

GPA-UNet: towards a novel UNet pipeline featuring Gated Post Adapter in optic disk and cup segmentation

DiaoZiJie*

Fudan University

22307130433@m.fudan.edu.cn

XiaShengJun*

Fudan University

22300240015@m.fudan.edu.cn

XieZhiKang*

Fudan University

22307110187@m.fudan.edu.cn

Abstract

Glaucoma, a leading cause of irreversible blindness, is primarily diagnosed through the cup-to-disk ratio (CDR), which relies on the accurate segmentation of optic cup (OC) and optic disc (OD) in retinal images. However, challenges such as limited dataset size, lack of diversity in imaging sources, and the absence of pre-trained U-Net models hinder effective OC/OD segmentation. To overcome these obstacles, we propose a novel U-Net-based pipeline comprising three stages: pre-training, fine-tuning, and post-processing. We enhance the GAMMA dataset by introducing the manually-inspected RefugeGAMMA dataset, which expands data diversity and quality. During fine-tuning, we employ data augmentation techniques and meticulously tune hyper-parameters to optimize model performance. The post-processing stage features a novel Gated PostAdaptor (GPA) module that dynamically refines segmentation results using a gating mechanism. Our pipeline achieved a score of 8.12756, ranking the third among all teams in the OC/OD segmentation challenge by Paddle. Our code will be available at <https://github.com/dzj441/FduCVPJ>.

1. Introduction

Glaucoma is one of the leading causes of irreversible blindness worldwide, mainly due to progressive damage to the optic nerve. The main clinical indicator for diagnosing glaucoma is the cup-to-disk ratio (CDR), which measures the relative size of the optic cup (OC) to the optic disc (OD) in fundus images. Accurate assessment of CDR allows ophthalmologists to identify patients at risk of glaucoma early, facilitating timely intervention and preventing significant vision loss. Therefore, a precise segmentation of OC and OD is essential for reliable CDR calculation and effective glaucoma management[9].

As shown in Fig. 1, the task of OC/OD segmenta-

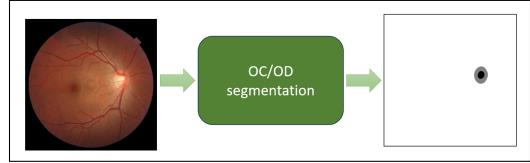


Figure 1. OC/OD segmentation. OC/OD segmentation involves turning fundus images into masks of OC (black) and OD (grey).

tion involves delineating these structures into corresponding masks within retinal images, which can be challenging due to variations in image quality, anatomical differences, and the presence of pathologies. Beyond these inherent difficulties, we face significant challenges related to data imbalance and scarcity. Fundus images are considered private patient data, which results in the GAMMA dataset [17] we utilized being relatively small and insufficient to meet the extensive training requirements of deep learning models. Additionally, the dataset comprises only images captured with Canon lenses, leading to a lack of diversity from other imaging sources and limiting the model’s generalization ability. Furthermore, due to the same privacy constraints and data limitations, there are no existing pre-trained UNet models for OC/OD segmentation. This absence necessitates training our UNet model from scratch, a process that is both time-consuming and computationally demanding.

To overcome these challenges, we propose a novel UNet-based pipeline for OC/OD segmentation. Our approach employs a three-stage process: pre-training, fine-tuning, and post-processing, as illustrated in Fig. 2. In the pre-training stage, we address the limited dataset by expanding it and implementing carefully designed data augmentation techniques. This enables us to pretrain a UNet backbone suitable for subsequent stages. Given that OC and OD occupy only a small proportion of fundus images, we initially apply a threshold-based method to perform rough segmentation, extracting the regions containing the OC and OD from the full fundus images. In the fine-tuning stage, we train the UNet model on these localized regions, meticulously

