Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15996
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Improving Optic Disc and Optic Cup Segmentation with Flip-Gamma Augmentation and SegFormer

Fitri Salamah[1)*], Erwin[2)], Anita Desiani[3)]
[1,2,3)] Universitas Sriwijaya, Palembang, Indonesia
[1)]fitrisalamah10@gmail.com, [2)]erwin@unsri.ac.id, [3)]anita_desiani@unsri.ac.id

Submitted : 13 March 2026 | **Accepted** : April 5, 2025 | **Published** : April 10, 2026

Abstract: The Cup-to-Disc Ratio (CDR) is widely used as a diagnostic indicator for glaucoma, although variations and irregularities can influence its accuracy in the Optic Disc (OD) and Optic Cup (OC). To overcome this challenge, automated image segmentation is used. However, image segmentation is challenged by image blurriness, noise, and uneven illumination, which can affect segmentation quality and increase the risk of misdiagnosis. To address these challenges, this study applies a combined Flip-Gamma Augmentation and SegFormer approach for OD and OC segmentation. Flip-Gamma augmentation increases image diversity and improves image quality by adjusting brightness and contrast. Meanwhile, the SegFormer uses a Transformer-based backbone and efficiently extracts multi-scale features to enhance segmentation performance. Experimental results on the Drishti-GS dataset show that applying Flip-Gamma (δ = 0.8, 0.9, 1.1, 1.2) is associated with improved segmentation performance across all classes, with sensitivity (90-99%), DSC (90-99%), IoU (82-99%), and ROC (94-99%), indicating consistent segmentation of OD, OC and background regions. Furthermore, a one-sided Mann-Whitney U test indicates differences in performance compared to other augmentation methods. These findings suggest that the proposed augmentation strategy is beneficial for segmentation on the Drishti-GS dataset. However, further validation on larger and more diverse datasets is required to assess generalizability.

Keywords: Augmentation, Glaucoma, Retinal image, SegFormer, Segmentation.

INTRODUCTION

Glaucoma is a significant cause of blindness worldwide, marked by optic nerve damage and progressive vision impairment that frequently goes undetected during the early stages (Kaushik et al. 2024). The Cup-to-Disc Ratio (CDR) is one of the important indicators used in glaucoma detection (Bian et al. 2020). It is calculated by comparing the diameters of the Optic Cup (OC) and the Optic Disc (OD) in retinal images. However, CDR estimation is highly dependent on the accuracy of OD and OC segmentation, which is prone to subjectivity and errors due to unclear structural boundaries and varying image quality (Liu et al. 2024). Therefore, automated segmentation is a key component in computer-based glaucoma detection systems (Gupta, Garg, and Agarwal 2022). Several studies have developed OD and OC segmentation methods using deep learning approaches and reported high performance, with Dice Similarity Coefficient (DSC) generally above 80% (Gagan et al. 2022; Lei et al. 2022; Liu et al. 2024). However, most of these studies still suffer from limitations in data collection and augmentation strategies. Several studies use datasets with limited size and low image variation and apply simple augmentations such as flipping or rotation without considering image quality enhancement. It may limit the model's generalizability when presented with low-quality images (Nuli, Desai, and Bhadri 2024; Yu et al. 2023).

Geometry-based augmentation techniques, such as flipping and rotation, have been widely used to increase data diversity (Kumar et al. 2023). While flipping preserves anatomical structures, rotation may distort important features if not applied carefully (Maiyanti et al. 2023). However, geometric augmentations only improve data variation and do not address image quality issues, such as blur, noise, and uneven illumination, which are common in fundus images (Sule, Viriri, and Abayomi 2020; Yu et al. 2023). Several studies have shown that poor image quality can degrade the performance of deep learning models in segmentation tasks (Jiang et al. 2020; Zhou et al. 2023). Although intensity-based techniques, such as gamma correction, can enhance image quality, the integration with geometric augmentation remains limited (Acharya and Giri 2020). Existing studies often address data

*name of corresponding

This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15996
e-ISSN : 2541-2019
p-ISSN : 2541-044X

diversity and image quality separately rather than in a unified framework, suggesting a potential research gap. Therefore, beyond data augmentation strategies, selecting an appropriate model architecture is equally important for achieving robust segmentation performance.

SegFormer is a recent Transformer-based segmentation model that uses Mix-FFN and multi-level feature aggregation to achieve accurate segmentation (Li et al. 2025; Xie et al. 2021). As a Transformer-based model lacking convolutional inductive biases, SegFormer requires high-quality and diverse training data to learn generalizable features (Mazraedoost 2024). Several studies have applied SegFormer in medical image segmentation. (Deng et al. 2024) used SegFormer for kidney tumor segmentation without any augmentation and achieved a DSC of 88%. (Yu et al. 2024) used SegFormer for periodontal bone segmentation using only flipping augmentation, resulting in a DSC of 74%. (Wu et al. 2024) applied SegFormer for ventricular segmentation without augmentation and achieved DSC of 69%. These results highlight the importance of data augmentation for improving segmentation performance, even with advanced models like SegFormer.

Based on several studies, there are several limitations in previous studies, namely: (1) the use of augmentation is still limited to geometric transformations, (2) the lack of exploration of the combination of geometric augmentation and image quality improvement, and (3) there are not many evaluations that directly compare various augmentation strategies within the same model framework, especially in OD and OC segmentation using SegFormer. Therefore, this study evaluates the effect of various augmentation strategies on the performance of OD and OC segmentation using the SegFormer model. The strategies compared include the use of original images (without augmentation), flipping augmentation, a combination of flipping with gamma correction at specific values (δ = 0.8 and 1.1), and broader combinations (δ = 0.8, 0.9, 1.1, and 1.2). This approach aims to systematically analyze how combinations of geometric and intensity-based augmentation affect model performance.

