A Phased Self-Teaching Framework for Hyperreflective Foci Segmentation in OCT

Zhi Wang[1,2], Donghan Wu[1,2], Huaying Hao[1], Le Cao[3], Bo Wu[3*], and Yitian Zhao[1,2*]

[1] Ningbo Institute of Materials Technology and Engineering, Chinese Academy of Sciences, Ningbo, China
[2] University of Chinese Academy of Sciences, Beijing, China
[3] Department of Neurology, West China Hospital, Sichuan University, Chengdu, Sichuan, China
yitian.zhao@nimte.ac.cn, dr.bowu@hotmail.com
*Corresponding authors.

Abstract. Idiopathic intracranial hypertension (IIH) requires longitudinal monitoring of disease activity and its retinal effects, while direct pressure assessment by lumbar puncture is invasive and impractical to repeat frequently. Optical coherence tomography (OCT) offers a non-invasive retinal adjunct, and hyperreflective foci (HRF) on OCT have emerged as a biomarker of clinical interest, making quantitative HRF readouts a candidate OCT-derived measure for investigation alongside pressure-based assessment. Accurate automated HRF segmentation, however, poses an asymmetric sensitivity–specificity dilemma beyond ordinary class imbalance: lesions occupy under 1% of a B-scan and are easily missed, while shifting toward sensitivity also admits speckle and vessel-like confounders under ambiguous annotations. We propose DiSS, a phased self-teaching scheme that, rather than competing the two goals within one objective, separates them in the training schedule on an unmodified backbone. A multi-timescale exponential moving average (EMA) teacher bank supplies progressively consolidated evidence; a first "cover" phase consolidates weak-lesion evidence via disjunctive (noisy-OR) fusion, and a second "suppress" phase applies soft-weighted residual modulation (SWRM) for evidence-guided spatial suppression and cyclic anchored snapshot self-teaching (CASST) for temporally anchored self-distillation to attenuate unsupported responses. Across two OCT datasets DiSS improves overlap, recall, and boundary metrics over the evaluated baselines; on an independent IIH cohort, DiSS-derived HRF readouts correlate with lumbar-puncture opening pressure more strongly than baseline-derived readouts (a directional, non-significant trend), and within-patient longitudinal changes in HRF burden track changes in neuroretinal layer thickness, while broader clinical validation is still required.

Keywords: Hyperreflective foci segmentation · Optical coherence tomography · Self-teaching · Exponential moving average

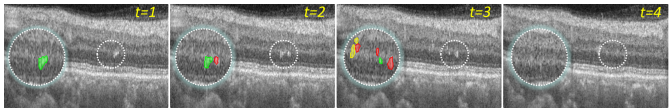

Fig. 1. Illustration of labeling inconsistency in adjacent OCT *B*-scans.

1 Introduction

Idiopathic intracranial hypertension (IIH) is characterized by raised intracranial pressure (ICP) and is often accompanied by papilledema; left unmonitored, it can lead to irreversible visual loss [17,9]. Its diagnosis and monitoring traditionally rely on ICP measurement, which is invasive and hard to repeat, and on fundus grading, which is subject to inter-observer variability [15,5]. Optical coherence tomography (OCT) has recently emerged as an important non-invasive adjunct: among OCT findings, hyperreflective foci (HRF)—small, punctate, highly reflective deposits—have been reported to correlate significantly with ICP, papilledema grading, and visual dysfunction, and are also observed in other retinal diseases such as diabetic retinopathy [1,4,2,16]. Accurate HRF segmentation therefore provides a more reliable basis for the auxiliary quantification and follow-up study of IIH.

