



HAL
open science

A New Vessel-Based Method to Estimate Automatically the Position of the Nonfunctional Fovea on Altered Retinography From Maculopathies

Aurélie Calabrèse, Vincent Fournet, Séverine Dours, Frédéric Matonti, Eric Castet, Pierre Kornprobst

► To cite this version:

Aurélie Calabrèse, Vincent Fournet, Séverine Dours, Frédéric Matonti, Eric Castet, et al.. A New Vessel-Based Method to Estimate Automatically the Position of the Nonfunctional Fovea on Altered Retinography From Maculopathies. *Translational vision science & technology*, 2023, 12, 10.1167/tvst.12.7.9 . hal-04159981

HAL Id: hal-04159981

<https://inria.hal.science/hal-04159981v1>

Submitted on 12 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Open licence - etalab

A New Vessel-Based Method to Estimate Automatically the Position of the Nonfunctional Fovea on Altered Retinography From Maculopathies

Aurélie Calabrèse^{1,2}, Vincent Fournet², Séverine Dours¹, Frédéric Matonti³⁻⁵, Eric Castet¹, and Pierre Kornprobst²

¹ Aix-Marseille Univ, CNRS, LPC, Marseille, France

² Université Côte d'Azur, Inria, France

³ Centre Monticelli Paradis d'Ophthalmologie, Marseille, France

⁴ Aix-Marseille Univ, CNRS, INT, Marseille, France

⁵ Groupe Almaviva Santé, Clinique Juge, Marseille, France

Correspondence: Aurélie Calabrèse, Aix-Marseille University, CNRS, LPC, 3 place Victor Hugo 13331 Marseille, Cedex 3, France. e-mail: aurelie.calabrese@univ-amu.fr

Received: December 19, 2022

Accepted: June 1, 2023

Published: July 7, 2023

Keywords: maculopathy; fovea detection; image processing; macular lesion; fundus image

Citation: Calabrèse A, Fournet V, Dours S, Matonti F, Castet E, Kornprobst P. A new vessel-based method to estimate automatically the position of the nonfunctional fovea on altered retinography from maculopathies. *Transl Vis Sci Technol.* 2023;12(7):9. <https://doi.org/10.1167/tvst.12.7.9>

Purpose: The purpose of this study was to validate a new automated method to locate the fovea on normal and pathological fundus images. Compared to the normative anatomic measures (NAMs), our vessel-based fovea localization (VBFL) approach relies on the retina's vessel structure to make predictions.

Methods: The spatial relationship between the fovea location and vessel characteristics is learnt from healthy fundus images and then used to predict fovea location in new images. We evaluate the VBFL method on three categories of fundus images: healthy images acquired with different head orientations and fixation locations, healthy images with simulated macular lesions, and pathological images from age-related macular degeneration (AMD).

Results: For healthy images taken with the head tilted to the side, the NAM estimation error is significantly multiplied by 4, whereas VBFL yields no significant increase, representing a 73% reduction in prediction error. With simulated lesions, VBFL performance decreases significantly as lesion size increases and remains better than NAM until lesion size reaches 200 degrees². For pathological images, average prediction error was 2.8 degrees, with 64% of the images yielding an error of 2.5 degrees or less. VBFL was not robust for images showing darker regions and/or incomplete representation of the optic disk.

Conclusions: The vascular structure provides enough information to precisely locate the fovea in fundus images in a way that is robust to head tilt, eccentric fixation location, missing vessels, and actual macular lesions.

Translational Relevance: The VBFL method should allow researchers and clinicians to assess automatically the eccentricity of a newly developed area of fixation in fundus images with macular lesions.

Introduction

Microperimeters are powerful tools that allow to take color fundus images of the retina while also mapping the visual field and measuring fixation stability. Therefore, these devices are of great interest when dealing with patients suffering from retinal defects,

and especially central field loss (CFL) induced by maculopathies, such as age-related macular degeneration (AMD). In such retinopathies, exact position of the fovea cannot be identified visually anymore as the darkest spot at the center of the macula. It also becomes impossible to locate it functionally with a fixation examination given the inability to fixate with the fovea, experienced by patients with CFL.^{1,2}

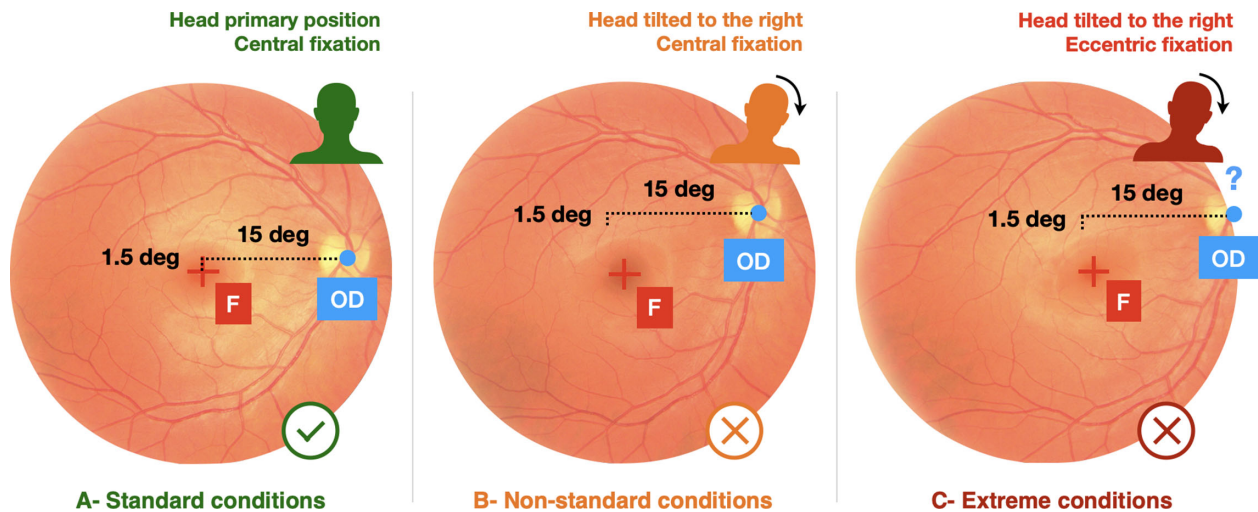


Figure 1. (A) Under “standard” conditions (with the head in primary position while fixating a central cross), the macula is located at the center of the image and normative anatomic measures (NAMs) described by Ref. 10 allow an accurate prediction of the fovea position (represented by the fixation cross). (B) In other conditions, such as a head tilt, rotation causes greater variation in the vertical distance between the optic disk and the fovea and NAM measures cannot predict fovea position correctly. (C) In other extreme conditions where fixation is eccentric (5 degrees in the right visual field), the optic nerve may not be entirely visible on the image, making the NAM estimation method inoperable.

Still, locating the exact position of the previously functional fovea in these patients with a central scotoma remains critical to estimate its distance from the newly developed area of fixation, called preferred retinal location (PRL). Indeed, the fovea-PRL distance has shown to be related to functional performance, such as fixation stability³ and reading speed.⁴ Moreover, it has recently been shown to be a predictor of PRL position change over the course of disease progression.⁵ Because fovea-PRL distance is a powerful measure to monitor and better understand the progression and functional impact of maculopathies,³⁻⁵ the exact position of the previously functional fovea must be estimated precisely in those patients with CFL.

In the literature, the most frequently used method to estimate the position of the fovea on pathological fundus images relies on normative anatomic measures (NAMs).^{4,6-8} According to the accepted measures, the mean foveal distance temporal to the middle of the optic disk is 15.5 ± 1.1 degrees horizontally and -1.5 ± 0.9 degrees vertically (Fig. 1A).^{3,9-11} However, using such normative measures presents several limitations. First, these measures represent a population average whereas the actual position of the fovea can vary dramatically from one individual to the other.¹² Second, these normative measures were estimated under “standard” testing conditions (with the head and trunk as closely aligned as possible, i.e. primary position of the head and central fixation target). In

the presence of a central scotoma, however, fixating on a specific target requires to use eccentric vision. In the case of a large scotoma, the eccentricity required to fixate can be so large that individuals may have to tilt their head and/or gaze to fixate. Therefore, they will not be able to maintain their head in the primary position, which will have a significant impact on the relative position of the different anatomic structures of the eyes (i.e. fovea, optic disk, and blood vessels). In addition to eccentric vision, abnormal ocular torsion, which has been observed in a number of strabismus conditions,¹³ can also amplify this phenomenon. For instance, the distance between the optic disk and the fovea is highly dependent on the eye and head orientation, as discussed in Ref. 10. This dependency is illustrated in Figure 1 where a unique healthy eye performed a fixation task under different conditions. Despite these strong limitations, and because it is easy to administer, this method remains the most used in clinical and research settings. Therefore, there is a need for new automated methods to estimate the position of the fovea on color fundus images where the morphology of the macular region has severely changed.

To address this need, two main types of automated image processing methods have emerged recently.¹⁴ The first one gathers anatomic structure-based methods, which use the main visible anatomic structures of a fundus image (optic disk, blood vessels, and macula) as landmarks to locate the macular region, and consequently the fovea. Whereas the simplest approach

is to detect the fovea as the center of the darkest circular region on the whole fundus image,^{15,16} most methods work in two stages: first, (i) the estimation of a region of interest (ROI) that most likely contains the fovea (using blood vessels or optic disk information), followed by (ii) the precise localization of the fovea within this ROI, using color intensity information.^{17–19} Despite their efficiency with healthy fundus images, these methods cannot be applied as such to cases where the morphology of the macula has severely changed.

As opposed to these methods, which only work on a case-by-case basis, methods using deep learning constitute a powerful alternative as they build relevant and representative features from large amounts of data. So far, they have been used for lesion, vessel, optic disk, and fovea detection^{20–24} (refer to Ref. 25 for a review). Overall, two main deep learning-based approaches are used: one considers fovea localization as a segmentation task, and the other as a regression task. With the segmentation approach, each pixel of the fundus image is classified by one or several convolutional neural networks into either the “fovea” or “non-fovea” categories.^{20,24,26} Because no local feature makes the fovea region distinct from the rest of the fundus in pathological images, this segmentation approach is not well-suited. The alternative strategy, which is to treat the fovea localization as a regression task,²⁷ allows to make predictions by building a regression network using all the features it extracts, even the ones relative to retinal regions far from the fovea. Yet, this method has always been trained and used with healthy fundus images, where the macular region could be used as a significant feature by the network.

Moving forward from these approaches, the goal of the present work is to propose a new method to predict the fovea position in pathological fundus images (where the fovea is not visible), using only characteristics of the blood vessels. Indeed, vessels are visible in nearly every fundus image, even in AMD cases, making the vascular structure the most robust landmark to the observed image alterations. As the vascular structure does not contain any direct visual information about the fovea location, we can use healthy fundus images where the position of the fovea can be clearly identified, serving us as a precise ground truth. However, locating the fovea from the vessel structure on a single retina image is virtually impossible because vessels show a large range of interindividual variability.^{28,29} To tackle this variability problem, while still taking advantage of the attractive properties of the vessel structure, we designed the vessel-based fovea localization (VBFL) method. This translational tool will allow researchers and eye care providers to estimate the exact position

of the nonfunctional fovea in patients with CFL, as well as the fovea-PRL distance. Such information will be crucial to: (1) monitor the progression of the disease and the resulting functional deficit, (2) better understand the underlying causes of this functional deficit, and (3) improve individualized rehabilitative care, especially for optometrists.^{3–5} After presenting the VBFL method in details, we will analyze its performance on healthy retina images taken under different conditions of head orientations and fixation positions (validation 1), as well as images with simulated lesions (validation 2) and pathological images (validation 3).