The main contributions of this study are: (1) providing a comparative evaluation of various augmentation strategies in the SegFormer model in the OD and OC segmentation, (2) analyzing the effect of the combination of flipping and gamma correction on segmentation quality, and (3) evaluating performance using accuracy, sensitivity, DSC, IoU, and ROC metrics. The results of this study are expected to provide a clearer understanding of the role of augmentation strategies in improving the performance and generalization ability of models for glaucoma detection.

LITERATURE REVIEW

The comparative summary of previous studies is presented in Table 1, highlighting differences in model architecture, augmentation strategies, dataset size, and limitations.

Table 1. The Comparative studies on OD and OC Segmentation

Study	Model	Dataset	Dataset Size	Augmentation	DSC (%)	Limitation
(Thakur and Juneja 2019)	Hybrid Model	Drishti-GS	101	None	89	No augmentation
(Al-Bander et al. 2018)	FC-DenseNet	Drishti-GS	101	Flipping, Translation	88	Only basic geometric augmentation
(Zhou et al. 2023)	U-Net	REFUGE	1200	Vertical flip, random crop	85	Only basic geometric augmentation
(Jiang et al. 2020)	JointRCNN	ORIGA	650	Flip, Rotation	86	Only basic geometric augmentation

Table 1 compares previous studies on OD and OC segmentation, highlighting differences in model architecture, augmentation strategies, dataset sizes, and their limitations. Hybrid models without augmentation on Drishti-GS achieved a DSC of 89% (Thakur and Juneja 2019), while UNet with basic flipping and translation on REFUGE achieved 85% (Zhou et al. 2023). Similarly, FC-DenseNet and JointRCNN applied simple geometric augmentations and reported DSCs of 88% and 86%, respectively (Al-Bander et al. 2018; Jiang et al. 2020). Most approaches rely on basic geometric transformations, such as flipping, rotating, and cropping, which increase data variation but do not address image quality issues. In fundus imaging, challenges such as noise, blur, and uneven illumination are common and can significantly impact segmentation performance. These factors are rarely addressed through intensity-based enhancement.

Furthermore, several studies are conducted on relatively small datasets, such as Drishti-GS, which may limit the model's ability to generalize to more diverse data (Al-Bander et al. 2018; Chandan and Thakur 2018). Even with larger datasets, the augmentation strategies remain relatively simple, suggesting that the reported performance may not fully reflect robustness across varying imaging conditions. From an architectural perspective, most

*name of corresponding

This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15996
e-ISSN : 2541-2019
p-ISSN : 2541-044X

methods are based on CNN models, which are effective in capturing local features but may have limitations in modeling global contextual information, an important aspect in OD and OC segmentation (Xie et al. 2021).

These limitations indicate that augmentation strategies combining geometric and intensity-based transformations have not been sufficiently explored, particularly within a consistent model framework. Therefore, this study evaluates the effect of different augmentation strategies, including geometric and intensity-based approaches, on OD and OC segmentation using the SegFormer architecture.

METHOD

The workflow of this study consists of six main stages: preprocessing, data splitting, data augmentation, training, testing, and evaluation. The overall workflow is illustrated in Figure 1.

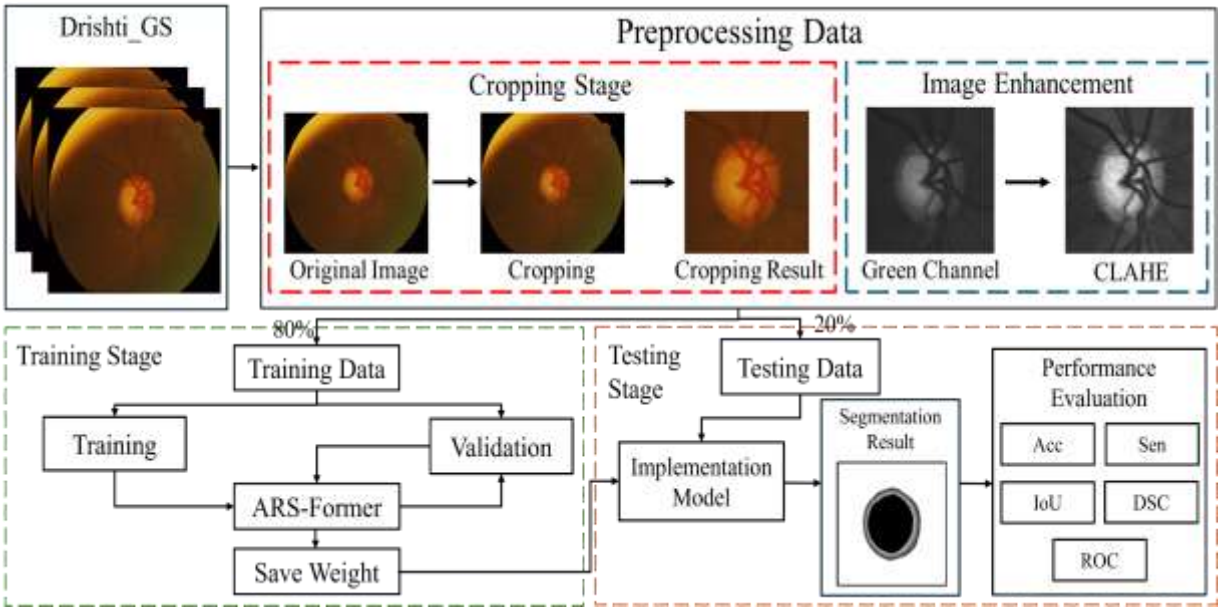


Figure 1. Workflow of the Proposed Method for OD and OC Semantic Segmentation

Data Description

This study utilizes fundus camera images obtained from the Drishti-GS dataset, developed by the International Institute of Information Technology, Hyderabad (IIIT-H), India (Sivaswamy et al. 2015). The dataset consists of 202 images, including 101 original fundus images and 101 corresponding ground-truth annotations, all with a resolution of 256 × 256 pixels in JPEG format. For retinal image segmentation, the ground truth masks define three classes: OD, OC, and background.

