Retinal Artery–Vein Segmentation via Attention-Guided W-Net and GAN-Based Boundary Refinement

Jing-Ming Guo[1], De-Yu Guu[2], Yih-Ping Luh[3], and Yi-Chong Zeng[4]
[1,2,4] National Taiwan University of Science and Technology, Taipei, Taiwan
E-mail: [1] jmguo@seed.net.tw, [2] M11207308@mail.ntust.edu.tw, [4] yichongzeng@gmail.com
[3] META AI INC., Hsinchu, Taiwan
E-mail: [3] luhyih19@gmail.com

***Abstract*— We present an attention-guided W-Net with GAN-based boundary refinement and an adaptive binary-to-multiclass fusion for retinal artery–vein segmentation. To ensure fairness, we adopt a unified training protocol that merges the training splits of DRIVE-AV and HRF-AV and applies identical augmentations across all variants. On DRIVE-AV, our model achieves a Dice score of 81.98 and an AUROC of 99.18; on HRF-AV, the AUROC is 98.94. Visual results show sharper vessel boundaries and fewer artery–vein confusions at crossings and in low contrast regions. The framework is lightweight and readily extensible to clinical screening workflows.**

I. Introduction

Retinal artery–vein segmentation underpins quantitative ophthalmic analysis but remains difficult on thin, low-contrast vessels and at crossings. Despite recent progress, CNN-based models are constrained by scarce and imbalanced pixel-wise annotations and by domain shifts across datasets that hinder generalization [1]. We propose a compact multi-stage architecture that couples attention-guided feature selection, adversarial boundary refinement, and an adaptive fusion from binary (vessel) to artery–vein labels, improving boundary fidelity and class disambiguation while keeping the model footprint practical.

II. The Proposed Method

Fig.1 illustrates the model architecture of the proposed method. This model adopts a multi-stage design that integrates attention-guided W-Net [2], GAN Boundary Refinement, and task-specific fusion to improve both vessel delineation and artery-vein discrimination. Attention-Guided W-Net is a two-stage encoder–decoder with attention gates in skip connections to suppress background and preserve thin structures. GAN Boundary Refinement is an adversarial refinement head aligned to ground truth topology to sharpen edges and mitigate over-smooth predictions. We implement an adaptive binary-to-multiclass fusion. A binary vessel

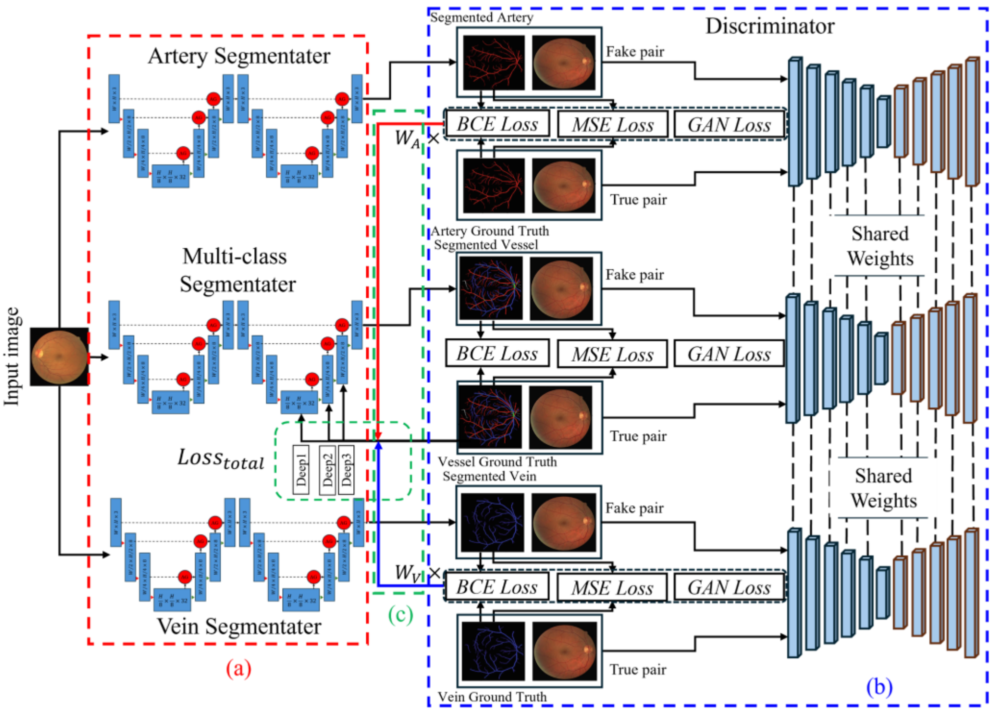

Fig. 1. Model architecture overview. The proposed multi-stage pipeline: (a) attention-guided W-Net backbone (red-dotted rectangle); (b) GAN-based boundary refinement (blue-dotted rectangle); (c) adaptive binary-to-multiclass fusion that transfers vessel cues to artery–vein decoding (green-dotted rectangle).

