

# Approximation of a pipeline of unsupervised retina image analysis methods with a CNN

**Citation for published version (APA):**

Heslinga, F., Pluim, J., Dasht Bozorg, B., Berendschot, T., Houben, A. J. H. M., Henry, R. M. A., & Veta, M. (2019). Approximation of a pipeline of unsupervised retina image analysis methods with a CNN. In E. D. Angelini, & B. A. Landman (Eds.), *Image Processing: SPIE Medical Imaging, 2019, San Diego, California, United States* [109491N] (Proceedings of SPIE; Vol. 10949). SPIE. <https://doi.org/10.1117/12.2512393>

**DOI:**

[10.1117/12.2512393](https://doi.org/10.1117/12.2512393)

**Document status and date:**

Published: 01/03/2019

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

# PROCEEDINGS OF SPIE

[SPIEDigitalLibrary.org/conference-proceedings-of-spie](https://SPIEDigitalLibrary.org/conference-proceedings-of-spie)

## Approximation of a pipeline of unsupervised retina image analysis methods with a CNN

Friso G. Heslinga, Josien P. W. Pluim, Behdad Dashtbozorg, Tos T. J. M. Berendschot, A. J. H. M. Houben, et al.

Friso G. Heslinga, Josien P. W. Pluim, Behdad Dashtbozorg, Tos T. J. M. Berendschot, A. J. H. M. Houben, Ronald M. A. Henry M.D., Mitko Veta, "Approximation of a pipeline of unsupervised retina image analysis methods with a CNN," Proc. SPIE 10949, Medical Imaging 2019: Image Processing, 109491N (15 March 2019); doi: 10.1117/12.2512393

**SPIE.**

Event: SPIE Medical Imaging, 2019, San Diego, California, United States

# Approximation of a Pipeline of Unsupervised Retina Image Analysis Methods with a CNN

Friso G. Heslinga<sup>a</sup>, Josien P.W. Pluim<sup>a</sup>, Behdad Dashtbozorg<sup>a</sup>, Tos T.J.M. Berendschot<sup>a,b</sup>,  
A.J.H.M. Houben<sup>c</sup>, Ronald M.A. Henry<sup>c</sup>, and Mitko Veta<sup>a</sup>

<sup>a</sup>Eindhoven University of Technology, Department of Biomedical Engineering, Eindhoven, the Netherlands

<sup>b</sup>Maastricht University Medical Center+, University Eye Clinic Maastricht, the Netherlands

<sup>c</sup>Maastricht University Medical Center+, Department of Internal Medicine and CARIM, the Netherlands

## ABSTRACT

A pipeline of unsupervised image analysis methods for extraction of geometrical features from retinal fundus images has previously been developed. Features related to vessel caliber, tortuosity and bifurcations, have been identified as potential biomarkers for a variety of diseases, including diabetes and Alzheimer's. The current computationally expensive pipeline takes 24 minutes to process a single image, which impedes implementation in a screening setting. In this work, we approximate the pipeline with a convolutional neural network (CNN) that enables processing of a single image in a few seconds. As an additional benefit, the trained CNN is sensitive to key structures in the retina and can be used as a pretrained network for related disease classification tasks. Our model is based on the ResNet-50 architecture and outputs four biomarkers that describe global properties of the vascular tree in retinal fundus images. Intraclass correlation coefficients between the predictions of the CNN and the results of the pipeline showed strong agreement (0.86 - 0.91) for three of four biomarkers and moderate agreement (0.42) for one biomarker. Class activation maps were created to illustrate the attention of the network. The maps show qualitatively that the activations of the network overlap with the biomarkers of interest, and that the network is able to distinguish venules from arterioles. Moreover, local high and low tortuous regions are clearly identified, confirming that a CNN is sensitive to key structures in the retina.

**Keywords:** Deep Learning, Retinal Image Analysis, Retinal Biomarkers, Distillation, Maastricht Study

## 1. INTRODUCTION

Retinal fundus imaging enables detailed visualization of the microvascular structure in the retina of the human eye. Geometrical properties of the retinal vessels have been identified as potential biomarkers for a variety of diseases, (e.g. systemic diseases such as diabetes,<sup>1</sup> eye diseases such as diabetic retinopathy<sup>2</sup> and neurological diseases such as Alzheimer's<sup>3</sup>). Early detection could benefit the disease management for some of these patients.

