This is the Pre-Published Version.

This version of the proceeding paper has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use(https://www.springernature.com/gp/open-research/policies/accepted-manuscript-terms), but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: http://dx.doi.org/10.1007/978-3-031-43990-2_61.

VF-HM: Vision Loss Estimation using Fundus Photograph for High Myopia

Zipei Yan[1,⋆], Dong Liang[2,⋆,✉], Linchuan Xu[1,✉], Jiahang Li[1], Zhengji Liu[2], Shuai Wang[3], Jiannong Cao[1], and Chea-su Kee[2]

[1] Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong
linch.xu@polyu.edu.hk

[2] School of Optometry, The Hong Kong Polytechnic University, Kowloon, Hong Kong
dong1.liang@connect.polyu.hk

[3] Department of Biomedical Engineering, Tsinghua University, Beijing, China

Abstract. High myopia (HM) is a leading cause of irreversible vision loss due to its association with various ocular complications including myopic maculopathy (MM). Visual field (VF) sensitivity systematically quantifies visual function, thereby revealing vision loss, and is integral to the evaluation of HM-related complications. However, measuring VF is subjective and time-consuming as it highly relies on patient compliance. Conversely, fundus photographs provide an objective measurement of retinal morphology, which reflects visual function. Therefore, utilizing machine learning models to estimate VF from fundus photographs becomes a feasible alternative. Yet, estimating VF with regression models using fundus photographs fails to predict local vision loss, producing stationary nonsense predictions. To tackle this challenge, we propose a novel method for VF estimation that incorporates VF properties and is additionally regularized by an auxiliary task. Specifically, we first formulate VF estimation as an ordinal classification problem, where each VF point is interpreted as an ordinal variable rather than a continuous one, given that any VF point is a discrete integer with a relative ordering. Besides, we introduce an auxiliary task for MM severity classification to assist the generalization of VF estimation, as MM is strongly associated with vision loss in HM. Our method outperforms conventional regression by 16.61% in MAE metric on a real-world dataset. Moreover, our method is the first work for VF estimation using fundus photographs in HM, allowing for more convenient and accurate detection of vision loss in HM, which could be useful for not only clinics but also large-scale vision screenings.

Keywords: Vision loss estimation · Visual field · Fundus photograph · Ordinal classification · Auxiliary learning.

⋆ Equal contribution

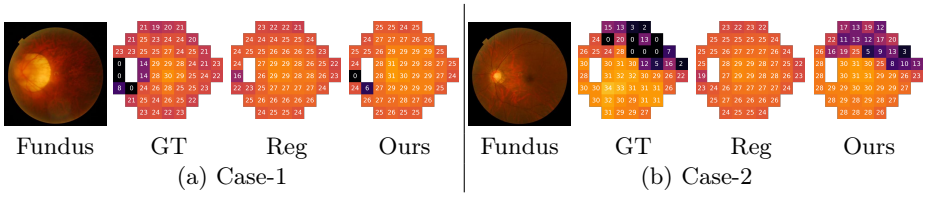


Fig. 1: Estimated VF from different methods using fundus. GT denotes the ground truth, Reg denotes the regression baseline, and Ours denotes our method.

1 Introduction

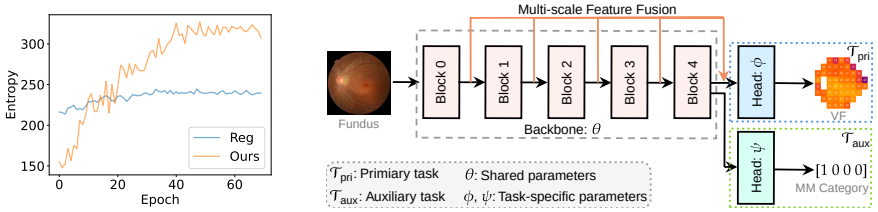
High myopia (HM) has become a global concern for public health, with its markedly growing prevalence [10] and its increased risk of irreversible vision loss and even blindness [25,16,14,27]. In brief, excessive axial elongation in HM eyes will produce mechanical stretching on the posterior segment of eyeballs, leading to various structural changes and HM-related complications, e.g., myopic maculopathy (MM), and consequently, functional changes, resulting in vision loss.

