Nonlinear Analysis: Modelling and Control, Vol. 29, No. 1, 96–110
https://doi.org/10.15388/namc.2024.29.33669

Vilnius University Press

Impact of eye fundus image preprocessing on key objects segmentation for glaucoma identification

Sandra Virbukaitė, Jolita Bernatavičienė

Institute of Data Science and Digital Technologies, Vilnius University,
Akademijos str. 4, Vilnius LT-08412, Lithuania
sandra.virbukaite@mif.vu.lt

Received: April 1, 2023 / **Revised:** November 3, 2023 / **Published online:** November 27, 2023

Abstract. The pathological changes in the eye fundus image, especially around Optic Disc (OD) and Optic Cup (OC) may indicate eye diseases such as glaucoma. Therefore, accurate OD and OC segmentation is essential. The variety in images caused by different eye fundus cameras makes the complexity for the existing deep learning (DL) networks in OD and OC segmentation. In most research cases, experiments were conducted on individual data sets only and the results were obtained for that specific data sample. Our future goal is to develop a DL method that segments OD and OC in any kind of eye fundus image but the application of the mixed training data strategy is in the initiation stage and the image preprocessing is not discussed. Therefore, the aim of this paper is to evaluate the image preprocessing impact on OD and OC segmentation in different eye fundus images aligned by size. We adopted a mixed training data strategy by combining images of DRISHTI-GS, REFUGE, and RIM-ONE datasets, and applied image resizing incorporating various interpolation methods, namely bilinear, nearest neighbor, and bicubic for image resolution alignment. The impact of image preprocessing on OD and OC segmentation was evaluated using three convolutional neural networks Attention U-Net, Residual Attention U-Net (RAUNET), and U-Net++. The experimental results show that the most accurate segmentation is achieved by resizing images to a size of 512×512 px and applying bicubic interpolation. The highest Dice of 0.979 for OD and 0.877 for OC are achieved on DRISHTI-GS test dataset, 0.973 for OD and 0.874 for OC on the REFUGE test dataset, 0.977 for OD and 0.855 for OC on RIM-ONE test dataset. Anova and Levene's tests with statistically significant evidence at $\alpha = 0.05$ show that the chosen size in image resizing has an impact on the OD and OC segmentation results, meanwhile, the interpolation method does influent OC segmentation only.

Keywords: image preprocessing, optic disc segmentation, optic cup segmentation, eye fundus images, deep neural network.

1 Introduction

Glaucoma is a group of eye conditions that damage the optic nerve and is the second most common eye disease causing blindness. Fortunately, the early identification and treatment of glaucoma can help preserve vision. In glaucoma examination, accurate segmentation

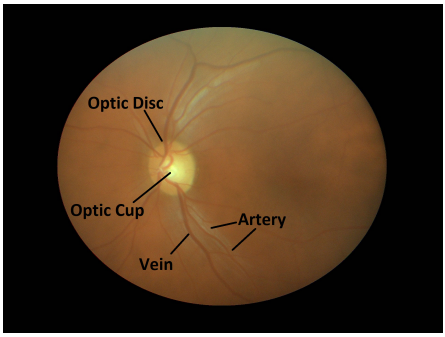


Figure 1. Eye fundus image.

of the key objects of the eye fundus image (Fig. 1), namely the optic cup (OC) and optic disc (OD) is essential. The OD occurs as a bright oval area, and the OC occurs as the brighter oval area in the center of the optic disc. The ratio of these objects, called Cup-to-Disc ratio (CDR) [7], may indicate damage to the optic nerve. The CDR of a healthy eye is about 0.3. The CDR of 0.4, $0.5-0.7$, and above 0.7 indicate mild glaucoma, moderate glaucoma, and severe glaucoma accordingly [16].

Eye health assessment for glaucoma examination, which is manual and time-consuming, can be automated with the help of computer-aided methods. One of the computer-aided methods is a convolutional neural network (CNN), which uses images of the same size. Depending on the eye fundus camera, the size of eye fundus images might vary in size of 1920×1440 px [28], 2144×1424 px [26], 2124×2056 px [29]. The images of size as such are too large to be directly fed into the CNN. To reduce the original image's dimensionality, various region of interest (ROI) extraction techniques are used. Depending on ROI extraction technique, the ROI is extracted of different sizes, which still can cause high computation power. Thus the resizing of the ROI image to a smaller dimension is necessary [1]. Image resizing involves interpolation, which comes in different methods, including bicubic, bilinear, nearest neighbor, and others. The other reason for image resizing is the size alignment of images of different datasets. Table 1 presents the typical eye fundus images from three publically available datasets, namely DRISHTI-GS [13], REFUGE [10] and RIM-ONE v.3 [27], and their ROI. As can be seen, the images in different datasets differ not only in size but also in parameters and resolution. The size of full images in the DRISHTI-GS dataset is 2045×1752 px, and the ROI size varies from 674×674 to 1060×1060 px. The size of full images in the REFUGE dataset is 2124×2056 px, ROI takes a size in a range from 408×408 to 616×616 px. The size of full images in RIM-ONE is 2144×1424 px, the ROI size varies from 456×456 to 890×890 px.