Methods

Vessel-Based Fovea Localization Method

The VBFL method can be decomposed into three steps (Fig. 2). First, it uses a large data set of fundus images to build statistical representations of the vascular structure at the population level, overcoming individual variations in the retinal vascular pattern, such as the asymmetry between the superior and inferior vascular arcades. These statistical representations are based on vessels density and direction (step 1). Then, it extracts the same representations of vessels' density and direction from a target image for which we want to locate the fovea position (step 2). Finally, it registers the statistical representations (where fovea position is known) with respect to the target representations to locate the fovea in this new fundus image. The VBFL algorithm is freely available as a stand-alone software and can be downloaded here: <https://team.inria.fr/biovision/vbfl>.

Step 1: Statistical Representation Maps of Vessel Density and Direction

To create a representation of the vessel structure at the population level, VBFL requires a large training set of healthy images where fovea position can be clearly identified. Here, we used the REFUGE data set³⁰ as the training set, referred to as D_{train} . It contains 840 healthy eye fundus images with a manually annotated fovea location and optic disk position. As a preliminary step, we extracted a vessel information map v_i for each single image u_i . To do so, we used the recent retinal vessel segmentation network SA-UNet,²³ a convolutional neural network designed for biomedical image segmentation. We used the weights provided by the authors, pretrained on the DRIVE data set.³¹ REFUGE and DRIVE images were acquired with different fundus cameras, all with a field of view of 45 degrees. The REFUGE data set was collected with

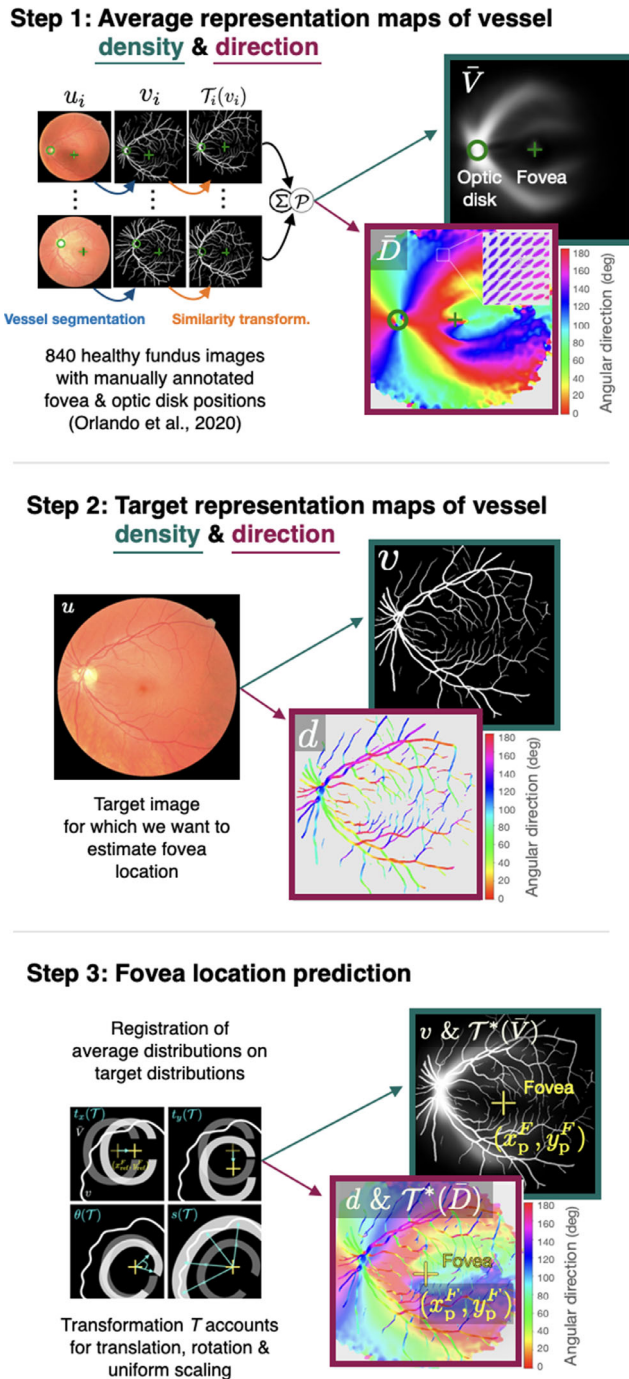


Figure 2. Step-by-step illustration of the VBFL method. Step 1: Statistical representations of vessels density and direction. Given a large dataset of fundus images where optic disk (circle) and fovea (cross) positions are known, images went through realignment, averaging (Σ) and postprocessing (P) steps, conducting to the average vessel density map \bar{V} , and the average vessel direction map \bar{D} . Step 2: For a target image u , estimation of the vessel density map v and direction map d . Step 3: Fovea position prediction through registration of average representations on the target representation. Transformation T of the average maps \bar{V} and \bar{D} accounts for translation, rotation and uniform scaling. Results of the target image registration are illustrated on the right: fovea position from the average maps is used as a predictor of the fovea position in the target image.

both a Zeiss Visucam 500 (resolution 2124×2056 pixels) and a Canon CR-2 (resolution 1634×1634 pixels). DRIVE images were taken with a Canon CR5 NM (resolution 768×584 pixels). To follow SA-UNet requirements, all images were resized to 592×592 pixels before they were fed to the network. This method outputs a binary map $v(x, y)$ indicating the presence or not of a vessel at each pixel (x, y) of the fundus image, so that $v(x, y) = 1$ if (x, y) belongs to a vessel, and $v(x, y) = 0$ otherwise. Aggregating these individual vessel maps v_i , we estimated two distribution representations: a vessel density map and a vessel direction map. The vessel density map (\bar{V}) aims at giving, for each position (x, y) , the likelihood to have a vessel passing through. The whole process is illustrated in Figure 2 - step 1 and detailed in Appendix A1. The vessel direction map (\bar{D}) aims at giving, for each pixel (x, y) , the most likely direction of a vessel between 0 and π . Details of its calculation are provided in the Appendix A2. The resulting direction \bar{D} map is shown in Figure 2 (step 1), where, for each pixel (x, y) , the main direction is represented with a color code, with tensors being represented as ellipses. Note that because small vessels show great variability, they make a relatively small contribution to the vessel maps, as opposed to the larger vessels that exhibit more reproducible structure. For that reason, our approach is essentially guided by the large vessels distribution whereas small vessels show little to no impact.

Step 2: Target Representation Maps of Vessel Density and Direction

Given a target fundus image u for which we want to predict fovea position, we extracted the same vessels representations as in step 1. Therefore, following the same procedures as described above, we estimated a target density map v and a target direction map d (see Fig. 2 - step 2).

Step 3: Fovea Location Prediction

Given the statistical representations (\bar{V} , \bar{D}) and target representations (v , d), as defined in steps 1 and 2, respectively, we register the statistical representations (where the fovea position is known) with respect to the target representations. Mathematically, the registration can be formulated as an optimization problem, which is detailed in Appendix A3. Once registered, the transformed fovea position is used as a prediction for the fovea position in the target fundus image u (see Fig. 2 - step 3).

Data

To test and validate the VBFL method, we used 3 types of fundus images: (1) 198 healthy images

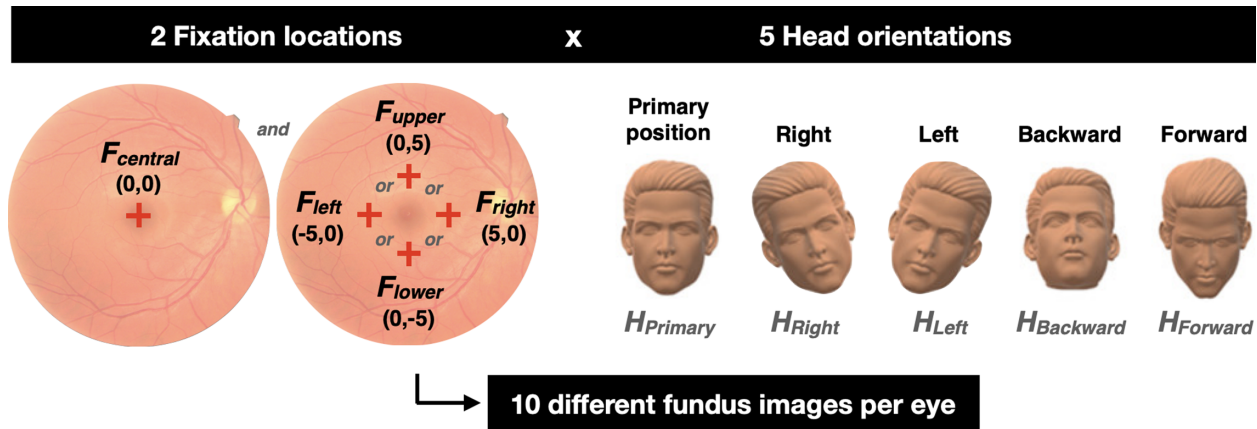


Figure 3. A combination of 5 head orientations and 2 fixation locations were used to acquire 10 different fundus images from each eye. The *left panel* shows the five possible fixation cross positions: in the center of the camera's field of view ($F_{\text{central}} = (0, 0)$) and 5 degrees of visual angle away from it, in the horizontal ($F_{\text{left}} = (-5, 0)$ and $F_{\text{right}} = (5, 0)$) or vertical direction ($F_{\text{lower}} = (0, -5)$ and $F_{\text{upper}} = (0, 5)$). The *right panel* shows the five head orientations tested: primary position (head looking straight ahead), tilted to the left and to the right by 30 degrees, forehead tilted forward and backward by 20 degrees. Each eye was tested both with the central fixation and one of the four eccentric locations. For each of them, all five head orientations were used successively in a random order.

acquired from our group, (2) these same 198 images with simulated lesions covering the macular region, and (3) 89 annotated images of eyes with AMD from the ADAM data set.³²

Healthy Fundus Images Acquisition

Nineteen normally sighted participants (11 female subjects) were recruited to collect healthy fundus images using a microperimeter MP-3 (Nidek Inc.). Subjects ranged in age from 19 to 40 years (mean = 26 ± 6 years) with no eye pathology or functional visual impairment, except for myopia. For each eye tested with the MP-3, the following additional information was also collected: subjective refraction, near visual acuity, and axial length (measured in mm using an IOL-Master, Zeiss Instruments). Corrected monocular visual acuity was 20/20. Axial length ranged from 22.68 to 28.77 with a mean value of 24.45 ± 1.72 . Informed consent was obtained from each participant after explanation of the nature and possible consequences of the study. The research was conducted in compliance with Inria's Operational Committee for the Evaluation of Legal and Ethical Risks (COERLE) and followed the tenets of the Declaration of Helsinki.