Automated HRF segmentation remains difficult [11,19]. HRF are small, low-contrast, and morphologically similar to speckle noise and other hyperreflective structures such as layer interfaces and vessel projections, which makes volumetric annotation particularly hard and inter-annotator differences substantial [20,10]. Many studies therefore model HRF on two-dimensional OCT *B*-scans; while this lowers the annotation burden, it does not remove the fundamental difficulty. HRF occupy less than 1% of a *B*-scan and are extremely sparse, while manual annotations are often inconsistent across adjacent slices. As illustrated in Fig. 1, anatomically similar structures may be annotated as HRF in one slice and labeled as background in a neighboring slice despite minimal structural change. Green marks HRF identified in the initial scan (t=1), red marks newly appearing HRF or lesion extensions in subsequent slices (t>1), and yellow marks vessel-projection artifacts that resemble true lesions; notably, at t=4 no HRF is labeled despite near-identical anatomy. Such inconsistency causes the same weak lesion to alternately appear as foreground and background during training, leading to substantial prediction fluctuations across epochs. Consequently, HRF segmentation poses an asymmetric sensitivity–specificity dilemma beyond ordinary class imbalance: weak lesions are easily missed, whereas increasing sensitivity tends to admit speckle noise and vessel-like confounders. Prior methods largely focus on enhancing feature representation through handcrafted features, focal priors, attention mechanisms, or multi-slice aggregation [22,21,18], yet still rely on a single end-to-end supervision paradigm that implicitly assumes reliable annotations, making it difficult to balance lesion coverage and confounder suppression.

Hence, we argue that HRF segmentation should not be formulated as a single optimization problem. Instead, it is more naturally viewed as a sequential

process involving two distinct objectives: coverage and suppression. The first objective, coverage, aims to preserve as many plausible lesion candidates as possible despite annotation uncertainty and temporal inconsistency, thereby reducing the risk of missing weak lesions. The second objective, suppression, aims to progressively eliminate unsupported responses and confounding structures after sufficient lesion evidence has been consolidated. Such a formulation better reflects both the characteristics of ultra-sparse HRF lesions and the way clinicians interpret OCT scans, where suspicious findings are first identified conservatively and subsequently refined through contextual verification.

Motivated by this observation, we propose DiSS for ultra-sparse HRF segmentation. This phased self-teaching framework first covers plausible lesions and then suppresses artifacts. Rather than introducing a new segmentation backbone, DiSS restructures how supervisory evidence is accumulated and consumed throughout training. During the cover stage, a multi-timescale EMA teacher ensemble (MTE) aggregates complementary temporal observations and consolidates weak lesion evidence across different temporal scales, thereby maximizing lesion coverage and reducing missed detections under inconsistent annotations. During the suppress stage, the accumulated evidence is consumed in two complementary ways. Spatially, soft-weighted residual modulation (SWRM) transforms teacher evidence into a guarded modulation field that emphasizes plausible lesion regions while progressively attenuating confounding structures. Temporally, cyclic anchored snapshot self-teaching (CASST) provides periodically refreshed temporal anchors that stabilize weak-lesion predictions and suppress abrupt epoch-to-epoch fluctuations. Together, MTE, SWRM, and CASST establish a phased evidence accumulation-and-consumption mechanism, through which lesion discovery and false-positive suppression are explicitly decoupled during training. As a result, DiSS effectively addresses the sensitivity–specificity dilemma arising from sparse lesions and unreliable supervision. Our main contributions are:

1. We propose DiSS, a two-stage *cover-then-suppress* self-teaching framework for ultra-sparse HRF segmentation whose core is not a new architecture but a principled training schedule: it decouples weak-lesion coverage from confounder suppression across stages and separates the *admission* and *consumption* of teacher evidence, alleviating the sensitivity–specificity imbalance without extra annotation or backbone modification.
2. We design two complementary mechanisms—SWRM (a guarded modulation field that focuses on plausible lesions, with an annotation-anchored soft constraint that lowers the risk of attenuating consolidated weak lesions) and CASST (a periodically refreshed snapshot designed to anchor, rather than follow, epoch-to-epoch fluctuations)—and validate each by ablation under a unified protocol.
3. We evaluate DiSS and all competing methods (re-trained under the same configuration) on two HRF datasets using identical five-fold splits and a unified inference-and-evaluation protocol: DiSS attains the highest mean Dice in the complete five-fold comparison, and on an independent IIH cohort its