Unsupervised image analysis methods for extraction of biomarkers predictive of the diseases mentioned above have been developed, related to vessel caliber,<sup>4</sup> tortuosity<sup>5</sup> and bifurcation features.<sup>6</sup> A pipeline that combines the validated algorithms has been assembled for the RetinaCheck project, a large-scale diabetes screening program in China.<sup>7</sup> The current pipeline is computationally expensive (taking 24 minutes to process an image) and could benefit substantially from advances in machine learning techniques to speed up the biomarker extraction.

This paper is a proof of principle for the approximation of part of the pipeline of retinal image analysis methods into a single convolutional neural network (CNN). The network is used to reproduce multiple biomarkers using a straightforward hard parameter sharing multi-target learning (MTL) approach. The network is fast and can be easily adapted to include different biomarkers.

---

Send correspondence to F.G. Heslinga. E-mail: f.g.heslinga@tue.nl, Telephone: +31 (0)40 247 55 37

## 1.1 Related work

The approximation concept described in this work is closely related to distillation of knowledge in neural networks as described by Hinton et al.<sup>8</sup> The authors show empirically that an ensemble of networks can be distilled into a single neural network. More generally, neural networks have mathematically been described as universal approximators.<sup>9</sup> We apply this concept for an ensemble of geometrical models that were used to calculate geometrical properties of the vascular tree of the retina. We use the output of the pipeline (which are global values for each image) as input for a CNN to approximate.

CNNs have been applied to color fundus images by Poplin et al.<sup>10</sup> to predict cardiovascular risk factors. In their experiments, multiple targets, including age, BMI and systolic blood pressure, were predicted with a single CNN. In the same work, cardiovascular events were predicted directly from the fundus images. To improve the event prediction, Poplin et al. used a MTL approach where additional variables were predicted simultaneously. In our work we also use a MTL strategy, hypothesizing that different retinal biomarkers are related.

## 2. METHODS

The color fundus images used for this research originate from the Maastricht Study\*, an observational prospective population-based cohort study. The rationale and methodology have been described elsewhere.<sup>11</sup> In brief, the study focuses on the cause, pathophysiology, complications, and comorbidities of type 2 diabetes mellitus (T2D) and is characterized by an extensive phenotyping approach. Eligible for participation were all individuals aged between 40 and 75 years and living in the southern part of the Netherlands. Participants were recruited through mass media campaigns and from the municipal registries and the regional Diabetes Patient Registry via mailings. Recruitment was stratified according to known T2D status, with an oversampling of individuals with T2D, for reasons of statistical efficiency. All presented experiments were conducted using 10,668 images from 2,782 subjects from the Maastricht study. The prevalences of T2D and impaired glucose tolerance among these subjects were 29.0% and 10.9%, respectively. The set comprises images of left and right eyes and are centered either on the fovea or on the optic disc. The images were resized to 384 by 384 pixels and channel-wise global contrast normalization was applied before further processing.

A selection of four biomarkers was chosen from an extensive list of features that were automatically calculated by the pipeline of methods described by ter Haar Romeny et al.<sup>7</sup> The biomarkers were chosen such that they represent key vessel geometries: vessel caliber, tortuosity and bifurcations. Descriptions of the biomarkers are given in Tab. 1. The asymmetry ratio values for 35 images were missing and therefore replaced with the mean of the asymmetry ratio values from the training set. All labels are scaled to have zero mean and unit standard deviation, based on the statistics of the labels from the training set.

Table 1. List of included biomarkers.

Label	Biomarker description
CRAE	Diameter of the retinal arterioles (Central Retinal Arteriolar Equivalent). Measured on the arterioles coursing through a standard area 0.5 to 1.0 disc diameters from the optic disc margin. <sup>12</sup>
CRVE	Diameter of the retinal venules (Central Retinal Venular Equivalent). Measured on the venules coursing through a standard area 0.5 to 1.0 disc diameters from the optic disc margin. <sup>12</sup>
Global tortuosity	Global tortuosity based on the mean of the exponential curve fits of all retinal vessels. <sup>5</sup>
Asymmetry ratio of bifurcations	Average ratio between the cross-sectional area of the smaller and larger branch following a bifurcation. <sup>13</sup>