*equal contribution

⁰see division of labor in Appendix

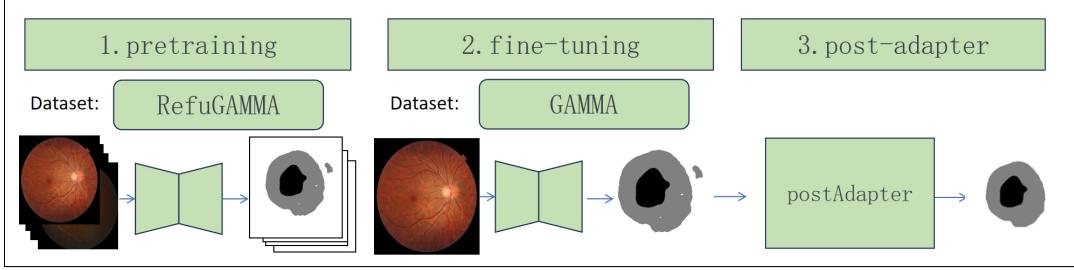


Figure 2. our pipeline.

tuning hyperparameters to enhance model performance. Finally, we develop a Gated PostAdaptor (GPA) module that utilizes a gating mechanism to adaptively refine the segmentation results produced by the UNet. Our pipeline achieved a score of 8.1276, securing third place among all teams in the OC/OD segmentation challenge organized by Paddle.

Our main contributions are summarized as follows:

1. We introduce a manually-inspected expanded dataset RefuGAMMA, enabling the training of a robust backbone for OC/OD segmentation.
2. We raise an effective and universal pipeline to solve the OC/OD segmentation challenge with high performance.
3. We propose a novel post-processing module, the Gated PostAdaptor (GPA), to further refine the accuracy and reliability of segmentation results, which has general applicability to OC/OD segmentation tasks.

2. Related Works

Segmenting optic cup (OC) and optic disc (OD) to determine the precise cup-to-disc ratio is a crucial step in glaucoma diagnosis. Existing methods can be broadly categorized into traditional and deep learning-based techniques. Traditional approaches include threshold-based and contour-based methods. Threshold-based methods segment images by applying specific threshold values [8], whereas contouring methods utilize an isotropic disc template as a prototype, adjusting it to the optic disc boundary by minimizing a local energy function[9]. However, these methods are often hindered by artifacts, inadequate lighting, and noise in fundus images.

In contrast, deep learning-based methods employ various backbone architectures to achieve OC/OD segmentation. For instance, [1] introduced a fully convolutional neural network (FCNN) with upsampling layers, a VGG-16 encoder, and a decoder to generate segmented images. Similarly, [21] proposed an ensemble learning architecture inspired by convolutional neural networks (CNNs) for OC and OD segmentation. These deep learning approaches have demonstrated superior performance by automatically

learning hierarchical features from datasets and thus effectively addressing many limitations of traditional methods.

2.1. Unet

The concept of skip-connections, also known as shortcuts or residuals, has become a fundamental technique in the design of neural networks. Introduced by ResNet [5], skip-connections redefine each layer as a residual learning function by incorporating an identity mapping that references the layer’s input. This architectural innovation simplifies the training of deeper networks and enhances accuracy by taking advantage of increased network depth. Drawing inspiration from ResNet, skip-connections have been extensively adopted in subsequent architectures [15]. For example, Inception-v4 [15] leverages skip-connections to significantly accelerate the training process of Inception networks. ResNeXt [19] integrates skip-connections with a meticulously designed homogeneous multi-branch architecture, while Res2Net [2] introduces hierarchical residual-like connections within a single residual block.

Building on the Fully Convolutional Network (FCN) framework [11], U-Net [13] incorporates similar long-path skip-connections between its contracting and expansive paths. This network design has achieved remarkable success in pixel-level prediction tasks such as semantic segmentation, instance segmentation, and image restoration, as skip-connections facilitate the transmission of detailed information that may be lost during the downsampling process. Due to its high performance, U-Net has been widely adopted as the backbone architecture for numerous medical image segmentation tasks [3, 4, 16, 18, 20], including optic cup and optic disc (OC/OD) segmentation.

3. Methodology

In this paper, our objective is to perform OC/OD segmentation in scenarios with limited dataset sizes. We have selected U-Net as our backbone architecture due to its proven effectiveness in medical image segmentation tasks. Our comprehensive pipeline is structured into three primary

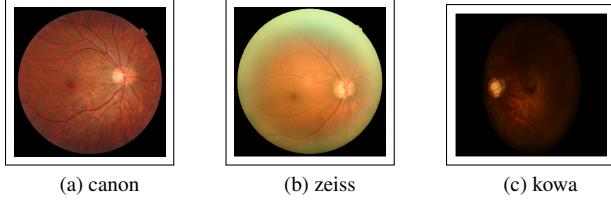


Figure 3. fundus images under different lens.

stages: pre-training, fine-tuning, and post-processing. This three-staged approach leverages pre-training on expanded datasets, optimizes model performance through fine-tuning, and enhances segmentation accuracy via GPA module.