Fundus images were collected for 21 eyes total (10 left eyes) with a standard flash intensity level of 9, while the other eyes of the subjects were patched. Prior to each image acquisition, a fixation examination was performed to allow for precise fovea location. Subjects were asked to fixate a single red cross presented for 5 seconds (size = 1 degree and thickness = 0.2 degrees). Using a combination of different head orien-

tations and fixation locations, a total of 198 images were acquired. Each eye was tested with two different fixation locations: a central fixation and an eccentric fixation (off by 5 degrees of visual angle), selected randomly among four positions: left, right, up, or down. For each fixation, 5 head orientations were tested: primary position (head looking straight ahead), tilted to the left by 30 degrees, tilted to the right by 30 degrees, forehead tilted forward by 20 degrees, and forehead tilted backward by 20 degrees (Fig. 3). Combining these 2 fixation positions and 5 head orientations in a random order, a total of 10 images (and their corresponding fixation examination) were acquired for each eye (except for subject [S]3, S4, and S12 who completed only 6 because of a time constraint).

Healthy Fundus Images Annotation

For each image, the actual position of the fovea (i.e. "ground truth" gt) was set at the (x, y) coordinates of the fixation cross presented during the fixation examination (in degrees of visual angle, noted (x_{gt}^F, y_{gt}^F)). Each image was also annotated using NAMs to derive its fovea position, noted as (x_{nam}^F, y_{nam}^F) . First, the center of the optic disk was annotated manually by one of the authors (V.F.), then, the (x, y) coordinates of the fovea location were derived using standard NAM, that is, 1.5 degrees inferior and 15.5 degrees temporal relatively to the optic disk center.¹⁰ Finally, the VBFL method was run on each image to predict its fovea location automatically, noted as (x_p^F, y_p^F) .

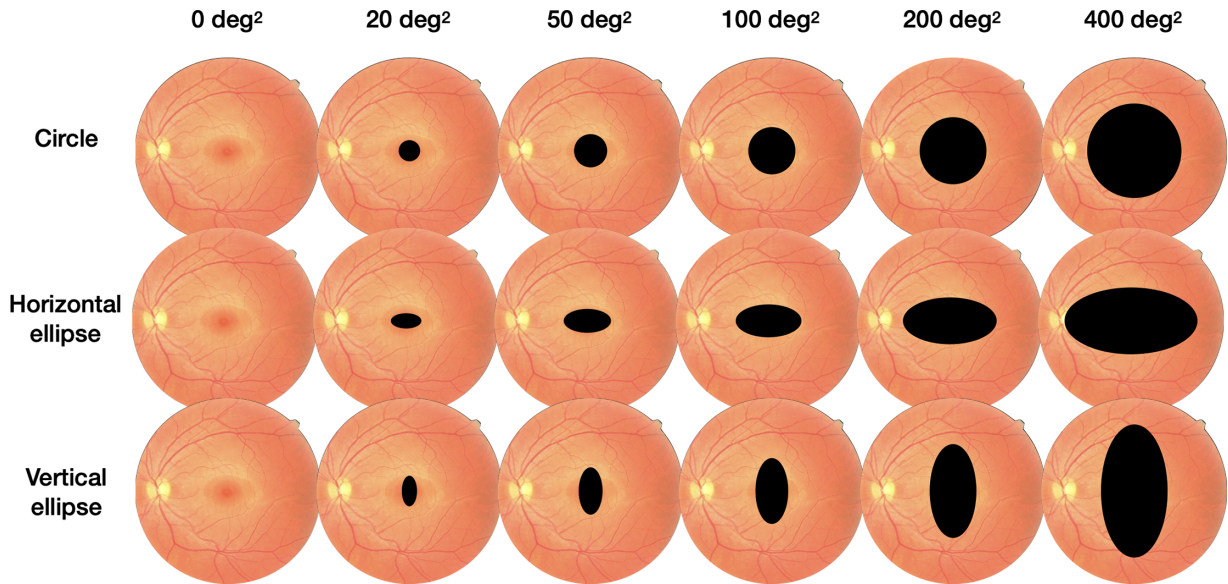


Figure 4. Illustration of the 18 different conditions of simulated macular lesion created for each retinal image. Rows correspond to the three shapes used: *circle*, *horizontal*, and *vertical* ellipse. Ellipses have a major axis twice as large as their minor axis. The columns represent the different sizes of masks: 0 (i.e. no mask), 20, 50, 100, 200, and 400 degrees², respectively.

Simulated Lesions on Healthy Fundus Images

In order to test the VBFL method on images with no informative features in the macular region, but for which the exact fovea location (x_{gt}^F, y_{gt}^F) could still be known, we simulated lesions on our set of healthy fundus images. To do so, we applied black masks over the fovea (x_i^F, y_i^F) on each vessel map v_i derived from the healthy images u_i . A total of three shapes (circle, horizontal, and vertical ellipse) and six sizes (0, 20, 50, 100, 200, and 400 degrees²) were combined to create 18 different masks, as illustrated in Figure 4. Each of these masks was applied to each of the 198 healthy fundus images, resulting in a total of 3564 images with simulated lesions, hiding more or less vessels depending on their shape and size. The VBFL method was run on each of these images to predict their fovea location automatically.

Pathological Fundus Images

Eighty-nine fundus images from eyes diagnosed with maculopathy and a manually annotated fovea position were used. These annotated images were part of the training data set released during the ADAM challenge, which goal was to evaluate automated algorithms for the detection of AMD in retinal fundus images.²² These 89 images included all forms and stages of AMD, showing representative retinal lesions, such as drusen's, exudates, hemorrhages, and/or scars. They were acquired using a Zeiss Visucam 500 fundus camera (resolution 2124 × 2056 pixels). Images were manually annotated by seven independent ophthal-

mologists who located the fovea position, as well as visible lesions. For each image, fovea's final coordinates were obtained by averaging the seven annotations and were quality-checked by a senior specialist. Similarly, manual pixel-wise annotations of the different lesion types (drusen, exudate, hemorrhage, and scar) were provided for each image. From the resulting lesion segmentation masks, we estimated for each image: its lesion type(s) (drusen, exudate, hemorrhage, and/or scar), and the size of each individual lesion (in degrees²), as well as its full lesion size (in degrees²) using the R packages for image processing, magick and countcolors (Fig. 5).

Prediction Error Estimation

Each method performance will be assessed using its error amplitude. Therefore, we used prediction results to compute measures of prediction error. For each healthy image, NAM estimation of the fovea position (x_{nam}^F, y_{nam}^F) was compared to the ground truth (x_{gt}^F, y_{gt}^F) , to derive a prediction error value $\epsilon_{nam,healthy}$, calculated as their Euclidean distance (the lower, the better). Similarly, VBFL predictions (x_p^F, y_p^F) were compared to the ground truth (x_{gt}^F, y_{gt}^F) , to derive a prediction error value $\epsilon_{p,healthy}$. For simulated images, only VBFL results of (x_p^F, y_p^F) were compared to the (x_{gt}^F, y_{gt}^F) to extract prediction error $\epsilon_{p,sim}$. For pathological images, VBFL predictions (x_p^F, y_p^F) were computed and compared to the annotated fovea position to derive a prediction error value $\epsilon_{p,AMD}$. In the

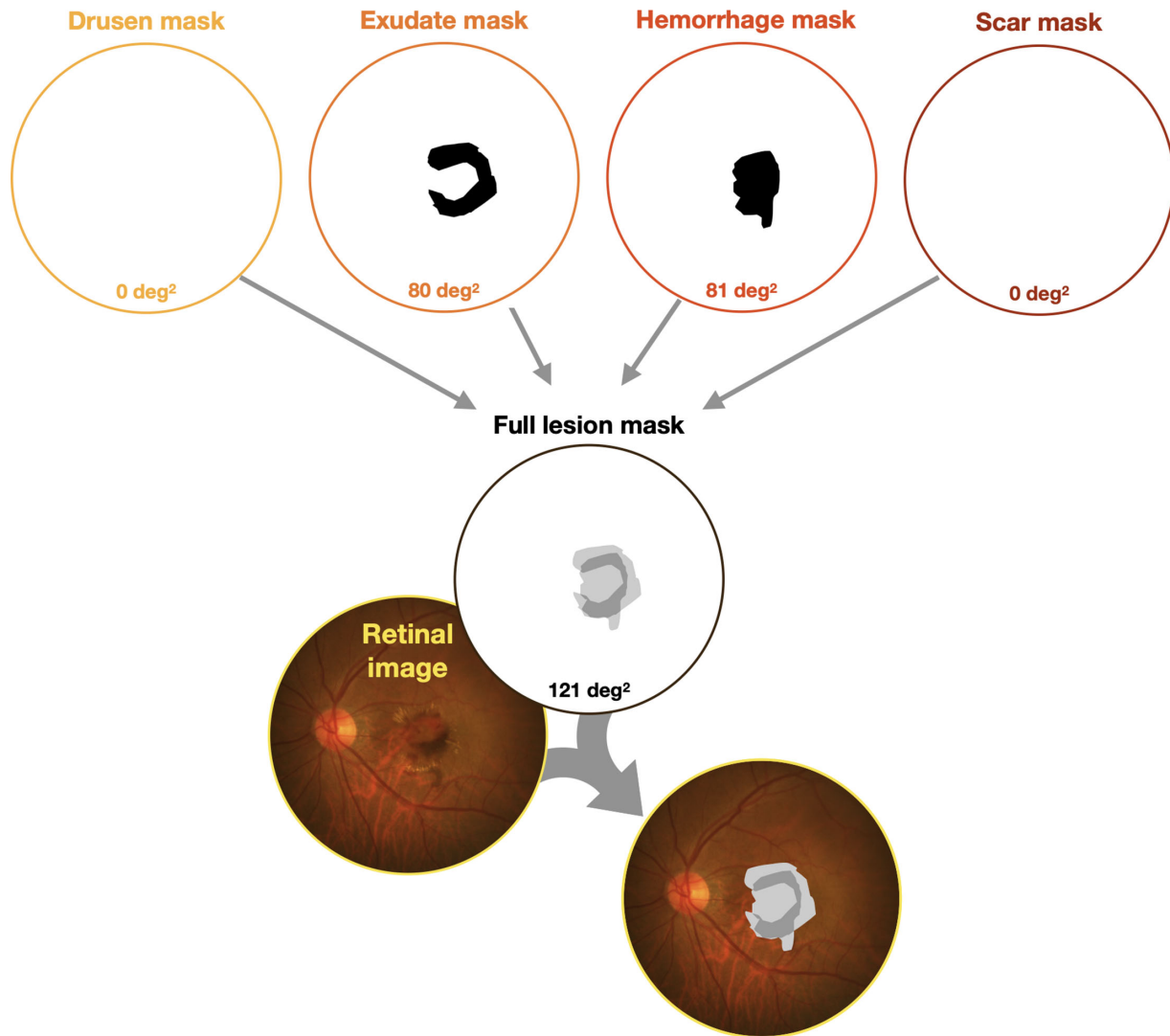


Figure 5. Lesion size estimation of the 89 annotated pathological images. With each retinal image, four lesion type masks were provided: drusen (yellow), exudate (orange), hemorrhage (red), and scar (brown). For each individual mask, we first estimated its size in square degrees. The example above shows an image with no drusen, exudate of 80 degrees², hemorrhage of 81 degrees², and no scar. Second, we combined all four individual masks to obtain a full lesion mask, for which size in square degrees was also estimated (121 degrees², here).

Results section, all prediction error values ($\in_{nam_{healthy}}$, $\in_{p_{healthy}}$, $\in_{p_{sim}}$, and $\in_{p_{AMD}}$) will be expressed in degrees of visual angle (DVA).