979-8-3315-7206-8/25/$31.00 ©2025 IEEE

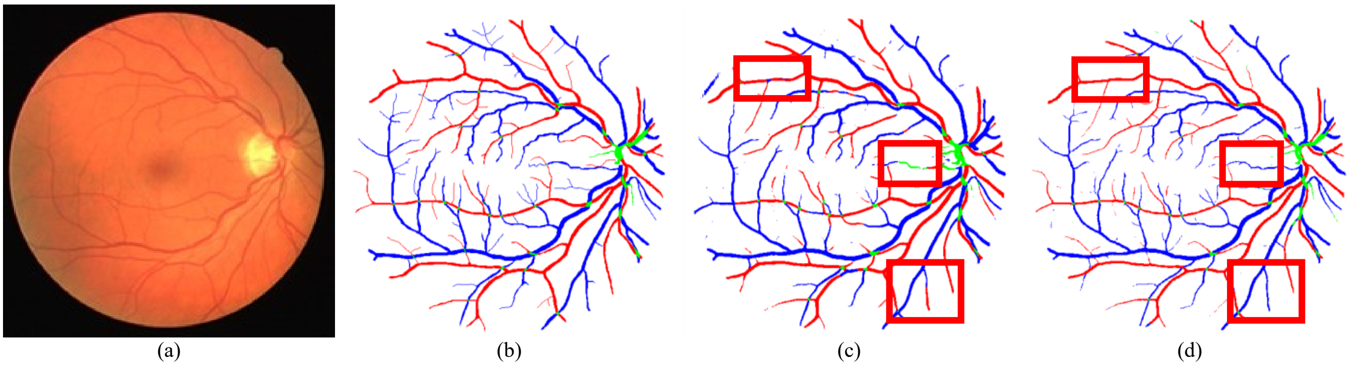

Fig. 2. Artery–vein Segmentation: (a) Original image, (b) ground truth, the segmentation results using (c) BMFN [3], and (d) the proposed method.

TABLE I. PERFORMANCE OF THE BMFN AND THE PROPOSED METHOD WITH VARIOUS COMPONENTS.

Components	DRIVE-AV			HRF-AV		
	Sensitivity	AUROC	DICE(F1)	Sensitivity	AUROC	DICE(F1)
BMFN [5]	75.56	93.76	79.75	69.11	94.44	75.86
BMFN+Wnet	75.24	95.79	80.29	70.13	95.49	76.81
BMFN+Attn Wnet	78.01	96.63	80.7	71.97	96.49	77.13
BMFN + Attn Went + Adaptive Loss-weighting	**79.65**	**99.18**	**81.98**	**72.76**	**98.94**	**77.22**

branch guides artery-vein decoding via dynamic weights and deep supervision, reducing artery-vein confusion at crossings. Adaptive loss weighting mechanism is also implemented, from ground truth masks we compute per-mini-batch artery/vein foreground ratios and scale the artery–vein sub-network losses inversely to these ratios, automatically compensating class imbalance without extra tuning. The model architecture employs standard convolutional backbones, Dice-style losses with class-aware weighting, and deep supervision at intermediate scales.

III. The Experiment Results

In this study, we have experimented two retinal artery–vein segmentation datasets, DRIVE-AV [3] and HRF-AV [4]. Under the same training pool and augmentations, our full model delivers consistent gains over the BMFN backbone. TABLE I lists the performance of the BMFN and the proposed method with various components. On DRIVE-AV, our method improves Sensitivity by 4.09%, AUROC by 5.42%, and Dice by 2.23%. On HRF-AV, the proposed method improves Sensitivity by 3.65%, AUROC by 4.50%, and Dice by 1.36%. These quantitative gains align with visual improvements—fewer artery–vein segmentation errors at crossings and crisper vessel boundaries—while attention gating and adaptive loss weighting contribute the bulk of the uplift beyond the backbone.

Fig.2 shows the retinal artery–vein segmentation results using BMFN [5] and the proposed method with various components. The red boxes highlight vessel crossings where baseline shows artery–vein mix-ups, while ours preserves topology and class labels; the red boxes show low contrast thin vessels recovered by our method.

IV. Conclusion

Under the same training pool and augmentations, our model outperforms the BMFN backbone. The proposed method improved the AUROC values by 5.42% and 4.50% on DRIVE-AV and HRF-AV datasets, respectively. Furthermore, it enhanced the Sensitivity by 4.09% and 3.65% on DRIVE-AV and HRF-AV datasets, respectively. These gains align with sharper boundaries and better thin-vessel recall from attention and adversarial refinement, while adaptive loss weighting re-balances artery–vein branches and reduces crossing-induced errors without extra tuning.

References

[1] Z. Liu, "Retinal vessel segmentation based on fully convolutional networks," arXiv preprint, arXiv:1911.09915, 2019.

[2] A. Galdran, A. Anjos, J Dolz, H Chakor, H. Lombaert, and I. B. Ayed, "State-of-the-art retinal vessel segmentation with minimalistic models," *Scientific Reports*, vol. 12, no. 1, p. 6174, Apr. 2022, doi: 10.1038/s41598-022-09675-y

[3] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Transactions on Medical Imaging*, vol. 23, no. 4, pp. 501–509, Apr. 2004, doi: 10.1109/TMI.2004.825627.

[4] R. Hemelings, B. Elen, I. Stalmans, K. Van Keer, P. De Boever, and M. B. Blaschko, "Artery-vein segmentation in fundus images using a fully convolutional network," *Computerized Medical Imaging and Graphics*, vol. 76, no. 101636, Sept. 2019.

[5] Y. Zhou, M. Xu, Y. Hu, H. Lin, J. Jacob, P. A. Keane, et al., "Learning to address intra-segment misclassification in retinal imaging," *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, 2021, pp. 482–492, doi: 10.1007/978-3-030-87193-2_46.