introduces group normalization and skip connection. The group normalization helps improve the models threshold towards noise and feature variation. While skip connections allow the model to access information from previous layers. This process is typically followed by activation and maxpooling to introduce non-linearity and reduce spatial dimensions in data respectively. An image set of size 216 × 64 was obtained after preprocessing and used as the input. A Conv2D operation is applied over the image with a kernel size of 3 × 3 containing 64 channels. After that, a max-pooling operation is performed, where it calculates the maximum pixel value of a selected group of pixels using a stride value of 2 and discards the rest, effectively, decreasing the image size to half of the original. This sequence of events is repeated until the lowest layer is reached and the image size is in the smallest dimension 27 × 8.

Now, the image travels through the expansion path. The structural make-up of this path remains similar to the contracting path, but with a key difference. Here, the max-pooling operation is replaced with an upscaling operation. The image is upscaled to double its existing size, decreasing the number of feature maps while increasing the dimensions. Consequently, it is concatenated with the image obtained from the same layer of the contraction side after all convolutional operations are performed on it. This is meaningful as it helps combine the features of both images to get a more precise prediction. After the final convolution block is reached, the image is reshaped to obtain the required result. The entire working pipeline of the SliceOCTNet can be seen in Fig. 6.

The developed approach is modular, so it lends itself highly to further customizability while retaining its original functions. It does not depend on the size of the input image, further amplifying its flexibility as a base. With relatively fewer training samples and time, it can produce results well above the baseline figure. Hence, it is ideal for biomedical image segmentation tasks. There were several modifications made to the existing shell of the U-shaped architecture to refine the workings of the model. Namely, the structure of every CNN layer has been altered to fit several components, i.e., a group normalization block, a ReLU activation function, and a skip connection that concatenates the output of the Conv2D block with the normalized block. Furthermore, a regularizer function is used to slowly decay the variance to prevent the overfitting of SliceOCTNet.

When an image is propagated from the input layer to the hidden layers of a CNN, the network generates a vector of activations. During training, the network's weights are iteratively updated to minimize loss, which can lead to changes in the distribution of activations. As a consequence, activations may become concentrated in specific areas, a phenomenon known as internal covariate shift (ICS). ICS can impact the model's ability to comprehend the interrelationships between different features in the dataset. To deal with this, normalization techniques were introduced into the model's design. Batch normalization is the first technique that was developed to coordinate changes between multiple layers by standardizing the activation of input parameters on every layer in a mini-batch format. Hence, the need for the successive layers to assume the weights is removed, and any update to the weights does not affect the distribution of inputs by a significant margin. This helps accelerate the process of training by requiring fewer epochs for completion.

It was observed by Yuxin et al. [34] that batch normalization requires a sufficiently large batch size to minimize errors, thus reducing its applicability on large datasets [35]. Yuxin et al. [34] introduced an alternative. Group normalization is a normalization technique that addresses concerns related to batch normalization by organizing the channel into different groups of similar size and is free of constraints presented by the batch size [34].

$$S_I = \{k \mid k_n = i_n, \left\lfloor \frac{k_c}{C/G} \right\rfloor = \left\lfloor \frac{i_c}{C/G} \right\rfloor\} \tag{2}$$

Eq. (2) denotes the function of group normalization. Here, S refers to the set of pixels using which the standard deviation and the mean are calculated. i refers to the four dimensional vector that stores information regarding the axes for indexing the features like batch axis

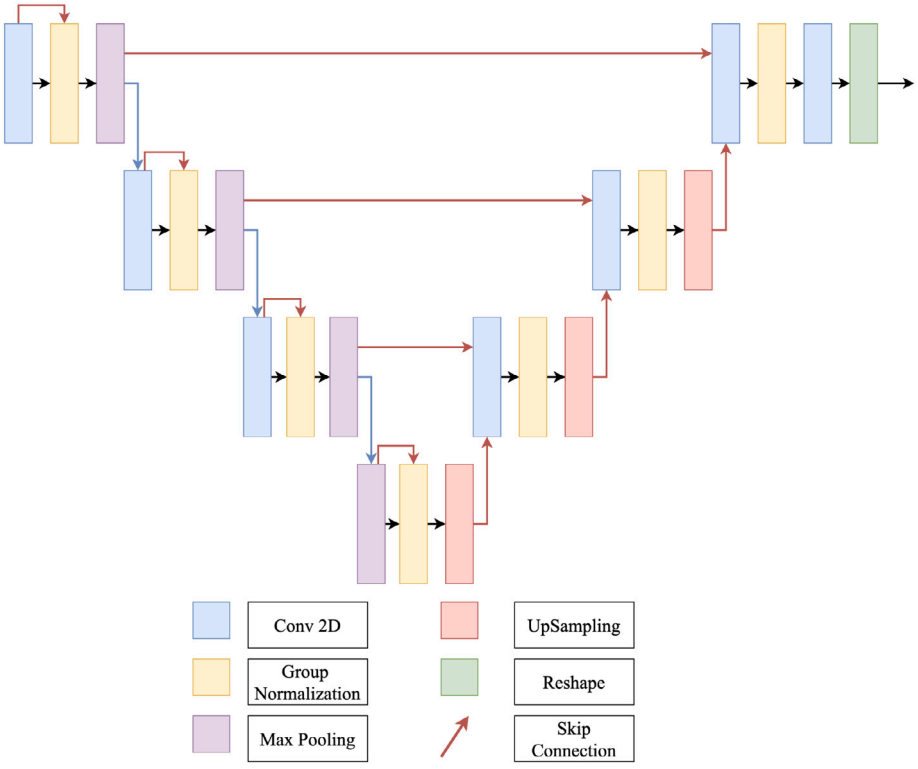

Fig. 4. An overview of the SliceOCTNet architecture.

