

In the format provided by the authors and unedited.

# Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning

**Ryan Poplin<sup>1,4</sup>, Avinash V. Varadarajan<sup>1,4</sup>, Katy Blumer<sup>1</sup>, Yun Liu<sup>1</sup>, Michael V. McConnell<sup>2,3</sup>,  
Greg S. Corrado<sup>1</sup>, Lily Peng<sup>1,4\*</sup> and Dale R. Webster<sup>1,4</sup>**


---

<sup>1</sup>Google Research, Google, Mountain View, CA, USA. <sup>2</sup>Verily Life Sciences, South San Francisco, CA, USA. <sup>3</sup>Division of Cardiovascular Medicine, Stanford School of Medicine, Stanford, CA, USA. <sup>4</sup>These authors contributed equally: Ryan Poplin, Avinash V. Varadarajan, Lily Peng and Dale R. Webster.  
\*e-mail: [lhpeng@google.com](mailto:lhpeng@google.com)

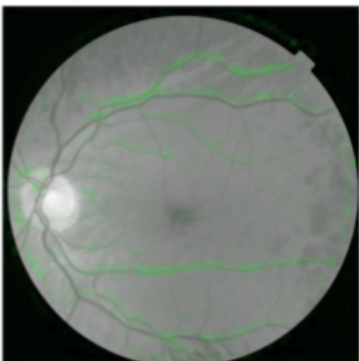
## Supplementary information

**A**

Original image



Masked image



**B**

This mask seems to be highlighting (check all that apply): \*

- ☐ Veins
- ☐ Venules
- ☐ Arteries
- ☐ Arterioles
- ☐ Vessel surroundings
- ☐ AV nicks
- ☐ Optic disc
- ☐ Optic disc edges
- ☐ Drusen
- ☐ Retinal pigment
- ☐ Image edges
- ☐ Nothing in particular
- ☐ Other: \_\_\_\_\_

(Optional) Comments

Your answer \_\_\_\_\_

**Fig. 1** | Screenshots of the image and mask shown to ophthalmologists (a) and the accompanying form (b) that ophthalmologists filled out to determine the features that were highlighted by the attention heat maps for various predictions.

Layer	Operation	Output size
Input	-	587x587x3
Conv/pool 1	3x3 conv (64 filters), 2x2 maxpool	293x293x64
Conv/pool 2	3x3 conv (128 filters), 2x2 maxpool	146x146x128
Conv/pool 3	3x3 conv (256 filters), 2x2 maxpool	73x73x256

Conv 4	2x2 conv (512 filters)	73x73x512
Alpha layer (saliency map)	-	73x73x1
Reverse maxpool 1	2x2 reverse maxpool	146x146x1
Reverse maxpool 2	2x2 reverse maxpool with padding	293x293x1
Reverse maxpool 3	2x2 reverse maxpool with padding	587x587x1

**Table 1** | Architectural details of the soft-attention neural network.