Preprocessing

Preprocessing is applied to enhance image quality before the segmentation process (Desiani et al. 2021). Preprocessing is applied to improve image quality prior to the segmentation process. Image cropping is performed to focus on the region of interest by removing irrelevant background areas (Desiani et al. 2024). Image enhancement techniques are employed to improve image quality, including green channel extraction and Contrast Limited Adaptive Histogram Equalization (CLAHE), which enhance the visibility of significant anatomical structures (Rudiansyah et al. 2024).

Data Splitting

Data splitting is the process of partitioning a dataset into training, validation, and testing subsets to ensure objective model evaluation (Yin et al. 2021). This study uses 101 original images and corresponding ground-truth annotations. The dataset is initially divided into training and testing sets with a ratio of 80:20. Subsequently, the training set is further split into training and validation subsets using the same ratio (80:20). This splitting process is applied consistently to both the original images and their corresponding ground truth to maintain data alignment. As a result, the final dataset consists of 65 training images, 16 validation images, and 20 test images.

Data Augmentation

Data augmentation is applied to increase data diversity and improve model generalization. In this study, augmentation is applied separately to the training, validation, and test sets to prevent data leakage, and the same

*name of corresponding

This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15996
e-ISSN : 2541-2019
p-ISSN : 2541-044X

transformations are applied to the corresponding ground truth to maintain spatial consistency. Four augmentation scenarios were evaluated: (1) no augmentation, (2) horizontal and vertical flipping, (3) flipping combined with gamma correction using δ = 0.8 and 0.9, and (4) flipping combined with gamma correction using δ = 0.8, 0.9, 1.1, and 1.2. These scenarios are designed to assess the impact of different augmentation levels on segmentation performance.

SegFormer Architecture

SegFormer is a transformer-based semantic segmentation architecture that employs a lightweight hierarchical Transformer as the encoder and a simple Multi-Layer Perceptron (MLP) as the decoder. Figure 2 illustrates the SegFormer-B0 architecture.

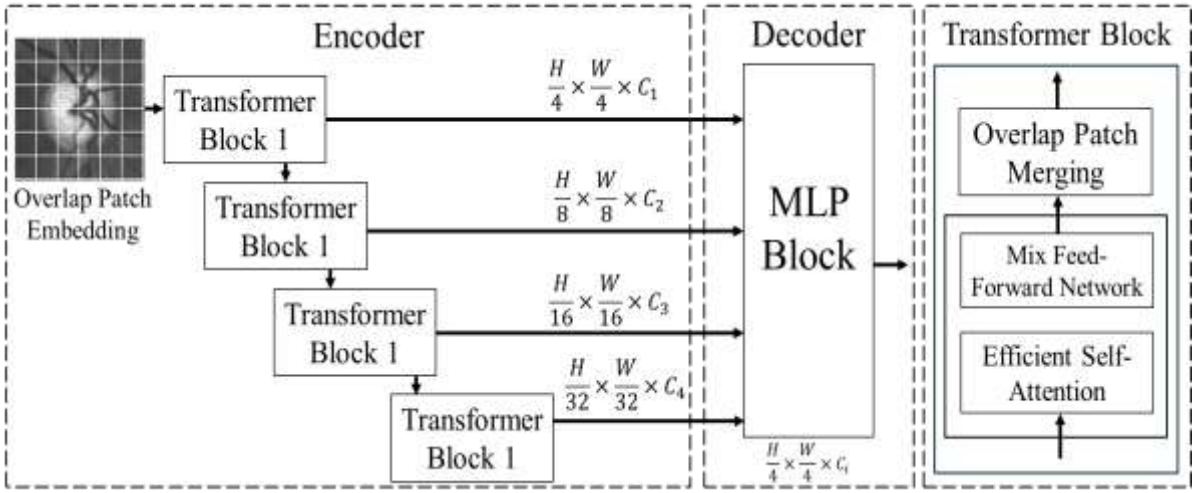


Figure 2. Illustration of SegFormer for Semantic Segmentation in OD and OC

Based on Figure 2, the SegFormer-B0 (with approximately 3.7 million parameters) consists of an encoder and a decoder. The encoder extracts hierarchical features using Overlap Patch Embedding, Transformer Blocks with Efficient Self-Attention and Mix-FFN, and Overlap Patch Merging to reduce resolution. The decoder fuses multi-level features via a Layer Structure, processes them with an MLP, and restores resolution via UpSampling. A softmax function produces segmentation into three classes: OD, OC, and background.

Configuration Training

The model was trained using the Adam optimizer with a learning rate of 0.0001 for 50 epochs, with a batch size of 32. The implementation was performed in PyTorch with CUDA 11.8, and the training process was accelerated on an NVIDIA RTX 3060 GPU with 12 GB of memory. In this study, the SegFormer-B0 model, with approximately 3.7 million parameters, was used.

Evaluation

The metrics used to measure model performance, including accuracy, sensitivity, DSC, and IoU, can be calculated using Equations (1) -(4) (Desiani et al. 2023; Yusro et al. 2019).

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (1)$$

$$Sen = \frac{TP}{TP + FN} \times 100\% \quad (2)$$

$$DSC = \frac{2\left(\frac{TP}{TP + FP}\right)\left(\frac{TP}{TP + FN}\right)}{\left(\left(\frac{TP}{TP + FP}\right)\right) + \left(\frac{TP}{TP + FN}\right)} \quad (3)$$

$$\text{IoU} = \frac{TP}{TP + FP + FN} \times 100\% \quad (4)$$

In OD and OC segmentation, True Positive (TP) refers to pixels in the OD or OC that the model correctly predicts as belonging to the OD or OC, respectively. False Negative (FN) represents OD or OC pixels that are misclassified as background. False Positive (FP) refers to background pixels incorrectly predicted as OD or OC,

*name of corresponding

This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15996
e-ISSN : 2541-2019
p-ISSN : 2541-044X

whereas True Negative (TN) indicates background pixels that are correctly identified as background (Rahmad, Suryanto, and Ramli 2020). The process of the proposed method is illustrated in Figure 2.