Accurate quantification of vision loss is integral to the early detection and timely treatment for MM and other HM-related complications [16]. Currently, the diagnosis of vision loss is made on the basis of visual field (VF) sensitivity by standard automated perimetry, which is a systematic metric and gold standard to quantify visual function [19]. However, measuring VF is prohibitively time-consuming and subjective as it highly requires patients' concentration and compliance during the test [12].

Conversely, imaging techniques, such as fundus photography (a.k.a., fundus), provide a relatively objective and robust measurement of the retinal morphology, which likely corresponds to the VF with an underlying "structure-function relationship" [32,27]. Actually, fundus is most commonly used for the diagnosis and evaluation of HM and its complications, in particular in rural and developing regions, with its lower cost and convenience of acquisition [17].

Therefore, utilizing machine learning models to estimate VF from fundus becomes a promising and feasible alternative for HM subjects in clinical practice. To the best of our knowledge, there is no existing approach to estimate VF from fundus. Some studies have been proposed to estimate the global indices (e.g., mean deviation) of VF from fundus [3,11], and others estimate VF using retinal thickness [18,4,28,30]. It is worth mentioning that, all these studies were conducted for the glaucoma population [3,11,18,4,28,30], in which most cases of visual abnormality or defect were likely glaucomatous. However, MM and other HM-related complications may lead to non-glaucomatous vision loss.

Actually, estimating VF with conventional regression [18] using fundus fails to predict local vision loss in our HM population, producing stationary nonsense predictions. As shown in Fig. 1, these predictions from regression exhibit a relatively similar and consistent pattern in most HM subjects, failing to cap-



(a) Entropy of feature space (b) An overview of the proposed method.

Fig. 2: (a) The entropy of feature space on training data during training progress from conventional regression (denoted by Reg) and our VF-HM. (b) An overview of our proposed method: VF-HM.

ture/learn the inter-subject variability and local defects of VF. And these predictions are very close to the mean value of VF in training data (see the supplementary material). The reason for such failure lies in regression's inability to learn high-entropy feature representations [31], which is further confirmed by measuring the entropy of feature representations, as marked in blue in Fig. 2a.

To tackle this challenge, we propose a novel method for estimating VF for HM using fundus, namely VF-HM. In general, VF-HM incorporates VF properties and is additionally regularized by an auxiliary task, thereby learning relatively high-entropy feature representations (see the orange line in Fig. 2a). In detail, we formulate VF estimation as an ordinal classification problem, where each VF point is interpreted as an ordinal variable rather than a continuous one, given that any VF point is a discrete integer with a relative ordering. Besides, we introduce an auxiliary task for MM severity classification to assist the generalization of VF estimation, because MM is strongly associated with vision loss in HM [21,16,7,32] and its symptom can be observed from the fundus directly. As a result, VF-HM significantly outperforms conventional regression and accurately predicts vision loss (see Fig. 1)

Our contributions are summarized as follows:

- We propose a novel method, VF-HM, for estimating VF from fundus for HM. VF-HM more accurately detects the local vision loss and significantly outperforms conventional regression by 16.61% in the MAE metric on a real dataset.
- VF-HM is the first work for VF estimation using fundus for HM, allowing for more convenient and cost-efficient detection of vision loss in HM, which could be useful for not only clinics but also large-scale vision screenings.

2 Problem Formulation

Let $\mathcal{D} = \{(\boldsymbol{x}_i, \boldsymbol{m}_i)\}$ denote the training set, where $\boldsymbol{x}_i \in \mathcal{X}$ denotes the fundus, $\boldsymbol{m}_i \in \mathcal{M}$ denotes its corresponded VF. And $\mathcal{A} = \{(\boldsymbol{x}_i, y_i)\}$ denotes the auxiliary

set, where $y_i \in \mathcal{Y}$ denotes the MM severity category of a given $\boldsymbol{x}_i$. The objective is to learn a model $f : \mathcal{X} \rightarrow \mathcal{M}$ by utilizing both $\mathcal{D}$ and $\mathcal{A}$. The novelty of this formulation is additionally utilizing the auxiliary set to improve the model's generalization. And challenges mainly come from the following two aspects. First, how to design the model f, as mentioned earlier, conventional regression fails to predict local vision loss. Second, how to properly utilize the auxiliary set to assist the generalization of f, as the auxiliary information is not always helpful during the learning progress, i.e., sometimes may interfere [6,5,22].