\*<https://www.demaastrichtstudie.nl/>

## 2.1 CNN models and experiments

The 10,668 images are split into sets for training ( $N = 6,352$  [60%]), validation ( $N = 2,130$  [20%]) and testing ( $N = 2,186$  [20%]). All images of a single patient are assigned to the same set. A neural network with a ResNet-50 architecture<sup>14</sup> is implemented in Keras<sup>15</sup> using a TensorFlow<sup>16</sup> backend. ResNet has been shown to have a high classification performance on ImageNet. It also is known for its excellent transfer learning capabilities,<sup>17</sup> which is an desirable property if this network is to be used as a pretrained network for future related tasks.

ResNet-50 contains 16 residual blocks that learn residual functions with reference to the layer inputs, instead of learning unreferenced functions. The last residual block is followed by a global average pooling layer and we replaced the output layer by a dense layer containing four nodes, corresponding to the four biomarkers. The loss function is defined as the sum of  $L_2$  losses for each biomarker and  $L_2$ -regularization is added for all trainable weights. The total number of trainable parameters is around 23M and He normal initialization<sup>18</sup> is used for setting the weights. Batches consist of randomly sampled and augmented images. Augmentation includes translation ( $\leq 38$  pixels), rotation ( $\leq 360^\circ$ ), horizontal and vertical reflection, intensity shift ( $\leq 20/256$ ), color shift ( $\leq 30/256$ ) and contrast shift ( $\leq 0.05$ ).

Weights are optimized using Adam optimization and training is done in two steps. First, the described model is trained for 16,950 iterations with a learning rate of  $1e-3$  and for 16,950 iterations with a learning rate of  $1e-4$ . Then, for visualization purposes (see below), the last two convolutional layers that downsize the spatial dimensions are adjusted so they no longer have a pooling effect (by setting the strides to 1). Second, training is continued for another 63,500 iterations, after which no improvement is seen on the validation set. For the second part of the training, the batch size is reduced to ensure that the larger activation tensors still fit in GPU memory. An overview of the learning rate schedule and batch size is shown in Tab. 2.

Table 2. Training details.

<b>iterations</b>	16950	16950	25400	25400	12700
<b>batch size</b>	28	28	10	10	10
<b>learning rate</b>	$1e-3$	$1e-4$	$1e-4$	$1e-5$	$1e-6$

For each biomarker, the performance of the model is evaluated on the test set by calculating the intraclass correlation coefficient (two-way mixed single measures)<sup>19</sup> between the labels and the predictions. At test time, the same augmentation settings are used to repeat the predictions for each image 30 times, which are then averaged for the final predictions.

To verify that the activations by the neural network indeed represent the anatomical locations of interest, we utilize Class Activation Maps (CAMs).<sup>20</sup> A CAM is constructed for each biomarker by multiplying the weights of the output layer with the activations that are inputs to the global average pooling layer. Since we removed two downscaling operations, the network now only downscales the x and y dimension 8 times, so the resulting CAM is of size 48 by 48 pixels.

After training the network, images from the test set are passed through. For each image, four CAMs are extracted and resized to the original image size. For visualization we use a 3-channel gray map of the original image and add positive activations and absolute values of the negative activations to the green channel and red channel respectively. Since we scaled the labels to have zero mean, local activations show whether a region positively or negatively contributes to the model predictions. For example, non-tortuous regions are expected to result in local negative activation, while tortuous regions are expected to give positive activations.

## 3. RESULTS

The prediction of the four biomarkers for all test images took 182 minutes on an Intel I7 CPU with a single GPU (Geforce GTX 1070), corresponding to a total analysis time of 5 seconds per image (as compared to 24 minutes for the original pipeline). The intraclass correlation coefficients (ICCs) between the model predictions and the targeted values are shown in Fig. 1. The ICCs indicate strong relations for CRAE (ICC = 0.86), CRVE (ICC

