

Supplementary for A New SLO Database and TrustDetector for Retinal Diseases

Yichen Hu¹, Chao Wang^{*2}, Weitao Song², Aleksei Tiulpin³, and Qing Liu^{*1,4}

¹ School of Computer Science, Central South University, China

² Xiangya Hospital of Central South University, China

³ Research Unit of Health Sciences and Technology, University of Oulu, Finland

⁴ Center for Machine Vision and Signal Analysis, University of Oulu, Finland

yichenhu@csu.edu.cn wangchao.csu@163.com

wsong1980@126.com {aleksei.tiulpin, qing.liu}@oulu.fi

Table 1: Disease class distribution of our Retina-SLO with regard to eyes and SLO images, respectively.

diseases	labels	images				eyes			
		train	validation	test	total	train	validation	test	total
ME	ME	257	40	38	335	124	23	20	167
	non-ME	5872	900	834	7606	2929	516	488	3933
	unclear	0	2	0	2	0	2	0	2
DR	DR	599	88	82	769	306	55	50	411
	non-DR	5530	847	790	7167	2747	481	458	3686
	unclear	0	7	0	7	0	5	0	5
glaucoma	glaucoma	302	41	42	385	160	28	24	212
	suspicious	165	24	25	214	104	16	16	136
	non-glaucoma	5662	722	715	7099	2789	414	415	3618
	unclear	0	155	90	245	0	83	53	136

Table 2: Multi-disease detection performance comparisons in terms of $F1/macro-F1$ and $Kappa$ on our Retina-SLO test set. 5-trials are conducted and the averages and standard deviations are reported.

Methods	$F1/macro-F1$				$Kappa$			
	ME	DR	glaucoma	mF1	ME	DR	glaucoma	mKappa
ResNet-50 [1]	43.58 \pm 3.63	69.94 \pm 1.45	54.90 \pm 2.39	56.14 \pm 1.26	40.78 \pm 3.75	66.72 \pm 1.70	40.98 \pm 4.56	49.49 \pm 2.02
SENet	46.97 \pm 3.99	68.11 \pm 3.65	55.81 \pm 1.29	56.96 \pm 2.61	44.31 \pm 4.14	64.70 \pm 4.18	42.44 \pm 2.28	50.48 \pm 2.67
Swin [3]	49.48 \pm 1.40	64.66 \pm 2.74	58.50 \pm 1.52	57.55 \pm 1.29	46.64 \pm 1.48	60.90 \pm 2.99	40.86 \pm 2.21	49.47 \pm 1.54
ConvNeXt V1 [4]	44.15 \pm 2.92	68.58 \pm 1.27	56.70 \pm 2.55	56.47 \pm 1.26	41.31 \pm 2.94	65.25 \pm 1.41	46.12 \pm 3.73	50.89 \pm 1.96
ConvNeXt V2 [5]	36.68 \pm 4.41	67.19 \pm 3.35	52.25 \pm 2.62	52.04 \pm 1.99	33.96 \pm 4.43	63.92 \pm 7.12	43.45 \pm 3.06	47.11 \pm 2.41
TrustDetector (ours)	48.82\pm1.07	70.10\pm1.36	57.96\pm2.07	58.96\pm0.89	46.26\pm1.17	66.93\pm1.50	47.81\pm1.52	53.67\pm0.88

* Co-corresponding authors

Table 3: Multi-disease detection performance comparisons in terms of *Acc* on our Retina-SLO test set. 5-trials are conducted and the averages and standard deviations are reported.

Methods	<i>Acc</i>			
	ME	DR	glacuoma	mAcc
ResNet-50 [1]	94.59 \pm 0.42	94.27 \pm 0.35	89.00 \pm 1.22	92.62 \pm 0.49
SENet [2]	94.82 \pm 0.33	93.90 \pm 1.02	89.82 \pm 1.21	92.85 \pm 0.50
Swin [3]	94.31 \pm 0.25	93.19 \pm 0.47	87.57 \pm 0.72	91.69 \pm 0.38
ConvNeXt V1 [4]	94.52 \pm 0.35	93.97 \pm 0.40	91.05 \pm 0.86	93.18 \pm 0.41
ConvNeXt V2 [5]	94.75 \pm 0.26	94.03 \pm 0.69	90.28 \pm 0.80	93.02 \pm 0.13
TrustDetector (ours)	95.09 \pm 0.32	94.27 \pm 0.29	90.89 \pm 0.52	93.42 \pm 0.30

Table 4: Multi-disease detection performance comparisons in terms of *Con* and *AccCon* on our Retina-SLO test set. 5-trials are conducted and the averages and standard deviations are reported.

Methods	<i>Con</i>				<i>AccCon</i>			
	ME	DR	glacuoma	mCon	ME	DR	glacuoma	mAccCon
ResNet-50 [1]	95.94 \pm 0.50	94.56 \pm 2.02	90.79 \pm 1.82	93.77 \pm 1.00	89.95 \pm 1.00	91.61 \pm 1.51	82.75 \pm 1.78	88.11 \pm 0.96
SENet [2]	96.77 \pm 0.98	94.28 \pm 1.74	91.32 \pm 2.29	94.13 \pm 1.02	91.61 \pm 0.76	90.88 \pm 1.91	84.66 \pm 3.39	89.05 \pm 1.33
Swin [3]	96.13 \pm 0.25	94.38 \pm 0.68	91.85 \pm 1.09	94.12 \pm 0.46	90.87 \pm 0.38	90.41 \pm 0.60	82.01 \pm 0.84	87.76 \pm 0.45
ConvNeXt V1 [4]	96.77 \pm 0.93	94.75 \pm 0.70	91.53 \pm 1.40	94.35 \pm 0.57	91.43 \pm 0.96	91.71 \pm 0.57	85.93 \pm 1.85	89.69 \pm 0.81
ConvNeXt V2 [5]	97.14 \pm 0.60	95.67 \pm 1.15	92.70 \pm 2.03	95.17 \pm 0.90	91.71 \pm 0.57	91.70 \pm 1.35	85.61 \pm 1.96	89.67 \pm 0.58
TrustDetector (ours)	97.14 \pm 0.83	95.48 \pm 0.76	92.59 \pm 0.65	95.07 \pm 0.47	92.35 \pm 0.62	92.63 \pm 0.73	86.77 \pm 0.75	90.58 \pm 0.47

References

1. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 770–778 (2016)
2. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 7132–7141 (2018)
3. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., Guo, B.: Swin transformer: Hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 10012–10022 (2021)
4. Liu, Z., Mao, H., Wu, C.Y., Feichtenhofer, C., Darrell, T., Xie, S.: A convnet for the 2020s. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 11976–11986 (2022)
5. Woo, S., Debnath, S., Hu, R., Chen, X., Liu, Z., Kweon, I.S., Xie, S.: Convnext v2: Co-designing and scaling convnets with masked autoencoders. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 16133–16142 (2023)