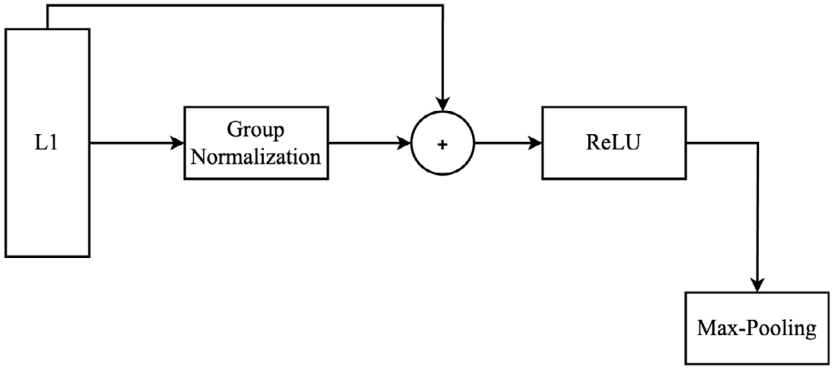

Fig. 5. Layer structure of the SliceOCTNet architecture.

(N), channel axis (C), spatial height (H), and width axes (W). In the case of batch normalization, the formula is $k_C = i_C$ meaning that the mean and standard deviation is calculated along the (N, H, W) axes. For instance normalization, the formula is $k_n = i_n$ allowing the instance to be normalized over (C, H, W) axes. Group normalization introduces a new parameter, G, which is the number of groups. The normalization is performed over (H, W) axes over a group of (C/G) channels. It becomes identical to layer normalization if G = 1 and instance normalization if $G = C$, that is, one group encompasses only a singular channel.

In addition to the internal covariate shift, there are further problems that may be observed due to the propagation of weights from layer to layer. The vanishing gradients problem is a well-documented obstacle faced by neural networks with a large number of layers. As each layer may update the weights proportional to the partial derivative of the previous layer, in some cases, the change in weights may be diminishingly small and may prevent any update in the weights. This may lead to the degradation of the performance of the model as a whole.

To deal with these, skip connections are introduced for use in residual blocks. The concept of residual blocks originates from ResNet whereby the result of a previous convolutional operation is added to another block present further deep in the layer. This allows the layer to retain meaningful updates as the image propagates through it and minimizes the problem of vanishing gradients. Here, the output of the Conv2D operation is concatenated with the group normalization output to retain changes in the weights. Furthermore, the long skip connections stretching from the encoder half to the decoder half were inserted to reduce the semantic gap. This improves the optimization of semantically similar feature maps by fusing the output from the previous convolution layer of the dense block to its corresponding up-sampled output layers of the lower dense block.

Due to the nature of multi-class segmentation problems, the strong predominance of one class over the other classes in terms of area occupied may lead to class imbalance issues, which remain a huge obstacle. Since the number of pixels to be segmented is extremely small as compared to the total number of pixels present in the image, the unwanted area can divert attention from areas of interest and may result in poor quality training. To deal with this, this study combines dice loss and categorical cross-entropy (CCE), represented by Eqs. (3) and (3)

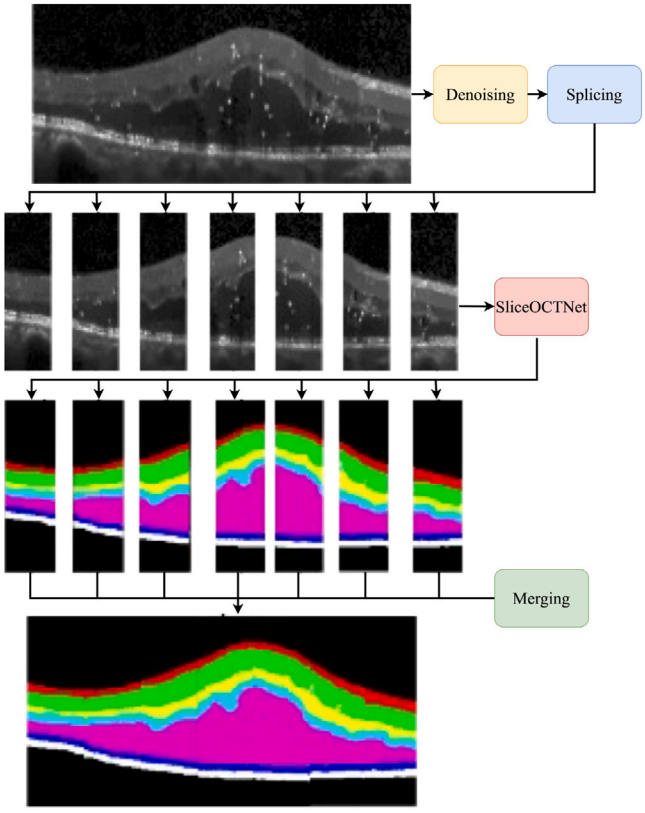

Fig. 6. Pipeline of the SliceOCTNet.

