

Deep learning for automatic diabetic retinopathy grading of ultra-widefield fundus images

Mhd Hasan Sarhan, M.Sc^{1,2}; Patricia Sha, OD^{3,4}; Mike Chen, OD^{3,4}; Mary K. Durbin, PhD³; Mehmet Yigitsoy, PhD¹; Abouzar Eslami, PhD¹

¹Carl Zeiss Meditec AG, Munich; ²Technical University of Munich, Munich; ³Carl Zeiss Meditec, Inc., Dublin, CA; ⁴Silicon Valley Eyecare, Santa Clara, CA

Poster # 1094 - A0052

PURPOSE

Ultra-widefield (UWF) fundus photography provides a powerful tool for monitoring the periphery of the retina compared to narrow field (NF) images. We study the effectiveness of data-driven deep learning techniques for automatic diabetic retinopathy (DR) grading in ultra-widefield fundus images using a **small dataset of annotated UWF images**.

METHODS

- 154 UWF fundus images were collected using CLARUS™ 500 (ZEISS, Dublin, CA). All images were graded for DR by an expert for ICDR disease severity scale. The data consisted of a mix of healthy and different DR severity levels images captured from 119 subjects.
- An automatic deep DR grading model using convolutional neural networks (CNN) [1] is trained with NF fundus images (35k images, 45° FOV). The model is then fine-tuned with UWF CLARUS images.
- The model is evaluated using 5-fold stratified cross validation. Each UWF image is divided into 4 non-overlapping tiles and the results are combined using the mean prediction value.
- We compare utilizing CLARUS data for fine-tuning the model with using the pre-trained model directly for evaluation without fine-tuning.

CONCLUSION

The scarcity of annotated UWF fundus images **can be addressed with transfer training** given the large number of available annotated NF fundus images and fine tuning with a small amount of annotated UWF images. This approach shows good generalization of an existing trained model without the need for large-scale data collection.

Email: mhd-hasan.sarhan@zeiss.com

Disclosures: MHS (E), MY (E), and AE (E): Carl Zeiss Meditec AG, Munich; PS (E), MC (E), MD (E): Carl Zeiss Meditec, Inc., Dublin

[1] M.H. Sarhan, S. Albarqouni, M. Yigitsoy, N. Navab, A. Eslami: Multi-scale Microaneurysms Segmentation Using Embedding Triplet Loss arXiv preprint (2019)

RESULTS

The results are shown in Table 1. The model reached an **accuracy of 84%, 97% and 94%** for DR grading, referable DR, and DR screening respectively. This is an improvement over using only NF images in all considered metrics and it is achievable using a small amount of UFW images.

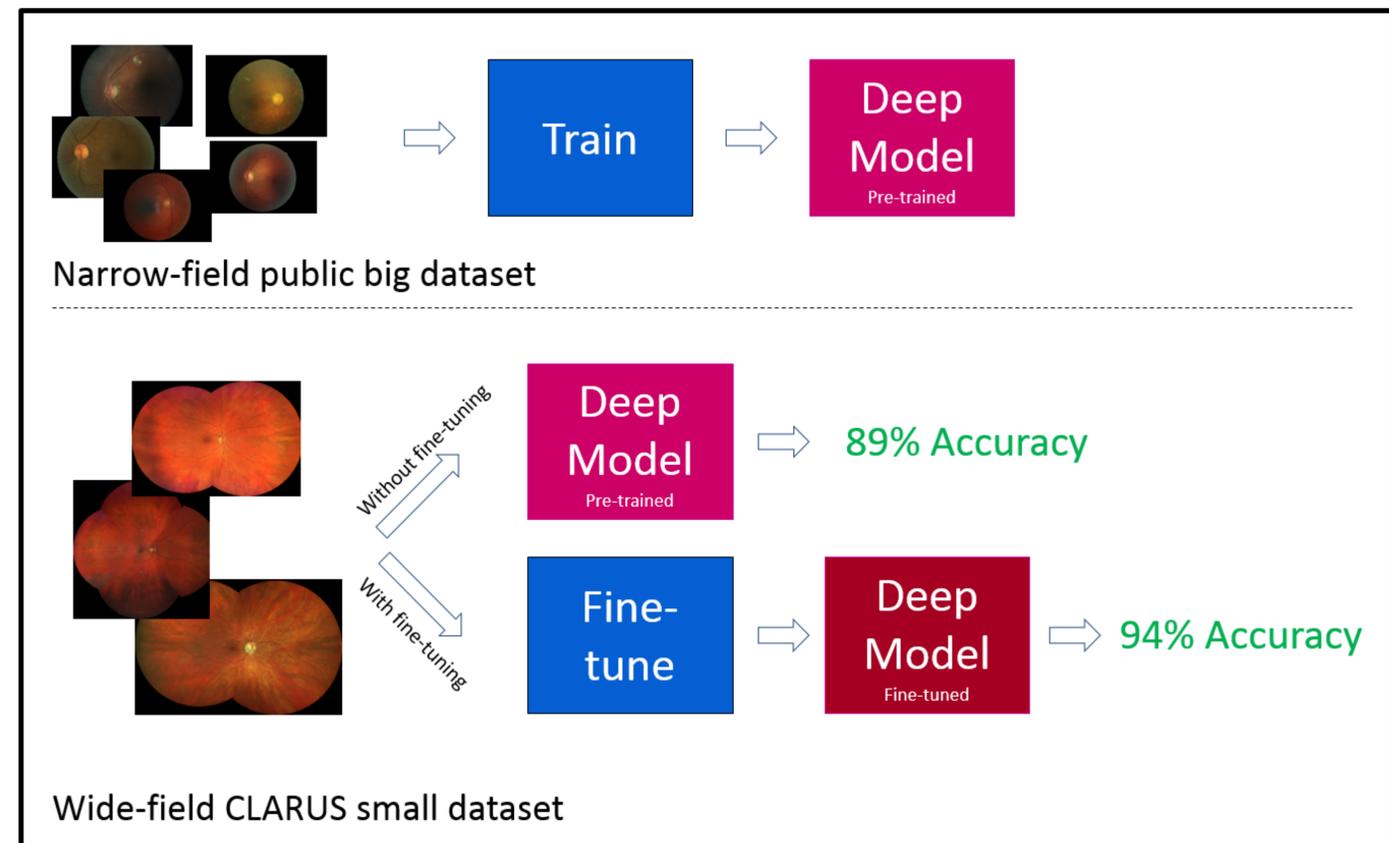


Figure 1. Schematic view of the fine-tuning approach

	With fine-tuning	κ -Score	Accuracy	Sensitivity	Specificity	AUC
5-levels Grading	✓	0.93 ± 0.04	0.84 ± 0.05	-	-	-
	-	0.90 ± 0.03	0.78 ± 0.05	-	-	-
Referable DR	✓	0.94 ± 0.02	0.97 ± 0.01	1.0 ± 0.0	0.94 ± 0.03	0.99 ± 0.01
	-	0.90 ± 0.03	0.96 ± 0.01	0.98 ± 0.03	0.94 ± 0.03	0.98 ± 0.01
DR Screening	✓	0.88 ± 0.06	0.94 ± 0.03	1.0 ± 0.0	0.98 ± 0.02	0.998 ± 0.002
	-	0.79 ± 0.09	0.89 ± 0.05	0.98 ± 0.03	0.98 ± 0.02	0.994 ± 0.006

Table 1. Comparison between using CLARUS data for fine-tuning and using CLARUS data directly for evaluation on the pre-trained model