3.1. Pre-training

Due to the sensitive nature of fundus images, there are no pre-trained U-Net models available for OC/OD segmentation. Our training dataset, GAMMA, consists of only 100 images, which is insufficient for effectively training a U-Net backbone on such a small scale. Additionally, the dataset exclusively contains images captured using Canon lenses, whereas the test set includes images from Canon, Zeiss, and KOWA lenses. This lack of diversity in imaging sources necessitates the expansion of our dataset to ensure the model’s generalization ability across different lens types.

To address this, we select the REFUGE dataset, a comprehensive collection designed for glaucoma classification tasks, which comprises 1,600 labeled fundus images captured with various lenses [12]. After manually inspecting the REFUGE dataset, we exclude 100 images of low quality, resulting in our enhanced dataset, RefugeGAMMA. This expanded dataset provides a more robust foundation for training, enabling the U-Net model to learn from a diverse set of high-quality images.

Furthermore, considering that OC and OD occupy only a small portion of fundus images, we adopt a two-stage segmentation approach. In the first stage, we employ an initial segmentation to isolate the regions likely containing the OC and OD from the entire fundus image, effectively reducing the presence of irrelevant areas and noise. This segmentation is based on brightness threshold, as the OC and OD regions typically exhibit the highest brightness levels within the images. In the second stage, we train the U-Net model on these localized regions to perform precise segmentation. We compared the performance of one-stage and two-stage segmentation methods, finding that the two-stage approach yields superior results by filtering out extraneous noise during the initial segmentation phase.

3.2. Fine-tuning

We conduct fine-tuning on the GAMMA dataset. However, the limited size of 100 images is insufficient for train-

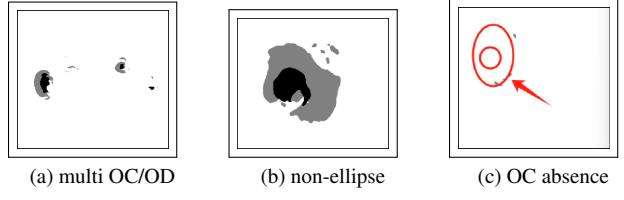


Figure 4. three main problems of segmentation results. (a) noise in fundus image may cause small mis-segmented OC/ODs. (b) the result doesn’t fit ellipse shape. (c) OC is absent but can be roughly predicted from the partial OD in result.

ing a U-Net. To address this, we meticulously design data augmentation strategies to increase the dataset size without compromising the image structure. Specifically, we apply random rotations and flips to preserve the aspect ratio, and introduce random Gaussian blurring to simulate low-resolution images of poorer quality.

Moreover, we carefully fine-tune several critical hyperparameters, primarily focusing on `reshape_size` and `batch_size`. Detailed experiments and results are presented in Section 4.3.

Finally, we develop a mixed loss function that dynamically adjusts during fine-tuning. Our approach incorporates Cross-Entropy (CE) loss, Focal loss [10], and Dice loss [14]. The CE loss maintains a fixed weight of 0.1. In the initial stages of training, Focal loss is assigned a higher weight to balance hard and easy samples. As training progresses, the weight of Dice loss increases to optimize the segmentation of edge regions. The dynamic weighting is adjusted following a cosine schedule.