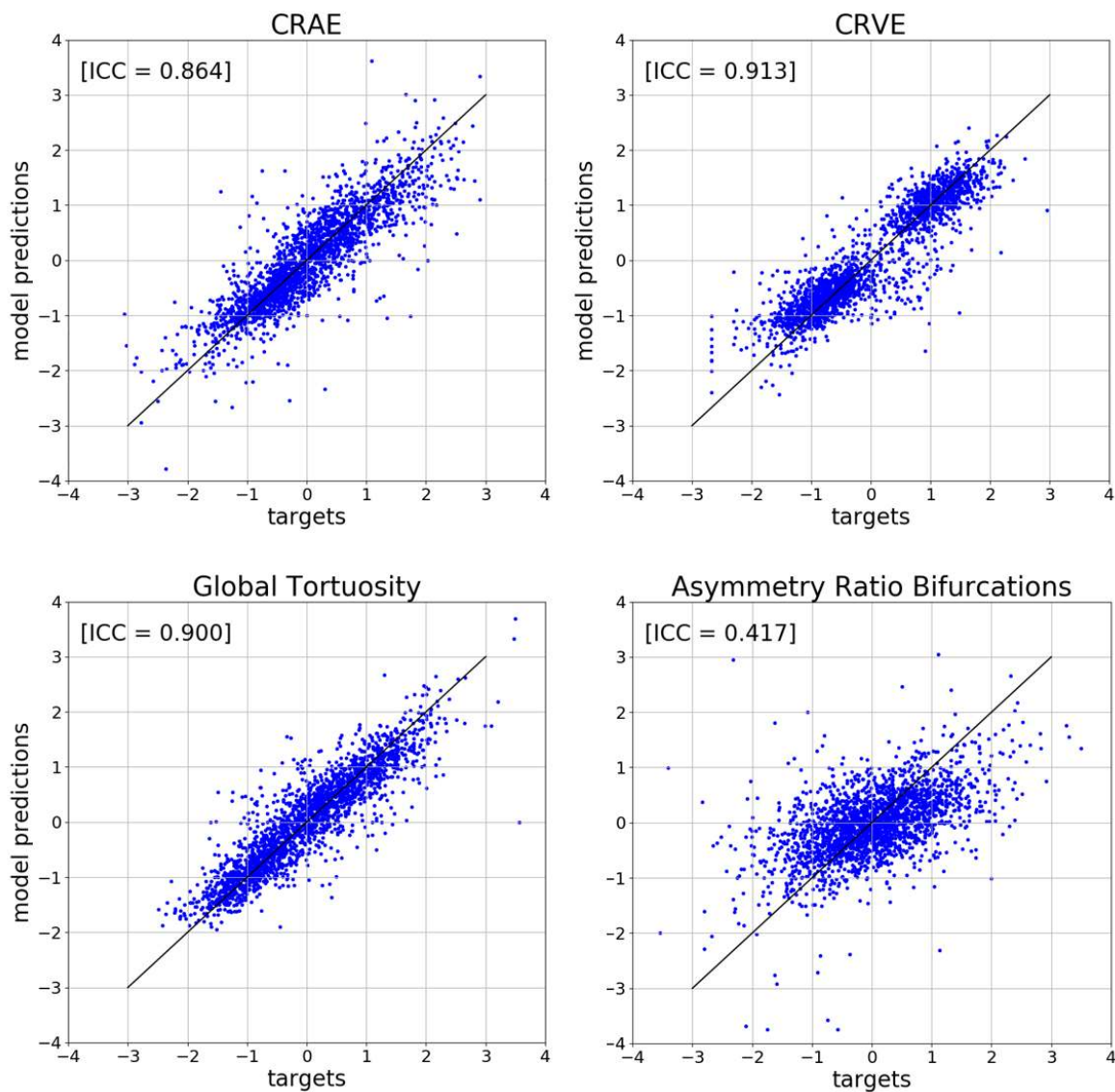


Figure 1. Comparing predicted values of the biomarkers with the labels from the original pipeline. The  $y = x$  line represents perfect agreement. The intraclass correlation coefficient is shown for each biomarker.

= 0.91) and global tortuosity (ICC = 0.90), whereas the correlation for the asymmetry ratio of bifurcations is considered moderate (ICC = 0.42). Without test-time augmentation, the analysis time was 0.07 seconds per image and the intraclass correlations decreased by 0.01-0.03 across all biomarkers.

