Predicting risk of cardiovascular disease using retinal OCT imaging

Predicting risk of cardiovascular disease using retinal OCT imaging

Cynthia Maldonado-Garcia[a], Rodrigo Bonazzola[a], Enzo Ferrante[b], Thomas H Julian[c,d],
Panagiotis I Sergouniotis[c,d,e,f], Nishant Ravikumar[a,k], Alejandro F Frangi[g,h,i,j,k]

[a]*Centre for Computational Imaging and Simulation Technologies in Biomedicine, School of Computing, University of Leeds, Leeds, UK*
[b]*Research Institute for Signals, Systems and Comp. Intelligence, sinc(i), CONICET-UNL, Argentina*
[c]*Division of Evolution, Infection and Genomics, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, United Kingdom*
[d]*Manchester Royal Eye Hospital, Manchester University NHS Foundation Trust, Manchester, United Kingdom*
[e]*Manchester Centre for Genomic Medicine, Saint Mary's Hospital, Manchester University NHS Foundation Trust, Manchester, United Kingdom*
[f]*European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Cambridge, United Kingdom*
[g]*Division of Informatics, Imaging and Data Sciences, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, United Kingdom*
[h]*School of Computer Science, Faculty of Science and Engineering, University of Manchester, Kilburn Building, Manchester, United Kingdom*
[i]*Christabel Pankhurst Institute, University of Manchester, Manchester, United Kingdom*
[j]*NIHR Manchester Biomedical Research Centre, Manchester Academic Health Science Centre, Manchester, United Kingdom*
[k]*Indicates joint senior authors.*

Abstract

Purpose: we investigated the potential of optical coherence tomography (OCT) as an additional imaging technique to predict future cardiovascular disease (CVD).

Design: Retrospective cohort study

Participants: We employed retinal optical coherence tomography (OCT) imaging data obtained from the UK Biobank. Data for 630 patients who suffered acute myocardial infarction (MI) or stroke within a 5-year interval after image acquisition, together with an equal number of participants without CVD (control group), were used to train our model (1260 subjects in total).

Methods: We utilised a self-supervised deep learning approach based on Variational Autoencoders (VAE) to learn low-dimensional (latent) representations of high-dimensional 3D OCT images and to capture distinct characteristics of different retinal layers within the OCT image. A Random Forest (RF) classifier was subsequently trained using the learned latent features and participant demographic and clinical data, to differentiate between patients at risk of CVD events (MI or stroke) and non-CVD cases.

Main Outcome Measures: Our predictive model, trained on multimodal data, was assessed based on its ability to correctly identify individuals likely to suffer from a CVD event (MI or stroke), within a 5-year interval after image acquisition.

Results: Our self-supervised VAE feature selection and multimodal Random Forest classifier differentiate between patients at risk of future CVD events and the control group with an AUC of 0.75, outperforming the clinically established QRISK3 score (AUC = 0.597). The choroidal layer visible in OCT images was identified as an important predictor of future CVD events using a novel approach to model *explanability*.

Conclusions: Retinal OCT imaging provides a cost-effective and non-invasive alternative to predict the risk of cardiovascular disease and is readily accessible in optometry practices and hospitals.

Keywords: Retinal Optical Coherence Tomography, Cardiovascular diseases

1. Introduction

Cardiovascular diseases (CVDs) continue to pose a significant global health challenge, affecting more than 500 million individuals worldwide. Specifically, in 2021, CVDs led to 20.5 million deaths. It is concerning to observe that up to 80% of premature myocardial infarction (MI) and stroke cases could potentially be prevented if detected early. Moreover, the burden of CVDs disproportionately impacts low- and middle-income countries, where nearly four out of every five CVD-related deaths worldwide occur [23]. Currently, tools like QRISK are utilized in primary care settings by healthcare professionals to pinpoint patients at higher risk of various CVDs. These tools are commonly employed during health assessments to assess a patient's risk based on factors such as demographic details (e.g., ethnicity, age, gender), clinical indicators (e.g., cardiac volume measurements, blood markers, indicators of obesity, etc.), and socioeconomic data [22]. Notably, early identification of individuals at risk is crucial since premature CVD is highly preventable. Effective primary prevention strategies can lead to a decrease in CVD mortality and morbidity, as demonstrated in several previous clinical studies [31, 20, 10].

The retinal and choroidal microvasculature have been shown to be sensitive indicators of systemic vascular conditions, such as diseases affecting the cerebral and coronary vasculature [11]. Therefore, examining the microvasculature in the back of the eye through retinal imaging presents an opportunity to detect individuals who may be at risk of common and serious cardiovascular diseases like stroke and heart attacks. Consequently, assessing the microvasculature visible in retinal imaging non-invasively provides a feasible method for identifying instances of microvascular dysfunction in the peripheral vasculature. This, in turn, could aid in recognizing individuals at risk of cardiovascular disease [2, 11, 30, 9]. Taking advantage of its non-invasive nature and cost-effectiveness, retinal imaging is increasingly being recognized as a valuable tool for early detection of cardiovascular disease, which is crucial for delivering preventive care promptly and devising effective treatment strategies [33]. Retinal imaging techniques (e.g., fundus photography, optical coherence tomography (OCT)) are commonly used in hospital eye clinics and most optometric practices, underscoring its potential as a widely accessible method for assessing cardiovascular disease risk. The role of the retina in systemic diseases is already well established [14, 16, 35], and it is typical for eye care providers to diagnose various systemic conditions based on the appearance of the retina, such as diabetes, hypertension, and atherosclerosis. With advancements in imaging technologies improving the characterization of retinal structure, there has been a growing body of research aiming to establish connections between retinal features and cardiovascular disease risk factors, as well as to develop predictive models for early identification of patients at risk of cardiovascular disease.