RESULT

Preprocessing

The preprocessing steps described in the methodology section were applied to all input images prior to segmentation. These steps include image cropping and image enhancement. Image cropping was performed to focus on the region of interest by removing irrelevant background, thereby making the optic disc (OD) and optic cup (OC) regions more prominent. Furthermore, image enhancement was performed using green-channel extraction and Contrast Limited Adaptive Histogram Equalization (CLAHE), which improved contrast and enhanced visibility of important anatomical structures.

Data Splitting

The data splitting procedure has been described in the methodology section. In this study, the dataset was divided into training and test sets at an 80:20 ratio, and the training set was further split 80:20 to obtain the validation set. As a result, the final dataset consisted of 65 training, 16 validation, and 20 test images.

Data Augmentation

The data augmentation procedure has been described in the methodology section. In this study, augmentation was applied after the data splitting process to increase data variability and improve model generalization. Augmentation was performed independently on the training, validation, and testing sets to prevent data leakage. Each transformation was consistently applied to both the input images and their corresponding ground truth masks to preserve spatial alignment. Four augmentation scenarios were evaluated: (1) no augmentation, (2) horizontal and vertical flipping, (3) flipping combined with gamma correction using δ = 0.8 and 0.9, and (4) flipping combined with gamma correction using δ = 0.8, 0.9, 1.1, and 1.2. The overall preprocessing workflow is illustrated in Figure 3.

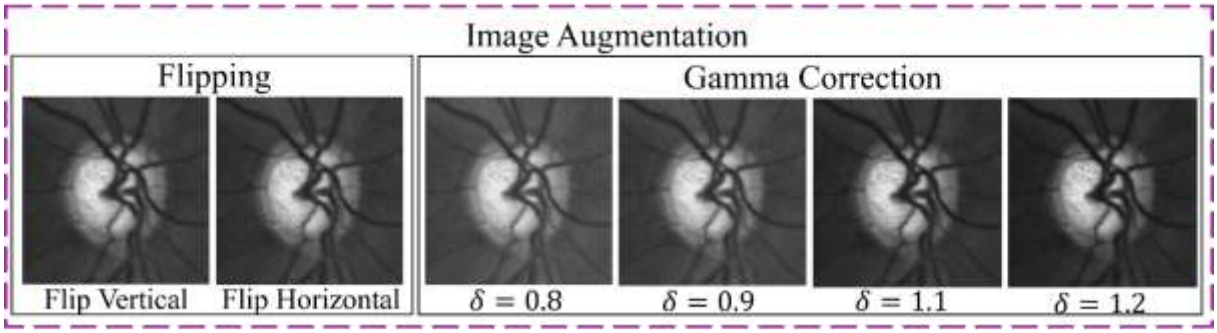


Figure 3. Examples of augmented images using flipping and gamma correction with different δ values.

The implementation of augmentation resulted in a significant increase in the number of samples across each dataset subset. Without augmentation, the dataset consisted of 65 training, 16 validation, and 20 testing images. When flipping was applied, the dataset size increased to 195 training, 48 validation, and 60 testing images. Furthermore, the combination of flipping and gamma correction using δ = 0.8 and 0.9 increased the dataset size to 390 training, 96 validation, and 120 testing images. In the extended configuration, where gamma correction with δ = 0.8, 0.9, 1.1, and 1.2 was applied, the dataset was further expanded to 780 training images, 192 validation images, and 240 testing images.

Training Stage

The training procedure has been described in the methodology section. In this study, model training and evaluation were conducted separately for each augmentation scenario to assess the impact of different data augmentation strategies on segmentation performance. For each scenario, the dataset was divided into training, validation, and test sets in the same proportions as previously described. The training process consisted of 50 epochs, during which the model's performance was monitored using accuracy and loss metrics. During training, a portion of the training data was used for validation to evaluate the model's generalization ability. The trends of accuracy and loss across epochs for each scenario are illustrated in Figure 4.

*name of corresponding

This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15996
e-ISSN : 2541-2019
p-ISSN : 2541-044X