respectively, into a customized loss function. Dice loss is a robust loss function that takes into account the overlap between local and robust features of the predictions and the ground truth. Categorical cross-entropy is commonly used in multiclass segmentation models because it can accurately determine the probability of an object belonging to one of several classes, all with a set probability.

$$L_{dice} = 1 - \frac{2 \times y \times \hat{y} + 1}{y + \hat{y} + 1} \tag{3}$$

$$L_{CCE} = \sum_{i}^{C} t_i log(f(s)) \tag{4}$$

$$L = \alpha \times L_{CCE} + \beta \times L_{dice} \tag{5}$$

In Eq. (3), y refers to ground truth and $\hat{y}$ refers to the predicted results. Whereas in Eq. (8), t_i and s_i refer to the ground truth of every class i in C. The softmax function $f(s_i)$ is used as the input to the logarithmic function to obtain categorical cross-entropy. Constants α and β in Eq. (5), adjust the influence of categorical cross-entropy and dice loss over the custom loss function. α was set to 0.7 while β was set to 0.4, the higher value on Categorical Cross-entropy helped the model to concentrate more on layer separation while dice coefficient loss helped maintain segmentation accuracy.

The activation function utilized in the layers is ReLU. It helps accelerate the training time. A sigmoid function is used while reshaping the input into the desired output. The optimization function of choice was the Adam optimizer which dynamically adapts the learning rate per parameter based on the value of changes in weights, making it a better fit over stochastic gradient descent. The evaluation was performed over a Google colab GPU. To optimize the results and reduce GPU memory usage, the model was trained repeatedly with varying batch sizes. The batch size for training was set to 20.

The model's performance was evaluated over a few standard metrics, namely the dice coefficient and mean pixel error. Their corresponding formulae are shown using Eqs. (6) and (7).

$$Dice = 2 \times \frac{|A \cap B|}{|A| + |B|} \tag{6}$$

$$PA = \frac{|A \cap B|}{|B|} \tag{7}$$

Here, A is the predicted output map, and B is the correct binary map. Using the custom U-shaped model and tailor-made loss function, our methodology obtains a high dice score of 0.941 for all seven layers.

4. Results

In the following section, the results produced from the developed approach will be examined. Figs. 7A and 7B illustrate the input image and its corresponding ground-truth with appropriate color delineations between the layers for visibility. Fig. 7C depicts the predictions of the SliceOCTNet model.

Attention U-Net [20], Long [12], and Fu et al. [36] comparisons have the compound layer (RNFL-IPL) which includes GCL. Liu [37], SGNet [26], SR-Net [38] comparisons have the compound layer (ISE–OS), instead of (OS-RPE) like the rest of the entries. Table 1 compares the average dice score of this work with several other state-of-the-art works that have been evaluated on the same dataset, i.e., Duke University's DME dataset. Whereas Table 2 shows how SliceOCTNet compares to the relevant state-of-the-art models in terms of pixel accuracy, refers to the ratio of correctly classified pixels to the total number of pixels [39]. For K+1 classes (K foreground classes and the background) the pixel accuracy is defined by Eq. (8).

$$Pixel\_Accuracy = \frac{Number\_of\_correctly\_classified\_pixels_{kclass}}{Total\_number\_of\_pixels} \tag{8}$$

Manual annotation was attempted for this study; however, a pre-annotated dataset proved most effective for training the model. Due to training in smaller batch sizes, the model was optimized to a greater extent. Strategically designed skip connections ensure that the features extracted in earlier layers are not lost in propagation. A distinct problem faced in the detection of layers is that the OCT B-scans depict certain intraretinal cysts with the same color code as the background area, making the detection of cysts tough. However, the proposed model is able to outperform RelayNet by a percentage difference of 4.34

5. Discussion

This study highlights some of the key problems faced when dealing with OCT B-scans and attempts to tackle them. OCT images are usually subject to a varying degree of noise, which skews the model into providing incorrect results. Hence, there is a need to emphasize proper pre-processing techniques to smooth the input. Typically, the presence of lesions and shadows emanating from nearby blood vessels may lead to inconsistent segmentation as they tend to blur retinal layer boundaries. The introduction of improved regularization steps using group normalization and concatenation at the layer level provides the model with a spatial outlook, the ability to distinctly identify different features of a multiclass dataset, increasing accuracy over other state-of-the-art CNN models that focus on pixel-wise classification. Visual comparison (refer Fig. 7 C–D) show that the predictive results are aligned with the inputs. The overall pixel accuracy, defined as the ratio of the number of correctly classified pixels to the total number of pixels in the image, of the model is 0.9562 and the average dice score is 0.941.