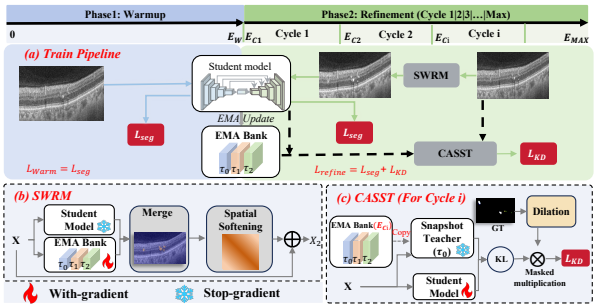

Fig. 2. The overall architecture of the proposed DiSS.

HRF readouts align with lumbar-puncture opening pressure (most strongly among the compared methods, though not statistically established) and their within-patient longitudinal changes track neuroretinal change, indicating potential clinical relevance.

2 Method

Let $x \in \mathbb{R}^{H\times W}$ be an OCT B-scan and $y \in \{0,1\}^{H\times W}$ its HRF mask. The student network f_θ is the standard nnU-Net 2D backbone, and its foreground probability is

$$p^{\mathrm{S}}(x) = \big[\mathrm{softmax}\, f_\theta(x)\big]_{\mathrm{fg}}. \tag{1}$$

DiSS introduces only training-time teachers and one frozen snapshot; all are removed at test time. The method has two phases split by E_{cov}: a cover phase builds teacher evidence, and a refinement phase consumes that evidence for spatial and temporal regularization.

2.1 Coverage-Oriented Teacher Evidence

EMA teachers and temporal ensembling are established regularizers [14,8,23]; DiSS adapts them to sparse HRF by using them as a recall-oriented candidate accumulator. The teacher bank stores $K = 3$ complete, non-trainable model copies that share the student's nnU-Net architecture, with momenta $\{0.8, 0.99, 0.999\}$ and activation epochs $\{50, 25, 0\}$. At its activation epoch, teacher k is initialized from the current student and subsequently updated after every student optimization step, without gradients, by

$$\phi_s^{(k)} = m_k \phi_{s-1}^{(k)} + (1 - m_k)\theta_s. \tag{2}$$

Let $\mathcal{A}_E$ be the active teachers at epoch E and $p_{k,E}^{\mathrm{T}}(x)$ their foreground maps generated online for the current input; the bank retains neither these maps nor

intermediate features. Because a weak lesion may only be visible at one timescale, we aggregate teacher maps with a soft disjunction rather than an average:

$$p_E^{\mathrm{T}}(x) = 1 - \prod_{k \in \mathcal{A}_E} \big(1 - p_{k,E}^{\mathrm{T}}(x)\big). \tag{3}$$

This noisy-OR is used as a coverage operator, not as a calibrated independent-probability model. It retains candidate support from any active timescale and intentionally tolerates some false positives for later refinement.

2.2 Evidence Consumption

SWRM turns the teacher evidence into a guarded residual modulation field. First, teacher and student evidence are merged by $p(x) = \max(p^{\mathrm{T}}(x), p^{\mathrm{S}}(x))$. A soft weight map is then built with a low-shifted sigmoid, dilation-like pooling, smoothing, and a non-zero floor:

$$\tilde{W} = \sigma\left(\frac{p(x) - t_{\mathrm{low}}}{\tau}\right), \quad W = w_{\min} + (1 - w_{\min})\,\mathrm{AvgPool}_{k_s}(\mathrm{MaxPool}_{k_d}(\tilde{W})). \tag{4}$$