Statistical Analysis

Statistical analyses were carried out using R software³³ with the following additional R packages: tidy, dplyr, stringr, nlme, emmeans, and ggplot2.

In validation 1, we analyzed the performance of VBFL on healthy images by comparing it to the performance of the NAM method for each of the 25 crossed-conditions (5 head orientations * 5 fixation positions). We fitted a liner mixed-effect (LME) model³⁴ with prediction error as a dependent variable. The following variables: “method,” “head orientation,” “fixation

location,” and “axial length of the eye” were set as fixed effects and “participants” was modeled as the random effect. Prediction error was transformed in natural logarithm (ln) units and eye axial length was centered around its mean.

In validation 2, we inspected the performance of VBFL on healthy images with simulated lesions of varying sizes and shapes by comparing it across all 18 conditions presented in Figure 4 (3 shapes * 6 sizes). We fitted an LME model with prediction error as dependent variable. The following variables: “lesion shape,” “lesion size,” “head orientation,” “fixation location,” and “axial length of the eye” were set as fixed effects, whereas “participants” was modeled as the random effect. Prediction error was transformed

in natural logarithm (ln) units and eye axial length was centered around its mean.

For both models, optimal model selection was performed using the Akaike's Information Criterion (AIC) and likelihood-ratio tests.³⁵ Significance of the fixed effects was estimated using t-values: absolute values of t-values larger than 2 were considered significant, corresponding to a 5% significance level in a 2-tailed test.^{36,37} In the Results section, fixed effects estimates are reported along with their t-values and 95% confidence intervals.³⁴

In validation 3, we inspected the performance of VBFL on pathological fundus images with AMD and report a descriptive analysis of average prediction error for all 89 images. Using simple linear regression models, we then estimated the effect of several parameters on prediction error: the density of vessels detected in the original image, the type of lesions (drusen, exudate, hemorrhage and/or scar), their sizes, and finally the size of the full lesion.

Results

Validation 1: Healthy Images

For the standard testing condition (head in primary position and central fixation), NAM fovea position estimation resulted in an average error of 1.03 degrees ($\exp(0.03)$, 95% confidence interval [CI] = 0.82, 1.30; Table 1). For this same condition, the error yielded by the VBFL estimation was significantly smaller, but only slightly, with an average value of 0.65 degrees ($\exp(0.03-0.47)$, $t = -2.9$, 95% CI = 0.50, 0.84). This difference represents a 37% error decrease with the VBFL method (see Fig. 5; top-left panel). Except for the two head orientations where the head is tilted to the side (left or right), these values remained constant for all other conditions, with an average prediction error ranging from 0.95 to 1.13 degrees (SD = 0.05) with NAM and from 0.71 to 1.18 degrees (SD = 0.14) with VBFL (Fig. 6).

For all conditions where the head was tilted to the left (i.e. all 5 fixation positions; see Fig. 6 column 4), the NAM estimation error was significantly multiplied by 4.48 ($\exp(1.5)$, $t = 12.77$, 95% CI = 3.59, 5.64) compared to the standard condition (head in primary position), reaching an average error value of 4.61 degrees (95% CI = 3.55, 5.88); that is a 348% increase. For the same conditions, the VBFL method on the other hand yielded no significant increase in error prediction, with an average value of 1.21 (95% CI = 0.93, 1.56). This significant difference represents

a 74% reduction in prediction error compared to the NAM method.

Similarly, with the head tilted to the right and across all fixation conditions (see Fig. 6 column 5), the NAM estimation error was significantly multiplied by 3.9 ($\exp(1.36)$, $t = 11.65$, 95% CI = 3.23, 5.26) compared to the standard condition (head in primary position), reaching an average error value of 4.02 degrees (95% CI = 3.08, 5.08); that is, a 290% increase. For the same conditions, the VBFL method on the other hand yielded no significant increase in error prediction, with an average value of 0.91 (95% CI = 0.70, 1.18), representing a 77% reduction in prediction error compared to the NAM method.

Individual values of eye axial length showed no significant effect on either the NAM or the VBFL method performance in any of the 25 conditions tested.

Validation 2: Simulated Lesions Images

Across all conditions, we found that VBFL prediction error increased significantly with lesion size, and that the amplitude of this effect was modulated by the position of the head (Fig. 7). For the standard testing condition (head in primary position and central fixation), prediction error was on average 0.6 degrees ($\exp(-0.503)$, $t = -5.83$, 95% CI = -0.67 to -0.33 ; Table 2) with no lesion. For each 1 degree increase in lesion size, error is multiplied significantly by 1.0023 ($\exp(0.0023)$, $t = 11.88$, 95% CI = 0.0019, 0.0027). In other words, increasing lesion size from 0 to 100 increases error by a factor of 1.26 ($\exp(0.0023)^{100}$), that is, a 26% increase, resulting in an average error prediction of 0.76 degrees. Following the same logic, prediction error reaches 0.96 degrees at 200 degrees², 1.21 at 300 degrees², and 1.57 at 400 degrees².

However, the amplitude of this effect was significantly larger when the head was tilted to the side, as shown by the significant interaction between lesion size and head orientation (see Table 2). As illustrated in Figure 7 - column 5, when the head is tilted to the right, increasing lesion size from 0 to 100 increases error by a factor of 1.38 ($\exp(0.0023 + 9.025e-04)^{100}$), that is, a 38% increase, resulting in an average error prediction of 0.83 degrees. Following the same logic, prediction error reaches 1.15 degrees at 200 degrees², 1.58 at 300 degrees², and 2.18 at 400 degrees².

Across all conditions, there was no significant effect of lesion shape, nor an interaction between size and shape, suggesting that no matter the lesion shape, its size always had the same effect on VBFL

Table 1. Fixed-Effects Estimates From the LME Model (Validation 1)

	Estimate	Std. Error	t Value	95% Confidence Interval
(Intercept)	0.03	0.12	0.25	[−0.2 to 0.26]
Method	−0.47	0.16	−2.9	[−0.77 to −0.16]
VBFL				
Method:Fixation location <i>NAM:F_Up</i>	0	0.16	−0.02	[−0.31 to 0.3]
Method:Fixation location VBFL:F_Up	0.43	0.16	2.67	[0.12 to 0.73]
Method:Fixation location <i>NAM:F_Down</i>	−0.08	0.14	−0.55	[−0.34 to 0.19]
Method:Fixation location <i>VBFL:F_Down</i>	0.11	0.14	0.83	[−0.15 to 0.38]
Method:Fixation location <i>NAM:F_Left</i>	−0.02	0.14	−0.11	[−0.29 to 0.26]
Method:Fixation location VBFL:F_Left	0.3	0.14	2.12	[0.03 to 0.57]
Method:Fixation location <i>NAM:F_Right</i>	0.02	0.13	0.13	[−0.23 to 0.27]
Method:Fixation location <i>VBFL:F_Right</i>	0.22	0.13	1.66	[−0.03 to 0.47]
Method:Head orientation <i>NAM:H_Forward</i>	0.04	0.12	0.31	[−0.19 to 0.26]
Method:Head orientation <i>VBFL:H_Forward</i>	0.09	0.12	0.75	[−0.14 to 0.31]
Method:Head orientation <i>NAM:H_Backward</i>	0.08	0.12	0.7	[−0.14 to 0.31]
Method:Head orientation <i>VBFL:H_Backward</i>	0.18	0.12	1.5	[−0.05 to 0.4]
Method:Head orientation NAM:H_Left	1.5	0.12	12.77	[1.28 to 1.73]
Method:Head orientation <i>VBFL:H_Left</i>	0.41	0.12	3.5	[0.19 to 0.64]
Method:Head orientation NAM:H_Right	1.36	0.12	11.65	[1.14 to 1.58]
Method:Head orientation <i>VBFL:H_Right</i>	0.13	0.12	1.1	[−0.1 to 0.35]

The dependent variable is log-transformed prediction error in degrees. The intercept estimate represents the log-transformed prediction error when all factors included in the model are at their reference level (NAM method; Primary head orientation; Central fixation). Interactions are represented by the symbol ":". Factors showing a significant effect on prediction error are in bold font.

performance. Lastly, as seen on [Figure 7](#), VBFL yields robust estimates (i.e. estimates smaller or comparable to NAM - purple line) for lesion size up to about 200 deg². For larger lesions, VBFL tends to yield larger prediction error than NAM, except for the two head tilted conditions (Fleft and Fright) for which VBFL prediction error remains smaller than NAM, no matter the lesion size.

Validation 3: Pathological Images

Over the 89 pathological images, average prediction error \pm SD was 2.83 degrees \pm 2.58. Overall, 64% of the images yield an error of 2.5 degrees or less, 23% yield an error comprised between 2.5 and 5 degrees and 13% yield an error ranging from 5 to 15 degrees. [Figure 8](#) shows this distribution, along