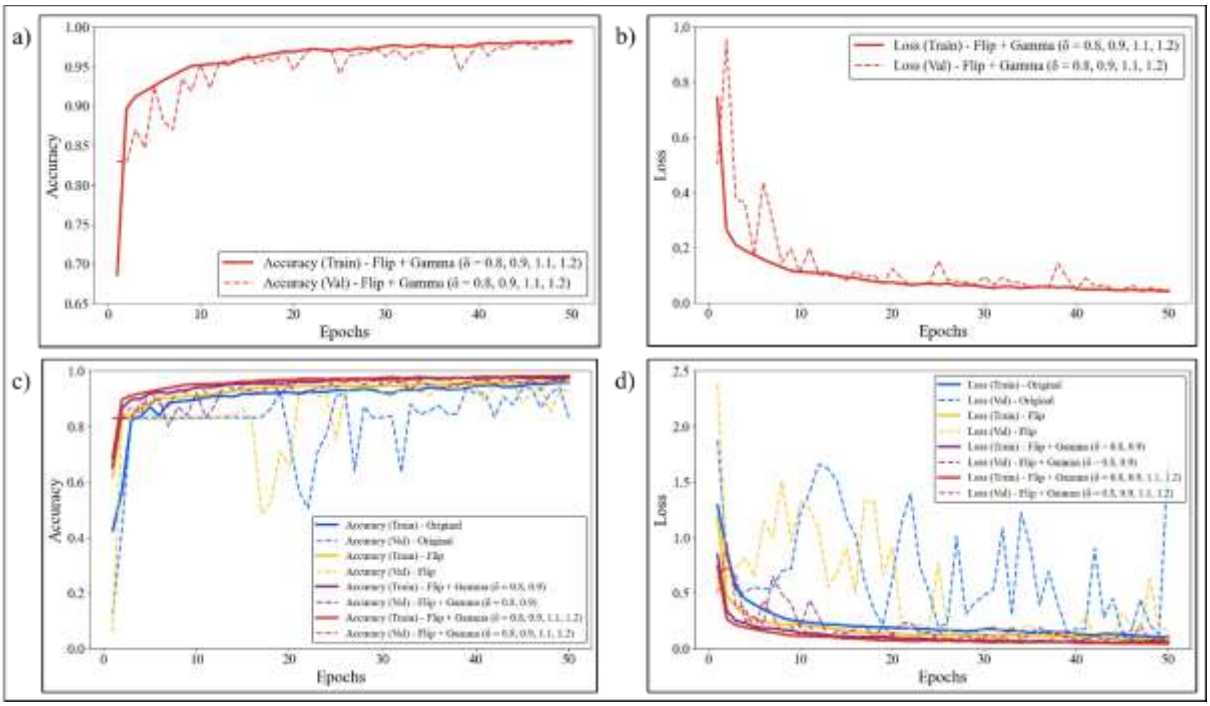


Figure 4. Training and validation performance: (a) accuracy and (b) loss for Flip-Gamma (δ = 0.8, 0.9, 1.1, 1.2), (c) accuracy comparison, and (d) loss comparison.

Figure 4 illustrates the training and validation performance across epochs for different augmentation scenarios. As shown in Figures 4(a) and 4(b), the model trained with Flip-Gamma (δ = 0.8, 0.9, 1.1, 1.2) demonstrates stable increases in accuracy and consistent decreases in loss, indicating effective convergence. Furthermore, the comparative results in Figures 4(c) and 4(d) show that augmentation strategies significantly influence model performance. The model without augmentation shows greater fluctuations in validation metrics, suggesting limited generalization. The flipping-only scenario improves stability. However, some inconsistencies remain. In contrast, the combination of flipping and gamma correction, particularly with extended gamma values, produces smoother accuracy curves and lower validation loss. This result indicates that incorporating both spatial and intensity variations enhances model robustness and improves generalization performance.

Testing Stage

In this study, model evaluation was conducted on the testing set for each augmentation scenario to assess the generalization performance of the trained models. The evaluation was performed using test data that was not used in the training or validation processes. It ensures that the results provide an unbiased assessment of the model's ability to segment unseen retinal images. The performance of each model was measured using quantitative metrics, including accuracy, sensitivity, DSC, IoU, precision, and recall, as shown in Figure 5.

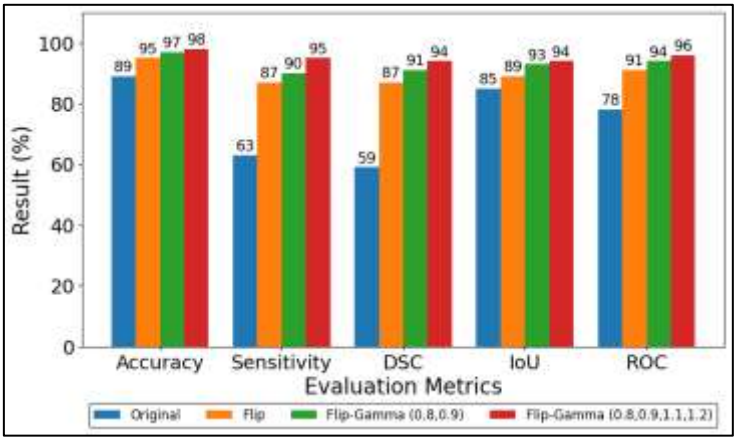


Figure 5. Comparison of segmentation performance across different augmentation methods.

*name of corresponding

This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15996
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Figure 5 shows a consistent improvement in SegFormer performance across all evaluation metrics as more advanced augmentation strategies are applied. The model trained on original images achieves reasonable accuracy (89%) but relatively low sensitivity and DSC, indicating limited ability to capture finer structures. The introduction of flipping significantly enhances performance by improving spatial generalization, while the addition of gamma correction further boosts results by incorporating intensity variations. The best performance is achieved with the Flip-Gamma (δ = 0.8, 0.9, 1.1, 1.2), yielding 98% accuracy and high sensitivity, DSC, IoU, and ROC values, demonstrating that the combination of spatial and multi-level intensity augmentation provides the most robust and generalizable representation for retinal image segmentation.

Evaluation

Based on the results presented in Figure 5, the Flip-Gamma configuration (δ = 0.8, 0.9, 1.1, 1.2) demonstrates the best overall performance among all augmentation strategies. Therefore, further analysis is conducted to examine segmentation performance at the class level and provide a more detailed understanding of the model's behavior across different anatomical regions. The class-wise segmentation results obtained using this augmentation strategy are shown in Figure 6.

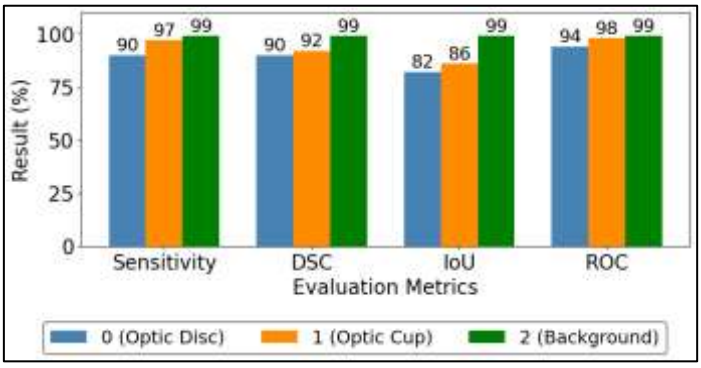
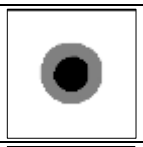
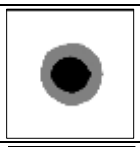
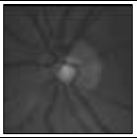
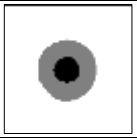
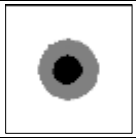
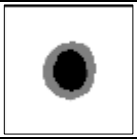
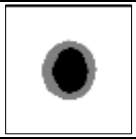