Numerous prior studies have explored the application of artificial intelligence (AI) in predicting both cardiovascular disease (CVD) risk factors and CVD events through retinal imaging [35]. For example, Poplin et al. [29] demonstrated for the first time that deep learning models could anticipate cardiovascular risk factors (such as age, gender, smoking status, blood pressure, body mass index (BMI), and HbA1c levels) using fundus photographs, leveraging information on anatomical features like retinal blood vessels and optic discs. They also established the feasibility of predicting major adverse cardiac events utilizing such images. Nusinovici et al. [27] created a deep learning system (RetiAGE) to forecast age as a risk factor for various diseases, including CVD. Their research revealed that the RetiAGE marker exhibited strong discriminatory power in predicting CVD mortality among individuals who passed away within six years of image capture in the Health Screening study and had a ten-year follow-up in the UKB Biobank data. Diaz-Pinto et al. [9] introduced an innovative method for identifying individuals at risk of myocardial infarction (MI) using retinal images and basic demographic details like age and gender as predictors. Their method involved a multi-stage deep learning framework incorporating a multichannel variational autoencoder (mcVAE) designed to learn a shared latent space representing both patients' retinal images and short axis cine-cardiac magnetic resonance (CMR) images. The trained mcVAE model was then used to generate cine-CMR scans solely from previously unseen retinal images. These synthesized cine-CMR scans, along with the patients' relevant demographic and clinical data, were input into a regression network to predict left ventricular end-diastolic volume (LVEDV) and left ventricular mass (LVM). Subsequently, a logistic regression model was trained using the predicted LVEDV and LVM values, alongside other features such as patient demographic and clinical information, to identify individuals at risk of future MI events. Cheung et al. [7] developed and validated a convolutional neural network named Singapore I vessel assessment using deep-learning system (SIVA-DLS) to measure retinal-vessel caliber from retinal photographs. The system aimed to estimate central retinal artery equivalent (CRAE) and central retinal vein equivalent (CRVE), which were independently linked to an elevated risk of incident CVD events, including stroke, MI, and CVD mortality. This

study underscores the potential of employing deep learning systems to evaluate CVD risk based on retinal calibre measurements.

While retinal fundus photographs offer a two-dimensional depiction of retinal vasculature, optical coherence tomography (OCT) with its high-resolution 3D imaging capabilities allows for quantitative assessment of the thickness and structure of different retinal layers and microvasculature, providing insights not possible with fundus photography alone. The potential of 3D OCT imaging lies in its ability to detect subtle abnormalities in retinal microstructure and microvasculature that may go unnoticed in 2D images, making it a valuable tool for identifying early disease indicators [11]. This advancement in OCT technology has transformed retinal imaging by enabling visualization of the chorioretinal microcirculation, which can serve as an early sign of microvascular disease. Studies have linked choroidal thinning to elevated diastolic blood pressure [34], and in diabetic patients, OCT analysis has revealed thinning of nerve fiber and ganglion cell layers even before the onset of clinically apparent retinopathy [26]. Farrah et al. [12] have emphasized the potential of combining deep learning techniques with OCT to predict cardiovascular disease risk, leveraging OCT's ability to provide detailed cross-sectional evaluation of retinal and choroidal microvasculature, which can reflect systemic conditions. Maldonado-Garcia et al. [15] introduced a binary classification network using a task-aware Variational Autoencoder (VAE) to learn a latent representation of patients' OCT images, demonstrating superior performance in predicting future myocardial infarction risk compared to standard convolutional neural network classifiers. Zhou et al. [37] developed RETFound, a model capable of identifying disease-related patterns and diagnosing ocular conditions from fundus photographs and OCT B-scans, showing promising results in predicting neurodegenerative diseases and cardiovascular issues. Retinal imaging offers a cost-effective and non-invasive means of assessing cardiovascular risk in various healthcare settings, though further research is necessary to establish its clinical relevance.

Contributions: This study introduces a predictive model that integrates various types of data, including features derived from 3D OCT imaging through a self-supervised deep neural network, along with patient demographic and clinical details. The aim is to detect individuals at risk of MI or stroke within five years after image capture. To the best of our knowledge, this is the pioneering investigation into the application of 3D OCT imaging and artificial intelligence for automatically forecasting patients vulnerable to adverse CVD incidents. The main contributions of this research are: (i) the development of a predictive model that merges multi-modal patient data (e.g., OCT imaging, demographic and clinical information); (ii) a thorough examination of predictive accuracy within different patient subcategories (classified by attributes like age and gender) to evaluate model equity and detect any inherent biases within patient subgroups; and (iii) an innovative method for enhancing model interpretability, which offers detailed localization of the features in retinal layers that have the most impact on accurately identifying patients at risk of adverse CVD events.