An annotation-anchored guard caps the weight outside a dynamic neighborhood of the ground truth,

$$W_{\mathrm{guard}}(u) = \begin{cases} W(u), & u \in \mathcal{N}_{r(E)}(y), \\ \min(W(u), w_{\mathrm{cap}}(E)), & u \notin \mathcal{N}_{r(E)}(y), \end{cases} \tag{5}$$

and the normalized field is injected as a bounded residual,

$$x_2 = x \odot \big(1 + \alpha_E \gamma \hat{W}_{\mathrm{guard}}\big). \tag{6}$$

SWRM therefore does not mask out candidate pixels; it only bounds the additional gain given to unsupported locations.

CASST supplies a temporal anchor. At the beginning of each cycle, the slow EMA teacher is frozen as a snapshot and held fixed until the next cycle. A masked asymmetric KL loss [7] aligns the student to this fixed target on regions selected from W_{guard} and a boundary band:

$$\mathcal{L}_{\mathrm{kd}}(x) = \frac{1}{\sum_u M(u) + \epsilon} \sum_u M(u) D_{\mathrm{KL}}\big(q^{\mathrm{snap}}(u) \,\|\, q^{\mathrm{S}}(u)\big), \tag{7}$$

where $q^{\mathrm{snap}} = \mathrm{softmax}(g^{\mathrm{snap}}/T)$ and $q^{\mathrm{S}} = \mathrm{softmax}(g^{\mathrm{S}})$. The final loss is standard supervised segmentation in the cover phase and, in refinement,

$$\mathcal{L}_{\mathrm{refine}} = \mathcal{L}_{\mathrm{seg}}(f_\theta(x_2), y) + \lambda_{\max} \rho(E) \mathcal{L}_{\mathrm{kd}}(x), \tag{8}$$

with a within-cycle ramp $\rho(E)$. The supervised and distillation terms update only the student; EMA teachers and snapshots receive no gradients.

3 Experiments and Results

Datasets and protocol. HRF-1 contains 27 patients, 46 eyes, 86 IIH OCT volumes (688 B-scans); HRF-2 contains 29 patients, 57 eyes, 57 diabetic-retinopathy OCT volumes (456 B-scans). Two experienced ophthalmologists annotated eight consecutive B-scans per volume, covering the typical two-to-four-slice HRF span. Five-fold cross-validation was performed within each development set at the OCT-volume rather than patient level, keeping all eight B-scans from each volume together and the test sets fixed. A third independent IIH cohort is used for clinical association analysis. All segmentation comparisons use identical preprocessing and augmentation, and the same nnU-Net training configuration: SGD with Nesterov momentum, initial learning rate 0.01, batch size 2, 300 epochs, and polynomial learning-rate decay. Competing methods are re-trained from scratch under this shared protocol and evaluated through one inference-and-metric pipeline. For lesion-level evaluation, two-dimensional connected components of at least one pixel are greedily matched one-to-one by intersection over union (IoU), with pairs at IoU ≥ 0.1 counted as true positives; lesion F1 is micro-averaged by pooling true positives, false positives, and false negatives over all test B-scans. DiSS uses one fixed configuration across datasets and folds: EMA momenta $\{0.8, 0.99, 0.999\}$ with activation epochs $\{50, 25, 0\}$ and refinement from epoch 80; SWRM uses $t_{\text{low}}=\tau=0.03$, $w_{\min}=0.1$, $\gamma=0.15$, skip probability 0.15, radius 7→3, and cap 1.0→0.5; CASST uses 30-epoch cycles, $R_{\text{ramp}}=25$, $\lambda_{\max}=0.2$, $T=1.5$, and $\tau_{\text{roi}}=0.25$.

Table 1. Segmentation comparison. Bold indicates best mean and underline second-best mean.