3 Proposed Method: VF-HM

In this section, we first present an overview of the proposed method. Then, we introduce the details of different components.

3.1 Overview

We present an overview of the proposed method in Fig. 2b. Specifically, the primary task (denoted by $\mathcal{T}_{\text{pri}}$) is the VF estimation and the auxiliary task is MM classification (denoted by $\mathcal{T}_{\text{aux}}$). Then, our method aims to solve $\mathcal{T}_{\text{pri}}$ with the assistance of $\mathcal{T}_{\text{aux}}$. We propose to parameterize the solution for $\mathcal{T}_{\text{pri}}$ and $\mathcal{T}_{\text{aux}}$ by two neural networks: $f(\cdot;\theta,\phi)$ and $g(\cdot;\theta,\psi)$, where they share the same backbone θ and have their own task-specific parameters ϕ and ψ. Thereafter, the overall objective function is formulated as follows:

$$\mathcal{L} = \mathcal{L}_{\text{pri}}(\theta,\phi) + \lambda\mathcal{L}_{\text{aux}}(\theta,\psi) \tag{1}$$

where $\mathcal{L}_{\text{pri}}$ and $\mathcal{L}_{\text{aux}}$ denote the loss function for $\mathcal{T}_{\text{pri}}$ and $\mathcal{T}_{\text{aux}}$, respectively. $\lambda \in (0,1]$ is a hyper-parameter to control the importance of $\mathcal{L}_{\text{aux}}$.

3.2 Primary Task: VF Estimation

The overall interest is *only* the primary task $\mathcal{T}_{\text{pri}}$, which is parameterized by $f(\cdot;\theta,\phi) : \mathcal{X} \rightarrow \mathcal{M}$. Specifically, we formulate $\mathcal{T}_{\text{pri}}$ as an *ordinal classification* (aka, *rank learning*) problem, where each VF point m_i^j represents an *ordinal variable/rank* rather than a continuous one. Such a formulation incorporates the distinct properties of VF, which include: 1) Discretization: $\forall m_i^j \in [0,40] \cap \mathbb{Z}$, that is, any VF value is a *positive discrete integer*. 2) Ordinalization: $m_i^0 \prec m_i^1 \prec ... \prec m_i^j$, there is a *relative order* among VF values. To achieve this goal, we extend the *ordinal variable/rank* into binary labels [13,2], i.e., $\boldsymbol{m}_i^j = [r_i^{j,1}, ..., r_i^{j,K-1}]^T$ where $r_i^{j,k} \in \{0,1\}$ indicates whether $\boldsymbol{m}_i^j$ exceeds k-th rank or not. To ensure rank-monotonic and guarantee prediction consistency, we utilize the *ordinal bias* [2]. In detail, the task-specific parameter ϕ contains independent bias for each ordinal variable. Thereafter, $\mathcal{T}_{\text{pri}}$ can be solved by the binary cross-entropy loss, which is defined as follows:

$$\mathcal{L}_{\text{pri}}(\theta,\phi) = \mathbb{E}_{(\boldsymbol{x}_i,\boldsymbol{m}_i)\in\mathcal{X}\times\mathcal{M}}[L_{\text{BCE}}(f(\boldsymbol{x}_i;\theta,\phi),\boldsymbol{m}_i)] \tag{2}$$

where $L_{\mathrm{BCE}}(\cdot)$ denotes the binary cross-entropy loss

In addition, we propose to reuse the features from different blocks, as they contain distinct spatial information. Specifically, we propose Multi-scale Feature Fusion (MFF) for aggregating features from different blocks. As highlighted in orange in Fig. 2b, MFF aggregates features from all blocks at the last in an addition operation. The detailed implementation is reported in Sec. 4.2.