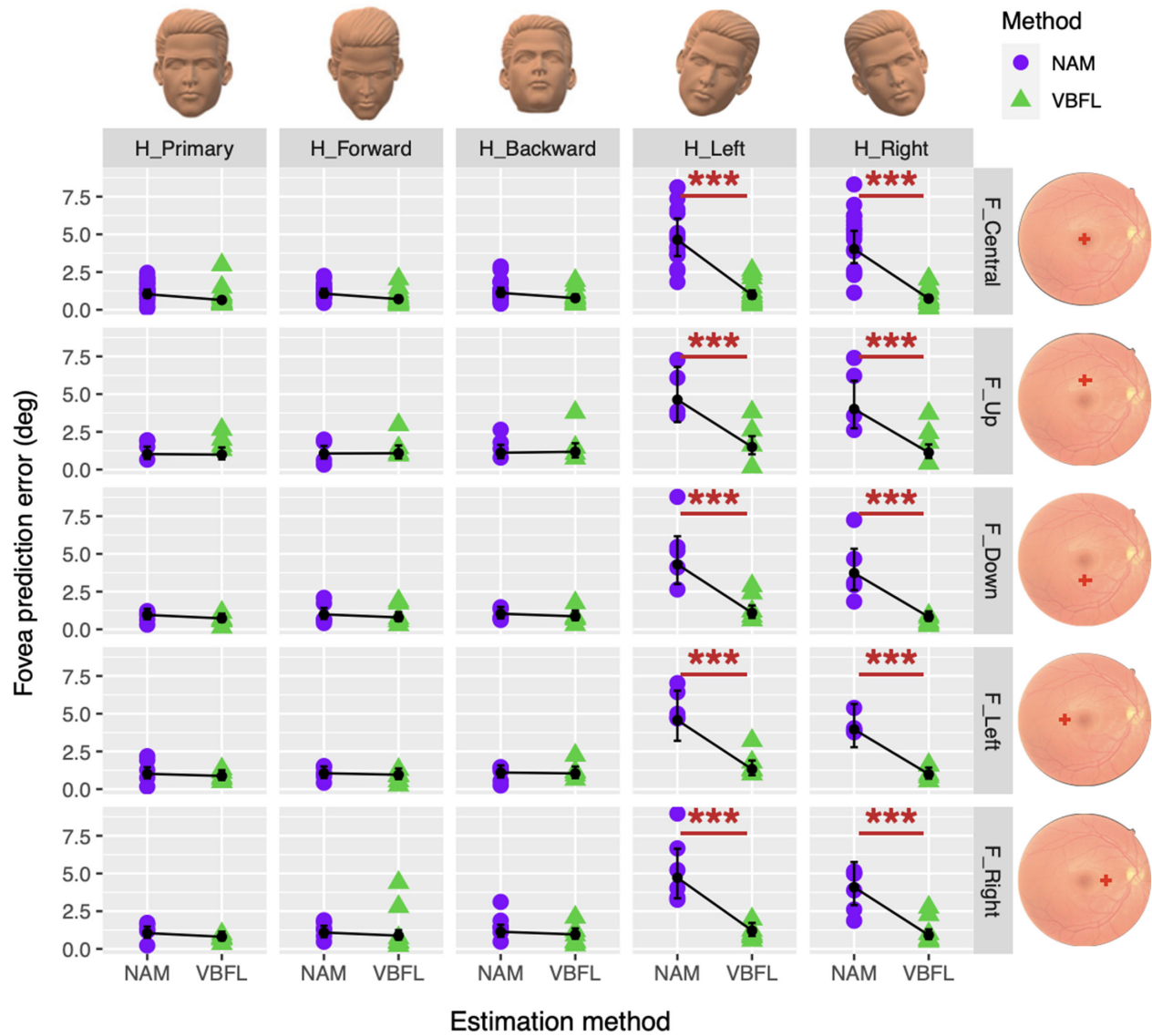


Figure 6. Effect of fixation location and head orientation on the performance of the NAM and VBFL methods in fovea position estimation. Columns show all five head conditions: Hprimary, Hforward, Hbackward, Hleft, and Hright. Rows show all five fixation conditions: Fcentral, Fup, Fdown, Fleft, and Fright. Each sub-panel shows the effect of estimation method NAM versus VBFL on the fovea position prediction error (in degrees). Raw data are represented with purple circles for the NAM method and green triangles for the VBFL method. Black circles represent the average estimates for each subgroup as given by the lme model. Error bars show their 95% confidence intervals.

with significant examples, for which the position of the annotated fovea (in green) and the VBFL prediction (in yellow), are given. For 29% of the images (26 out of 89), fovea estimation was found to be excellent, ranging from 0 to 1.5 degrees in prediction error. We conclude that the VBFL method seems to work efficiently for different stages of macular degeneration, namely macular scarring (see Fig. 8A), drusen (see Fig. 8B), dry AMD (see Fig. 8C), wet AMD (see Fig. 8D), or a combination of both (see Fig. 8E). For 35% of the images (31 out of 89), prediction

error was slightly larger, ranging from 1.5 to 2.5 (see Fig. 8F). For the remaining 36%, however, (32 images out of 89) the VBFL method is not as robust, with the prediction error as large as 14.76 degrees. For those specific cases, we noted some consistent patterns. First, for images showing uneven brightness with darker regions, that does not allow proper vessel segmentation. In such cases, individual vessel map v seem quite incomplete, resulting in a mis-estimation of the fovea (see Figs. 8G, 8H, 8J). Another recurrent issue, is the incomplete representation of the optic disk

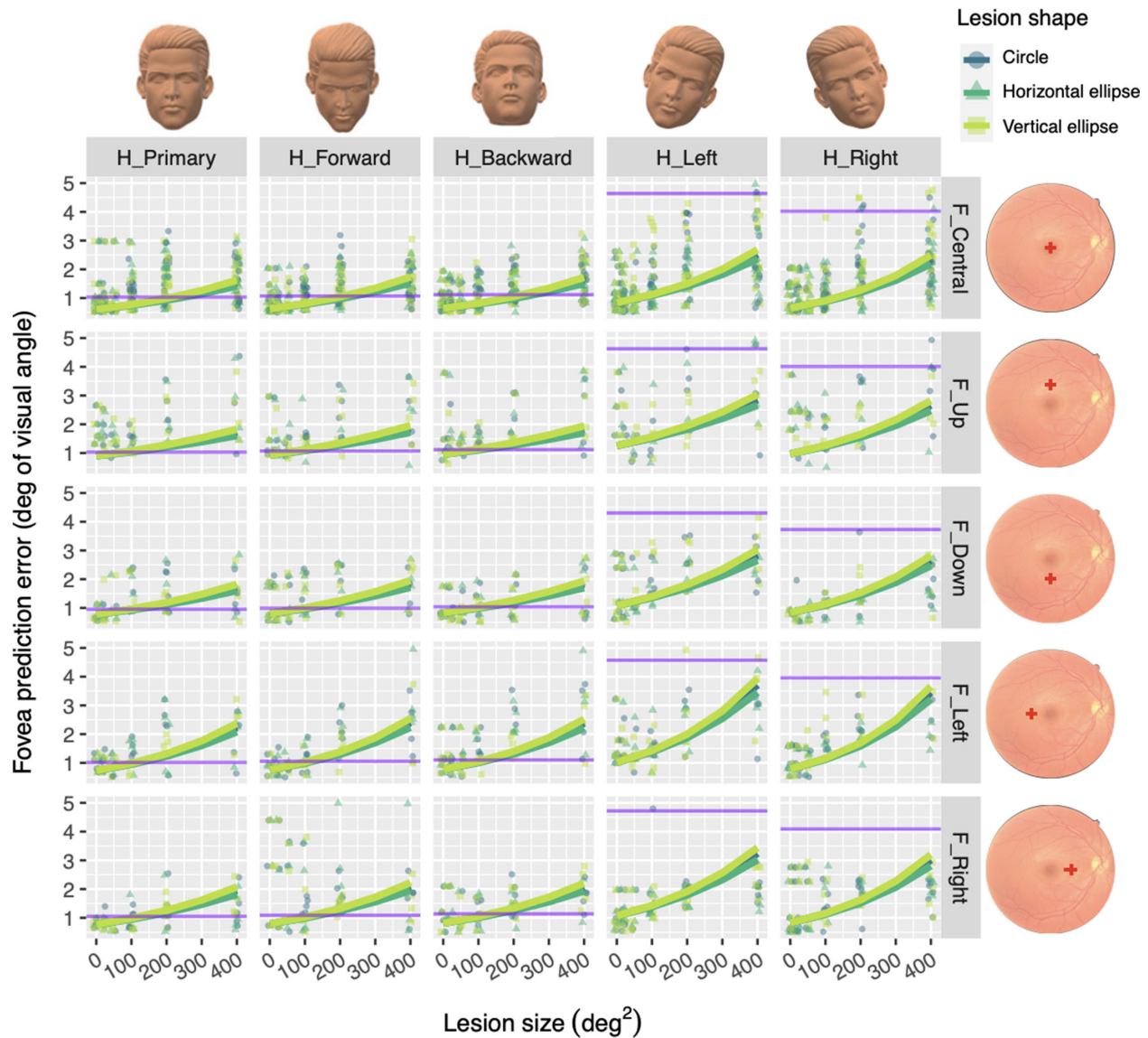


Figure 7. Effect of lesion size and shape on the VBFL method performance, for all conditions of fixation location and head orientation. The columns show all five head orientation conditions: Hprimary, Hfrontward, Hbackward, Hleft, and Hright. The rows show all five fixation conditions: Fcentral, Fup, Fdown, Fleft, and Fright. Each sub-panel shows the effect of lesion size and lesion shape (color-coded in shades of green) on the VBFL fovea position prediction error (in degrees). Raw data are represented, respectively, with dark green circles for circular lesions, medium green triangles for horizontal elliptical lesions, and light green squares for vertical elliptical lesions. Fitted green lines represent the effect of size for each type of lesion shape, as estimated by the lme model. On each sub-panel, the purple line shows mean error prediction produced by NAM estimation, as measured in validation 1 for healthy images.

due to a strong focus on the macular region when shooting the photograph. If so, the method cannot extract the emerging of vessels, nor the lower and upper vessel branches properly, resulting in a very large prediction error (see Figs. 8I, 8K). When inspecting extrapolated vessel maps v_i , we found a significant effect of the percentage of vessels detected on the overall accuracy of the method's prediction: the higher the density in the vessel map, the smaller the prediction error. For instance, for a map contain-

ing 10% of pixels representing a vessel, which corresponds to the average density of vessels extracted in the 198 healthy images, prediction error shows an average value of 1.6 degrees ($t = 4.6$, 95% CI = 1.3, 2.0). For each 1% decrease in vessel density, prediction error increased by a factor of 0.86, reaching a mean value of 2.2 at 8%, 2.9 at 6%, and 3.9 degrees at 4% density.

To further try and characterize images for which our method provides poor prediction estimates, we

Table 2. Fixed-Effects Estimates From the LME Model (Validation 2)

	Estimate	Std. Error	t Value	95% Confidence Interval
(Intercept)	−0.503	0.0862	−5.8371	[−0.6738 to −0.3318]
Lesion size	0.0023	2e-04	11.8813	[0.0019 to 0.0027]
Fixation location	0.398	0.0628	6.3337	[0.2756 to 0.5219]
<i>F_Up</i>				
Fixation location	0.2509	0.0536	4.6853	[0.1458 to 0.3554]
<i>F_Down</i>				
Fixation location	0.1896	0.0559	3.3891	[0.0803 to 0.2991]
<i>F_Left</i>				
Fixation location	0.2372	0.0517	4.584	[0.136 to 0.3383]
<i>F_Right</i>				
Head orientation	0.0323	0.0479	0.6755	[−0.0613 to 0.1259]
<i>H_Forward</i>				
Head orientation	0.06	0.0479	1.2529	[−0.0336 to 0.1535]
<i>H_Backward</i>				
Head orientation	0.3306	0.0482	6.8529	[0.2362 to 0.4249]
<i>H_Left</i>				
Head orientation	0.0723	0.0479	1.5115	[−0.0213 to 0.1659]
<i>H_Right</i>				
Lesion size:Fixation location	-7e-04	3e-04	−2.5166	[−0.0012 to −2e-04]
<i>F_Up</i>				
Lesion size:Fixation location	-3e-04	3e-04	−1.2989	[−8e-04 to 2e-04]
<i>F_Down</i>				
Lesion size:Fixation location	5e-04	3e-04	1.9278	[0 to 0.001]
<i>F_Left</i>				
Lesion size:Fixation location	0	2e-04	0.1145	[−4e-04 to 5e-04]
<i>F_Right</i>				
Lesion size:Head orientation	1e-04	3e-04	0.3619	[−4e-04 to 6e-04]
<i>H_Forward</i>				
Lesion size:Head orientation	0	3e-04	0.0043	[−5e-04 to 5e-04]
<i>H_Backward</i>				
Lesion size:Head orientation	4e-04	3e-04	1.7155	[−1e-04 to 9e-04]
<i>H_Left</i>				
Lesion size:Head orientation	9e-04	3e-04	3.5575	[4e-04 to 0.0014]
<i>H_Right</i>				

The dependent variable is log-transformed prediction error in degrees. The intercept estimate represents the log-transformed prediction error when all factors included in the model are at their reference level (categorical variables) or at 0 (continuous variable). Reference levels are respectively: “H_{primary}” for the factor “Head orientation” and “F_{central}” for “Fixation location.” Interactions are represented by the symbol “:”. Factors showing a significant effect on prediction error are in bold font.

inspected the effect of lesion size and lesion type on fovea prediction error. Despite a slightly noticeable increase in prediction error above 100 deg², we found no significant effect of full lesion size on the error amplitude. Similarly, we found no significant effect of lesion type on prediction error, no matter their individual sizes.