	HRF-1					HRF-2				
Method	Dice↑	Jac.↑	Rec.↑	ASD↓	Lesion F1↑	Dice↑	Jac.↑	Rec.↑	ASD↓	Lesion F1↑
U-Net [12]	0.6907	0.5496	<u>0.7097</u>	13.3302	0.7083	0.6683	0.5320	0.6401	22.0037	0.7085
nnU-Net [6]	0.7164	0.5763	0.6663	<u>11.0986</u>	0.7456	0.6723	0.5388	0.5948	17.2747	0.6901
SANet [21]	0.6953	0.5593	0.6596	12.9510	0.7586	0.6445	0.5081	0.5685	16.7661	0.6940
CCSNet [19]	<u>0.7225</u>	<u>0.5826</u>	0.6955	14.3866	<u>0.7676</u>	0.6756	0.5453	0.6124	21.3881	0.6953
nnWNet2D [25]	0.6985	0.5597	0.6546	14.9922	0.7561	0.6779	0.5493	0.6277	16.8164	0.7103
HAFNet [24]	0.7132	0.5753	0.7008	14.1340	0.7586	<u>0.6977</u>	<u>0.5682</u>	<u>0.6528</u>	17.3674	0.7150
MUSENet [18]	0.7073	0.5671	0.6992	12.5855	0.7494	0.6400	0.5122	0.6028	<u>15.6184</u>	<u>0.7185</u>
DiSS	**0.7338**	**0.5945**	**0.7439**	**10.3971**	**0.7735**	**0.7112**	**0.5816**	**0.6957**	**12.2111**	**0.7245**

Segmentation performance. Table 1 compares DiSS with U-Net, nnU-Net, recent medical segmentation networks, and the HRF-specific MUSENet. DiSS achieves the best mean Dice, Jaccard, recall, and ASD on both datasets. Relative to nnU-Net, Dice improves from 0.7164 to 0.7338 and Jaccard from 0.5763 to 0.5945 on HRF-1; on HRF-2, Dice improves from 0.6723 to 0.7112 and Jaccard from 0.5388 to 0.5816. ASD decreases from 11.0986 to 10.3971 and from 17.2747 to 12.2111, respectively. DiSS also achieves the highest lesion F1 on HRF-1 (0.7735) and HRF-2 (0.7245), indicating the best balance among the compared methods between lesion detection and false-positive lesion control.

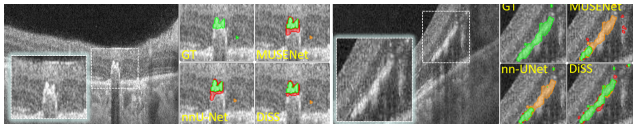

Fig. 3. Qualitative comparison: green, correct; orange, under-segmentation; red, over-segmentation. DiSS gives more coherent masks in low-contrast and artifact-affected regions.

Table 2. Ablation study. MTE is the shared evidence source; SWRM and CASST are toggled as evidence-consumption paths.

	Comp.		HRF-1			HRF-2		
ID	S	C	Dice↑	Recall↑	ASD↓	Dice↑	Recall↑	ASD↓
A0 Base	–	–	0.7164	0.6663	11.0986	0.6723	0.5948	17.2747
A1 w/o S	–	✓	0.7001	0.7108	13.7023	0.6141	0.5742	22.7454
A2 w/o C	✓	–	0.6990	0.7577	15.0766	0.6208	0.6442	23.8532
A3 Full	✓	✓	**0.7338**	0.7439	**10.3971**	**0.7112**	**0.6957**	**12.2111**

Ablation. Table 2 shows that using only one evidence-consumption path is insufficient. Removing SWRM or CASST increases the sensitivity–specificity imbalance: recall can rise in one setting, but Dice and ASD degrade on both datasets. The full model gives the strongest Dice/ASD balance, supporting the need for spatial and temporal evidence consumption after coverage-oriented teacher construction.