Chiu et al. [17] conducted benchmarking on the publicly released Duke DME dataset. Subsequently, this dataset found use among numerous researchers, some of whom incorporated a private dataset alongside the Duke DME dataset [25,40]. Various data splits were employed for training and testing, with subsequent benchmarking of the results.

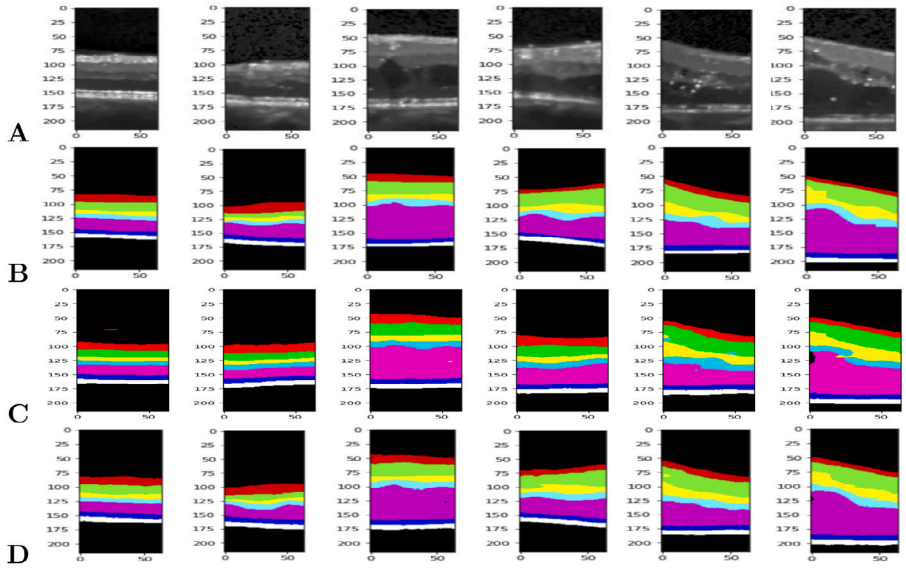

Fig. 7. Sample OCT-B scan images. (**A**) Input OCT-B scan images, (**B**) Ground truth for the images in **A**, (**C**) Output segmentation from the base U-Net model for the images in **A**, (**D**) Output segmentation from the SliceOCTNet for the images in **A**.

Table 1
Layer-wise comparison of dice score with state-of-the-art models.

Model	RNFL	GCL-IPL	INL	OPL	ONL-ISM	ISE	ISE-OS/OS-RPE	Average
U-Net [11,25]	0.86	0.91	0.83	0.81	0.91	0.90	0.83	0.86
Long [12]	0.84	–	0.72	0.71	0.88	0.88	0.86	0.81
Attention U-Net [20]	0.90	–	0.90	0.78	0.76	0.93	0.89	0.85
Chiu et al. [17]	0.86	0.89	0.77	0.72	0.87	0.85	0.82	0.82
Fu et al. [36]	0.92	–	0.87	0.83	0.95	0.89	.86	0.88
Wang et al. [40]	0.86	0.90	0.78	0.78	0.94	0.90	0.86	0.85
Swin-Unet [41,42]	0.86	0.90	0.84	0.80	0.91	0.89	0.82	0.86
Chakravarty et al.[43]	0.84	0.93	0.87	0.80	0.94	0.86	0.85	0.87
DRUNET [25,44]	0.85	0.88	0.72	0.73	0.92	0.89	0.84	0.83
MGU-NET [25]	0.87	0.91	0.81	0.79	0.94	0.90	0.87	0.87
Wang et al. [45]	0.89	0.93	0.83	0.82	0.95	0.9	0.88	
ReLayNet [19]	0.90	0.94	0.87	0.84	0.93	0.92	0.90	0.90
SGNet [46]	–	–	0.89	0.85	0.94	0.93	0.91	–
DelNet [32]	–	–	0.92	0.90	0.97	0.95	0.91	–
TransUNet [41]	0.91	0.94	0.87	0.85	0.96	0.91	0.89	0.90
Liu et al.[37]	0.90	0.94	0.89	0.86	0.94	0.92	0.91	0.91
RLMENet [47]	–	0.91	0.85	0.91	0.91	–	0.95	–
SliceOCTNet (proposed)	0.93	0.96	0.92	0.91	0.97	0.95	0.94	0.94

Table 2
Pixel accuracy comparison with relevant models.