In most research cases, applying deep learning-based methods for OD and OC segmentation, the experiments are conducted on individual data sets only and the results are obtained for that specific data sample [4, 9, 22]. The application of the mixed training strategy based on different eye fundus image datasets [24] is in the initiation stage and authors are focusing on the architecture development of deep learning networks. Though

Table 1. Images comparison of different datasets.

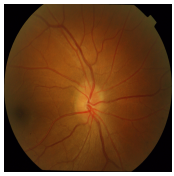
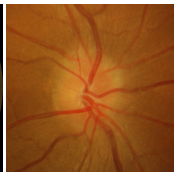
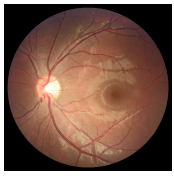
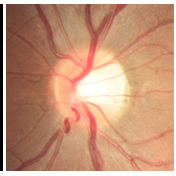
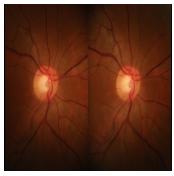
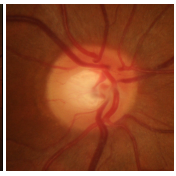
Dataset	Image	ROI	Image size, px	Size range of ROIs, px
DRISHTI-GS			2045×1752	$674 \times 674 - 1060 \times 1060$
REFUGE			2124×2056	$408 \times 408 - 616 \times 616$
RIM-ONE			2144×1424	$456 \times 456 - 890 \times 890$

image resizing for images of different datasets alignment is necessary, the impact of different image resizing techniques on the performance of deep-learning-based methods for OD and OC segmentation has not been explored yet.

In this research, we applied a mixed data strategy by combining eye fundus images of different datasets DRISHTI-GS, REFUGE, and RIM-ONE, and aim to evaluate the impact on the OD and OC segmentation results depending on applied image size alignment methods such as image resizing using:

- Different image size;
- Different interpolation method.

The subsequent sections of this paper are organized as follows: Section 2 presents the related work of convolutional neural networks including image preprocessing; Section 3 describes the materials and methods used in our experiments; Section 4 presents our achieved results for OD and OC segmentation; Section 5 puts forward the research conclusions.

2 Related work

Convolutional neural networks have recently gained popularity in automated medical image segmentation. The researchers, inspired by the traditional U-Net architecture [11], which works well for binary segmentation, are continuing to improve it to achieve higher OD and OC segmentation accuracy (Table 2). Here, ROI extraction accompanied by resizing is one of the main steps in image preprocessing. Liu et al. [6] proposed a deep learning-based model named densely connected depthwise separable convolution network

Table 2. Comparison between different U-Net modifications for OD and OC segmentation with the DRISHTI-GS, REFUGE and RIM-ONE datasets by Dice.

References	Method	DRISHTI-GS		REFUGE		RIM-ONE	
		OD	OC	OD	OC	OD	OC
Zhu et al. [24]	U-Net	0.96	0.84	0.96	0.88	0.82	0.68
Tulsani et al. [15]	Improved UNET++	0.97	0.95	–	–	–	–
Zhu et al. [24]	U-Net++	0.96	0.83	0.96	0.88	0.82	0.65
Jiang et al. [4]	Multi-Path Recurrent U-Net	0.98	0.89	–	–	–	–
Shyamalee et al. [9]	Attention U-Net	–	–	–	–	0.99	0.97
Zhao et al. [22]	Attention U-Net incorporating transfer learning	0.96	0.88	–	–	–	–

(DDSC-Net). In this method, firstly the OD was located. After that, the OD and OC were segmented jointly according to the interesting areas. In ROI extraction, a Circular Hough Transform was applied for the center and radius of the OD calculation. The cropped ROI was resized to a size of 240×240 px.

For U-NET modification presented in [12] the images were cropped by OD area, such that their diameter occupied a fifth part of an image-side length. The ROI was resized into a size of 128×128 px by bilinear interpolation.

The U-Net modification with the removed dropout layer proposed by Juneja et al. [5] consisted of two different convolutional neural networks for separate OD and OC segmentation. The merge layers were incorporated for the merge of previous and upcoming layers in terms of the loss reduction of data across various layers. The retinal fundus images were cropped to size 512×512 px and resized to the size 128×128 px.

Yi et al. [21] developed a particle swarm optimization-based method, named SePSO, aimed to segment the clear boundary of OD and reduce the interference of blood vessels and OC. Here, by changing the position of the particles, the deformation of the segmented contour was completed. The optimal contour was solved by an iterative update. Additionally, the constraint equations for particle position and velocity were optimized. The extracted ROI was resized to 256×256 px by a Lanczos interpolation over an 8×8 px neighborhood.

In the U-Net-based coarse-to-fine deep learning framework for OD segmentation proposed by Wanga et al. [18], the pixel intensities and vessel density information were integrated for the OD localization and segmentation. Particularly, a vessel density map was developed to characterize the spatial relationship between OD and retinal vessels. This helped to exclude false positive objects in the whole image region. In ROI preparation, the dark background was excluded, the field of view of size 512×512 px was cropped and resized to a dimension of 256×256 px.

Xiong et al. [20] addressed the limitation of the supervised machine learning-based training, which requires the pixel-level annotation of the OD masks. To overcome this limitation, a weak label-based Bayesian U-Net using Hough transform-based annotations was proposed. In ROI extraction, the Gaussian blur was applied that obtained the brightest pixels in the images. With the help of the Canny edge detector, the edges within the ROI were utilized and expanded by image dilation to make it detectable with the Hough transform.