Our final loss function is defined as follows:

$$L = \lambda_1 \cdot L_{CE} + \lambda_2 \cdot L_{focal} + \lambda_3 \cdot L_{dice}$$

$$\text{s.t. } \begin{cases} \lambda_1 + \lambda_2 + \lambda_3 = 1, \\ \lambda_1 = 0.1, \\ \lambda_2 \text{ and } \lambda_3 \text{ follow a cosine schedule.} \end{cases} \quad (1)$$

3.3. Gated Post Adapter (GPA) module

Although the fine-tuned model accurately segments most OCs and ODs, several issues persist, as is shown in Fig. 4. Firstly, noise in fundus image can cause multiple OCs and ODs within a single output. Secondly, lacking the inductive bias that OC and OD are of elliptical shapes makes segmentation results deviate from the expected elliptical morphology. Lastly, for lower-resolution images, the model may only partially segment out OC and OD. In such case, the partial segments can be utilized to predict the coarse location of OC and OD. We propose Gated Post Adapter(GPA) module to mitigate these problems, the structure of which is illustrated in Fig. 5.

Inspired by the concept of Non-Maximum Suppression (NMS) [7], the GPA module incorporates a Non-Maximum

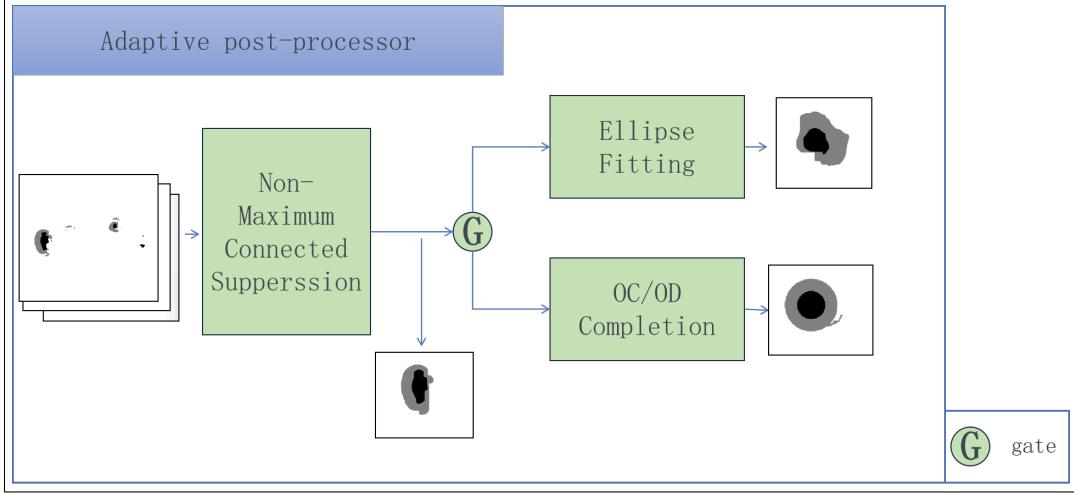


Figure 5. the structure of GPA module.

Connected Suppression (NMCS) mechanism. This approach involves detecting connected components of OC and OD and retaining only the largest connected component as the final segmentation result, thereby filtering out noisy segmentation outputs. Additionally, we observe that the majority of segmentation results only need ellipse fitting to better approximate the elliptical morphology of OC and OD. However, for a small subset of low-resolution images where the segmentation is partial, OC/OD completion becomes necessary. Drawing inspiration from gating mechanisms in Long Short-Term Memory (LSTM) networks [6], we introduced a gating mechanism defined as follows:

$$G = \begin{cases} 1, & \text{if } A(S) < 1500 \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where G is the gating variable, $A(S)$ represents the area of the segmentation result after NMCS. A value of $G = 1$ indicates that the segmentation occupies less than 1500 pixels, suggesting a likely absence of complete OC/OD regions, thereby triggering OC/OD completion. Conversely, a value of $G = 0$ implies sufficient segmentation coverage, warranting only ellipse fitting. This mechanism is illustrated in the following equation:

$$\text{Result} = G \cdot EF(X) + (1 - G) \cdot Complete(X) \quad (3)$$

where G is the gating variable and X is NMCS result. EF and $Complete$ denotes ellipse fitting and OC/OD completion respectively.

4. Experiments

4.1. basic settings

Since we joint the OC/OD segmentation challenge by Paddle, our deep learning framework is based on paddle. We carry out our experiments on GAMMA dataset [17] and train our model with 1 NVIDIA Tesla V100 32G GPU.

We employ Dice coefficient as the primary evaluation metric for assessing the segmentation performance of OC and OD. The Dice coefficient is defined as:

$$\text{Dice} = \frac{2|X \cap Y|}{|X| + |Y|} \quad (4)$$

where X represents the set of ground truth segmented target pixels, and Y represents the set of predicted segmented target pixels.