2. Methodology

2.1. Database

In this research, we utilized retinal OCT imaging data sourced from the UK Biobank, captured using the Topcon 3D OCT 1000 Mark 2 system. OCT is a non-invasive imaging method that uses light waves to create detailed images of eye structures such as the retina, choroid, and optic nerve. The UK Biobank contains a vast array of health-related information from more than 500,000 participants in the UK, encompassing genetics, demographics, clinical measurements, lifestyle aspects, and medical imaging. A total of 83,940 and 84,431 participants underwent retinal imaging for their right and left eyes, respectively. To ensure only high-quality images were included, we automatically evaluated image quality using a quality index (QI) detailed in a prior study [32]. This QI is a globally accurate quality assessment algorithm derived from the intensity ratio, which is based on a histogram covering the entire image, and the tissue signal ratio, indicating the ratio of highly reflective pixels to less reflective ones. By applying a quality score threshold, 14,573 images for the right eye and 20,873 images for the left eye were excluded, leaving 69,367 and 63,558 remaining images, respectively. Among these, we identified 2,448 (left eye) and 2,228 (right eye) images from participants who had experienced a stroke or MI event, referred to as CVD+ participants. However, only images from the left eye of 875 participants and the right eye of 791 participants were taken before the CVD event. Additionally, we excluded a total of 131 patients with diabetes and/or cardiomyopathy for the left eye and 121 patients for the right eye. A visual representation of the participant selection and exclusion criteria used to establish the cohort for this study is depicted as a STROBE diagram in Fig.1.
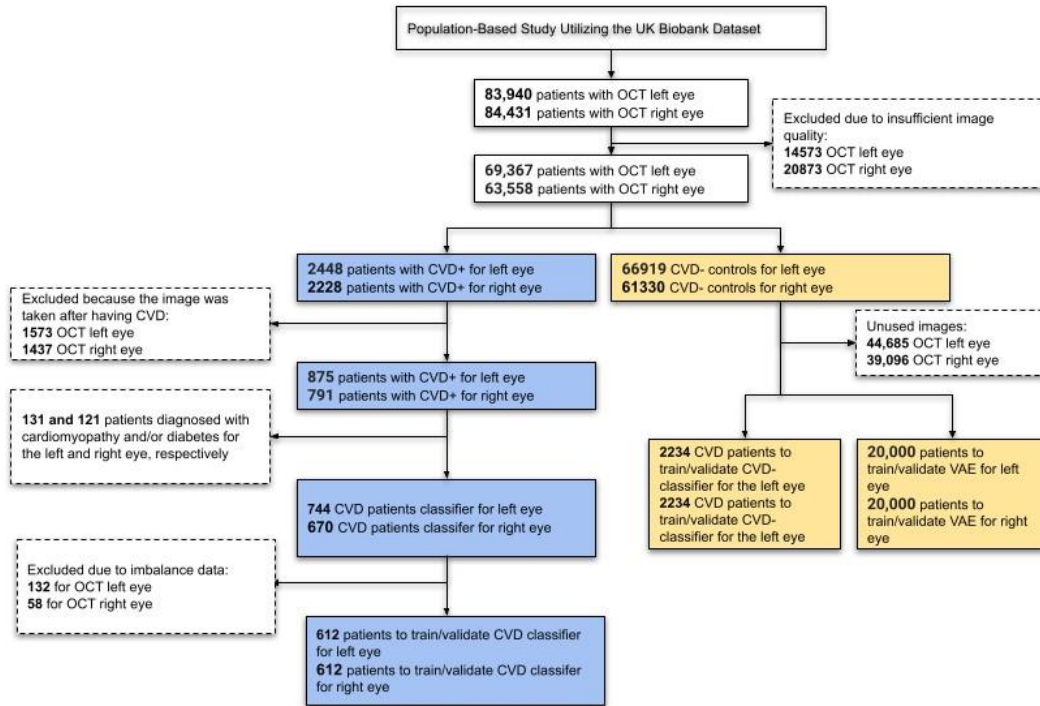
Figure 1: STROBE flow chart describing participant inclusion and exclusion criteria applied to define the study cohort.

The size of the CVD+ group was determined based on the application of specific inclusion and exclusion criteria, as outlined in the STROBE diagram in Fig.1, resulting in a final count of 612 participants with OCT images of both eyes. For the non-CVD or CVD- group, 2234 participants were randomly selected for OCT images of both eyes. The essential patient characteristics used to match the CVD+ and CVD- groups included demographic factors and clinical measurements, which are detailed in Table 1. The average age of individuals with and without CVD was 60.78 ± 6.47 years, showing no significant difference between the two groups. The majority of participants in the UK Biobank cohort were of white ethnicity, with similar proportions in both groups. The average body mass index (BMI) was 28.31 ± 4.45 kg/m^2 for those with CVD and 27.43 ± 4.33 kg/m^2 for those without. In terms of blood pressure readings, individuals with CVD had a systolic blood pressure (SBP) of 147.26 ± 19.57 mm Hg, while those without CVD had an SBP of 145.1 ± 18.75 mm Hg. The diastolic blood pressure (DBP) was 84.75 ± 10.23 mm Hg for individuals with CVD and 83.22 ± 9.73 mm Hg for those without. The mean level of haemoglobin A1c (HbA1c) was 36.52 ± 4.32 mmol/L for individuals with CVD and 36.59 ± 6.61 for those without. A notable percentage of participants reported being current alcohol consumers, accounting for 90.69% of individuals with CVD and 91.83% of those without. These participant characteristics for both groups are summarized in Table 1.