Discussion

In this paper, we present a novel method to detect automatically the fovea in both normal and pathological fundus images. This VBFL method relies on the anatomic structure of the vessels to estimate the

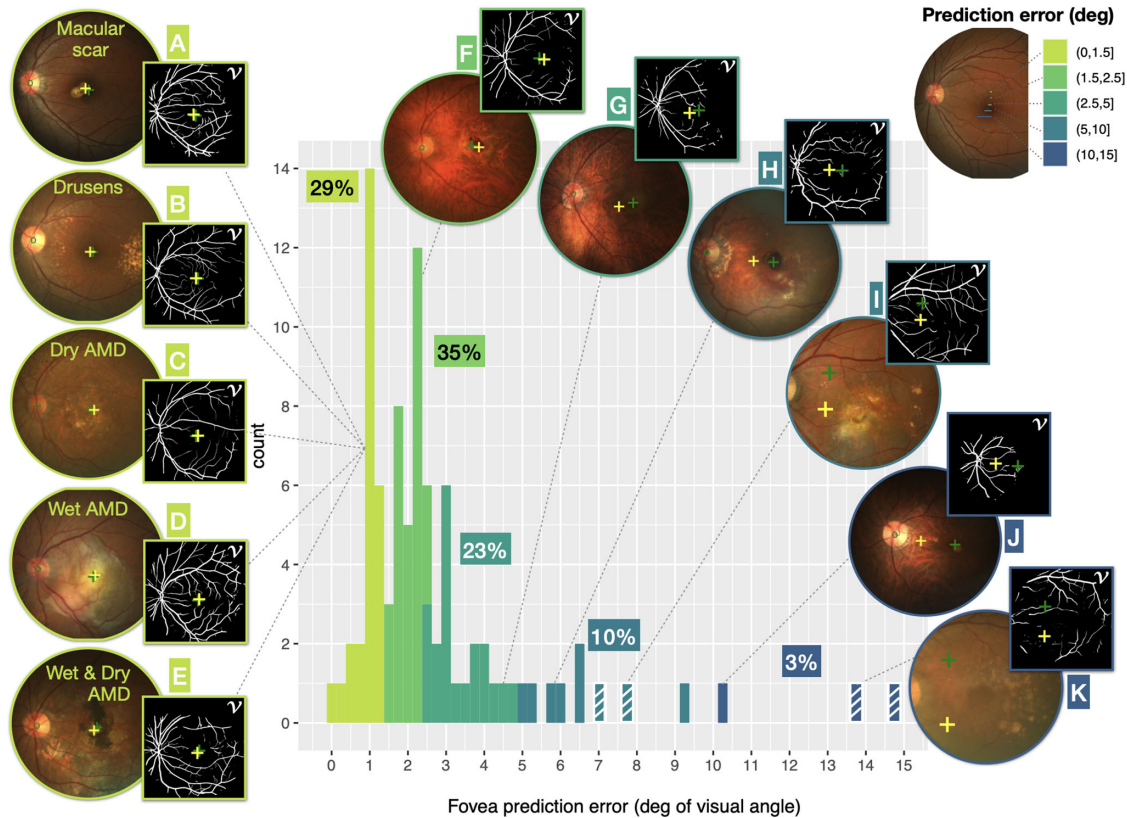


Figure 8. Distribution of VBFL prediction error for 89 pathological fundus images. Fovea prediction error (in degrees of visual angle) is color coded from *light green* (smaller error) to *dark blue* (larger error). Representative fundus images are shown, along with their individual vessel map v . For each image, position of the annotated fovea is represented with a *green plus sign*; fovea position estimated by the VBFL method is shown with a *green plus sign*. *Oblique white line* patterns mark images for which the optic disk was not entirely visible.

position of the fovea on a given fundus retinal image after estimating both its vessel density map and its vessel direction map. This new image processing tool should greatly benefit both researchers and eye care providers working on CFL and its impact on visual function. Because it will improve fovea location estimation in normal and pathological eye images, it will allow to study further the implication of PRL characteristics on the remaining performance of patients with CFL. It will also give optometrists valuable insight toward the specific deficit and compensatory strategies developed by their patients, therefore allowing more personalized care.

In order to validate our approach, we applied it on three different fundus image datasets: nonpathological images from healthy eyes (validation 1), nonpathological images with a simulated central lesion (validation 2), and pathological images with maculopathy (validation 3).

For nonpathological images, we found that the VBFL method gives accurate results, with prediction error smaller than 1 degree of visual angle. Hence, our

method appears to be slightly more precise than the commonly used method that relies on NAMs. Furthermore, in case of atypical data acquisition conditions, such as eccentric fixation or head tilt, VBFL predictions remain accurate, whereas the NAM method may provide off predictions, by up to 7 degrees. In this sense, the proposed method may be considered as an improvement over the literature standard. However, it remains to be determined how frequent those extreme data acquisition conditions may be encountered in a clinical environment.

In our sample of nonpathological eyes, we included three eyes with high myopia, for which axial length was greater than 26 mm.³⁸ In these eyes, higher axial length can cause a distorted retinal shape, resulting in an altered geometry of the apparent vessels.³⁹ Therefore, we expected our method to yield greater fovea localization error in those eyes with high axial length. However, we did not find a significant effect of eye axial length on the VBFL prediction error. Because our sample size of high myopic eyes was quite small, it remains to be tested in a greater sample of eyes present-

ing high myopia, whether the VBFL method can still be considered reliable on extreme values of axial length.

Although the VBFL method was found to be efficient on normal fundus images, it must also be reliable with impaired images to be deemed purposeful. Therefore, we also tested it on normal images for which the macular region was occluded by a simulated lesion varying in size and shape. The point of this validation, was to test our method on images with missing central features, while still knowing precisely the actual position of the fovea to assess predictions precision. We found that, in the presence of macular visual masks that make the vessel structure nonvisible in the central part of the images, VBFL predictions remain acceptable up to mask areas of 200 degrees². This cutoff value covers 75% of reported scotoma sizes in both wet and dry AMD,⁴ making our method widely efficient. After that limit, however, because hidden vessel features cannot be extracted properly, the method becomes unreliable. We conclude that the most informative vessel features for the VBFL method lay beyond 8 degrees from the fovea. This value being considerably larger than the mean radius of the foveal avascular zone, which ranges between 0.7 and 1.2 degrees in healthy subjects,⁴⁰ VBFL performance should not be affected by interindividual differences in the size of the foveal avascular zone.

However, given that fundus images are resized during step 2, small vessels around the macular area may not be detected anymore by the system, potentially influencing the limits of the most informative area of 8 degrees+ from the fovea. It remains to be tested whether larger resolution images would bring these limits closer to the fovea.

We had hypothesized that elliptical lesions/masks, which hide both upper and lower vessel branches if oriented vertically, or the vessels emergence near the optic disk if oriented horizontally, would be more detrimental than circular lesions. However, we found no significant effect of lesion/mask shape and VBFL predictions remained constant no matter the shape of the simulated lesion applied. This unanticipated result is a strong asset of the VBFL method.

When evaluated on pathological images, the VBFL method remains somewhat robust with a prediction error in fovea position of 2.5 degrees or less in 64% of the images. This evaluation also revealed some clear limitations. First, VBFL suffers from the optic disk absence, because the emerging vessel branches become not visible either. This was the case for images taken with a strong focus on the macular region, which may not arise frequently in clinical settings. Second, VBFL gave poor estimation for dark images. Although this point may not be critical in optimal

clinical settings where eyes are dilated and fundus images taken with a powerful ophthalmologic tool, it may become critical under other circumstances. For instance, smartphone-based fundus imaging, which couples a smartphone with a retinal camera, now allows for mobile and inexpensive fundus examination.^{41,42} Despite the poorer image quality they convey, it remains useful, especially in low-resource settings. It remains to be tested whether VBFL can accurately estimate the position of the fovea under such quality images with very low resolution.

The pathological images we used in validation 3 were made available specifically for the ADAM challenge, whose aim was to make artificial intelligence (AI) research teams compete during the International Symposium on Biomedical Imaging. The challenge objective was to evaluate and compare automated algorithms for the detection of AMD on a common dataset of retinal fundus images.³² Results were published recently²² and show that all methods proposed for fovea localization were based on machine learning (segmentation and/or regression) relying on neural networks. However, comparing their results to ours is quite difficult for several reasons. First of all, their estimation was based on both normal and pathological images. Therefore, the reported value of prediction error average (in pixels only) may appear better than ours given that the majority of their images were nonpathological (311 normal against 89 pathological images). Furthermore, because they reported one single average value of prediction error for the whole dataset, it is impossible to compare our results on a case-by-case basis. Because we already stated that VBFL is not robust with absent optic disk or dark images, it would have been useful to compare our results only for cases where VBFL can be reliable to assess its efficiency. However, in Ref. 22, it is reported that participating teams can get better fovea localization results when the macular region has higher contrast and cleaner texture, regardless of AMD or non-AMD images, which is in line with our own conclusion. Second, they used a combination of methods (at least 2 or more), and then combined them by taking the best estimate for each image. Last, none of these approaches are described in detail, which does not allow us to compare them in detail with the VBFL approach.

One critical point regarding our validation of the method with pathological images, is that prediction accuracy is estimated by comparing the method's prediction to the "actual" fovea location. However, in the case of the ADAM images, fovea location was annotated manually by ophthalmologists, with no clear rules on how they performed their annota-

tion. Therefore, assessing the accuracy of the method depends highly on the quality of the annotation and one can hardly know how trustful it is. This methodological limitation should also be considered when looking at the results from the ADAM challenge itself.²²

Finally, in the first step of our method, we only used images from young subjects to build our statistical representations of retina vessels. Given that retinal vessels exhibit changes with aging (most notably a decrease in density, especially around the fovea^{43,44}), future instances of VBFL should provide increased age range representations by including older adults' images in the training set. Thanks to this individualized approach, the VBFL model would allow to apply an age-appropriate statistical representation map of vessel density and direction, taking into account potential changes in the density, direction or both of the vascular structure across the life span.

Acknowledgments

The authors would like to thank Julien Kaplan, Mateó Borlat and Nicolas Lanoux for designing the user interface of the VBFL method.

Supported by ANR (grant ANR-20-CE19-0018); PACA Region Council; and Monticelli-Paradis Ophthalmology Center (Marseille – France) – Co-funding of Microperimeter with PACA Region in 2019.