Figure 6. Class-wise segmentation performance using Flip-Gamma (δ = 0.8, 0.9, 1.1, 1.2)

Figure 6 presents the class-wise segmentation performance of the model using the best augmentation strategy, Flip-Gamma (δ = 0.8, 0.9, 1.1, 1.2). Overall, the background class achieves the highest performance across all evaluation metrics, with values close to 99% for sensitivity, DSC, IoU, and ROC. This result indicates that the model is highly effective at segmenting large, homogeneous regions. The OC class also shows strong performance, particularly in sensitivity and DSC, suggesting that the model can accurately capture the cup region despite its relatively low contrast. In contrast, the OD exhibits comparatively lower IoU, indicating that boundary delineation in this region remains more challenging. These results highlight that while the proposed augmentation strategy significantly improves segmentation performance across all classes, variations in anatomical structure size and complexity still influence the overall accuracy. This section presents an example of the model's segmentation results compared with the corresponding ground truth labels, as shown in Table 2.

Table 2. Visual Comparison of OD and OC Segmentation Results Between Ground Truth and Predictions

Original Image	GroundTruth	Prediction
		
		
		

*name of corresponding

This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15996
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Table 2 presents a qualitative comparison between the original fundus images, ground truth masks, and the corresponding model predictions for optic disc (OD) and optic cup (OC) segmentation. The results demonstrate that the proposed method accurately captures the shapes and boundaries of both OD and OC regions, with predictions closely matching the ground-truth annotations. Across different samples, the model shows consistent performance in segmenting both large and small structures, indicating its robustness. Minor differences in boundary smoothness are observed. However, the overall segmentation results remain visually consistent with the ground truth, supporting the quantitative improvements reported in the evaluation metrics.

The results highlight the effectiveness of combining Flip-Gamma (δ = 0.8, 0.9, 1.1, 1.2) in improving segmentation accuracy. To statistically validate this improvement, a one-sided Mann-Whitney U test was conducted on four evaluation metrics: accuracy, sensitivity, IoU, and DSC. In this study, the unit of observation corresponds to per-image performance scores obtained from the test datasets, where each augmentation method was evaluated on different test samples (N_1 = 20, N_2 = 60, N_3 = 120, and N_4 = 240 for each method). All test samples were drawn from the same data distribution to ensure comparability. The one-sided Mann-Whitney U test was selected as a nonparametric method that does not assume normality. The results yielded p-values < 0.05 across all metrics, indicating that the Flip-Gamma (δ = 0.8, 0.9, 1.1, 1.2) configuration outperforms other augmentation strategies. The statistical analysis using the one-sided Mann-Whitney U test is reported in Table 3.

Table 3. The one-sided Mann-Whitney U Test Results Between Flip-Gamma (δ = 0.8, 0.9, 1.1, 1.2) and Other Augmentation Method

Metric	Method	U-Statistic	P-Value	Significance
Acc	Original	2329	$5.229 \times e^{-14}$	Significant
Acc	Flip	2206	$2.24333 \times e^{-11}$	Significant
Acc	Flip-Gamma (δ = 0.8, 0.9)	1754	0.000259225	Significant
Sen	Original	2322	$7.5243 \times e^{-14}$	Significant
Sen	Flip	2199	$3.10097 \times e^{-11}$	Significant
Sen	Flip-Gamma (δ = 0.8, 0.9)	1743	0.000342885	Significant
IoU	Original	2310	$1.39666 \times e^{-13}$	Significant
IoU	Flip	2182	$6.74277 \times e^{-11}$	Significant
IoU	Flip-Gamma (δ = 0.8, 0.9)	1717	0.000650029	Significant
DSC	Original	2309	$1.47009 \times e^{-13}$	Significant
DSC	Flip	2181	$7.05505 \times e^{-11}$	Significant
DSC	Flip-Gamma	1716	0.000665818	Significant

The results presented in Table 4 indicate that the one-sided Mann-Whitney U test shows that the Flip-Gamma (δ = 0.8, 0.9, 1.1, 1.2) method provides significantly better performance than other augmentation methods (original, flip, and flipgamma) on all evaluation metrics: accuracy, sensitivity, IoU, and DSC, with a p-value <0.05 for each test. This finding indicates that using Flip-Gamma (δ = 0.8, 0.9, 1.1, 1.2) can significantly improve model performance.

DISCUSSIONS

The proposed method is compared with several existing studies using widely adopted evaluation metrics, including Accuracy (Acc), Sensitivity (Sen), Dice Similarity Coefficient (DSC), Intersection over Union (IoU), and ROC. These metrics are used to measure and analyze the segmentation performance of each approach. A detailed comparison is presented in Table 5.

Table 5. Comparison Several Studies with the Proposed Method

Methods	Acc (%)	Sen (%)	DSC (%)	IoU (%)	Roc (%)
FC DenseNet (Al-Bander et al. 2018)	96	58	53	55	-
U-Net++ (Tulsani et al. 2021)	85	85	-	-	-
GDCSegNet (Zhu et al. 2021)	-	-	93	89	-
PSBN (Shah et al. 2019)	-	-	91	85	-
SegFormer (Proposed Method)	**98**	**95**	**94**	**94**	**96**

Table 5 compares the proposed method with previous studies, including FC-DenseNet, U-Net++, GDCSegNet, and PSBN. The proposed method achieves higher performance across most reported metrics, particularly in

*name of corresponding

This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

e-ISSN : 2541-2019
p-ISSN : 2541-044X

accuracy, sensitivity, DSC, and IoU, compared to the results reported in these studies. This improvement is likely due to gamma correction, which enhances contrast and improves visibility of vascular structures, while the selected gamma values (δ = 0.8, 0.9, 1.1, 1.2) enable the model to learn across varying intensity conditions. However, these comparisons are based on reported results in the literature and should be interpreted as indicative rather than definitive. Additionally, the limited size of fundus datasets and potential variations across data sources may affect generalization. Future work will focus on validating the proposed method across larger, more diverse datasets.