We used an age-sex-matched cohort for controls in CVD and non-CVD scenarios with a ratio of 1:3, respectively. One significant benefit of employing an age-sex-matched study cohort is that it helps mitigate the effects of confounders that might heavily influence the predictive model. Machine learning models are prone to capture spurious correlations (i.e. to learn shortcuts) between predictors and targets, such as, for example, linking age or gender to the presence of pathology, unless care is taken when defining the predictors and study cohort [5]. Figure 2 illustrates the

Characteristics	CVD+	CVD-
Number of subjects	612	2234
Age: Mean (s.d), years	60.78 (6.47)	60.78 (6.47)
Gender: F, M %	29.74, 70.26	29.74, 70.26
Ethnicity, %	90.18 White, 4.26 Mixed, 3.93 Asian or Asian British, 0.33 Black or Black British, 0.16 Chinese, 1.15 Other ethnic group	89.22 White, 4.25 Mixed, 4.41 Asian or Asian British, 0.82 Black or Black British, 0.49 Chinese, 0.82 Other ethnic group
BMI: Mean, kg/m^2	28.31 (4.45)	27.43 (4.33)
SBP: Mean, mm Hg	147.26 (19.57)	145.1 (18.75)
DBP: Mean, mm Hg	84.75 (10.23)	83.22 (9.73)
HbA1c: Mean, mmol/mol	36.52 (4.32)	36.59 (6.61)
Alcohol consumption: N, P, C, NA %	3.59, 5.72, 90.69, 0	3.92, 3.92, 91.83, 0.33

Table 1: Characteristics of patients in the CVD+ and CVD- sets. N, Never. P, Previous. C, Current. NA, Not answer
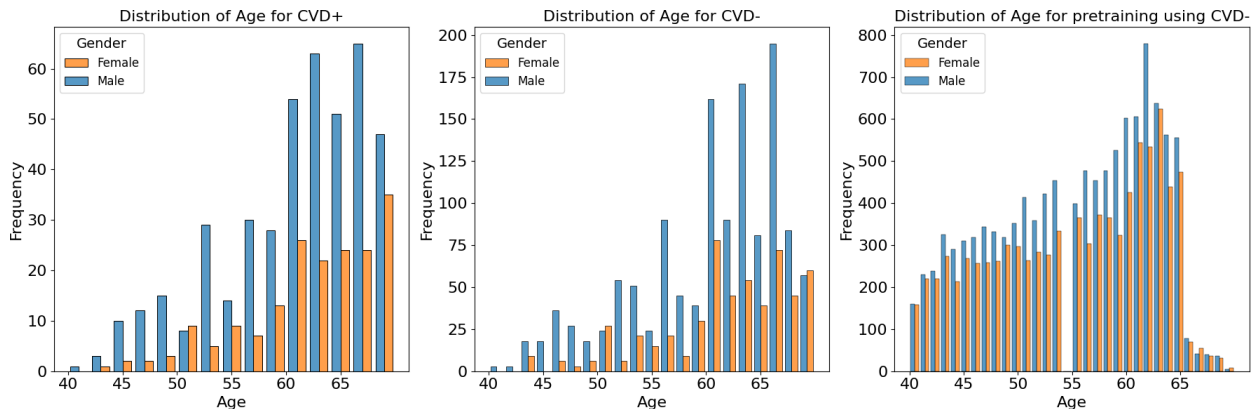

Figure 2: Distribution of age-sex cohort match. M, Male. F, Female. The left histogram illustrates the total of CVD+ labeled data used solely for the classification task. The middle histogram shows the total of CVD- for the classification task, while the right histogram illustrates CVD- used for the pretraining task.

age and sex in the CVD+ and CVD- patient groups, which were used to train and evaluate the predictive model proposed in this study. The construction of the metadata incorporated eight clinical variables, namely sex, age, HbA1c, systolic and diastolic blood pressure, alcohol consumption, and body mass index. The decision was made to omit the smoking variable from the study analysis because a significant number of participants did not provide responses to the relevant questionnaire item. From now on, the term "metadata" will be used to describe patient details, including demographic and clinical history information.

2.2. Framework of Self-Supervised Feature Selection VAE and Multimodal Random Forest Classification

This study proposes a predictive model for classifying patients into CVD+ and CVD- categories, comprising a Variational Autoencoder (VAE) [19] to extract features in a self-supervised manner from retinal OCT images, and a

Random Forest (RF) classifier which combines the former with patient metadata and uses the resulting set of multi-modal features as predictors. The proposed model consists of two stages, a self-supervised feature extraction stage, and a subsequent classification stage. A schematic diagram of the overall predictive framework is shown in Figure 3.

2.2.1. Self-supervised VAE

In the first stage of the proposed model, a VAE is used to learn latent representations for B-scan OCT images. VAEs are a widely used type of generative model in neural networks, comprising a pair of networks or network branches that are trained together, called the encoder and decoder networks/branches. Given some input data ($\mathbf{x}$), the encoder is designed to approximate the posterior distribution of the latent variables ($q_\phi(\mathbf{z}|\mathbf{x})$), under some assumed prior distribution ($p(\mathbf{z})$) over the latent variables (typically, a multivariate Gaussian prior is used, that is, $p(\mathbf{z})$ ( )), while the decoder is trained to reconstruct the input data by sampling from the approximated posterior distribution ($p_\vartheta(\hat{\mathbf{x}} \mid \mathbf{z})$. In other words, the encoder network maps inputs to low-dimensional latent representations, and the decoder network acts as the generative model. The approximation of the true but intractable posterior distribution is obtained by maximising the lower bound of the evidence (ELBO), which can be expressed as follows:

$$\textbf{ELBO} = \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})} \left[\log p_\vartheta(\mathbf{x} \mid \mathbf{z}) \right] - D_{KL} \left(q_\phi(\mathbf{z} \mid \mathbf{x}) || p(\mathbf{z}) \right) \tag{1}$$