Jiang et al [4] proposed a combination of a recurrent neural network and a convolutional neural network, named Multi-Path Recurrent U-Net. The proposed architecture was constructed of coding and decoding paths. The coding path consisted of four layers of convolutional blocks. In the decoding path, the convolutional blocks were located at the nodes of each layer. The skip connections were placed between the corresponding coding and decoding layers. This made the network to be able to keep low-level features for the final prediction. The proposed method used two types of operations, namely average pooling and maximum pooling. The usage of maximum pooling operation helped to retain more texture information.

Gao et al. [2] proposed a network named Recurrent Fully Convolution Network (RFC-Net) for automatic joint OD and OC segmentation aiming to minimize the loss of spatial information. The usage of the recurrent convolution layer ensured feature representation for OD and OC segmentation tasks through feature collection. The ROI was prepared by taking images of different sizes in pixels, including 400×400, 500×500, 550×550, 600×600, 650×650, 700×700, 750×750, 800×800, 850×850, and 900×900 px based on the center point of the OD by applying the YOLOv2 model. The images were scaled to the size of 512×512 px.

3 Materials and methods

The high-level view of the applied methodology, including training data of different datasets preparation by applying ROI extraction, augmentation, and resizing, training and evaluation strategies of different convolutional neural networks are illustrated in Fig. 2. A detailed description of this pipeline is presented Sections 3.1–3.4. In Section 3.1, we provide a description of the datasets used in our experiments. Section 3.2 presents the applied image preprocessing techniques. In Section 3.3, we describe the convolutional neural networks used in this paper, the CNNs parameters and training approach, and implementation details. Section 3.4 provides the description of metrics used for CNN performance evaluation and the test datasets.

3.1 Datasets description

REFUGE [10] is a public dataset consisting of 1200 fundus images originally, which are saved in JPG image format and with ground truth and clinical glaucoma labels, which are saved in BMP image format. The dataset is split $1 : 1 : 1$ into 3 subsets equally for training, validation, and testing, stratified to have equal glaucoma presence percentage. The training set with a total of 400 color fundus images captured by a Zeiss Visucam 500 fundus camera of resolution 2124×2056 px is provided together with the corresponding glaucoma status and the unified manual pixel-wise ground truths. The testing dataset consists of 800 color fundus images captured by a Canon CR-2 camera of size 1634×1634 px and is further split into 400 validation set images and 400 test set images. In our experiments, the images of validation and testing subsets were used only.

RIM-ONE v.3 [27] is a public dataset containing 159 stereo eye fundus images, which are taken by Nidek AFC-210 camera and saved in JPEG image format with a resolution

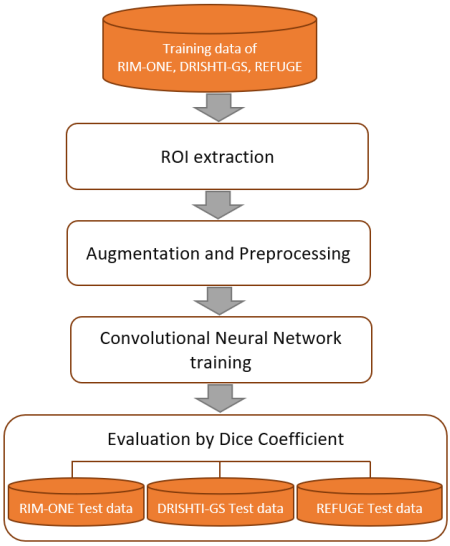


Figure 2. Flow diagram of applied methodology.

of 2144×1424 px. The OD of each image has been segmented by 2 experts in ophthalmology to create the ground truth, which is saved in PNG image format.

DRISHTI-GS [13] is a public dataset consisting of 101 images, which are divided into 50 training and 51 testing images. All the images have been marked by 4 eye experts with varying clinical experience. All images were taken centered on OD with a Field-Of-View (FOV) of 30 degrees and saved in the PNG uncompressed image format with a resolution of 2045×1752 px. Ground truth was collected from data experts with varying clinical experience of 3, 5, 9, and 20 years, respectively.

From the literature review, it was observed that the same dataset is usually used for training and validating deep learning-based models. This way of learning leads to the adaptation of the model to the same type of images. To overcome this issue, we aim to use as many images of different quality and sizes as possible for deep learning-based model training. Since there are few studies where the mixed-dataset approach is applied, for our research we combined the images from these three different datasets into one training dataset and evaluated CNN's ability in the OD and OC segmentation on the mixed dataset. Also, we evaluated the impact of image preprocessing on the OD and OC segmentation results.

3.2 Preprocessing

At the image preprocessing stage, due to different image formats in different datasets described in Section 3.1, the image format was aligned by converting the images and their ground truths to PNG uncompressed image format. To increase the diversity and prevent overfitting, the number of images in each dataset was amplified to 1000 by applying

various image augmentation techniques, such as image rotation by an angle of rotation from $0°$ to $45°$, zooming by 20%, and horizontal and vertical flipping. The different distributions in datasets have been aligned by including the same number of images from each dataset in the training dataset.