CONCLUSION

Based on experimental results across four augmentation scenarios on the Drishti-GS dataset, the use of Flip-Gamma (δ = 0.8, 0.9, 1.1, 1.2) improved segmentation performance, both overall and class-wise (OD and OC), compared to other methods. Statistical analysis using a one-sided Mann-Whitney U test indicates differences in performance across the evaluated metrics. However, this study is limited by a relatively small dataset and the absence of external validation or cross-dataset evaluation. Therefore, these findings are specific to the evaluated scenarios and dataset, and further validation across larger, more diverse datasets is required to assess generalizability.

ACKNOWLEDGMENT

The authors thank the master's program in Computer Science, Faculty of Computer Science, Universitas Sriwijaya, for supporting our research.

REFERENCES

Acharya, Aditya, and A. Venkat Giri. 2020. "Contrast Improvement Using Local Gamma Correction." Pp. 110–14 in *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*. IEEE.

Al-Bander, Baidaa, Bryan Williams, Waleed Al-Nuaimy, Majid Al-Taee, Harry Pratt, and Yalin Zheng. 2018. "Dense Fully Convolutional Segmentation of the Optic Disc and Cup in Colour Fundus for Glaucoma Diagnosis." *Symmetry* 10(4):87. doi: 10.3390/sym10040087.

Bian, Xuesheng, Xiongbiao Luo, Cheng Wang, Weiquan Liu, and Xiuhong Lin. 2020. "Optic Disc and Optic Cup Segmentation Based on Anatomy Guided Cascade Network." *Computer Methods and Programs in Biomedicine* 197:105717. doi: 10.1016/j.cmpb.2020.105717.

Chandan, and Ritula Thakur. 2018. "An Intelligent Model for Indian Soil Classification Using Various Machine Learning Techniques." *International Journal of Computational Engineering Research* 08(9):33–41.

Deng, Yani, Xin Liu, Lianhe Shao, Kai Wang, Xihan Wang, and Quanli Gao. 2024. "Kidney Tumor Segmentation Based on DWR-SegFormer." *Electronics* 13(16):3226. doi: 10.3390/electronics13163226.

Desiani, Anita, Yuli Andriani, Indri Ramayanti, Sigit Priyanta, Bambang Suprihatin, Chairu Nisa Apriyani, and Muhammad Arhami. 2024. "RIB-NET as Modification of CNN Architecture for Semantic Segmentation Of Optic Disc and Optic Cup." *Biomedical Engineering: Applications, Basis and Communications* 36(06). doi: 10.4015/S1016237224500364.

Desiani, Anita, Member Erwin, Bambang Suprihatin, Sugandi Yahdin, Ajeng I. Putri, and Fathur R. Husein. 2021. "Bi-Path Architecture of CNN Segmentation and Classification Method for Cervical Cancer Disorders Based on Pap-Smear Images." *International Journal of Computer Science* 48.

Desiani, Anita, Sigit Priyanta, Indri Ramayanti, Bambang Suprihatin, Muhammad Gibran Al-Filambany, and Fitri Salamah. 2023. "Improved U-Net Performance with Augmentation for Retinal Optic Segmentation." Pp. 284–89 in *International Conference on Informatics, Multimedia, Cyber and Information Systems (ICIMCIS)*.

Gagan, J. H., Harshit S. Shirsat, Yogish S. Kamath, Neetha I. R. Kuzhuppilly, and J. R. Haris. Kumar. 2022. "Automated Optic Disc Segmentation Using Basis Splines-Based Active Contour." *IEEE Access* 10(August):88152–63. doi: 10.1109/ACCESS.2022.3199347.

Gupta, Neeraj, Hitendra Garg, and Rohit Agarwal. 2022. "A Robust Framework for Glaucoma Detection Using CLAHE and EfficientNet." *Visual Computer* 38(7):2315–28. doi: 10.1007/s00371-021-02114-5.

Jiang, Yuming, Lixin Duan, Jun Cheng, Zaiwang Gu, Hu Xia, Huazhu Fu, Changsheng Li, and Jiang Liu. 2020. "JointRCNN: A Region-Based Convolutional Neural Network for Optic Disc and Cup Segmentation." *IEEE Transactions on Biomedical Engineering* 67(2):335–43. doi: 10.1109/TBME.2019.2913211.

Kaushik, Meenakshi, Prabhakar Tiwari, Tanuj Dada, and Rima Dada. 2024. "Beyond the Optic Nerve: Genetics, Diagnosis, and Promising Therapies for Glaucoma." *Gene* 894(September 2023):147983. doi: 10.1016/j.gene.2023.147983.

Kumar, Abhinav, Anshul Sharma, Amit Kumar Singh, Sanjay Kumar Singh, and Sonal Saxena. 2023. "Data Augmentation for Medical Image Classification Based on Gaussian Laplacian Pyramid Blending with a Similarity Measure." *IEEE Journal of Biomedical and Health Informatics* PP. doi: 10.1109/JBHI.2023.3307216.

Lei, Haijun, Weixin Liu, Hai Xie, Benjian Zhao, Guanghui Yue, and Baiying Lei. 2022. "Unsupervised Domain

*name of corresponding

This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Adaptation Based Image Synthesis and Feature Alignment for Joint Optic Disc and Cup Segmentation." *IEEE Journal of Biomedical and Health Informatics* 26(1):90–102. doi: 10.1109/JBHI.2021.3085770.