Disclosure: **A. Calabrèse**, None; **V. Fournet**, None; **S. Dours**, None; **F. Matonti**, None; **E. Castet**, None; **P. Kornprobst**, None

References

1. Tarita-Nistor L, Brent MH, Steinbach MJ, González EG. Fixation stability during binocular viewing in patients with age-related macular degeneration. *Invest Ophthalmol Vis Sci*. 2011;52(3):1887–1893.
2. Tarita-Nistor L, Gill I, González EG, Steinbach MJ. Fixation stability recording: How long for eyes with central vision loss? *Optom Vis Sci*. 2017;94(3):311–316.
3. Tarita-Nistor L, Gonzalez EG, Markowitz SN, Steinbach MJ. Fixation characteristics of patients with macular degeneration recorded with the Mp-1 microperimeter. *Retina*. 2008;28(1):125–133.
4. Calabrèse A, Bernard J-B, Hoffart L, et al. Wet versus dry age-related macular degeneration in patients with central field loss: Different effects on maximum reading speed. *Invest Ophthalmol Vis Sci*. 2011;52(5):2417–2424.
5. Tarita-Nistor L, Mandelcorn MS, Mandelcorn ED, Markowitz SN. Effect of disease progression on the PRL location in patients with bilateral central vision loss. *Transl Vis Sci Technol*. 2020;9(8):47.
6. Ahuja AK, Yeoh J, Dorn JD, et al. Factors affecting perceptual threshold in Argus II retinal prosthesis subjects. *Transl Vis Sci Technol*. 2013;2(4):1.
7. Gomes NL, Greenstein VC, Carlson JN, et al. A comparison of fundus autofluorescence and retinal structure in patients with Stargardt disease. *Invest Ophthalmol Vis Sci*. 2009;50(8):3953–3959.
8. Vullings C, Verghese P. Mapping the binocular scotoma in macular degeneration. *J Vis*. 2021;21(3):9.
9. Reinhard J, Messias A, Dietz K, et al. Quantifying fixation in patients with Stargardt disease. *Vis Res*. 2007;47(15):2076–2085.
10. Rohrschneider K. Determination of the location of the fovea on the fundus. *Invest Ophthalmol Vis Sci*. 2004;45(9):3257–3258.
11. Timberlake GT, Sharma MK, Grose SA, Gobert DV, Gauch JM, Maino JH. Retinal location of the preferred retinal locus relative to the fovea in scanning laser ophthalmoscope images. *Optom Vis Sci*. 2005;82(3):177–185.
12. Nair AA, Liebenthal R, Sood S, et al. Determining the location of the fovea centralis via en-face SLO and cross-sectional OCT imaging in patients without retinal pathology. *Transl Vis Sci Technol*. 2021;10(2):25.
13. Kang H, Lee SJ, Shin HJ, Lee AG. Measuring ocular torsion and its variations using different nonmydriatic fundus photographic methods. *PLoS One*. 2020;15(12):1–11.
14. Besenczi R, Tóth J, Hajdu A. A review on automatic analysis techniques for color fundus photographs. *Comput Struct Biotechnol J*. 2016;14:371–384.
15. Singh J, Joshi GD, Sivaswamy J. Appearance-based object detection in colour retinal images. In *2008 15th IEEE International Conference on Image Processing*; San Diego, CA, USA: IEEE; 2008; pp 1432–1435.
16. Sinthanayothin C, Boyce JF, Cook HL, Williamson TH. Automated localisation of the optic disc, fovea, and retinal blood vessels from

- digital colour fundus images. *Br J Ophthalmol*. 1999;83(8):902–910.
17. Li H, Chutatape O. Automated feature extraction in color retinal images by a model based approach. *IEEE Trans Biomed Eng*. 2004;51(2):246–254.
 18. Sagar AV, Balasubramanian S, Chandrasekaran V. Automatic detection of anatomical structures in digital fundus retinal images. Conference: Proceedings of the IAPR Conference on Machine Vision Applications (IAPR MVA2007), May 18, 2007, Tokyo, Japan. Available at: https://www.researchgate.net/publication/221280680_Automatic_Detection_of_Anatomical_Structures_in_Digital_Fundus_Retinal_Images#:~:text=This%20paper%20proposes%20a%20novel%20system%20for%20the,extraction%20of%20blood%20vessels%20and%20localization%20of%20macula;2007.
 19. Sekhar S, Al-Nuaimy W, Nandi AK. Automated localisation of optic disk and fovea in retinal fundus images. In *2008 16th European Signal Processing Conference*; 2008; pp 1–5.
 20. Al-Bander B, Al-Nuaimy W, Williams BM, Zheng Y. Multiscale sequential convolutional neural networks for simultaneous detection of fovea and optic disc. *Biomed Signal Process Control*. 2018;40:91–101.
 21. An C, Wang Y, Zhang J, Bartsch D-UG, Freeman WR. Fovea localization neural network for multimodal retinal imaging. *Proc SPIE 11511 Appl Mach Learn*. 2020;2020:196–202.
 22. Fang H, Li F, Fu H, et al. ADAM challenge: Detecting age-related macular degeneration from fundus images. *IEEE Trans Med Imaging*. 2022;41:2828–2847.
 23. Guo C, Szemenyei M, Yi Y, Wang W, Chen B, Fan C. SA-UNet: Spatial attention U-Net for retinal vessel segmentation. *ArXiv200403696 Cs Eess preprint* 2020. Available at: <https://arxiv.org/abs/2004.03696>.
 24. Tan JH, Acharya UR, Bhandary SV, Chua KC, Sivaprasad S. Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network. *J Comput Sci*. 2017;20:70–79.
 25. Li T, Bo W, Hu C, et al. Applications of deep learning in fundus images: A review. *ArXiv210109864 Cs Eess preprint* 2021.
 26. Kamble R, Samanta P, Singhal N. Optic disc, cup and fovea detection from retinal images using U-Net++ with EfficientNet encoder. In *Ophthalmic Medical Image Analysis*. Fu H, Garvin MK, MacGillivray T, Xu Y, Zheng Y, Eds.; Cham, Switzerland: Lecture Notes in Computer Science; Springer International Publishing; 2020; pp. 93–103.
 27. Xie R, Liu J, Cao R, et al. End-to-end fovea localisation in colour fundus images with a hierarchical deep regression network. *IEEE Trans Med Imaging*. 2021;40(1):116–128.
 28. Mutlu F, Leopold IH. The structure of human retinal vascular system. *Arch Ophthalmol*. 1964;71(1):93–101.
 29. Semerád L, Drahanský M. Retinal vascular characteristics. In *Handbook of Vascular Biometrics*. Uhl A, Busch C, Marcel S, Veldhuis R, Eds.; Cham, Switzerland: Springer International Publishing; 2020; pp 309–354.
 30. Orlando JI, Fu H, Barbosa Breda J, et al. REFUGE challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs. *Med Image Anal*. 2020;59:101570.
 31. Staal J, Abramoff MD, Niemeijer M, Viergever MA, van Ginneken B. Ridge-based vessel segmentation in color images of the retina. *IEEE Trans Med Imaging*. 2004;23(4):501–509.
 32. Fu HA. Automatic detection challenge on age-related macular degeneration. *IEEE Dataport*. 2020. Available at: <https://iee-dataport.org/documents/adam-automatic-detection-challenge-age-related-macular-degeneration>. <https://dx.doi.org/10.21227/dt4f-rt59>.
 33. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2020.
 34. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using Lme4. *J Stat Softw*. 2015;67(1):1–48.
 35. Zuur AF, Ieno EN, Elphick CS. A protocol for data exploration to avoid common statistical problems. *Methods Ecol Evol*. 2010;1(1):3–14.
 36. Baayen RH, Davidson DJ, Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. *J Mem Lang*. 2008;59:390–412.
 37. Gelman A, Hill J. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York, NY: Cambridge University Press; 2007.
 38. Flitcroft DI, He M, Jonas JB, et al. IMI – Defining and classifying myopia: A proposed set of standards for clinical and epidemiologic studies. *Invest Ophthalmol Vis Sci*. 2019;60(3):M20–M30.
 39. Asano S, Yamashita T, Asaoka R, et al. Retinal vessel shift and its association with axial length elongation in a prospective observation in Japanese junior high school students. *PLoS One*. 2021;16(4):e0250233.
 40. Chui TYP, Zhong Z, Song H, Burns SA. Foveal avascular zone and its relationship to foveal pit shape. *Optom Vis Sci*. 2012;89(5):602–610.

41. Bolster NM, Giardini ME, Livingstone IA, Bastawrous A. How the smartphone is driving the eye-health imaging revolution. *Expert Rev Ophthalmol.* 2014;9(6):475–485.
42. Wintergerst MWM, Jansen LG, Holz FG, Finger RP. A novel device for smartphone-based fundus imaging and documentation in clinical practice: Comparative image analysis study. *JMIR Mhealth Uhealth.* 2020;8(7):e17480.
43. Wei Y, Jiang H, Shi Y, et al. Age-related alterations in the retinal microvasculature, microcirculation, and microstructure. *Invest Ophthalmol Vis Sci.* 2017;58(9):3804–3817.
44. Kornzweig AL, Eliasoph I, Feldstein M. Retinal vasculature in the aged. *Bull N Y Acad Med.* 1964;40(2):116–129.
45. Tschumperle D, Deriche R. Vector-valued image regularization with PDEs: A common framework for different applications. *IEEE Trans Pattern Anal Mach Intell.* 2005;27(4):506–517.
46. *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations.* Aubert G., Kornprobst P., Eds.; New York, NY: Applied Mathematical Sciences; Springer; 2006. https://doi.org/10.1007/978-0-387-44588-5_3.
47. Weickert J, Werdegang W, Zucker SW, et al. Anisotropic diffusion in image processing. Research Gate. 1998, 184. Available at: https://www.researchgate.net/publication/202056812_Anisotropic_Diffusion_In_Image_Processing/link/54494eb60cf2f6388081425a/download.
48. Basser PJ, Mattiello J, Lebihan D. Estimation of the effective self-diffusion tensor from the NMR spin echo. *J Magn Reson B.* 1994;103(3):247–254.
49. Powell MJD. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Comput. J.* 1964;7(2):155–162.

Appendix

Appendix A1: Vessel Density Map Estimation

The vessel density map (\bar{V}) aims at giving, for each position (x, y) , the likelihood to have a vessel passing through.

First, we transformed each vessel map v_i so that their fovea and optic disc position be aligned with a reference fovea position (x_{ref}^F, y_{ref}^F) and optic disc position $(x_{ref}^{OD}, y_{ref}^{OD})$. This consists in computing an exact similarity transformation, accounting for transla-

tion, rotation, and uniform scaling. They are described by four parameters: $\mathcal{T} = (t_x, t_y, \theta, s)$, where t_x represents translation along the horizontal axis, t_y represents translation along the vertical axis, θ represents rotation around the reference fovea location (x_{ref}^F, y_{ref}^F) , and s represents uniform scaling around the reference fovea location (x_{ref}^F, y_{ref}^F) . This simple transformation allows for the perfect alignment of two pairs of points, denoted by \mathcal{T}_i , such that:

$$\mathcal{T}_i(x_i^{OD}, y_i^{OD}) = (x_{ref}^{OD}, y_{ref}^{OD}),$$

$$\text{and } \mathcal{T}_i(x_i^F, y_i^F) = (x_{ref}^F, y_{ref}^F).$$

Applying this transformation \mathcal{T}_i to each vessel map v_i , we obtained the resulting aligned vessel maps noted \tilde{v}_i . Next, we computed a first density estimate \bar{V}_0 by averaging the warped vessel maps \tilde{v}_i as follows:

$$\bar{V}_0 = \frac{1}{|D_{train}|} \sum_{u_i \in D_{train}} \tilde{v}_i,$$

where $|D_{train}|$ is the cardinal of D_{train} . The final average density \bar{V} was then obtained after applying an anisotropic smoothing operator followed by a simple image enhancement.⁴⁵

Appendix A2: Vessel Direction Map Estimation

The vessel direction map (\bar{D}) aims at giving, for each pixel (x, y) , the most likely direction of a vessel between 0 and π . This value is obtained by averaging vessel direction values from all aligned vessel maps \tilde{v}_i within a given region around (x, y) . To perform this averaging operation properly, we used the notion of structure tensors, which have been widely used in the domains of anisotropic diffusion^{46,47} or diffusion magnetic resonance imaging (MRI).⁴⁸ Hence, we estimated average vessel directions by averaging locally the vectors orthogonal to the gradients of our vessels maps v . To do so, we first define a local vessel direction by $t = (\nabla v_\sigma)^\perp$, where $v_\sigma = k_\sigma * v$ (i.e. after convolving v with a Gaussian kernel k_σ of variance σ which make the direction less sensitive to noise). Then, we define the vessel direction tensor by:

$$s(v) = k_\rho * (t t^T), \quad (1)$$

where the Gaussian kernel k_ρ corresponds to the size of the neighborhood region around (x, y) used to average directions. Overall, the eigen elements of the tensor $s(v)$ are informative about vessels direction distributions. For instance, in regions where vessels follow a consistent direction, tensors are characterized by

$\lambda_1 \gg \lambda_2 \approx 0$ and e_1 indicates the average direction. In regions containing multiple directions (e.g. within the optic disk) or bifurcations, tensors are characterized by $\lambda_1 \approx \lambda_2$. Finally, in regions with low vessel density (e.g. around the fovea), tensors are characterized by $\lambda_1 \approx \lambda_2 \approx 0$.