The ROI was extracted by cropping the area of OD. The cropped ROIs varied in size depending on the original fundus image sizes, including from 674×674 to 1060×1060 px in the DRISHTI-GS dataset, from 408×408 to 616×616 px in REFUGE dataset, from 456×456 to 890×890 px in RIM-ONE dataset. The variety of cropped image sizes was aligned by applying the image resizing to sizes of 128×128, 256×256, and 512×512 px incorporating the three most common interpolation methods [3], such as bilinear, nearest neighbor, and bicubic.

In *bilinear* interpolation [3], the weighted average of the four nearest neighbors is computed to produce the output:

$$f(x,y) = \frac{1}{(x_2 - x_1)(y_2 - y_1)} [x_2 - xx - x_1] \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} y_2 & y \\ y & y_1 \end{bmatrix},$$

where x_1, y_1, x_2, y_2 – coordinates of interpolation base points, $Q_{11}, Q_{12}, Q_{21}, Q_{22}$ – values of these points, x, y – the point at which we want to interpolate.

In *nearest neighbor* interpolation [3], instead of weighted average value computation, the shortest distance to the neighbor pixels is computed:

$$f(x,y) = F\big(\lfloor x + 0.5 \rfloor, \lfloor y + 0.5 \rfloor\big),$$

where F is the samples of an image f, (x,y) – the point at which we want to interpolate by nearest neighbor interpolation.

Bicubic interpolation [3] uses the weighted average of the sixteen nearest neighbors to produce the output:

$$f(x,y) = \sum_{i=0}^{3} \sum_{j=0}^{3} a_{ij} x^i y^j,$$

where a – unknown coefficients, (x,y) – the point at which we want to interpolate by bicubic interpolation.

3.3 Training and implementation

A literature review reveals that most studies have been conducted using images of the same dataset for convolutional neural network training and validation. In the case when the CNN was trained on images from one data set and validated on images from another dataset [17], the OD and OC segmentation results dropped significantly. However, the improvements in the OD and OC segmentation results were noticeable by training a convolutional neural network on a mixed-dataset consisting of two different datasets, particularly DRISHTI-GS and RIM-ONE (Table 3). Inspired by this mixed-data strategy, in our study, aiming to improve CNN's ability to work on as many different images, we

Table 3. Results of Dice for OD and OC segmentation on images resized to 256×256 px of DRISHTI-GS, RIM-ONE, and mixed-dataset.

Train datasets	Test datasets			
	DRISHTI-GS		RIM-ONE	
	OD	OC	OD	OC
DRISHTI-GS	0.97	0.88	0.77	0.57
RIM-ONE	0.77	0.57	0.96	0.88
Mixed-dataset	0.94	0.84	0.91	0.83

extended the mixed-data approach by constructing training datasets of images from three datasets, namely DRISHTI-GS, RIM-ONE, and REFUGE. The different distributions in datasets have been aligned by including the same number of images from each dataset in the training dataset. Due to the variety of image sizes caused by different datasets, the images were resized to the following commonly [2, 5, 18] used sizes in px:

- 128×128, 256×256, 512×512 by bilinear interpolation;
- 128×128, 256×256, 512×512 by nearest neighbor interpolation;
- 128×128, 256×256, 512×512 by bicubic interpolation.

The three latest U-Net modifications, particularly Attention U-Net [9], Residual Attention U-Net (RAUNet) [8], and UNet++ [23] with significant improvement in object segmentation have been chosen to be trained on these mixed-datasets for OD and OC segmentation.

Attention U-Net [9] overcomes the limitation of traditional U-Net architecture of a lack of feature information in the initial layers. The proposed architecture consists of the encoder, decoder, and attention gate at the skip connection of each level. With attention modules, the network focuses more on important spatial regions within an image and less on those of lower importance. Each convolution layer is followed by a rectified linear units (ReLU) activation and batch normalization.

UNet++ [23] is a nested and dense skip connections-based architecture. Compared with traditional U-Net, the skip pathways of UNet++ are redesigned to connect the two sub-networks and use deep supervision. In the encoder, the feature maps undergo a dense convolution block whose number of convolution layers depends on the pyramid level. Because of the nested skip pathways, the proposed architecture generates full-resolution feature maps at multiple semantic levels, which are liable to deep supervision.

Residual Attention U-Net (RAUNet) [8] is encoder-decoder-based architecture, which aims to get high-resolution masks. The encoder consists of pre-trained ResNet34 to extract semantic features. With this help, the model size is reduced while inference speed is improved. The decoder consists of a new augmented attention module(AAM) to fuse multi-level features and capture global context. Transposed convolution is applied to carry out upsampling for acquiring refined edges.

The CNN's training was performed on a single GPU machine [25] with 1 TB of RAM in the Keras and TensorFlow frameworks by applying an early stopping technique to reduce unnecessary training time. The Adam optimizer and binary cross-entropy loss function were used during the training of these three convolutional neural networks,

namely Attention U-Net, UNet++, and Residual Attention U-Net (RAUNet). The parameters, namely learning rate, batch size, and the dropout rate for each convolutional neural network were searched separately by applying KerasTuner framework. In the UNet++ case, the training time in seconds per step according to the image size 512×512, 256×256, and 128×128 px are 250, 130, and 43, respectively. To train the RAUNet it took 347, 173, and 63 seconds per step according to the image size 512×512, 256×256, and 128×128 px, respectively. In the Attention U-Net case, the training time in seconds per step according to the image size 512×512, 256×256, and 128×128 px are 392, 189, and 67.