Fig. 2 shows two examples of color fundus images with corresponding class activations maps for the four biomarkers. In both examples can be seen that the regions that contribute to the final predictions (either positively or negatively) overlap strongly with the vascular tree. For Example A, the CAM showing CRVE highlights the dark colored venules close to the optic disc in green, indicating that many regions positively contribute to the CRVE prediction. The normalized CRVE label of this example is indeed relatively large, meaning that the venules around the optic disc are relatively wide. The CAM showing tortuosity for example A contains both green and red regions. Positive activations strongly overlap with curved parts of the vessels, while negative activations overlap with straight parts of vessels. The normalized global tortuosity label for this example is close to 0.

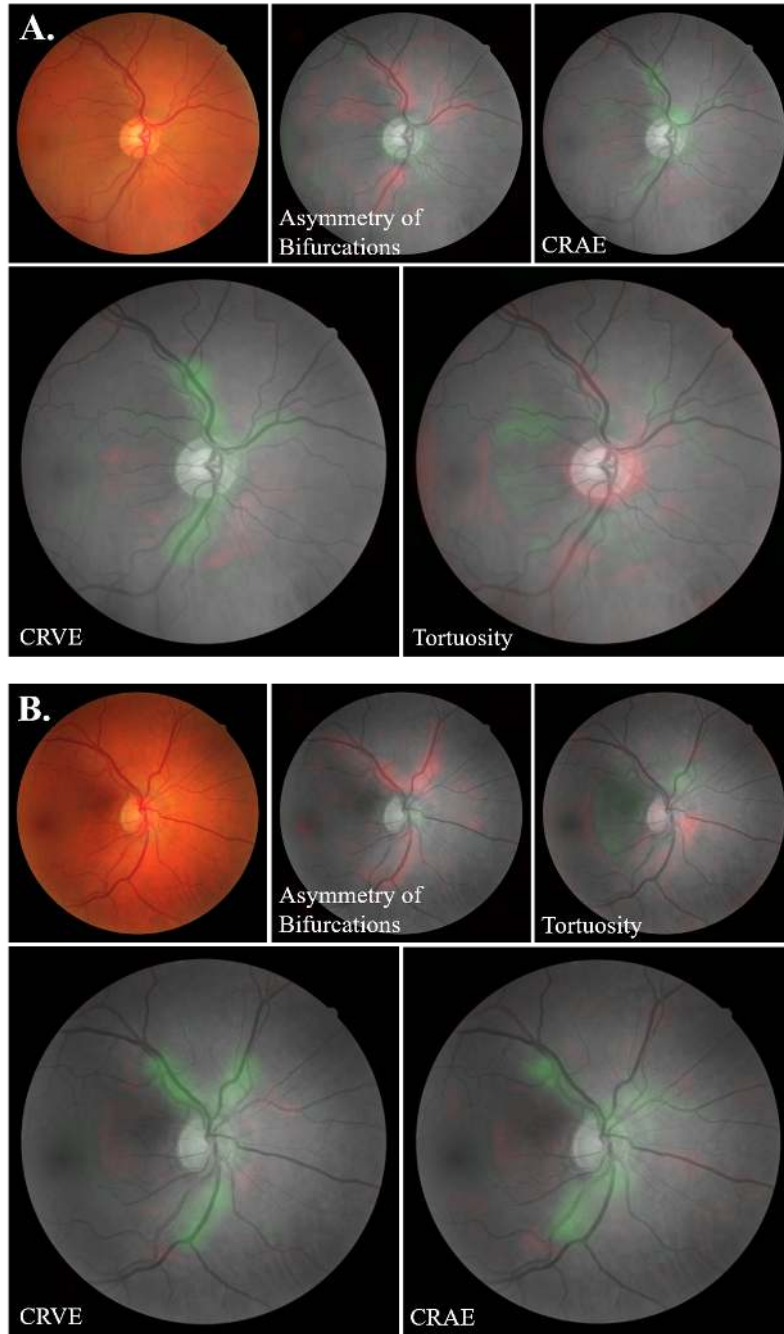


Figure 2. Two examples (A and B) are shown with CAMs corresponding to the four biomarkers. Green indicates a region that positively contributes to the final predictions of the model, while red regions correspond to negative contributions to the final predictions. Negative intensities were multiplied by 3, since we found it to improve visibility

Example B shows that the model indeed seems to differentiate between arterioles and venules. In the bottom left CAM (showing CRVE) the dark colored venules are highlighted, while in the bottom right CAM (showing CRAE) the more light colored arterioles are highlighted. The normalized CRAE and CRVE labels are strongly

positive for this example, indicating relatively large vessels. Interestingly, these examples also clearly show that the model predictions are mainly based on the regions close to the optic disc. This corresponds with the measures of the original pipeline, where the CRAE and CRVE are measured in the area between 0.5 and 1.0 disc diameters from the optic disc margin.