In addition to Dice coefficient, we also measure the Mean Absolute Error (MAE) of the vertical cup-to-disc ratio (vCDR) between predicted results and ground truth.

The final evaluation score integrates the Dice coefficients of the OC/OD segmentation, as well as the MAE of vCDR:

$$\text{Score} = 2.5 \cdot \text{Dice}_{\text{cup}} + 3.5 \cdot \text{Dice}_{\text{disc}} + 0.4 \cdot \frac{1}{\text{vCDR} + 0.1} \quad (5)$$

4.2. Loss Function

In OC/OD segmentation, addressing class imbalance is essential. After extensive experiments, we determine the following loss function combinations:

Table 1. Ablation Study Results

Settings	$Dice_{disc}$ (\uparrow)	$Dice_{cup}$ (\uparrow)	$vCDR$ (\downarrow)	Score (\uparrow)
Baseline	0.81429	0.57403	0.15791	5.83603
Baseline + Pretraining	0.86521	0.77031	0.06991	7.30814
Baseline + Pretraining + Tuning	0.92669	0.84192	0.05492	7.93019
Baseline + Pretraining + Tuning + Post-Processor (full image)	0.92452	0.8445	0.0475	8.05888
Baseline + Pretraining + Tuning + Post-Processor (localized image)	0.93250	0.85249	0.04638	8.12756

Baseline Comparison

While the paddle official baseline use Dice loss alone, we observe that a ratio of 0.9 Dice loss and 0.1 Cross-Entropy (CE) loss produce better segmentation results compared to using Dice loss alone (Omitted here for brevity. Detailed Comparison can be found in Supplementary Material: Table 2).

Focal Loss and Dynamic Weighting

To further enhance performance:

- **Early Stages:** Adding Focal Loss with a higher weight improved segmentation on challenging samples, but slightly reduced smoothness.
- **Later Stages:** Increasing Dice Loss weight refined edge smoothness.
- **Dynamic Adjustment:** A cosine decay schedule with smoothing dynamically balanced Dice and Focal Loss throughout training, improving stability and edge smoothness.

Detailed results for static and dynamic weighting schemes are in Supplementary Material Table 3.

Key Observations

1. Dynamic weighting schedules yielded the best edge refinement in later training stages.
2. A fixed CE loss ratio (0.1) provided consistent stability throughout training.
3. Adaptive ratios with a wider range for Dice and Focal Loss improved overall segmentation performance.

4.3. Hyper-parameter Tuning

We evaluate two key hyper-parameters: reshape size and batch size, analyzing their impact on $Dice_{disc}$, $Dice_{cup}$ and $vCDR$.

4.3.1 Reshape Size

Resizing input images affects performance and computational cost. As summarized in Table 4 (in supplementary material), a reshape size of 800x1200 provide the best overall score.

4.3.2 Batch Size

Batch size significantly influenced performance. As shown in Table 5 (in supplementary material), a batch size of 3 achieve the highest $Dice_{disc}$ and $Dice_{cup}$ values, while a batch size of 2 offers the best overall score with competitive metrics. What's more, lower batch size highly improves the training efficiency.

4.4. Ablation Study

We perform an ablation study to assess the impact of key components, i.e. pre-training, fine-tuning, post-processing, and the rough initial segmentation. As shown in Table 1, each component contributed incrementally to the overall improvement.

Pre-training provides a robust initialization, leading to notable gains in $Dice_{disc}$ and $Dice_{cup}$. Fine-tuning further optimizes these features, improving dice coefficients and reducing $vCDR$ error. Post-processing refines segmentation results by smoothing predictions, yielding a higher overall score and further reducing $vCDR$. Finally, the full pipeline is retrained on the coarsely segmented localized fundus images, achieving the best score, since the initial rough segmentation filters out some noise in dataset.

The full model achieves a score of 8.12756, demonstrating the critical importance of combining these strategies for optimal performance.