Internal design choices. Table 3 restores the component-level ablations. The slow EMA anchor gives the best Dice/ASD trade-off on both datasets, supporting the use of a long-memory snapshot for CASST. In SWRM, neither merge nor softening alone explains the final behavior; the full combination gives the strongest overall Dice and ASD, indicating that candidate-preserving evidence fusion and soft residual weighting are complementary.

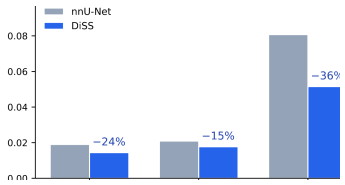

Fig. 4. Last-50-epoch HRF-1 validation pseudo-Dice volatility; lower is more stable. DiSS reduces late-training standard deviation, epoch-to-epoch change, and worst jump versus nnU-Net.

Training stability. Consistent with the interval-fixed anchoring intended by CASST, DiSS is steadier during the converged phase than nnU-Net: late-

Table 3. Detailed component ablations. Bold and underline denote the best and second-best mean values, respectively.

(a) Anchor selection in CASST.

	HRF-1			HRF-2		
Model	Dice↑	Recall↑	ASD↓	Dice↑	Recall↑	ASD↓
EMA Fast	0.6773	**0.7806**	15.4355	<u>0.6578</u>	<u>0.6629</u>	22.1920
EMA Medium	0.6941	<u>0.7748</u>	14.9539	0.6421	0.6303	25.5133
EMA Slow	**0.7338**	0.7439	**10.3971**	**0.7112**	**0.6957**	**12.2111**
Student	<u>0.7089</u>	0.7689	<u>13.8313</u>	0.6414	0.6419	<u>16.3731</u>

(b) Operations in SWRM.

	HRF-1			HRF-2		
Condition	Dice↑	Recall↑	ASD↓	Dice↑	Recall↑	ASD↓
Full	**0.7338**	0.7439	**10.3971**	**0.7112**	**0.6957**	**12.2111**
w/o Softening	<u>0.7028</u>	<u>0.7612</u>	14.8366	<u>0.6395</u>	0.6179	<u>18.7337</u>
w/o Merge	0.6947	0.7383	15.5868	0.6478	<u>0.6283</u>	26.0347
w/o Both	0.7001	0.7108	13.7023	0.6141	0.5742	22.7454

training validation pseudo-Dice standard deviation is reduced by 24%, the mean epoch-to-epoch change by 15%, and the worst single-epoch jump by 36% (Fig. 4).

Clinical alignment. On the independent IIH cohort, HRF count and burden are computed per eye and averaged by patient. With patient-level Spearman correlation [13] and 5000 bootstrap resamples [3], DiSS-derived burden/count show the strongest associations with lumbar-puncture opening pressure among the compared readouts ($\rho = 0.28/0.26$, versus CCSNet 0.21/0.21 and nnU-Net 0.19/0.16; Fig. 5a), although overlapping confidence intervals indicate only a directional trend. In the paired follow-up subset ($n = 33$), change in DiSS-derived HRF burden correlates with changes in GCL+IPL ($\rho = 0.54$), INL ($\rho = 0.51$), and RNFL ($\rho = 0.49$; Fig. 5b), with FDR-corrected significance.

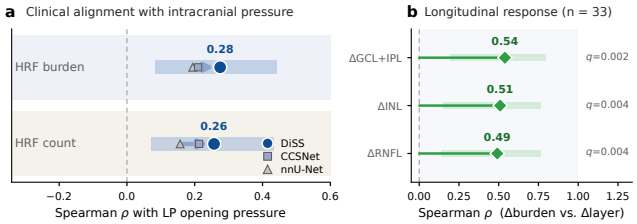

Fig. 5. Clinical alignment of DiSS-derived HRF readouts. (a) Patient-level Spearman correlations with opening pressure; shaded intervals are DiSS 95% bootstrap CIs. (b) In paired follow-up patients ($n = 33$), ΔHRF burden correlates with ΔGCL+IPL, ΔINL, and ΔRNFL; bars show bootstrap 95% CIs and q values are FDR-corrected.