Layer	U-Net	Re LayNet	DRU-Net	MGU-Net [25]	SR-Net [38]	BE-SemiNet [28]	Kumar et al. [48]	Slice OCTNet
RNFL	83.1	81.8	82.1	84.6	94	90	83.3	90.2
GCL	66.9	68.7	67.9	70.3	–	–	68.5	74.6
IPL	72.1	70.9	71.4	74.4	–	–	73.9	78.2
INL	76.0	77.9	77.8	81.0	88	0.85	80.7	84.0
OPL	78.3	79.7	79.5	82.0	0.95	93	91.7	86.4
ONL	91.4	90.3	91.9	91.9	87	86	85.84	93.4
IS/OS	86.2	86.6	85.7	87.0	87	86	85.84	90.5
RPE	83.1	82.3	81.9	89.2	90	87	84.6	91.9

Cross-validation techniques like two-fold [36,45], three-fold [40], four-fold [48], five-fold [43], and others [28,38,46,48] were utilized. Some studies opted for random splits, such as the 60:20:20 ratio. Additionally, a few studies performed data splitting after augmentation [41,47], raising concerns about potential data leakage into the testing phase and subsequent results reliability issues. In our study, we adopted a train-test split commonly employed in analogous methodologies [19, 32]. This approach ensures fair comparisons by conducting the split before augmentation, promoting more reliable and unbiased results.

The performance of the proposed SliceOCTNet surpasses that of recent state-of-the-art methods, including BE-SemiNet [28], SGNet [26], MGU-NET [25]. The majority of prior studies enhance performance by incorporating additional layers or employing multiple networks. In contrast, the proposed SliceOCTNet achieves improvement without necessitating an increase in network parameters. This enhancement is attributed to the utilization of an advanced preprocessing technique, an improved training methodology featuring enhanced regularization, and the implementation of a customized loss function.

Table 3
Results of ablation study conducted on the proposed SliceOCTNet model.

Model	RNFL	GCL-IPL	INL	OPL	ONL-ISM	ISE	ISE-OS/OS-RPE	Average
U-Net	0.86	0.91	0.83	0.81	0.91	0.90	0.83	0.86
U-Net + Group normalization	0.90	0.94	0.90	0.84	0.94	0.93	0.92	0.91
U-Net + Group normalization + Custom loss	0.91	0.94	0.91	0.88	0.94	0.94	0.92	0.92
SliceOCTNet (U-shaped network + Group normalization+ Custom loss + Skip connection)	0.93	0.96	0.92	0.91	0.97	0.95	0.94	0.94

The retinal layers are generally parallel to each other but have varying thicknesses. Combined with the fact that the grayscale OCT images have low features, it is important to reduce semantic gaps between the encoder and the decoder whilst training the segmentation model. This is facilitated by skip connections in conjunction with the developed model architecture, which encapsulates a global view of high level semantic features that would typically be lost in a traditional CNN. The presence of an improved loss function aids the model in producing classification outputs that are consistent with the ground truths. An ablation analysis was conducted to assess the contributions of individual modules within the SliceOCTNet. The outcomes of the different methods explored are summarized in Table 3. To establish a baseline for segmented images, a well-established basic U-Net [20] model was employed. It is noteworthy that the segmented layers exhibited noisy boundaries, as depicted in Fig. 7, featuring extra pixels in regions of noise. It is important to note that this pixelation is not a consequence of poor image quality but rather stems from the inherent noise present in OCT images. This observation holds true for all datasets under investigation, including the one selected for this study.

The incorporation of group normalization within SliceOCTNet significantly contributed to noise reduction and minimizing variations among features. There are more false positives in the OCT image dataset when this normalization technique is not used, as seen in the base U-Net model. This is because the base U-Net model has trouble understanding how multiple features from different classes are connected. Additionally, a custom loss function was implemented, combining weighted CCE and weighted dice loss. The CCE facilitated the distinction among diverse layers within the OCT image, while the weighted aspect addressed the imbalance between the background and individual classes. The dice loss played a role in penalizing false positives in smaller regions, thereby enhancing overall performance. Moreover, custom skip connections contributed to improving spatial awareness at each layer and capturing short-range dependencies. These collective modifications led to an improvement in the overall retinal layer segmentation performance.

Utilizing group normalization techniques in conjunction with strategically placed skip connections has allowed the SliceOCTNet to provide a clear and distinct result, which may help further the technology of automated semantic segmentation. However, there remain limitations that may be explored in later work. SliceOCTNet is heavily dependent on annotated datasets for training, which require a substantial investment of time by professionals in the field. A larger set of annotated data may deliver a better result. Furthermore, an adaptable weight function that can pinpoint and focus on the areas delivering low accuracy in feature mapping may enhance the training process. It could be observed that the proposed approach does not lead to an increase in the training parameters for the baseline U-Net model.

In this study, the speckle noise was reduced by applying the non-local means algorithm. However, shadowing effects from the presence of lesions and blood vessels in the retina remain a challenge, suggesting the potential for an alternative approach to more effectively mitigate this issue. The size of the dataset used in the study was limited by availability, but the automated approach delivers an improved prediction by focusing on the model development process. This helps reduce the reliability of domain experts required to set up a machine learning or graph based approach. A more comprehensive database may likely deliver results of increasingly higher precision and is one of the approaches that may be explored in the future.