5. Conclusion

In this study, we address the critical task of OC and OD segmentation in fundus images. Faced with challenges such as limited dataset size, lack of diversity in imaging sources, and the absence of pre-trained U-Net models, we develop a novel U-Net-based pipeline comprising three stages: pre-training, fine-tuning, and post-processing. The pre-training stage mitigates the constraints of a small and homogeneous dataset. The fine-tuning stage improves the backbone model's performance. The post-processing stage dynamically refines segmentation outputs.

Future work may explore the integration of additional fundus features and further optimization of the GPA module to enhance segmentation performance. Additionally, expanding the dataset to include more diverse imaging conditions and patient demographics could further improve the model's generalization ability and clinical applicability.

References

- [1] Venkata Gopal Edupuganti, Akshay Chawla, and Amit Kale. Automatic optic disk and cup segmentation of fundus images using deep learning. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 2227–2231, 2018. [2](#)
- [2] Shang-Hua Gao, Ming-Ming Cheng, Kai Zhao, Xin-Yu Zhang, Ming-Hsuan Yang, and Philip Torr. Res2net: A new multi-scale backbone architecture. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(2):652–662, 2021. [2](#)
- [3] Xiaolin Gou, Chuanlin Liao, Jizhe Zhou, Fengshuo Ye, and Yi Lin. Fif-unet: An efficient unet using feature interaction and fusion for medical image segmentation, 2024. [2](#)
- [4] Nima Hassanpour and Abouzar Ghavami. Deep learning-based bio-medical image segmentation using unet architecture and transfer learning, 2023. [2](#)
- [5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015. [2](#)
- [6] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997. [4](#)
- [7] Jan Hosang, Rodrigo Benenson, and Bernt Schiele. Learning non-maximum suppression, 2017. [3](#)
- [8] Ahsan Khawaja, Tariq M. Khan, Mohammad A. U. Khan, and Syed Junaid Nawaz. A multi-scale directional line detector for retinal vessel segmentation. *Sensors*, 19(22), 2019. [2](#)
- [9] J. R. Harish Kumar, Aditya Kumar Pediredla, and Chandra Sekhar Seelamantula. Active discs for automated optic disc segmentation. In *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pages 225–229, 2015. [1, 2](#)
- [10] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection, 2018. [3](#)
- [11] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3431–3440, 2015. [2](#)
- [12] José Ignacio Orlando, Huazhu Fu, João Barbosa Breda, Karel van Keer, Deepti R. Bathula, Andrés Diaz-Pinto, Ruogu Fang, Pheng-Ann Heng, Jeyoung Kim, JoonHo Lee, Joonseok Lee, Xiaoxiao Li, Peng Liu, Shuai Lu, Balamurali Murugesan, Valery Naranjo, Sai Samarth R. Phaye, Sharath M. Shankaranarayana, Apoorva Sikka, Jaemin Son, Anton van den Hengel, Shujun Wang, Junyan Wu, Zifeng Wu, Guanghui Xu, Yongli Xu, Pengshuai Yin, Fei Li, Xiulan Zhang, Yanwu Xu, and Hrvoje Bogunović. Refuge challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs. *Medical Image Analysis*, 59:101570, 2020. [3](#)
- [13] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015. [2](#)
- [14] Carole H. Sudre, Wenqi Li, Tom Vercauteren, Sébastien Ourselin, and M. Jorge Cardoso. *Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations*, page 240–248. Springer International Publishing, 2017. [3](#)
- [15] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alex Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning, 2016. [2](#)
- [16] Qiyuan Tian, Zhuoyue Wang, and Xiaoling Cui. Improved unet brain tumor image segmentation based on gsconv module and eca attention mechanism, 2024. [2](#)
- [17] Junde Wu, Huihui Fang, Fei Li, Huazhu Fu, Fengbin Lin, Jiongcheng Li, Lexing Huang, Qinji Yu, Sifan Song, Xinxing Xu, Yanyu Xu, Wensai Wang, Lingxiao Wang, Shuai Lu, Huiqi Li, Shihua Huang, Zhichao Lu, Chubin Ou, Xifei Wei, Bingyuan Liu, Riadh Kobbi, Xiaoying Tang, Li Lin, Qiang Zhou, Qiang Hu, Hrvoje Bogunovic, José Ignacio Orlando, Xiulan Zhang, and Yanwu Xu. Gamma challenge:glaucoma grading from multi-modality images, 2022. [1, 4](#)
- [18] Ruiqiang Xiao and Zhuoyue Wan. Gaei-unet: Global attention and elastic interaction u-net for vessel image segmentation, 2023. [2](#)
- [19] Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. Aggregated residual transformations for deep neural networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5987–5995, 2017. [2](#)
- [20] Wenhui Zhu, Xiwen Chen, Peijie Qiu, Mohammad Farazi, Aristeidis Sotiras, Abolfazl Razi, and Yalin Wang. Selfreg-unet: Self-regularized unet for medical image segmentation, 2024. [2](#)
- [21] Julian Zilly, Joachim M. Buhmann, and Dwarikanath Mahapatra. Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation. *Computerized Medical Imaging and Graphics*, 55: 28–41, 2017. Special Issue on Ophthalmic Medical Image Analysis. [2](#)