The loss function utilised for training the proposed self-supervised VAE comprises two key elements: (1) the loss of mean square error (MSE) L_{MSE}, detailed in Equation 3, which evaluates the discrepancy in reconstruction between the original data ($\mathbf{x}_i$) and the reconstructed data ($\hat{\mathbf{x}}_i$), and (2) the loss of Kullback-Leibler divergence L_{KL}, illustrated in Equation 4. KL divergence quantifies the dissimilarity between the learnt latent distribution and a previously specified distribution ($p(\mathbf{z})$), which, in this scenario, is a multivariate Gaussian distribution. By minimising the KL divergence, the model is incentivised to shape a latent space that adheres to the target Gaussian distribution. The integration of these two loss components steers the VAE towards the dual objective of reducing reconstruction errors and aligning the learnt latent distribution with the intended prior distribution.

$$\mathsf{L}_{VAE} = \mathsf{L}_{MSE} + \beta \mathsf{L}_{KL} \tag{2}$$

$$\mathsf{L}_{MSE} = \frac{1}{N} \sum_{i=1}^{N} (\mathbf{x}_i - \hat{\mathbf{x}}_i)^2 \tag{3}$$

$$\mathsf{L}_{KL} = \frac{1}{2} \left(1 + \log \sigma^2_i - \sigma^2_i - \mu^2_i \right) \tag{4}$$

We trained independent VAEs for each eye, to learn unique latent features from the OCT images. Subsequently, a classifier used these learnt features to predict the probability of an individual's prospective CVD incidence.

2.2.2. Classification

Using the features acquired from the VAE in the previous stage, we trained a Random Forest (RF) classifier to distinguish between individuals in the CVD+ and CVD- categories, as illustrated in Figure 3. The input for this process consists of the latent vector representation of each OCT image generated by the VAE for each eye, which is merged with a vector containing the relevant patient information. Random forests are a type of ensemble machine learning method that involves multiple decision trees, each of which is trained on a randomly selected subset of training data [4]. Using the power of numerous decision trees and incorporating random feature selection, this ensembling technique enhances the generalisability of the predictive model to new data by reducing model variance by averaging predictions from the trees in the ensemble. RFs have been widely applied in medical settings for both classification and regression tasks, including in previous studies related to CVD diagnosis [18, 36]. One notable advantage of RFs compared to other classification algorithms is their ability to easily handle multimodal data that include various data types (such as categorical, ordinal, and continuous). Decision trees within the ensemble operate independently, and their combined predictions are aggregated to produce the final RF prediction for a given input using majority voting for classification tasks.
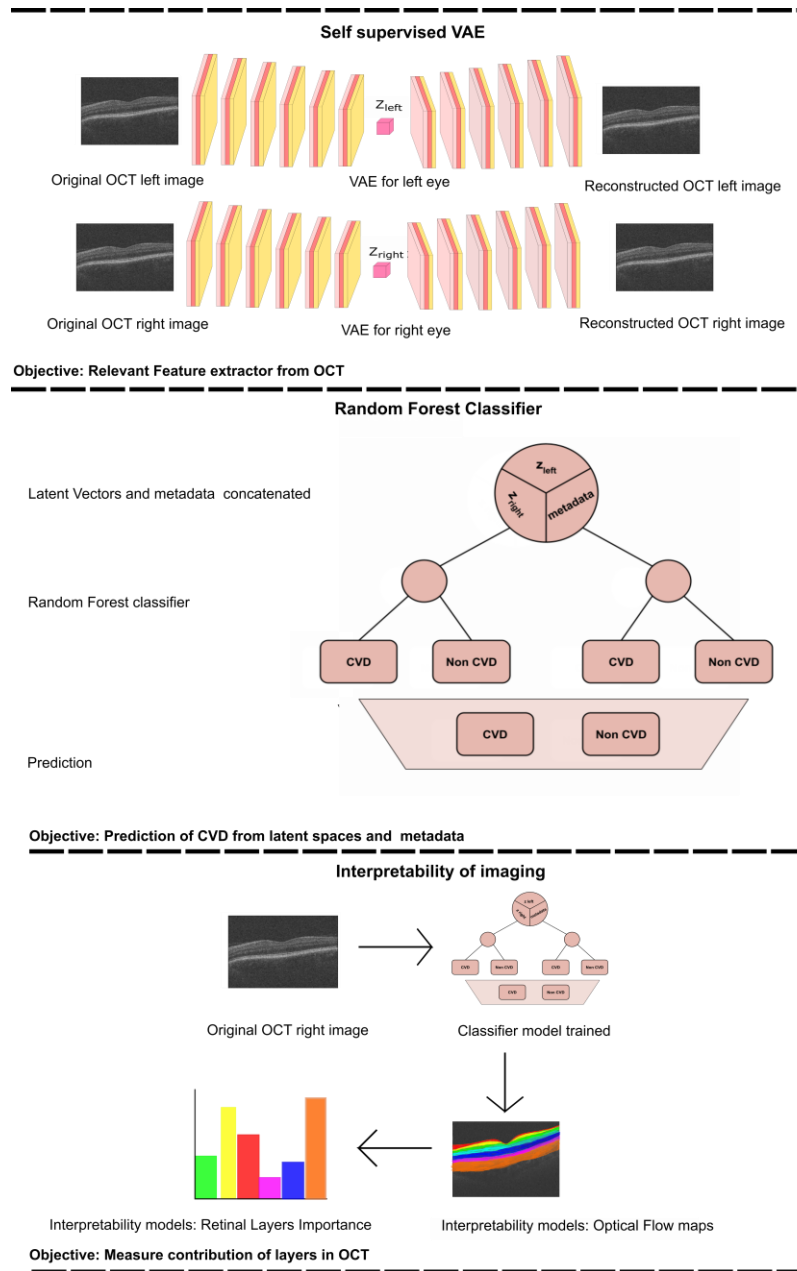
Figure 3: The provided workflow diagram illustrates the comprehensive process of training the Variational Autoencoder (VAE) and subsequently using it to acquire the latent vectors (upper section). These latent vectors are then combined with metadata and serve as inputs to the Random Forest (RF) classifier (middle section). Finally, we perform an interpretability analysis by perturbing the most relevant features, reconstructing the corresponding image and computing the optical flow between the perturbed reconstructions (lower section). z_{left} represents the latent vector obtained from the training of the VAE for the left eye. Z_{right} corresponds to the latent vector acquired from training the VAE for the right eye.