3.4 Metrics and evaluation

The trained Attention U-Net, RAUNet, and UNet++ were tested on 50 test images of each database, meaning DRISHTI-GS, REFUGE, and RIM-ONE, separately. The performance of OD and OC segmentation was evaluated by the Dice coefficient (Dice), which is used in most cases [4, 15, 24] to describe the similarity between the two images.

$$\text{Dice} = \frac{2|S \cap L|}{|S| + |L|},$$

where, S – the result based on segmentation, L – the ground truth label.

The statistical significance at level α = 0.05 of the OD and OC segmentation results were evaluated using statistical methods such as the analysis of variance (ANOVA) test [14] and Levene's test [19].

ANOVA is a test of the hypothesis that is appropriate to compare the means of a continuous variable in two or more comparison groups.

Levene's test assesses the hypothesis of the equality of variances for a variable in two or more comparison groups.

4 Results

The three different convolutional neural networks, particularly Attention U-Net, UNet++, and Residual Attention U-Net were trained on mixed-dataset.

In this section, we present the obtained OD and OC segmentation results, the calculated statistics to evaluate the significance of the metric, and a few examples of cases of segmented OD and OC by one of three CNNs.

The obtained Dice score of the optic disc and optic cup segmentation by applying different methods, particularly Attention U-Net, RAUNet, and UNet++ to be trained on mixed-datasets consisting of differently preprocessed images are presented in Table 4. The results indicate that the most accurate segmentation is achieved using eye fundus images resized to a size of 512×512 px by bicubic interpolation. However, the CNN training on images of size 512×512 px requires approximately 2.1 and 5.5 times more computational power than training on images of size 256×256 and 128×128 px, respectively. In a comparison of Convolutional Neural Networks, the best segmentation

Table 4. Dice of OD and OC segmentation by different interpolation methods and image size.

CNN and Dataset		512 × 512 px		256 × 256 px		128 × 128 px	
		OD	OC	OD	OC	OD	OC
		BILINEAR					
UNet++	REFUGE	0.963	0.855	0.954	0.844	0.935	0.829
	RIM-ONE	0.962	0.811	0.938	0.728	0.927	0.715
	DRISHTI-GS	0.964	0.855	0.599	0.849	0.944	0.824
RAUNet	REFUGE	0.950	0.827	0.939	0.804	0.910	0.781
	RIM-ONE	0.948	0.808	0.929	0.791	0.908	0.777
	DRISHTI-GS	0.950	0.826	0.937	0.809	0.912	0.787
Attention U-Net	REFUGE	0.966	0.864	0.965	0.855	0.938	0.842
	RIM-ONE	0.963	0.820	0.961	0.809	0.938	0.782
	DRISHTI-GS	0.967	0.860	0.961	0.859	0.947	0.845
		NEAREST NEIGHBOR					
UNet++	REFUGE	0.961	0.858	0.949	0.848	0.934	0.822
	RIM-ONE	0.961	0.836	0.937	0.769	0.913	0.659
	DRISHTI-GS	0.965	0.867	0.958	0.839	0.952	0.800
RAUNet	REFUGE	0.954	0.838	0.935	0.823	0.912	0.808
	RIM-ONE	0.946	0.811	0.927	0.805	0.903	0.810
	DRISHTI-GS	0.953	0.834	0.936	0.825	0.916	0.796
Attention U-Net	REFUGE	0.968	0.865	0.966	0.852	0.942	0.848
	RIM-ONE	0.963	0.839	0.962	0.825	0.939	0.798
	DRISHTI-GS	0.968	0.872	0.966	0.869	0.953	0.854
		BICUBIC					
UNet++	REFUGE	**0.964**	**0.862**	**0.957**	**0.857**	**0.944**	**0.849**
	RIM-ONE	**0.965**	**0.846**	**0.951**	**0.827**	**0.936**	**0.793**
	DRISHTI-GS	**0.970**	**0.873**	**0.966**	**0.859**	**0.956**	**0.836**
RAUNet	REFUGE	**0.951**	**0.846**	**0.949**	**0.831**	**0.931**	**0.815**
	RIM-ONE	**0.949**	**0.830**	**0.937**	**0.829**	**0.928**	**0.819**
	DRISHTI-GS	**0.956**	**0.853**	**0.944**	**0.834**	**0.922**	**0.818**
Attention U-Net	REFUGE	**0.973**	**0.874**	**0.969**	**0.873**	**0.950**	**0.859**
	RIM-ONE	**0.977**	**0.855**	**0.967**	**0.828**	**0.952**	**0.813**
	DRISHTI-GS	**0.979**	**0.877**	**0.968**	**0.870**	**0.959**	**0.862**

results are achieved using Attention U-Net, where the highest Dice of 0.979 for OD and 0.877 for OC are observed on the DRISHTI-GS test dataset. Therefore, in the significance evaluation of the obtained OD and OC segmentation results, statistics, namely the mean and variance were calculated for the Dice score obtained using Attention U-Net on images resized to a size of 512×512 px and presented in Table 5.

Based on a p-value with statistically significant evidence at $\alpha = 0.05$ of Anova and Levene's tests, there is no significant difference between Dice results obtained on test data of different datasets using images resized by different interpolation methods for OD segmentation. However, the impact of the interpolation methods used in image resizing exists for OC segmentation.