6. Conclusions

Segmenting retinal layers using OCT is of significant importance in comprehending and diagnosing different eye conditions. In this work, a novel SliceOCTNet U-Net based model is developed, aimed at enhancing the segmentation of retinal OCT images and isolating distinct layers of the retina. SliceOCTNet has demonstrated distinct advancements in precisely identifying layers of the retina, however, the underlying technology has the potential to be utilized in various applications. It can provide new insight into isolating pathological markers of different retinal disorders, which is a crucial step for an accurate diagnosis of the disease. It can see relevance in autonomous machine learning tasks that are conducted in conjunction with layer detection, such as assessing the spread of DME and pinpointing whether it may be classified as serous retinal detachment (SRD), diffuse retinal thickening (DRT), or cystoid macular edema (CME). The most effective way of treating eye disorders lies in early detection, prevention, and treatment. Hence, it is imperative that retinal segmentation be a fast and accurate process. SliceOCTNet takes a step further in achieving the goal of a reliable and speedy automated process that may be used by both domain experts and non-technical users.

The findings of this research have the potential to provide healthcare practitioners with a valuable asset for precisely delineating retinal layers, allowing for tailored and specific treatment strategies for individual patients. Although the current model shows encouraging outcomes, there is an opportunity for the advancement of SliceOCTNet by exploring the potential integration of a spatial and channel attention mechanism. Additionally, employing state-of-the-art optimization techniques could further fine-tune the hyperparameters of SliceOCTNet.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

References

[1] James G. Fujimoto, Costas Pitris, Stephen A. Boppart, Mark E. Brezinski, Optical coherence tomography: An emerging technology for biomedical imaging and optical biopsy, Neoplasia (ISSN: 1476-5586) 2 (1) (2000) 9–25.

[2] Thomas Neyer, Ahmed ElTanboly, Agustina C. Palacio, Marwa Ismail, Andy Switala, Ahmed Soliman, Amir Hajrasouliha, Amir Hadayer, Douglas Kenneth Sigford, Ayman El-Baz, et al., A novel automated method for the objective quantification of retinal layers based on spectral domain optical coherence tomography (sd-oct) imaging reveals sequential changes in the normal retina with age, Invest. Ophthalmol. Vis. Sci. 57 (12) (2016) 5943.

[3] Aditya Tripathi, Preetham Kumar, Akshat Tulsani, Pavithra Kodiyalbail Chakrapani, Geetha Maiya, Sulatha V. Bhandary, Veena Mayya, Sameena Pathan, Raghavendra Achar, U. Rajendra Acharya, Fuzzy logic-based system for identifying the severity of diabetic macular edema from OCT B-Scan images using DRIL, HRF, and cystoids, Diagnostics 13 (15) (2023).

[4] Aditya Tripathi, Preetham Kumar, Veena Mayya, Akshat Tulsani, Generating OCT B-Scan DME images using optimized generative adversarial networks (GANs), Heliyon 9 (8) (2023).

[5] Suchetha Manikandan, Rajiv Raman, Ramachandran Rajalakshmi, S. Tamilselvi, R. Janani Surya, Deep learning-based detection of diabetic macular edema using optical coherence tomography and fundus images: A meta-analysis, Indian J. Ophthalmol. 71 (5) (2023) 1783–1796.

[6] Atsuya Miki, Felipe A. Medeiros, Robert N. Weinreb, Sonia Jain, Feng He, Lucie Sharpsten, Naira Khachatryan, Na'Ama Hammel, Jeffrey M. Liebmann, Christopher A. Girkin, Pamela A. Sample, Linda M. Zangwill, Rates of retinal nerve fiber layer thinning in glaucoma suspect eyes, Ophthalmology 121 (7) (2014) 1350–1358.

[7] Lyvia Zhang, Elon H.C. Van Dijk, Enrico Borrelli, Serena Fragiotta, Mark P. Breazzano, OCT and OCT angiography update: Clinical application to age-related macular degeneration, central serous chorioretinopathy, macular telangiectasia, and diabetic retinopathy, Diagnostics 13 (2) (2023).

[8] Neslihan Dilruba Koseoglu, Andrzej Grzybowski, T.Y. Alvin Liu, Deep learning applications to classification and detection of age-related macular degeneration on optical coherence tomography imaging: A review, Ophthalmol. Therapy 12 (5) (2023) 2347–2359.

[9] Mateo Gende, Victor Mallen, Joaquim de Moura, Beatriz Cordon, Elena Garcia-Martin, Clara I. Sanchez, Jorge Novo, Marcos Ortega, Automatic segmentation of retinal layers in multiple neurodegenerative disorder scenarios, IEEE J. Biomed. Health Inf. (2023) 1–12.

[10] Pratul P. Srinivasan, Leo A. Kim, Priyatham S. Mettu, Scott W. Cousins, Grant M. Comer, Joseph A. Izatt, Sina Farsiu, Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images, Biomed. Opt. Express 5 (10) (2014) 3568–3577.

[11] Olaf Ronneberger, Philipp Fischer, Thomas Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241.

[12] Jonathan Long, Evan Shelhamer, Trevor Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.