2.3. Experiment Details

All experiments were carried out with an NVIDIA Tesla M60 GPU. The model was trained using PyTorch (v1.10.2) and a grid search strategy was used with five-fold cross-validation to determine the best hyperparame-

ters. The data set was divided into training, validation, and test sets in a ratio of 6: 2: 2. The encoder and decoder networks were constructed with six 2D convolution layers each; the encoder used Rectified Linear Unit (ReLU) activations, while Leaky Rectified Linear Unit (LeakyReLU) activations were used in the decoder (see the upper section of Figure 3). The training of our model occurred in two phases, initially engaging in self-supervised learning for each eye independently to acquire latent representations, which were then utilised to initialise the subsequent fine-tuning phase.

During the classification phase, we conducted a thorough investigation into the impact of latent representations derived from the OCT images of the left and right eyes, along with patient metadata, on the predictive task. This was accomplished by generating seven distinct datasets from the same group of 2846 patients, each comprising different combinations of data sources: (i) latent representations from the left eye only (LE); (ii) latent representations from the right eye only (RE); (iii) latent representations from both eyes (BE); (iv) metadata only (MTDT); (v) left eye with metadata (LE-MTDT); (vi) right eye with metadata (RE-MTDT); and (vii) both eyes with metadata (BE-MTDT). Random Forest classifiers were trained separately on each of these seven datasets, as depicted in Figure 3 for the BE-MTDT dataset. Finally, the optimal hyperparameter values of RF classifiers were determined through a combination of grid search and empirical experimentation, to identify the best performing RF model for each specific dataset. We divided the dataset into training, validation, and test sets, following a distribution ratio of 5 : 2 : 3, respectively (resulting in, 1882 patients in the training set and 964 patients in the held-out unseen test set). Grid search was performed using five-fold cross-validation, while an independent, unseen test set remained fixed throughout all experiments to evaluate all trained classifiers fairly. Furthermore, a feature selection method using Recursive Feature elimination (RFE) was employed to mitigate overfitting and train the model with the most relevant variables for classification. In the majority of cases, the model was trained with the top 10 most significant features, with the exception of the RE case, where, we used only 5 variables to avoid overfitting.

To evaluate the effectiveness of our model, we compared predictive performance against the QRISK3 algorithm, the current gold standard used by healthcare professionals / cardiologists to assess the patient's risk of stroke or heart attack (acute myocardial infarction), in a 10-year period. The QRISK score was calculated within the specified test dataset, following the methodological guidelines outlined in [22]. The evaluation of the QRISK3 score involved entering essential variables such as the cholesterol-to-HDL ratio, age, SBP, standard deviation of SBP, smoking status, BMI, Townsend score, sex, weight, height, and ethnicity based on available data. In cases where certain information was missing, we consistently represented these gaps as '0' when evaluating the QRISK3 scores on the test dataset. For our classification task, we evaluated the model performance using a range of metrics. Accuracy, precision, sensitivity, and specificity were determined by calculating true positives, true negatives, false positives, and false negatives (using a classification probability threshold of $t = 0.5$, i.e. if the predicted probability is ≥ 0.5, the patient is classified as CVD+, else as CVD-). The receiver operating characteristic (ROC) curve was constructed by computing the true positive rate and false positive rate. The area under the ROC curve (AUROC) was then employed as a performance measure to assess both our model and the QRISK3 algorithm.

2.4. Model Explainability

Predictive models based on machine/deep learning algorithms, proposed for identifying risk of disease from medical imaging often fail to report both 'local' and 'global' explanations for the model's predictions. This is especially prevalent in the case of deep learning-based approaches that are often treated as black boxes, with little information provided on the mechanism by which models arrive at specific predictions. Local explanations provide insights to individual decisions/predictions of the model. For example, this may involve identifying specific input variables/regions of an image that had the most influence on the model's prediction for that instance. On the other hand, global explanations describe the model's behaviour across predictions for all instances in all classes of interest. Specifically, global explanations provide information on the most common discriminative features identified by the model for all instances in each class of interest. Providing both local and global explanations of model behaviour is essential for developing responsible AI in healthcare applications, as it can help identify systematic biases in data and mitigate for the same (e.g., learning of 'short-cuts' is a common issue encountered in the application of deep learning-based methods for predictive tasks using medical images), build trust in AI systems by improving transparency in model decision making, and may even provide new insights to previously known associations between image-derived phenotypes and the presence or progression of diseases. Therefore, in this study, we employ distinct techniques to provide both local and global explanations for the proposed predictive model.