As the Attention U-Net with applied bicubic interpolation in image resizing demonstrated the highest optic disc and optic cup segmentation results, this approach was used in the calculation of the statistical significance of sizes, meaning 128×128, 256×256, and 512×512 px, used in image resizing. Anova and Levene's test analysis showed that

Table 5. Statistics of Dice for OD and OC segmentation on images of size 512×512 px using Attention U-net.

Statistics	p-value*		BILINEAR		NEAREST NEIGHBOR		BICUBIC	
	OD	OC	OD	OC	OD	OC	OD	OC
				DRISHTI-GS				
Mean	0.124	0.001	0.974	0.830	0.975	0.817	0.979	0.856
Variance	0.144	0.002	0.002	0.015	0.002	0.018	0.001	0.013
				REFUGE				
Mean	0.185	0.003	0.966	0.838	0.967	0.847	0.973	0.856
Variance	0.177	0.001	0.000	0.006	0.000	0.005	0.000	0.002
				RIM-ONE				
Mean	0.271	0.000	0.972	0.815	0.973	0.759	0.977	0.837
Variance	0.158	0.001	0.000	0.013	0.000	0.033	0.000	0.010

*p-value at significance level $\alpha = 0.05$ by ANOVA test for means comparison and by Levene's test for variances comparison.

Table 6. Statistics of Dice for OD and OC segmentation on images resized by bicubic interpolation to different sizes using Attention U-net.

Statistics	p-value*		512×512 px		256×256 px		128×128 px	
	OD	OC	OD	OC	OD	OC	OD	OC
				DRISHTI-GS				
Mean	0.002	0.002	0.979	0.866	0.964	0.865	0.955	0.845
Variance	0.002	0.003	0.001	0.013	0.001	0.016	0.001	0.044
				REFUGE				
Mean	0.002	0.001	0.973	0.836	0.968	0.860	0.950	0.865
Variance	0.002	0.006	0.000	0.005	0.003	0.004	0.001	0.007
				RIM-ONE				
Mean	0.005	0.001	0.969	0.847	0.956	0.820	0.949	0.807
Variance	0.003	0.001	0.002	0.013	0.004	0.032	0.005	0.052

*p-value at significance level $\alpha = 0.05$ by ANOVA test for means comparison and by Levene's test for variances comparison.

the chosen size in image resizing does influence the OD and OC segmentation results (Table 6).

We combined the test data of DRISHTI-GS, REFUGE, and RIM-ONE datasets into three datasets of size range groups as $[400 \times 400, 550 \times 550)$ px, $[550 \times 550, 700 \times 700)$ px, and $[700 \times 700, 1060 \times 1060]$ px by extracted ROI size. Each group consisted of 57, 43, and 50 images accordingly. The histogram in Fig. 3 demonstrates the average of achieved Dice for optic disc (gray column) and optic cup (orange column) segmentation by applying bicubic interpolation in image resizing to a size of 512×512 px from the initial ROI size. It is observed that in the case of OD segmentation, the initial size where from ROI was resized to 512×512 px, has no impact on optic disc segmentation results. However, ROI images are sensitive to resizing for OC segmentation. The better quality of the optic cup is extracted from images of higher resolution.

Some visual test results obtained using Attention U-Net on DRISHTI-GS dataset images resized by different interpolation methods, including bilinear, nearest neighbor,

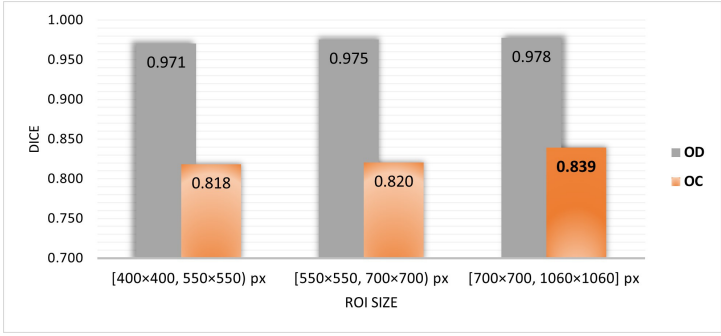


Figure 3. Dice score distribution of OD (gray column) and OC (orange column) segmentation on images resized by bicubic interpolation from initial ROI size to the size of 512×512 px.

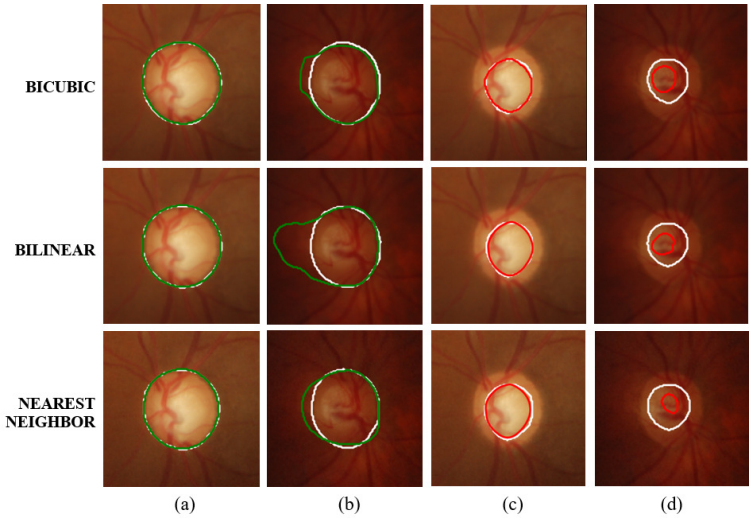


Figure 4. Optic disc and Optic cup segmentation results by different interpolation methods using Attention U-Net. (a) The best case of OD segmentation, (b) The worst case of OD segmentation, (c) The best case of OC segmentation, (d) The worst case of OC segmentation. The white circle indicates ground truth.