3.3 Auxiliary Task: MM Classification

The auxiliary task $\mathcal{T}_{\mathrm{aux}}$ is introduced *only* to assist the generalization of $\mathcal{T}_{\mathrm{pri}}$. Specifically, $\mathcal{T}_{\mathrm{aux}}$ is to predict MM severity category y_i from fundus $\boldsymbol{x}_i$, which is parameterized by $g(\cdot;\theta,\psi) : \mathcal{X} \to \mathcal{Y}$. MM is highly correlated to vision loss [21,16,7,32], and its symptom can be observed from the fundus directly. According to its increasing severity, MM can be classified into five categories [26], i.e., $C_0 \prec C_1 ... \prec C_4$. Therefore, we also interpret the MM category as the *ordinal variable/rank*. Similar to the label extension in $\mathcal{T}_{\mathrm{pri}}$, we extend the MM catergory into binary labels $\boldsymbol{y}_i = [r_1, r_2, r_3, r_4]^T$. The loss function $\mathcal{L}_{\mathrm{aux}}$ for solving $\mathcal{T}_{\mathrm{aux}}$ is also the binary cross-entropy, which is defined as follows:

$$\mathcal{L}_{\mathrm{aux}}(\theta,\psi) = \mathbb{E}_{(\boldsymbol{x}_i,\boldsymbol{y}_i)\in\mathcal{X}\times\mathcal{Y}}[L_{\mathrm{BCE}}(g(\boldsymbol{x}_i;\theta,\psi),\boldsymbol{y}_i)] \tag{3}$$

However, the $\mathcal{T}_{\mathrm{aux}}$ is not always helpful for $\mathcal{T}_{\mathrm{pri}}$ because of the negative transfer [6,22,5]. The negative transfer refers to a problem that sometimes $\mathcal{T}_{\mathrm{aux}}$ becomes harmful for $\mathcal{T}_{\mathrm{pri}}$. Specifically, let $\nabla_\theta\mathcal{L}$ denote the gradient of Eq.(1) in terms of the shared parameters θ, and it can be decomposed as follows:

$$\nabla_\theta\mathcal{L} = \nabla_\theta\mathcal{L}_{\mathrm{pri}} + \lambda\nabla_\theta\mathcal{L}_{\mathrm{aux}} \tag{4}$$

$\mathcal{T}_{\mathrm{aux}}$ becomes harmful for $\mathcal{T}_{\mathrm{pri}}$, when the *cosine similarity* between $\nabla_\theta\mathcal{L}_{\mathrm{pri}}$ and $\nabla_\theta\mathcal{L}_{\mathrm{aux}}$ becomes negative [6], i.e., $\cos(\nabla_\theta\mathcal{L}_{\mathrm{aux}}, \nabla_\theta\mathcal{L}_{\mathrm{pri}}) < 0$. Negative transfer is observed in our setting when optimizing Eq.(1) directly, as illustrated in Fig. 3a.

Following [6], we mitigate negative transfer by refining $\nabla_\theta\mathcal{L}_{\mathrm{aux}}$. Specifically, we adapt the weighted cosine simiarily to refine $\nabla_\theta\mathcal{L}_{\mathrm{aux}}$, which is defined as follows:

$$\nabla_\theta\mathcal{L}_{\mathrm{aux}} = \max\left(0, \cos(\nabla_\theta\mathcal{L}_{\mathrm{aux}}, \nabla_\theta\mathcal{L}_{\mathrm{pri}})\right) \cdot \nabla_\theta\mathcal{L}_{\mathrm{aux}} \tag{5}$$

4 Experiments

In this section, we conduct experiments on a clinic-collected real-world dataset to evaluate the performance of our proposed method[4].

[4] Our code is available at `https://github.com/yanzipei/VF-HM`

Table 1: Main results. 'K-fold' denotes performance from K-fold cross-validation on training data. 'Test' denotes performance on test data (pre-trained on training data). (↓) denotes the lower value indicates better performance. RT-(·) denotes different retinal thicknesses. And the better results are **bold-faced**.