To provide local explanations of the behaviour of the model and elucidate how the model uses OCT imaging-derived features to classify instances in the CVD+ group, we first selected the best performing RF classifier according to all classification metrics used to evaluate and compare all investigated models. Subsequently, based on the selected RF classifier, we identified the latent variable derived from the OCT image with the highest importance assigned by the RF, which we denoted z_{max}. To visually assess the regions of the retina in OCT images that contribute significantly to the prediction of CVD, we propose a novel optical flow-based latent traversal approach that evaluates the impact of perturbing the most important latent feature z_{max} on subsequent reconstructions of OCT images. Specifically, given an image $\mathbf{x}$, we compute the corresponding latent vector $\mathbf{z}$ using the trained encoder network. Next, we perturb only the dimension z_{max} of the calculated latent vector and reconstruct the corresponding image. This perturbation is performed by multiplying the latent dimension z_{max} by the standard deviation of this latent component calculated throughout the training population, defined as σ_{max}. The remaining latent variables in the computed latent vector are left unchanged, resulting in a perturbed latent vector ($\hat{\mathbf{z}}$). Finally, we reconstruct the input OCT image ($\hat{\mathbf{x}}$) using the perturbed latent vector $\hat{\mathbf{z}}$.

To visualise the regions in the reconstructed OCT images affected by the altered latent vector $\hat{\mathbf{z}}$, we examine the variances between the initial image $\mathbf{x}$ and the reconstructed image $\hat{\mathbf{x}}$ derived from $\hat{\mathbf{z}}$. In this context, we calculate the optical flow between these images using the Lucas-Kanade algorithm [24]. The resulting optical flow, showing the magnitude of the displacement vector for the moving pixels, was then superimposed on the original image, as shown in Figure 7. The optical flow forms a vector field that indicates how the pixels between the images (that is, $\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}$) change due to the latent traversal from $\mathbf{z}$ to $\hat{\mathbf{z}}$. The estimated vector field between $\mathbf{x}$ and $\hat{\mathbf{x}}$ helps to visually illustrate the regions in the OCT image that were altered by modifying the latent component z_{max}. This aids in pinpointing the areas of the image influenced by alterations to the crucial latent variable for accurately classifying a patient's CVD risk based on their OCT image(s), and consequently, helps to understand which retinal areas are informative for distinguishing between the CVD+ and CVD- patient groups.

To provide global explanations of the behaviour of the model, we calculate the importance of the characteristics assigned to each characteristic by the RF in each predictive model investigated. As mentioned previously, the RFs in each predictive model were trained using a reduced set of characteristics identified by RFE. Feature importance is calculated in RFs as the average Gini information gain for any given feature, calculated across all decision trees in the forest. Feature importance values estimated by all seven classifiers investigated in this study, across all instances in the test set, are summarised as bar plots in Figure 5. Additionally, we calculate the relative importance of the type / channel of data used as inputs/predictors in this study, namely, OCT images of the left and right eye and patient metadata, for the task of distinguishing between the CVD+ and CVD- groups (summarised in Figure 6.

3. Results

3.1. Classification Performance

As discussed previously, we trained and evaluated the performance of several RF classifiers, where each classifier was trained and evaluated independently using seven different combinations of data types obtained from the same set of patients (comprising CVD+ and CVD- groups). Specifically, the datasets used were LE, RE, BE, LE-MTDT, RE-MTDT, BE-MTDT and MTDT. Henceforth, for brevity, we refer to classifiers trained and evaluated on these datasets as LE-RF, RE-RF, BE-RF, LE-MTDT-RF, RE-MTDT-RF, BE-MTDT-RF, and MTDT-RF. The performance of all seven classifiers was evaluated and compared using the same unseen test set (which contains 964 patient data, 834 CVD- and 130 CVD+), and using the same set of evaluation metrics outlined in Section 2.3. The rationale for comparing all seven classifiers against each other was to - (i) assess whether combining latent features learnt from OCT images of both eyes (BE) provided greater discriminative power than using those from either left (LE) or right eye (RE) alone; (ii) compare the discriminative power of OCT image-derived latent features against patient metadata; and (iii) evaluate the discriminative power gained by enriching OCT image-derived latent features with patient metadata.