and bicubic are presented in Fig. 4. The green circle indicates the segmented optic disc, the red circle indicates the segmented optic cup, and the white circle indicates ground truth. The first column (a) shows the best case of optic disc segmentation, the second column (b) shows the worst case of optic disc segmentation, the third column (c) shows the best case of optic cup segmentation, and the fourth column (d) shows the worst case of optic cup segmentation. The best showcases of optic disc and optic cup segmentation do not demonstrate the difference between applied different interpolation methods in the image resizing process. However, the visual difference in worst cases for optic disc and optic cup segmentation is obvious. The network mostly fails in object segmentation when bilinear interpolation is applied in ROI image resizing.

5 Conclusions

In order to realize the future goal of developing a deep learning-based method to accurately segment optic disc and optic cup in any kind of eye fundus image, in this paper, we adopted a mixed data strategy by constructing a training dataset of different databases, including DRISHTI-GS, REFUGE, and RIM-ONE and evaluated the impact of image preprocessing required for the alignment of different size images using Attention U-Net, Residual Attention U-net, and U-Net++. The variety in fundus images of each database caused by different fundus cameras was aligned by applying image preprocessing techniques such as image resizing to sizes of 128×128, 256×256, and 512×512 px by bicubic, bilinear, and nearest neighbor interpolation methods.

From the experimental results, we can conclude that image preprocessing has an impact on OD and OC segmentation results. The significance evaluation of experimental results showed that there is a significant difference between Dice results obtained using images resized by different interpolation methods for optic cup segmentation, though optic disc segmentation results are not affected by the different interpolation methods used in image resizing.

Whereas the applied size in pixels, namely 128×128, 256×256, and 512×512 px in image resizing has a statistically significant impact on both objects, namely optic disc and optic cup segmentation. The most accurate segmentation is achieved by resizing images to a size of 512×512 px and applying the bicubic interpolation method. We would like to note that the CNN training on images of size 512×512 px requires approximately 2.1 and 5.5 times more computational power than training on images of size 256×256 and 128×128 px, respectively.

Comparing the convolutional neural networks, the best segmentation results are observed in training Attention U-Net on images of size 512×512 px. The highest Dice of 0.979 for OD and 0.877 for OC are achieved on DRISHTI-GS test dataset, 0.973 for OD and 0.874 for OC on the REFUGE test dataset, 0.977 for OD and 0.855 for OC on RIM-ONE test dataset.

The results of our experiments, where the most accurate OD and OC segmentation was achieved by resizing images to a size of 512×512 than 256×256 or 128×128 px, lead us to explore image resolution enhance methods such as super-resolution imaging in our future work.

Acknowledgment. We are thankful for the HPC resources provided by the `IT Research Center of Vilnius University`.

References

1. O.J. Afolabi, G.P. Mabuza-Hocquet, F.V. Nelwamondo, B.S. Paul, The use of U-Net Lite and extreme gradient boost (XGB) for glaucoma detection, *IEEE Access*, **9**:47411–47424, 2021.
2. J. Gao, Y. Jiang, H. Zhang, F. Wang, Joint disc and cup segmentation based on recurrent fully convolutional network, *PLoS ONE*, 2020, `https://doi.org/10.1371/journal.pone.0238983`.

3. D. Han, Comparison of commonly used image interpolation methods, in *Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013)*, Atlantis Press, 2013, pp. 1556–1559, https://doi.org/10.2991/iccsee.2013.391.

4. Y. Jiang, F. Wang, J. Gao, S. Cao, Multi-path recurrent U-Net segmentation of retinal fundus image, *Appl. Sci.*, **10**(11):3777, 2020, https://doi.org/10.3390/app10113777.

5. M. Juneja, S. Singh, N. Agarwal, S. Bali, S. Gupta, N. Thakur, P. Jindal, Automated detection of glaucoma using deep learning convolution network (G-net), *Multimedia Tools Appl.*, **79**: 15531–15553, 2020, https://doi.org/10.1007/s11042-019-7460-4.

6. B. Liu, D. Pan, H. Song, Joint optic disc and cup segmentation based on densely connected depthwise separable convolution deep network, *BMC Med. Imaging*, **21**, 2021, https://doi.org/10.1186/s12880-020-00528-6.

7. P. Liu, C.T. Tran, B. King, R. Fang, CADA: Multi-scale Collaborative Adversarial Domain Adaptation for unsupervised optic disc and cup segmentation, *Neurocomputing*, **469**:209–220, 2022, https://doi.org/10.1016/j.neucom.2019.05.039.

8. Z.-L. Ni, G.-B. Bian, X.-H. Zhou, Z.-G. Hou, X.-L. Xie, C. Wang, Y.-J. Zhou, R.-Q. Li, Z. Li, RAUNet: Residual attention U-Net for semantic segmentation of cataract surgical instruments, in *Neural Information Processing. ICONIP 2019*, Lect. Notes Comput. Sci., Vol. 11954, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-36711-4_13.