Method	Modality	K-fold(K=5)			Test		
		RMSE (↓)	MAE (↓)	SMAPE (↓)	RMSE (↓)	MAE (↓)	SMAPE (↓)
Regression	RT-(a)	4.94 ± 0.23	3.12 ± 0.05	13.47 ± 0.16	-	-	-
Regression	RT-(b)	4.80 ± 0.17	3.04 ± 0.12	13.21 ± 0.36	-	-	-
Regression	RT-(c)	4.86 ± 0.22	3.13 ± 0.18	13.42 ± 0.57	-	-	-
Regression	Fundus	4.62 ± 0.07	2.95 ± 0.07	12.94 ± 0.32	4.28 ± 0.03	2.89 ± 0.06	12.13 ± 0.30
Ours(λ=0.1)	Fundus	**4.44 ± 0.27**	**2.78 ± 0.10**	**12.50 ± 0.26**	**3.69 ± 0.03**	**2.41 ± 0.04**	**11.38 ± 0.14**

4.1 The Studied Data

The studied data comes from a HM population, including 75 patients, each with diagnosis information for both eyes. For each eye, there are one fundus, VF, and MM severity category. Specifically, the fundus is captured in colorful mode, the VF is measured in the 24-2 mode with 52 effective points, and MM category is labeled by registered ophthalmologists. Besides, 34 patients (i.e., 68 eyes) have SD-OCT scans in the macular region. For these SD-OCT scans, we extract the retinal thickness with the pre-trained model [20] in order to compare our method to conventional regression using retinal thickness. According to whether the eye has SD-OCT scans or not, we divide the whole data into a training set and a test set. Specifically, the training and test data contain 68 eyes (from 34 patients) and 82 eyes (from 41 patients), respectively. It is worth mentioning that the training data and test data do not have the same patient. Besides, in the following K-fold cross-validation experiments, we split the training data based on the patient's ID to ensure that there is no information leakage.

4.2 Experimental Setup

Data pre-processing. We choose the *left* eye pattern as our base. For fundus, VF and retinal thickness are not in the *left* eye pattern, we convert them using the horizontal flip.

Data augmentation. Following [1], we consolidate a set of data augmentations for both fundus and retinal thickness, respectively. The details are reported in the supplementary material. Different from applying *all* [1] augmentations during training, we utilize the *TrivialAugment* [15] instead, which randomly selects one from the given data augmentations, generating more diverse augmented data.

Evaluation methods. For quantitative evaluation, we utilize three metrics [18,33,29,3,4]: RMSE, MAE and SMAPE. For qualitative evaluation, we visualize two representative predictions on the test set, and more visualized results are presented in the supplementary material.

Table 2: Ablation study on main components. OC denotes the ordinal classification baseline. MFF denotes multi-scale feature fusion. AUX denotes the auxiliary task. MNT denotes mitigating negative transfer from Eq.(5).

OC	MFF	AUX	MNT	RMSE (↓)	MAE (↓)
✓	✓	✓	✓	**3.69 ± 0.03**	**2.41 ± 0.04**
✓	✓	✓		3.74 ± 0.02	2.46 ± 0.03
✓	✓			3.73 ± 0.04	2.45 ± 0.02
✓				3.77 ± 0.02	2.49 ± 0.03

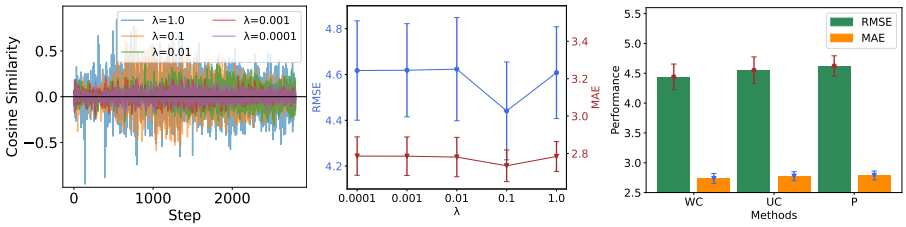
Baseline methods. We mainly compare our method to conventional regression that estimates VF from fundus. Besides, for a more comprehensive comparison, we also compare our approach to conventional regression using different retinal thicknesses. In detail, we consider three variants: (a) the combination of GCIPL, RNFL and RCL [33], (b) the combination of GCIPL and RNFL [18], (c) only RNFL [4]. Due to the limited data, we compare our method to conventional regression using the above thickness by K-fold cross-validation on training data.