[13] Mona Kathryn Garvin, Michael David Abramoff, Xiaodong Wu, Stephen R. Russell, Trudy L. Burns, Milan Sonka, Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomography images, IEEE Trans. Med. Imaging 28 (9) (2009) 1436–1447.

[14] Stephanie J. Chiu, Xiao T. Li, Peter Nicholas, Cynthia A. Toth, Joseph A. Izatt, Sina Farsiu, Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation, Opt. Express 18 (18) (2010) 19413–19428.

[15] Seifedine Kadry, Ayman Abdallah, Chibli Joumaa, On the optimization of Dijkstra's algorithm, in: Informatics in Control, Automation and Robotics, Springer, 2011, pp. 393–397.

[16] Yu-Ying Liu, Mei Chen, Hiroshi Ishikawa, Gadi Wollstein, Joel S. Schuman, James M. Rehg, Automated macular pathology diagnosis in retinal OCT images using multi-scale spatial pyramid and local binary patterns in texture and shape encoding, Med. Image Anal. 15 (5) (2011) 748–759.

[17] Stephanie J. Chiu, Michael J. Allingham, Priyatham S. Mettu, Scott W. Cousins, Joseph A. Izatt, Sina Farsiu, Kernel regression based segmentation of optical coherence tomography images with diabetic macular edema, Biomed. Opt. Express 6 (4) (2015) 1172–1194.

[18] Karthik Gopinath, Samrudhdhi B. Rangrej, Jayanthi Sivaswamy, A deep learning framework for segmentation of retinal layers from OCT images, in: 2017 4th IAPR Asian Conference on Pattern Recognition, ACPR, IEEE, 2017, pp. 888–893.

[19] Abhijit Guha Roy, Sailesh Conjeti, Sri Phani Krishna Karri, Debdoot Sheet, Amin Katouzian, Christian Wachinger, Nassir Navab, ReLayNet: Retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks, Biomed. Opt. Express 8 (8) (2017) 3627–3642.

[20] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y. Hammerla, Bernhard Kainz, et al., Attention u-net: Learning where to look for the pancreas, 2018, arXiv preprint arXiv:1804.03999.

[21] Donghuan Lu, Morgan Heisler, Da Ma, Setareh Dabiri, Sieun Lee, Gavin Weiguang Ding, Marinko V. Sarunic, Mirza Faisal Beg, Cascaded deep neural networks for retinal layer segmentation of optical coherence tomography with fluid presence, 2019, arXiv preprint arXiv:1912.03418.

[22] Ivana Zadro Matovinovic, Sven Loncaric, Julian Lo, Morgan Heisler, Marinko Sarunic, Transfer learning with U-net type model for automatic segmentation of three retinal layers in optical coherence tomography images, in: 2019 11th International Symposium on Image and Signal Processing and Analysis, ISPA, IEEE, 2019, pp. 49–53.

[23] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[24] Zeyu Fu, Yang Sun, Xiangyu Zhang, Scott Stainton, Shaun Barney, Jeffry Hogg, William Innes, Satnam Dlay, MPG-net: Multi-prediction guided network for segmentation of retinal layers in OCT images, in: 2020 28th European Signal Processing Conference, EUSIPCO, IEEE, 2021, pp. 1299–1303.

[25] Jiaxuan Li, Peiyao Jin, Jianfeng Zhu, Haidong Zou, Xun Xu, Min Tang, Minwen Zhou, Yu Gan, Jiangnan He, Yuye Ling, et al., Multi-scale GCN-assisted two-stage network for joint segmentation of retinal layers and discs in peripapillary OCT images, Biomed. Opt. Express 12 (4) (2021) 2204–2220.

[26] Kaidong Li, Nina Y. Wang, Yiju Yang, Guanghui Wang, Sgnet: A super-class guided network for image classification and object detection, in: 2021 18th Conference on Robots and Vision, CRV, IEEE, 2021, pp. 127–134.

[27] Mousa Moradi, Yu Chen, Xian Du, Johanna M. Seddon, Deep ensemble learning for automated non-advanced AMD classification using optimized retinal layer segmentation and SD-OCT scans, Comput. Biol. Med. (ISSN: 0010-4825) 154 (2023) 106512.

[28] Ye Lu, Yutian Shen, Xiaohan Xing, Chengwei Ye, Max Q.-H. Meng, Boundary-enhanced semi-supervised retinal layer segmentation in optical coherence tomography images using fewer labels, Comput. Med. Imaging Graph. (ISSN: 0895-6111) 105 (2023) 102199.

[29] Mateo Gende, Joaquim de Moura, José Ignacio Fernández-Vigo, José María Martínez-De-la Casa, Julián García-Feijóo, Jorge Novo, Marcos Ortega, Robust multi-view approaches for retinal layer segmentation in glaucoma patients via transfer learning, Quant. Imaging Med. Surg. 13 (5) (2023) 2846–2859.

[30] Antoni Buades, Bartomeu Coll, Jean-Michel Morel, Non-local means denoising, Image Process. Line 1 (2011) 208–212.