In both examples, the CAMs that show the asymmetry of bifurcations do overlap with the vascular tree, but do not seem to be very specific to regions with bifurcations.

#### 4. DISCUSSION AND CONCLUSION

Key biomarkers, originally extracted with a pipeline of unsupervised geometrical models, can be reproduced with the predictions of a single CNN. The network is fast and can easily be expanded to include other biomarkers.

Visual verification, using class activation maps for the four biomarkers, suggests that the attention of the network overlaps with geometrical structures that the biomarkers represent and thus not only identify confounding features. Not all CAMs that we produced were as clear as the ones presented in this paper and the model perhaps uses related features to improve the predictions. For example, we noticed that the biomarkers are mutually correlated. This insight is likely to be picked up in a multi-target learning approach as presented here.

The relatively low correlation for the ratio of bifurcations indicates that it is difficult to precisely approximate bifurcation information with the current experiments. This could be because of the loss of information due to downsizing of the fundus images. For future research, the original image size (3744 by 3744) or intermediate dimensions could be used, although memory requirements should be kept in mind. Additional validation can be done by testing on an external dataset.

The chosen biomarkers represent the main retinal microvascular structures. The distilled model is thus sensitive to features of the retina that have been related to several diseases and can be considered pretrained for related classification tasks. A potential application of the pretrained network could be to improve the classification of diabetes Type 2 and prediabetes directly from fundus images,<sup>21</sup> since both diseases are associated with microvascular dysfunction.<sup>22</sup> Also, for a similar objective, retinal biomarkers could be used as additional targets in a multi-target approach.

#### ACKNOWLEDGMENTS

This research is financially supported by the TTW Perspectief program and Philips Research. The authors have no conflicts of interests to report. This work has not been submitted for publication anywhere else.

The color fundus images used in the research originate from the Maastricht Study. The Maastricht Study was supported by the European Regional Development Fund via OP-Zuid, the Province of Limburg, the Dutch Ministry of Economic Affairs (grant 31O.041), Stichting De Weijerhorst (Maastricht, The Netherlands), the Pearl String Initiative Diabetes (Amsterdam, The Netherlands), CARIM School for Cardiovascular Diseases (Maastricht, The Netherlands), Stichting Annadal (Maastricht, The Netherlands), Health Foundation Limburg (Maastricht, The Netherlands) and by unrestricted grants from Janssen-Cilag B.V. (Tilburg, The Netherlands), Novo Nordisk Farma B.V. (Alphen aan den Rijn, The Netherlands), and Sanofi-Aventis Netherlands B.V. (Gouda, The Netherlands).

#### REFERENCES

- [1] Zhang, J., Dashtbozorg, B., Huang, F., Berendschot, T., and ter Haar Romeny, B., "Analysis of retinal vascular biomarkers for early detection of diabetes," *Lecture Notes in Computational Vision and Biomechanics* **27**, 811–7 (2017).
- [2] Habib, M., Al-Diri, B., Hunter, A., and Steel, D., "The association between retinal vascular geometry changes and diabetic retinopathy and their role in prediction of progression - an exploratory study," *BMC Ophthalmology* **14**(1), 1–11 (2014).