Implementation details. We utilize the ResNet-18 [8] as the backbone. For the regression baseline, we use only one *linear* layer at last. For our method, we use the combination of *Conv2D*, *BatchNorm2D* and *ReLU* as the classification head for $\mathcal{T}_{\text{pri}}$. For the MFF, we utilize the above classification head to aggregate features from different blocks. Note that the features from earlier blocks have relatively large features, thus we use *AdaptiveAvgPooling2D* to perform downsampling first. For $\mathcal{T}_{\text{aux}}$, we use only one *linear* layer as the classifier. For a fair comparison, we train all methods with the same training configurations. Specifically, we train the models with 80 training epochs and the SGD optimizer, where the batch size is set to 32, the learning rate is set to 0.01, momentum is set to 0.9 and L2 weight decay is set to $1e^{-4}$. Besides, we utilize a cosine learning rate decay [9] to adjust the learning rate per epoch. Finally, we fix all input resolutions to 384×384 for both training and evaluation. All experiments are run independently with four seeds: 0, 1, 2, and 3. As for hyper-parameters, we search them on training data with K-fold cross-validation.

4.3 Experimental Results

Main results. Table 1 reports the performance of our method and baselines. In general, our method achieves the best performance compared to these baselines. Specifically, compared to conventional regression using fundus, our method outperforms it by 13.79% and 16.61% according to the RMSE and MAE metric on test data. Besides, our method achieves better performance than baselines using different retinal thicknesses.

Visualization of predictions. As shown in Fig. 1, we visualize predictions from methods using fundus on two representative cases. Specifically, conventional regression fails to predict local vision loss, as its predictions share a similar and



(a) Negative transfer (b) Impact of λ (c) Mitigate negative transfer

Fig. 3: Visualization of (a) Negative transfer when optimizing Eq.(1) directly, (b) Impact of hyper-parameter λ, and (c) Different methods for mitigating the negative transfer.

consistent pattern for both cases. In contrast, predictions from our method are more precise, revealing the local vision loss. More visualized results are presented in the supplementary material.

4.4 Ablation Study

To get a better understanding of the effectiveness of the main components in our proposed method, we conduct a series of ablation studies.

Effectiveness of main components. We first examine the effectiveness of the main components by ablating them. The results are reported in Table 2. In general, we can observe that all components can improve performance except AUX. Specifically, AUX denotes solely introducing the auxiliary task, which brings a degradation, because of the existence of negative transfer. Meanwhile, with the help of Eq.(5), the negative transfer can be mitigated. Besides, we observe these main components allow the model to learn high-entropy feature representations, thereby improving the model's performance [31]. More details are reported in the supplementary material.

Impact of hyper-parameter λ. We study the impact of the hyper-parameter λ with K-fold cross validation on training data. We choose $\lambda \in \{1.0, 0.1, 0.01, 0.001, 0.0001\}$. According to the results shown in Fig. 3b, we observe that $\lambda = 0.1$ achieves the best performance.

Different methods for mitigating the negative transfer. We consider three alternatives to refine the auxiliary gradient for mitigating the negative transfer: (1) weighted cosine (WC) similarity [6] (2) unweighted cosine (UC) similarity [6] (3) projection (P) [22]. For a fair comparison, we set $\lambda = 0.1$, then conduct experiments on training data with K-fold cross-validation. As shown in Fig. 3c, and we observe that (1) WC achieves the best performance.

5 Conclusion

In this work, we propose VF-HM for estimating VF from fundus for HM, which is the first work for VF estimation in HM; and it provides a more convenient

and cost-effective way to detect HM-related vision loss. The major limitations include: first, our sample size is limited; second, we utilize both eyes from one patient as two independent inputs, which ignores their similarity; third, we only include the MM severity as the auxiliary information. Future work could be conducted as follows. First, collecting more data from different clinical sites. Second, modeling the relationship between both eyes from the same patient [33]. Third, exploring more auxiliary information. Besides, studying how to adapt our method to different domains is a crucial problem [24], as we seek to improve the generalizability. In addition, exploring VF prediction with the missing modalities [23]: either fundus or thickness is another interesting direction.

Acknowledgements This work was supported by the Research Institute for Artificial Intelligence of Things, The Hong Kong Polytechnic University, HK RGC Research Impact Fund No. R5060-19; and the Centre for Myopia Research, School of Optometry; the Research Centre for SHARP Vision (RCSV), The Hong Kong Polytechnic University; and Centre for Eye and Vision Research (CEVR), InnoHK CEVR Project 1.5, 17W Hong Kong Science Park, HKSAR. We thank Drs Rita Sum and Vincent Ng for their guidance on data analysis of clinical population; and Prof. Ruihua Wei for external validation of the model.