[31] T. Guru Pradeep Reddy, Kandiraju Sai Ashritha, T.M. Prajwala, G.N. Girish, Abhishek R. Kothari, Shashidhar G. Koolagudi, Jeny Rajan, Retinal-layer segmentation using dilated convolutions, Adv. Intel. Syst. Comput. 1022 AISC (2020) 279–292.

[32] B.N. Anoop, Rakesh Pavan, G.N. Girish, Abhishek R. Kothari, Jeny Rajan, Stack generalized deep ensemble learning for retinal layer segmentation in optical coherence tomography images, Biocybern. Biomed. Eng. 40 (4) (2020) 1343–1358.

[33] Karamjeet Singh, Sukhjeet Kaur Ranade, Chandan Singh, Method noise based two stage nonlocal means filtering approach for Gaussian noise reduction, in: Proceedings of Sixth International Conference on Soft Computing for Problem Solving, Springer, 2017, pp. 178–187.

[34] Yuxin Wu, Kaiming He, Group normalization, in: Proceedings of the European Conference on Computer Vision, ECCV, 2018, pp. 3–19.

[35] Sergey Ioffe, Christian Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, in: International Conference on Machine Learning, PMLR, 2015, pp. 448–456.

[36] Yang Sun, Zeyu Fu, Scott Stainton, Shaun Barney, Jeffry Hogg, William Innes, Satnam Dlay, Automated retinal layer segmentation of OCT images using two-stage FCN and decision mask, in: 2019 IEEE International Symposium on Signal Processing and Information Technology, ISSPIT, IEEE, 2019, pp. 1–6.

[37] Xiaoming Liu, Jun Cao, Shaocheng Wang, Ying Zhang, Man Wang, Confidence-guided topology-preserving layer segmentation for optical coherence tomography images with focus-column module, IEEE Trans. Instrum. Meas. 70 (2020) 1–12.

[38] Yufan He, Aaron Carass, Yihao Liu, Bruno M. Jedynak, Sharon D. Solomon, Shiv Saidha, Peter A. Calabresi, Jerry L. Prince, Deep learning based topology guaranteed surface and MME segmentation of multiple sclerosis subjects from retinal OCT, Biomed. Opt. Express 10 (10) (2019) 5042–5058.

[39] Jonathan Long, Evan Shelhamer, Trevor Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 07-12-June-2015, 2015, pp. 431–440.

[40] Jie Wang, Zhe Wang, Fei Li, Guoxiang Qu, Yu Qiao, Hairong Lv, Xiulan Zhang, Joint retina segmentation and classification for early glaucoma diagnosis, Biomed. Opt. Express 10 (5) (2019) 2639–2656.

[41] Zhijun Gao, Zhiming Wang, Yi Li, A novel intraretinal layer semantic segmentation method of fundus OCT images based on the TransUNet network model, Photonics 10 (4) (2023).

[42] Zhijun Gao, Lun Chen, Research on semantic segmentation method of macular edema in retinal OCT images based on improved swin-unet, Electronics 11 (15) (2022).

[43] Arunava Chakravarty, Jayanthi Sivaswamy, A supervised joint multi-layer segmentation framework for retinal optical coherence tomography images using conditional random field, Comput. Methods Programs Biomed. 165 (2018) 235–250.

[44] Sripad Krishna Devalla, Prajwal K. Renukanand, Bharathwaj K. Sreedhar, Giridhar Subramanian, Liang Zhang, Shamira Perera, Jean-Martial Mari, Khai Sing Chin, Tin A. Tun, Nicholas G. Strouthidis, Tin Aung, Alexandre H. Thiéry, Michaël J.A. Girard, DRUNET: A dilated-residual U-net deep learning network to segment optic nerve head tissues in optical coherence tomography images, Biomed. Opt. Express 9 (7) (2018) 3244–3265.

[45] Yiqian Wang, Carlo Galang, William R. Freeman, Alexandra Warter, Anna Heinke, Dirk-Uwe G. Bartsch, Truong Q. Nguyen, Cheolhong An, Retinal OCT layer segmentation via joint motion correction and graph-assisted 3D neural network, IEEE Access 11 (2023) 103319–103332.

[46] Xiaoming Liu, Jun Cao, Tianyu Fu, Zhifang Pan, Wei Hu, Kai Zhang, Jun Liu, Semi-supervised automatic segmentation of layer and fluid region in retinal optical coherence tomography images using adversarial learning, IEEE Access 7 (2019) 3046–3061.

[47] Jun Wu, Shuang Liu, Zhitao Xiao, Fang Zhang, Lei Geng, Joint segmentation of retinal layers and macular edema in optical coherence tomography scans based on RLMENet, Med. Phys. 49 (11) (2022) 7150–7166.

[48] Arunodhayan Sampath Kumar, Tobias Schlosser, Holger Langner, Marc Ritter, Danny Kowerko, Improving OCT image segmentation of retinal layers by utilizing a machine learning based multistage system of stacked multiscale encoders and decoders, Bioengineering (ISSN: 2306-5354) 10 (10) (2023).