Given the tensor definition, we derived an average vessel direction tensor across D_{train} as follows. First, for each aligned vessel maps \tilde{v}_i , we estimated its vessel direction tensor $s(\tilde{v}_i)$. Second, we accumulated direction information across D_{train} by summing each vessel direction tensor pixel-wise to obtain the average vessel direction tensor:

$$\bar{S}(x, y) = \sum_{u_i \in D_{train}} s(\tilde{v}_i)(x, y), \quad \forall (x, y).$$

Each of these average tensors can be represented as an ellipse whose axes direction and size represent its eigenvectors and eigenvalues. The eigenvector e_1 ($\lambda_1 \geq \lambda_2 \geq 0$) indicates the orientation minimizing the gray-value fluctuations, that is, the direction of the vessels. Here, we only keep the direction information:

$$\bar{D}(x, y) = \arccos(e_{1,x}(x, y))$$

where $\bar{D} \in [0, \pi]$.

Appendix A3: Registration of the Statistical Representations With Respect to the Target Representations

Given the statistical representations (\bar{V}, \bar{D}) and target representations (v, d) , defined in steps 1 and 2, respectively, we register the statistical representations (where fovea position is known) with respect to the target representations. Mathematically, the registration can be formulated as an optimization problem, which aims at minimizing an energy E with respect to a transformation \mathcal{T} . This translates into:

$$\begin{aligned} \inf_{\mathcal{T}} E(\mathcal{T}), \text{ with } E(\mathcal{T}) \\ = E_V(v, \mathcal{T}(\bar{V})) + \eta E_D(d, \mathcal{T}(\bar{D})). \end{aligned} \quad (2)$$

The first term E_V penalizes an incorrect alignment of \bar{V} , with respect to v . If they are correctly aligned, areas of high density in $\mathcal{T}(\bar{V})$ should correspond to areas containing vessels in v , and conversely, areas of low density in $\mathcal{T}(\bar{V})$ should mostly correspond to empty areas in v . To ensure this, we choose a simple weighted mean squared error:

$$E_V(\mathcal{T}) = \sum_{x,y} w_V(x, y) (v(x, y) - \mathcal{T}(\bar{V})(x, y))^2, \quad (3)$$

where the weight $w_V(x, y)$ allows to give more importance to the macular region surrounding the fovea which is poor in terms of large visible vessels (see [Appendix A4](#)).

The second term E_D penalizes an incorrect alignment of \bar{D} , with respect to d . Images will be correctly aligned at a point (x, y) if the directions $d(x, y)$ and $\mathcal{T}(\bar{D}) + \theta(\mathcal{T})$ are close, modulo π . Note that $\theta(\mathcal{T})$ has to be added because we do rotations of a function representing directions. So, we defined the term E_D by:

$$\begin{aligned} E_D(\mathcal{T}) = \sum_{x,y} w_D(x, y) \sin^2(d(x, y) \\ - (\mathcal{T}(\bar{D})(x, y) + \theta(\mathcal{T}))), \end{aligned} \quad (4)$$

where the precise definition of the weight w_D is given in [Appendix A4](#).

The optimization was performed using Powell's method,⁴⁹ which is classically used in multi-modal registration. Thus, given the optimal transform \mathcal{T}^* minimizing criteria (8), the predicted fovea position (x_p^F, y_p^F) was given by: $(x_p^F, y_p^F) = \mathcal{T}^*(x_{ref}^F, y_{ref}^F)$. Although more complex transformations were considered (e.g. ones that are able to reproduce the 3D morphology of the eye), they would have resulted in longer computational time and greater risks to converge on local minima. Therefore, we decided to choose similarity transform as a good compromise between the real transformation we want to approximate and the convergence of our method.

Appendix A4: Definition of the Weights w_V and w_D Inside Energy E (2)

The weights defined in terms (3) and (4) are intended to help the registration approach to find the best solution faster.

For both weights, we first use the notion of "region-based" weight, denoted by $\lambda_{\mathcal{T}}(x, y)$, which aims at giving more importance to specific regions as defined in [Figure 9](#). First, given a vessel map v , we define its domain of definition Λ (see [Fig. 9A](#)). The domain Λ contains all positions (x, y) where fundus image is defined. Thus, it contains the vessel information, and there is no information outside Λ . Similarly, we define the domain of definition of the average density \bar{V} , denoted by $\bar{\Lambda}$ (see [Fig. 9B](#)). Here, we chose $\bar{\Lambda}$ as the smallest domain containing at least 50% of the domains $T_i(\Lambda_i)$, for all $u_i \in D_{train}$, where T_i is the transformation used to warp v_i onto \bar{V} , and Λ_i is the domain of u_i . Note that it is the same domain to consider for the distribution of directions \bar{D} . Finally, considering the average density \bar{V} , let us define a circular domain

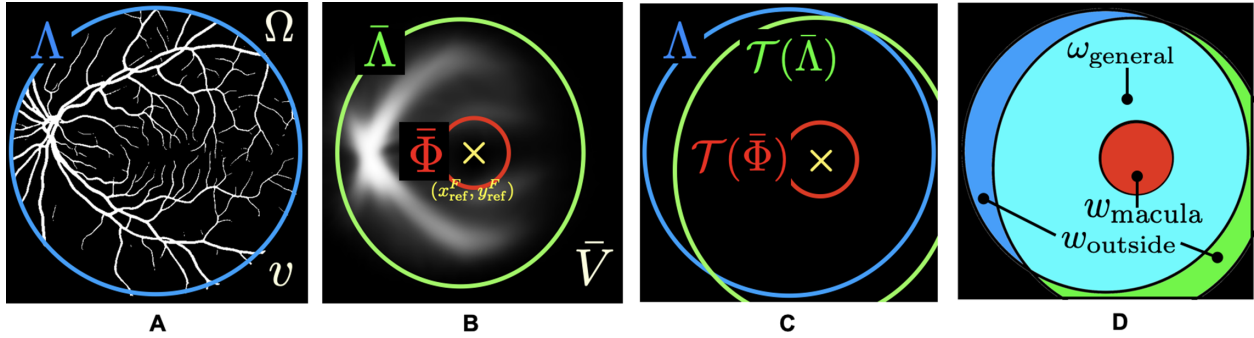


Figure 9. Definition of the region-based weight $\lambda_{\mathcal{T}}(x, y)$. **(A)** Shows the domain of definition Λ of v . **(B)** Shows the domain of definition $\bar{\Lambda}$ of \bar{V} and the macular zone $\bar{\Phi}$. **(C)** Shows the transformed domains attached to the average representation and the domain of definition of the target representation. **(D)** Shows the weight values according to the different regions.

$\bar{\Phi}$ centered around the fovea reference position (x_{ref}^F, y_{ref}^F) corresponding to the macular zone. We chose the size of this domain so that it covers the area around the fovea where there is no visible vessel (on average). Indeed, in this region surrounding the fovea, there are some macular vessels, but they are branches of the vessels of the temporal sector which divide rapidly into smaller components to create a very dense but small-diameter vascular network. In addition, the most central area (500–600 microns central, called the foveal avascular zone) has no retinal vessels at all, and it is vascularized only by the choroid.

Given these definitions, when minimizing energy E , the principle is to compare the transformed average representations (i.e. $T(\bar{V})$ and $T(\bar{D})$) with the target representations (i.e. v and d). As such, given a transformation T , we compare information defined in the transformed domain of definition ($T(\bar{\Lambda})$) and the transformed macular zone ($T(\bar{\Phi})$) with information defined in the target domain of definition (Λ). These domains are shown together in Figure 9C, which allows to define for types of regions defining the region-based weight (see Fig. 9C):

- In the region corresponding to the macular zone of the average representation ($T(\bar{\Phi})$), because we want to predict the fovea position in the target image, we choose a high weight here (denoted by w_{macula}) to encourage the alignment of the macular zones from the target image and the distribution.
- At the intersection of the two domains of definition ($T(\bar{\Lambda}) \cap \Lambda$), that is, where information is available for both sides (average and target representation), we choose a weight $w_{general}$ such that $w_{general} < w_{macula}$.
- Where information is partly available, that is, for positions (x, y) such that $(x, y) \in \Lambda$ but $(x, y) \notin T(\bar{\Lambda})$ (or vice versa), it is not possible to estimate

relevant energy because information is missing for either the average representation or the target representation. In these regions, we chose a weight $w_{outside}$ such that $w_{outside} < w_{general}$. This weight should be the lowest because it is not necessarily the case that both domains of definition should match. This is because the domain Λ is fixed and always the same. It is not related to the information shown inside. However, we found that this term was useful to avoid the algorithm diverging, especially at the early stages of the optimization, because it prevents the two domains of definition from being too distinct.

- Finally, for the positions outside the domains of definition such that $(x, y) \notin T(\bar{\Lambda}) \cup \Lambda$, the value of the weight does not matter because there is no information in this region for both target image and average distribution (so that the energy is zero).

Given this definition of the region-based weight, we can directly define the weight w_V by:

$$w_V(x, y) = \frac{\lambda_{\mathcal{T}}(x, y)}{C_{\mathcal{T}}}$$

where $C_{\mathcal{T}}$ is a normalisation coefficient ($C_{\mathcal{T}} = \sum_{x,y} \lambda_{\mathcal{T}}(x, y)$).

Similarly, we define w_D by:

$$w_D = \frac{\lambda_{\mathcal{T}}(x, y) \xi(x, y)}{C_{\mathcal{T}} \max(\xi)}$$

where an additional weight ξ based on saliency map has been added to account for regions where direction $d(x, y)$ is not well-defined (e.g. regions with no vessels or with many directions due to vessels crossings). We defined the weight ξ by: $\xi(x, y) = v(x, y)(\lambda_1(x, y) - \lambda_2(x, y))$, where $\lambda_1(x, y)$ and $\lambda_2(x, y)$ are, respectively, the higher and lower eigenvalues of the vessel tensor

$s(x, y)$. This weighted saliency map is minimal in two cases: (i) when (x, y) does not belong to a vessel (because of the multiplication by $v(x, y)$), (ii) when (x, y) belongs to a vessel but $\lambda_1(x, y) \approx \lambda_2(x, y)$, that is, when there are multiple directions in the neighborhood of $(x,$

$y)$. It is maximal when (x, y) belongs to a vessel, and $\lambda_1(x, y) \gg \lambda_2(x, y) \approx 0$, that is, along linear portions of vessels. This is where we can trust directions and we want to compare them.