References

1. Bar-David, D., Bar-David, L., Soudry, S., Fischer, A.: Impact of data augmentation on retinal oct image segmentation for diabetic macular edema analysis. In: MICCAI. pp. 148–158 (2021)
2. Cao, W., Mirjalili, V., Raschka, S.: Rank consistent ordinal regression for neural networks with application to age estimation. Pattern Recognition Letters **140**, 325–331 (2020)
3. Christopher, M., Bowd, C., Belghith, A., Goldbaum, M.H., Weinreb, R.N., Fazio, M.A., Girkin, C.A., Liebmann, J.M., Zangwill, L.M.: Deep learning approaches predict glaucomatous visual field damage from oct optic nerve head en face images and retinal nerve fiber layer thickness maps. Ophthalmology **127**(3), 346–356 (2020)
4. Datta, S., Mariottoni, E.B., Dov, D., Jammal, A.A., Carin, L., Medeiros, F.A.: Retinervenet: Using recursive deep learning to estimate pointwise 24-2 visual field data based on retinal structure. Scientific Reports **11**(1), 1–10 (2021)
5. Dery, L.M., Dauphin, Y.N., Grangier, D.: Auxiliary task update decomposition: The good, the bad and the neutral. In: ICLR (2021)
6. Du, Y., Czarnecki, W.M., Jayakumar, S.M., Farajtabar, M., Pascanu, R., Lakshminarayanan, B.: Adapting auxiliary losses using gradient similarity. arXiv preprint arXiv:1812.02224 (2018)
7. Hayashi, K., Ohno-Matsui, K., Shimada, N., Moriyama, M., Kojima, A., Hayashi, W., Yasuzumi, K., Nagaoka, N., Saka, N., Yoshida, T., et al.: Long-term pattern of progression of myopic maculopathy: A natural history study. Ophthalmology **117**(8), 1595–1611 (2010)
8. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: CVPR. pp. 770–778 (2016)