Li, Tianping, Xiaolong Yang, Zhenyi Zhang, Zhaotong Cui, and Zhou Maoxia. 2025. "Mix-Layers Semantic Extraction and Multi-Scale Aggregation Transformer for Semantic Segmentation." *Complex & Intelligent Systems* 11(1):36. doi: 10.1007/s40747-024-01650-6.

Liu, Yong, Jin Wu, Yuanpei Zhu, and Xuezhi Zhou. 2024. "Combined Optic Disc and Optic Cup Segmentation Network Based on Adversarial Learning." *IEEE Access* 12(June):104898–908. doi: 10.1109/ACCESS.2024.3435552.

Maiyanti, Sri Indra, Anita Desiani, Syafrina Lamin, Puspitashati, Muhammad Arhami, Nuni Gofar, and Destika Cahyana. 2023. "Rotation-Gamma Correction Augmentation on Cnn-Dense Block for Soil Image Classification." *Applied Computer Science* 19(3):96–115.

Mazraedoost, Sargol. 2024. *Computers in Biology and Medicine*.

Nuli, Uday A., Shrinivas D. Desai, and Gururaj N. Bhadri. 2024. "Study on Limited Data Problem - Semantic Segmentation of Retinal Fundus Images." *Procedia Computer Science* 233:782–92. doi: 10.1016/j.procs.2024.03.267.

Rudiansyah, Anita Desiani, Dian Palupi Rini, Lucky Indra Kesuma, Fitri Salamah, and Silfani Cahaya Putri. 2024. "Median Filter and U-Net Architecture for Robust Segmentation Nucleus and Cytoplasm on Pap Smear." Pp. 779–84 in *2024 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*. IEEE.

Shah, Shivam, Nikhil Kasukurthi, and Harshit Pande. 2019. "Dynamic Region Proposal Networks For Semantic Segmentation In Automated Glaucoma Screening." Pp. 578–82 in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE.

Sivaswamy, Jayanthi, Subbaiah Krishnadas, Arunava Chakravarty, Gopal Joshi, and Ujjwal. 2015. "A Comprehensive Retinal Image Dataset for the Assessment of Glaucoma from the Optic Nerve Head Analysis." *JSM Biomed Imaging Data Pap* 2.

Sule, Olubunmi Omobola, Serestina Viriri, and Abdultaofeek Abayomi. 2020. "Effects of Image Enhancement Techniques on CNNs Based Algorithms for Segmentation of Blood Vessels: A Review." Pp. 1–6 in *2020 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD)*. IEEE.

Thakur, Niharika, and Mamta Juneja. 2019. "Optic Disc and Optic Cup Segmentation from Retinal Images Using Hybrid Approach." *Expert Systems with Applications* 127:308–22. doi: 10.1016/j.eswa.2019.03.009.

Tulsani, Akshat, Preetham Kumar, and Sumaiya Pathan. 2021. "Automated Segmentation of Optic Disc and Optic Cup for Glaucoma Assessment Using Improved UNET++ Architecture." *Biocybernetics and Biomedical Engineering* 41(2):819–32. doi: 10.1016/j.bbe.2021.05.011.

Wu, Hanqiong, Gangrong Qu, Zhifeng Xiao, and Fan Chunyu. 2024. "Enhancing Left Ventricular Segmentation in Echocardiography with a Modified Mixed Attention Mechanism in SegFormer Architecture." *Heliyon* 10(15):e34845. doi: 10.1016/j.heliyon.2024.e34845.

Xie, Enze, Wenhai Wang, Zhiding Yu, Anima Anandkumar, Jose M. Alvarez, and Ping Luo. 2021. "SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers."

Yin, Rui, Zihan Luo, Pei Zhuang, Zhuoyi Lin, and Chee Keong Kwoh. 2021. "VirPreNet: A Weighted Ensemble Convolutional Neural Network for the Virulence Prediction of Influenza A Virus Using All Eight Segments." *Bioinformatics* 37(6):737–43. doi: 10.1093/bioinformatics/btaa901.

Yu, Hanwen, Xin Ye, Wanjing Hong, Rui Shi, Yi Ding, and Chengcheng Liu. 2024. "A Cascading Learning Method with SegFormer for Radiographic Measurement of Periodontal Bone Loss." *BMC Oral Health* 24(1):325. doi: 10.1186/s12903-024-04079-y.

Yu, Ying, Chunping Wang, Qiang Fu, Renke Kou, Fuyu Huang, Boxiong Yang, Tingting Yang, and Mingliang Gao. 2023. "Techniques and Challenges of Image Segmentation: A Review." *Electronics* 12(5):1199. doi: 10.3390/electronics12051199.

Yusro, M., E. Suryana, K. Ramli, D. Sudiana, and K. M. Hou. 2019. "Testing The Performance of A Single Pole Detection Algorithm Using The Confusion Matrix Model." P. 77066 in *Journal of Physics: Conference Series*. Vol. 1402. IOP Publishing.

Zhou, Wei, Jianhang Ji, Yan Jiang, Jing Wang, Qi Qi, and Yugen Yi. 2023. "EARDS: EfficientNet and Attention-Based Residual Depth-Wise Separable Convolution for Joint OD and OC Segmentation." *Frontiers in Neuroscience* 17. doi: 10.3389/fnins.2023.1139181.

Zhu, Qianlong, Xinjian Chen, Qingquan Meng, Jiahuan Song, Gaohui Luo, Meng Wang, Fei Shi, Zhongyue Chen, Dehui Xiang, Lingjiao Pan, Zuoyong Li, and Weifang Zhu. 2021. "GDCSeg-Net: General Optic Disc and Cup Segmentation Network for Multi-Device Fundus Images." *Biomedical Optics Express* 12(10):6529. doi: 10.1364/BOE.434841.

*name of corresponding

This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.