Visual-acuity change was not associated with ΔHRF; these exploratory results warrant prospective confirmation.

4 Conclusion

DiSS uses EMA teachers, SWRM, and CASST to separate HRF coverage from artifact suppression, improving segmentation on two OCT datasets. In an independent IIH cohort, its readouts show the strongest but not significantly superior association with opening pressure, and burden changes track neuroretinal change. Broader validation is needed before deployment.

Acknowledgments. This work is supported by the National Science Foundation Program of China (62422122, 62272444, U20A20386), the Key Research and Development Program of Zhejiang Province (2024C03204), the Ningbo 'KeChuang Yongjiang 2035' Key R&D Program (2024Z230).

Ethics Statement All images were acquired with regulatory approvals and patient consent as appropriate, following the Declaration of Helsinki.

Disclosure of Interests. The authors declare no competing interests.

References

1. Bolz, M., Schmidt-Erfurth, U., Deak, G., Mylonas, G., Kriechbaum, K., Scholda, C., et al.: Optical coherence tomographic hyperreflective foci: a morphologic sign of lipid extravasation in diabetic macular edema. Ophthalmology **116**(5), 914–920 (2009). https://doi.org/10.1016/j.ophtha.2008.12.039
2. Cao, L., Wang, H., Kwapong, W.R., Wan, J., Yan, Y., Liu, G., et al.: Macular exudate in idiopathic intracranial hypertension affects outer retina and visual acuity. BMJ Open Ophthalmology **9**(1), e001810 (2024). https://doi.org/10.1136/bmjophth-2024-001810
3. Efron, B.: Bootstrap methods: another look at the jackknife. In: Breakthroughs in statistics: Methodology and distribution, pp. 569–593. Springer (1992). https://doi.org/10.1007/978-1-4612-4380-9_41
4. Fragiotta, S., Abdolrahimzadeh, S., Dolz-Marco, R., Sakurada, Y., Gal-Or, O., Scuderi, G.: Significance of hyperreflective foci as an optical coherence tomography biomarker in retinal diseases: characterization and clinical implications. Journal of ophthalmology **2021**(1), 6096017 (2021). https://doi.org/10.1155/2021/6096017
5. Horton, J.C.: Idiopathic intracranial hypertension. New England Journal of Medicine **393**(14), 1409–1419 (2025). https://doi.org/10.1056/NEJMra2404929
6. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods **18**(2), 203–211 (2021). https://doi.org/10.1038/s41592-020-01008-z
7. Kullback, S., Leibler, R.A.: On information and sufficiency. The annals of mathematical statistics **22**(1), 79–86 (1951). https://doi.org/10.1214/aoms/1177729694