9. He, T., Zhang, Z., Zhang, H., Zhang, Z., Xie, J., Li, M.: Bag of tricks for image classification with convolutional neural networks. In: CVPR. pp. 558–567 (2019)
10. Holden, B.A., Fricke, T.R., Wilson, D.A., Jong, M., Naidoo, K.S., Sankaridurg, P., Wong, T.Y., Naduvilath, T.J., Resnikoff, S.: Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. Ophthalmology **123**(5), 1036–1042 (2016)
11. Lee, J., Kim, Y.W., Ha, A., Kim, Y.K., Park, K.H., Choi, H.J., Jeoung, J.W.: Estimating visual field loss from monoscopic optic disc photography using deep learning model. Scientific reports **10**(1), 1–10 (2020)
12. Lewis, R.A., Johnson, C.A., Keltner, J.L., Labermeier, P.K.: Variability of quantitative automated perimetry in normal observers. Ophthalmology **93**(7), 878–881 (1986)
13. Li, L., Lin, H.: Ordinal regression by extended binary classification. In: NeurIPS. pp. 865–872 (2006)
14. Lin, F., Chen, S., Song, Y., Li, F., Wang, W., Zhao, Z., Gao, X., Wang, P., Jin, L., Liu, Y., et al.: Classification of visual field abnormalities in highly myopic eyes without pathologic change. Ophthalmology **129**(7), 803–812 (2022)
15. Müller, S.G., Hutter, F.: Trivialaugment: Tuning-free yet state-of-the-art data augmentation. In: ICCV. pp. 754–762 (2021)
16. Ohno-Matsui, K., Wu, P.C., Yamashiro, K., Vutipongsatorn, K., Fang, Y., Cheung, C.M.G., Lai, T.Y., Ikuno, Y., Cohen, S.Y., Gaudric, A., et al.: Imi pathologic myopia. Investigative ophthalmology & visual science **62**(5), 5–5 (2021)
17. Panwar, N., Huang, P., Lee, J., Keane, P.A., Chuan, T.S., Richhariya, A., Teoh, S., Lim, T.H., Agrawal, R.: Fundus photography in the 21st century—a review of recent technological advances and their implications for worldwide healthcare. Telemedicine and e-Health **22**(3), 198–208 (2016)
18. Park, K., Kim, J., Lee, J.: A deep learning approach to predict visual field using optical coherence tomography. PLOS ONE **15**(7), 1–19 (2020)
19. Phu, J., Khuu, S.K., Yapp, M., Assaad, N., Hennessy, M.P., Kalloniatis, M.: The value of visual field testing in the era of advanced imaging: Clinical and psychophysical perspectives. Clinical and Experimental Optometry **100**(4), 313–332 (2017)
20. Roy, A.G., Conjeti, S., Karri, S.P.K., Sheet, D., Katouzian, A., Wachinger, C., Navab, N.: Relaynet: Retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. Biomedical Optics Express **8**(8), 3627–3642 (2017)
21. Silva, R.: Myopic maculopathy: A review. Ophthalmologica **228**(4), 197–213 (2012)
22. Vivien: Learning through auxiliary tasks. `https://vivien000.github.io/blog/journal/learning-though-auxiliary_tasks.html`
23. Wang, S., Yan, Z., Zhang, D., Wei, H., Li, Z., Li, R.: Prototype knowledge distillation for medical segmentation with missing modality. In: ICASSP (2023)
24. Wang, S., Zhang, D., Yan, Z., Zhang, J., Li, R.: Feature alignment and uniformity for test time adaptation. In: CVPR (2023)
25. Wong, T.Y., Ferreira, A., Hughes, R., Carter, G., Mitchell, P.: Epidemiology and disease burden of pathologic myopia and myopic choroidal neovascularization: An evidence-based systematic review. American Journal of Ophthalmology **157**(1), 9–25.e12 (2014)
26. Xiao, O., Guo, X., Wang, D., Jong, M., Lee, P.Y., Chen, L., Morgan, I.G., Sankaridurg, P., He, M.: Distribution and severity of myopic maculopathy among highly myopic eyes. Investigative ophthalmology & visual science **59**(12), 4880–4885 (2018)

27. Xie, S., Kamoi, K., Igarashi-Yokoi, T., Uramoto, K., Takahashi, H., Nakao, N., Ohno-Matsui, K.: Structural abnormalities in the papillary and peripapillary areas and corresponding visual field defects in eyes with pathologic myopia. Investigative Ophthalmology & Visual Science **63**(4), 13–13 (2022)
28. Xu, L., Asaoka, R., Kiwaki, T., Murata, H., Fujino, Y., Matsuura, M., Hashimoto, Y., Asano, S., Miki, A., Mori, K., et al.: Predicting the glaucomatous central 10-degree visual field from optical coherence tomography using deep learning and tensor regression. American Journal of Ophthalmology **218**, 304–313 (2020)
29. Xu, L., Asaoka, R., Kiwaki, T., Murata, H., Fujino, Y., Yamanishi, K.: Pami: A computational module for joint estimation and progression prediction of glaucoma. In: KDD. pp. 3826–3834 (2021)
30. Xu, L., Asaoka, R., Murata, H., Kiwaki, T., Zheng, Y., Matsuura, M., Fujino, Y., Tanito, M., Mori, K., Ikeda, Y., et al.: Improving visual field trend analysis with oct and deeply regularized latent-space linear regression. Ophthalmology Glaucoma **4**(1), 78–88 (2021)
31. Zhang, S., Yang, L., Mi, M.B., Zheng, X., Yao, A.: Improving deep regression with ordinal entropy. In: ICLR (2023)
32. Zhao, X., Ding, X., Lyu, C., Li, S., Liu, B., Li, T., Sun, L., Zhang, A., Lu, J., Liang, X., et al.: Morphological characteristics and visual acuity of highly myopic eyes with different severities of myopic maculopathy. Retina **40**(3), 461–467 (2020)
33. Zheng, Y., Xu, L., Kiwaki, T., Wang, J., Murata, H., Asaoka, R., Yamanishi, K.: Glaucoma progression prediction using retinal thickness via latent space linear regression. In: KDD. pp. 2278–2286 (2019)