GPA-UNet: towards a novel UNet pipeline featuring Gated Post Adapter in optic disk and cup segmentation

Supplementary Material

6. Detailed Experiment Results

Table 2 illustrates the performance of the model under different combinations of DiceLoss and CELoss ratios.

Table 2. Performance Comparison of Different Loss Ratios

DiceLossRatio	CELossRatio	Dice _{disc} (↑)	Dice _{cup} (↑)	vCDR (↓)	Score (↑)
1.0	0.0	0.91941	0.83565	0.0594	7.81653
0.9	0.1	0.92195	0.83382	0.05869	7.83198
0.1	0.9	0.92454	0.8353	0.06096	7.80919
0.0	1.0	0.91594	0.82784	0.05836	7.8013

Table 3 shows the results of the model with and without dynamic adjustments to the loss weights.

Table 3. Results with Dynamic Loss Weight Adjustments

Dynamic?	DiceLoss	FocalLoss	CELoss	Dice _{disc} (↑)	Dice _{cup} (↑)	vCDR (↓)	Score (↑)
No	0.5	0.4	0.1	0.91421	0.83732	0.06313	7.74511
Yes	0.5-0.79	0.4-0.11	0.1	0.92228	0.83785	0.06225	7.79232
No	0.7	0.2	0.1	0.91086	0.8335	0.05815	7.80108
Yes	0.7-0.85	0.2-0.05	0.1	0.92137	0.84195	0.05947	7.83789
Yes	0.35-0.7	0.6-0.25	0.05	0.92022	0.83785	0.05778	7.85062
Yes	0.19-0.85	0.76-0.1	0.05	0.92502	0.8411	0.05931	7.85116
Yes	0.1-0.95	0.95-0.1	0.05	0.92781	0.84299	0.05939	7.86431

Table 4. Performance Comparison with Different Reshape Sizes

ReshapeSize	Dice _{disc} (↑)	Dice _{cup} (↑)	vCDR (↓)	Score (↑)
1400x2100	0.91177	0.81391	0.06357	7.67136
1200x1800	0.92662	0.84051	0.05983	7.84704
1000x1500	0.92195	0.83382	0.05869	7.83198
800x1200	0.92779	0.83837	0.05889	7.86070
700x1050	0.92424	0.84286	0.05893	7.85878

Table 5. Performance Comparison with Different Batch Sizes

BatchSize	Dice _{disc} (↑)	Dice _{cup} (↑)	vCDR (↓)	Score(↑)
1	0.91610	0.83776	0.05704	7.84777
2	0.92669	0.84192	0.05492	7.93019
3	0.92781	0.84299	0.05939	7.86431
4	0.92771	0.83865	0.05781	7.87831

7. Division of Labor

Most work is done by all three of us. The following is only a coarse division of labor:

- DiaoZiJie
 - 1. analyze related works
 - 2. put forward dataset RefuGAMMA
 - 3. pre-train the backbone
 - 4. help tune the hyper-parameters
 - 5. presentation
 - 6. write the final paper
- XiaShengJun
 - 1. put forward the GPA module
 - 2. implement the GPA module
 - 3. help tune the hyper-parameters
 - 4. help presentation
- XieZhiKang
 - 1. fine tuning reshape size and batch size
 - 2. put forward the dynamic loss
 - 3. implement the dynamic loss
 - 4. help presentation
 - 5. help write the final paper