- [3] Frost, S., Kanagasalingam, Y., Sohrabi, H., Vignarajan, J., Bourgeat, P., Salvado, O., Villemagne, V., Rowe, C., Macaulay, S. L., Szoek, C., Ellis, K., Ames, D., Masters, C., S.Rainey-Smith, Martins, R., and the AIBL Research Group, "Retinal vascular biomarkers for early detection and monitoring of alzheimers disease," *Translational Psychiatry* **3**(2), e233–8 (2013).
- [4] Dashtbozorg, B., Mendonca, A., Penas, S., and Camphillo, A., "Computer-aided diagnosis system for the assessment of retinal vascular changes," *Proceedings OMIA 2014*, 6328–31.
- [5] Bekkers, E., Zhang, J., Duits, R., and ter Haar Romeny, B., "Curvature based biomarkers for diabetic retinopathy via exponential curve fits in SE(2)," *Proceedings OMIA 2015* **27**, 113–20.
- [6] Abbasi-Sureshjani, S., Smit-Ockeloen, I., Bekkers, E., Dashtbozorg, B., and ter Haar Romeny, B., "Automatic detection of vascular bifurcations and crossings in retinal images using orientation scores," *IEEE ISBI* **13**, 189–92 (2016).
- [7] ter Haar Romeny, B., Bekkers, E., Zhang, J., Abbasi-Sureshjani, S., Huang, F., Duits, R., Dashtbozorg, B., Berendschot, T., Smit-Ockeloen, I., Eppenhof, K., Feng, J., Hannink, J., Schouten, J., Tong, M., Wul, H., van Triest, H., Zhu, S., Chen, D., He, W., Xu, L., Hand, P., and Kang, Y., "Brain-inspired algorithms for retinal image analysis. machine vision and applications," *Machine Vision and Applications* **27**(8), 1117–1135 (2016).
- [8] Hinton, G., Vinyals, O., and Dean, J., "Distilling the knowledge in a neural network," *arXiv:1503.02531*, 1–9 (2015).
- [9] Hornik, K., Stinchcombe, M., and White, H., "Multilayer feedforward networks are universal approximators," *Neural Networks* **2**(5), 359 – 366 (1989).
- [10] Poplin, R., Varadarajan, A., Blumer, K., Liu, Y., McConnell, M., Corrado, G., Peng, L., and Webster, D., "Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning," *Nature Biomedical Engineering* **2**, 158–64 (2018).
- [11] Schram, M., Sep, S., van der Kallen, C., Dagnelie, P., Koster, A., Schaper, N., Henry, R., and Stehouwer, C., "The maastricht study: An extensive phenotyping study on determinants of type 2 diabetes, its complications and its comorbidities," *European Journal of Epidemiology* **29**(6), 439–451 (2014).
- [12] Huang, F., Dashtbozorg, B., Zhang, J., Yeung, A., Berendschot, T. T., and ter Haar Romeny, B. M., "Validation study on retinal vessel caliber measurement technique," *Lecture Notes in Computational Vision and Biomechanics* **27**, 818–826 (2017).
- [13] Al-Diri, B., Hunter, A., Steel, D., and Habib, M., "Manual measurement of retinal bifurcation features," *Annual International Conference of the IEEE Engineering in Medicine and Biology* (2010).
- [14] He, K., Zhang, X., Ren, S., and Sun, J., "Deep residual learning for image recognition," *arXiv:1512.03385v1*, 1–12 (2015).
- [15] "Keras," (2015). Software available from keras.io.
- [16] "TensorFlow: Large-scale machine learning on heterogeneous systems," (2015). Software available from tensorflow.org.
- [17] Kornblith, S., Shlens, J., and Le, Q. V., "Do better ImageNet models transfer better?," *CoRR* **abs/1805.08974** (2018).
- [18] He, K., Zhang, X., Ren, S., and Sun, J., "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," *CoRR* **abs/1502.01852** (2015).
- [19] McGraw, K. and Wong, S., "Forming inferences about some intraclass correlation coefficients," *Psychological Methods* **1**(1), 30–46 (1996).
- [20] Zhou, B., Khosla, A., Lapedriza, À., Oliva, A., and Torralba, A., "Learning deep features for discriminative localization," *CoRR* **abs/1512.04150** (2015).
- [21] Abbasi-Sureshjani, S., Dashtbozorg, B., ter Haar Romeny, B., and Fleuret, F., "Exploratory study on direct prediction of diabetes using deep residual networks," *Lecture Notes in Computational Vision and Biomechanics* **27**, 797–802 (2017).
- [22] B.M. Sørensen, A. H., Berendschot, T., Schouten, J., Kroon, A., van der Kallen, C., Henry, R., Koster, A., Sep, S., Dagnelie, P., Schaper, N., Schram, M., and Stehouwer, C., "Prediabetes and type 2 diabetes are associated with generalized microvascular dysfunction: The maastricht study," *Circulation* **134**(18), 1339–52 (2016).