8. Laine, S., Aila, T.: Temporal ensembling for semi-supervised learning. In: International Conference on Learning Representations (ICLR) (2017)
9. Markey, K.A., Mollan, S.P., Jensen, R.H., Sinclair, A.J.: Understanding idiopathic intracranial hypertension: mechanisms, management, and future directions. The Lancet Neurology **15**(1), 78–91 (2016). https://doi.org/10.1016/S1474-4422(15)00298-7
10. Monemian, M., Ghaderi Daneshmand, P., Rakhshani, S., Rabbani, H.: A new texture-based labeling framework for hyper-reflective foci identification in retinal optical coherence tomography images. Scientific Reports **14**(1), 22933 (2024). https://doi.org/10.1038/s41598-024-73927-2
11. Okuwobi, I.P., Fan, W., Yu, C., Yuan, S., Liu, Q., Zhang, Y., et al.: Automated segmentation of hyperreflective foci in spectral domain optical coherence tomography with diabetic retinopathy. Journal of medical imaging **5**(1), 014002 (2018). https://doi.org/10.1117/1.JMI.5.1.014002
12. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241. Springer (2015). https://doi.org/10.1007/978-3-319-24574-4_28
13. Spearman, C.: The proof and measurement of association between two things. The American Journal of Psychology **15**(1), 72–101 (1904). https://doi.org/10.2307/1412159
14. Tarvainen, A., Valpola, H.: Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. In: Advances in Neural Information Processing Systems (NeurIPS). pp. 1195–1204 (2017)
15. Tata, G., Kisabay, A., Gokcay, F., Celebisoy, N.: Idiopathic intracranial hypertension: Are there predictors for visual outcome or recurrences? Clinical Neurology and Neurosurgery **183**, 105378 (2019). https://doi.org/10.1016/j.clineuro.2019.105378
16. Viggiano, P., Boscia, G., Giannaccare, G., Santoro, M., Petrara, G., Borriello, C., et al.: Topographical quantification of hyper-reflective foci may predict the development of macular atrophy in patients with neovascular age-related macular degeneration. Investigative Ophthalmology & Visual Science **65**(14), 45 (2024). https://doi.org/10.1167/iovs.65.14.45
17. Wakerley, B.R., Mollan, S.P., Sinclair, A.J.: Idiopathic intracranial hypertension: update on diagnosis and management. Clinical Medicine **20**(4), 384–388 (2020). https://doi.org/10.7861/clinmed.2020-0232
18. Wang, X., Ma, Y., Guo, X., Zheng, Y., Zhang, J., Liu, Y., et al.: A hyper-reflective foci segmentation network for OCT images with multi-dimensional semantic enhancement. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 645–655. Springer (2024). https://doi.org/10.1007/978-3-031-72378-0_60
19. Wang, X., Zhang, Y., Ma, Y., Kwapong, W.R., Ying, J., Lu, J., et al.: Automated evaluation of retinal hyperreflective foci changes in diabetic macular edema patients before and after intravitreal injection. Frontiers in Medicine **10**, 1280714 (2023). https://doi.org/10.3389/fmed.2023.1280714
20. Wei, J., Yu, S., Du, Y., Liu, K., Xu, Y., Xu, X.: Automatic segmentation of hyperreflective foci in OCT images based on lightweight DBR network. Journal of Digital Imaging **36**(3), 1148–1157 (2023). https://doi.org/10.1007/s10278-023-00786-0

21. Yao, C., Zhu, W., Wang, M., Zhu, L., Huang, H., Chen, H., et al.: SANet: a self-adaptive network for hyperreflective foci segmentation in retinal OCT images. In: Medical Imaging 2021: Image Processing. vol. 11596, p. 115962Y. SPIE (2021). https://doi.org/10.1117/12.2580699
22. Zhang, B., Ma, L., Zhao, H., Hao, Y., Fu, S., Wang, H., et al.: Automatic segmentation of hyperreflective dots via focal priors and visual saliency. Medical Physics **49**(11), 7025–7037 (2022). https://doi.org/10.1002/mp.15848
23. Zhang, L., Song, J., Gao, A., Chen, J., Bao, C., Ma, K.: Be your own teacher: Improve the performance of convolutional neural networks via self distillation. In: IEEE/CVF International Conference on Computer Vision (ICCV). pp. 3712–3721 (2019). https://doi.org/10.1109/ICCV.2019.00381
24. Zhang, Y., Bao, W., Yang, Y., Wan, W., Xiao, Q., Zou, X.: HAFNet: Hierarchical attention fusion network for infrared small target detection. IEEE Transactions on Geoscience and Remote Sensing **63**, 1–16 (2025). https://doi.org/10.1109/TGRS.2025.3607732
25. Zhou, Y., Li, L., Lu, L., Xu, M.: nnWNet: Rethinking the use of transformers in biomedical image segmentation and calling for a unified evaluation benchmark. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 20852–20862 (2025). https://doi.org/10.1109/CVPR52734.2025.01942