9. O. Oktay, J. Schlemper, L. Le Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N.Y. Hammerla, B. Kainz, B. Glocker, D. Rueckert, Attention U-Net: Learning Where to Look for the Pancreas, 2018, https://doi.org/10.48550/arXiv.1804.03999.

10. J.I. Orlando, H. Fu, J.B. Breda, K. van Keer, D.R. Bathula, A. Diaz-Pinto, R. Fang, P.A. Heng, J. Kim, J. Lee, REFUGE Challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs, *Med. Image Anal.*, **59**:101570, 2020, https://doi.org/10.1016/j.media.2019.101570.

11. O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in *Proceedings of Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Lect. Notes Comput. Sci., Vol. 9351, Springer, Cham, 2015, pp. 234–241, https://doi.org/10.1007/978-3-319-24574-4_28.

12. A. Sevastopolsky, Optic disc and cup segmentation methods for glaucoma detection with modification of U-Net convolutional neural network, *Pattern Recognit. Image Anal.*, **27**:618–624, 2017, https://doi.org/10.1134/S1054661817030269.

13. J. Sivaswamy, S.R. Krishnadas, G.D. Joshi, M. Jain, A.U.S. Tabish, Drishti-GS: Retinal image dataset for optic nerve head (ONH) segmentation, in *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), Beijing, China, 2014*, IEEE, 2014, pp. 53–56, https://doi.org/10.1109/ISBI.2014.6867807.

14. L. Sullivan, Hypothesis testing – analysis of variance (ANOVA), 2023, https://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704_hypothesistesting-anova/bs704_hypothesistesting-anova_print.html.

15. A. Tulsani, P. Kumar, S. Pathan, Automated segmentation of optic disc and optic cup for glaucoma assessment using improved UNET++ architecture, *Biocybern. Biomed. Eng.*, **41**: 819–832, 2021, https://doi.org/10.1016/j.bbe.2021.05.011.

16. S. Virbukaitė, J. Bernatavičienė, Deep learning methods for glaucoma identification using digital fundus images, *Baltic J. Modern Computing*, **8**(4):520–530, 2020, https://doi.org/10.22364/bjmc.2020.8.4.03.

17. S. Virbukaitė, J. Bernatavičienė, Image resizing impact on optic disc and optic cup segmentation, *Computer Science Research Notes*, **3201**:306–309, 2022, https://doi.org/10.24132/CSRN.3201.39.

18. L. Wanga, H. Liua, Y. Lub, H. Chenc, J. Zhangc, J. Pua, A coarse-to-fine deep learning framework for optic disc segmentation in fundus images, *Biomed. Signal Process. Control*, **51**: 82–89, 2019, https://doi.org/10.1016/j.bspc.2019.01.022.

19. S.A. Wind, Levene's homogeneity of variance test, 2023, https://doi.org/10.4135/9781506326139.

20. H. Xiong, S. Liu, R.V. Sharan, E. Coiera, S. Berkovsky, Weak label based Bayesian U-Net for optic disc segmentation in fundus images, *Artif. Intell. Med.*, **126**:102261, 2022, https://doi.org/10.1016/j.artmed.2022.102261.

21. J. Yi, Y. Ran, G. Yang, Particle swarm optimization-based approach for optic disc segmentation, *Entropy*, **24**(6):796, 2022, https://doi.org/10.3390/e24060796.

22. X. Zhao, S. Wang, J. Zhao, H. Wei, M. Xiao, N. Ta, Application of an attention U-Net incorporating transfer learning for optic disc and cup segmentation, *Signal Image Video Process.*, **15**:913–921, 2021, https://doi.org/10.1007/s11760-020-01815-z.

23. Z. Zhou, M. Siddiquee, N. Tajbakhsh, J. Liang, UNet++: A nested U-Net architecture for medical image segmentation, in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLMIA ML-CDS 2018 2018*, Lect. Notes Comput. Sci., Vol. 11045, Springer, Cham, 2018, https://doi.org/10.1007/978-3-030-00889-5_1.

24. Q. Zhu, X. Chen, Q. Meng, J. Song, G. Luo, M. Wang, F. Shi, Z. Chen, D. Xiang, L. Pan, Z. Li, W. Zhu, GDCSeg-Net: General optic disc and cup segmentation network for multi-device fundus images, *Biomed. Opt. Express*, **12**:6529–6544, 2021.

25. ITAPC, Information Technology Research Center of Vilnius University, https://mif.vu.lt/lt3/en/about/structure/it-research-center.

26. Nidek AFC-210 camera, Digital Eye Center, 2023, https://www.digitaleyecenter.com/product/nidek-afc-210-fundus-camera-refurbished/.

27. RIME-ONE v.3 Dataset, Medical Image Analysis Group, 2023, http://medimrg.webs.ull.es/research/retinal-imaging/rim-one/.

28. Smartscope M5 PRO camera, Ophthalmic Instruments, 2021, https://www.ophthalmic.com.sg/product/smartscope-m5-pro/.

29. VISUCAM 500 from ZEISS camera technical specification, Carl Zeiss Meditec AG, 2010, http://www.knsentp.com/userfiles/file/Eye-department/New-Fundus-Camera-Visucam-500.pdf.