The performance of all seven classifiers on the unseen test set is summarised in Table 3. These results show that the BE-MTDT-RF classifier consistently outperformed all six other classifiers, indicating that combining information from OCT images of both eyes (ie, learnt latent representations) with patient metadata was more informative in distinguishing between the CVD+ and CVD- groups. In terms of evaluating the effectiveness of OCT image-derived characteristics and metadata information for classifying CVD+ and CVD- patients, results for the BE-MTDT-RF, LE-MTDT-RF, RE-MTDT-RF and MTDT-RF classifiers indicate that combining latent features learnt from OCT images
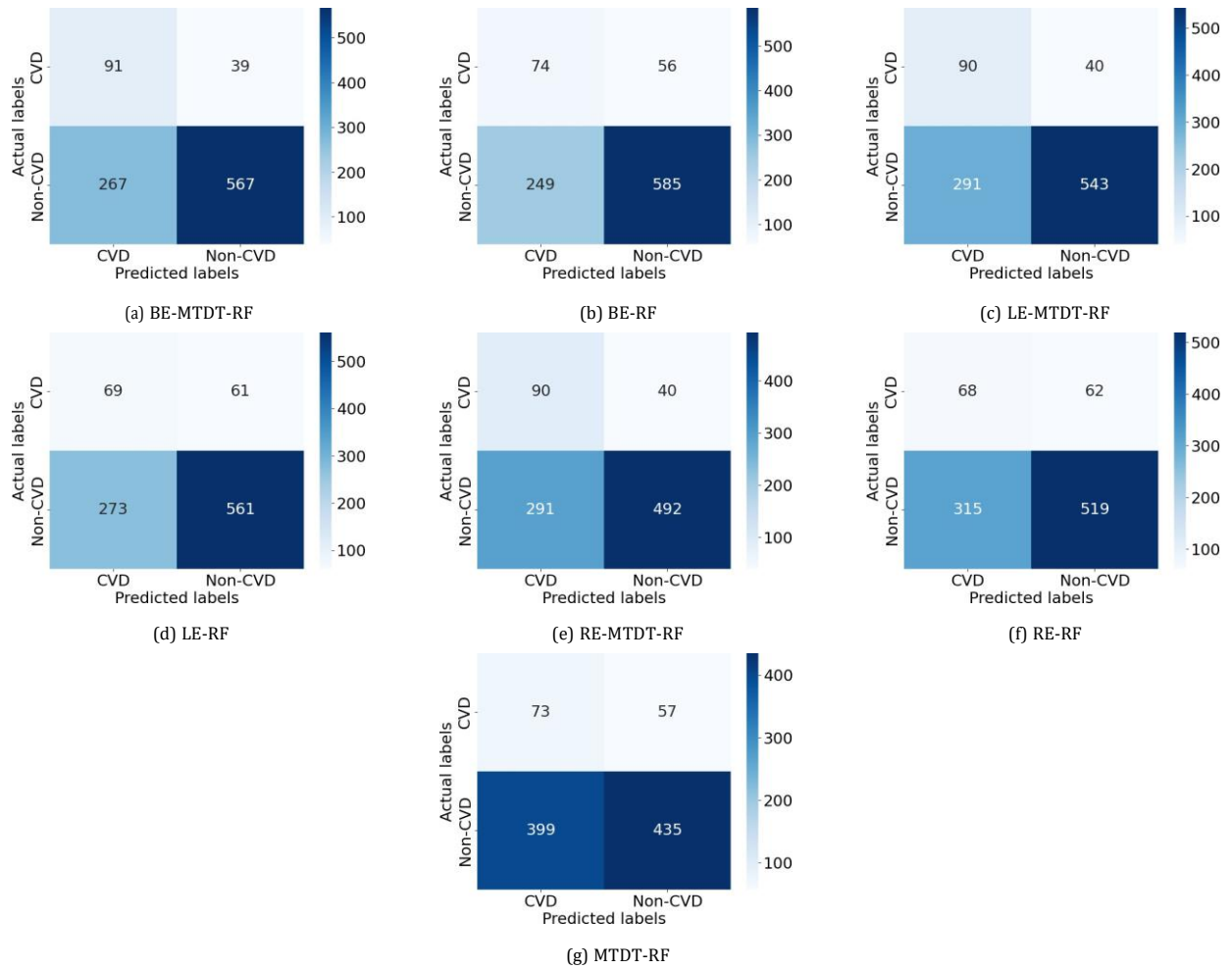
Figure 4: Confusion matrix for all seven classifiers.

Classification metrics				
Modality	Accuracy	Sensitivity	Specificity	AUC
LE-MTDT-RF	0.62	0.7	0.63	0.72
RE-MTDT-RF	0.66	0.7	0.59	0.7
BE-MTDT-RF	**0.69**	**0.71**	**0.68**	**0.75**
MTDT-RF	0.52	0.52	0.46	0.54
LE-RF	0.62	0.56	0.58	0.64
RE-RF	0.54	0.52	0.54	0.62
BE-RF	0.6	0.57	0.61	0.67

Table 2: Predictive analysis of cardiovascular disease (CVD) metrics utilizing UK Biobank data across seven distinct cases.

with patient metadata (i.e. BE-MTDT-RF, LE-MTDT-RF, RE-MTDT-RF), consistently improves classification performance relative to using patient metadata alone (MTDT-RF). Specifically, the BE-MTDT-RF classifier performed the best, achieving an accuracy of (0.69), sensitivity (0.71), specificity (0.68), and AUC score of (0.75). Furthermore, combining latent features from RE or LE OCT images with metadata (ie LE-MTDT-RF and RE-MTDT-RF) achieved 3‑19% improvements in all classification metrics relative to the MTDT-RF classifier. cNotably, the MTDT-RF classifier exhibited the lowest sensitivity (0.52) and AUC score (0.54) of all seven classifiers investigated. The BE-RF, RE-RF, and LE-RF classifiers also consistently outperformed the MTDT-RF classifier in all classification metrics; however, they performed worse than their counterparts that combined OCT image-derived characteristics with patient metadata (ie BE-MTDT-RF, LE-MTDT-RF, RE-MTDT-RF). A significant observation in the results is that including both eyes was advantageous for both cases, BE-MTDT-RF and BE-RF, compared to their counterparts, LE-MTDT-RF and RE-MTDT-RF, RE-RF, and LE-RF, respectively. Furthermore, in both scenarios, with/without metadata, the left eye consistently provided improved classification performance compared to the right eye. This observation is in concordance with the global explanations of models' predictions summarised in Figure 5, wherein features attributed to the left eye were found to be more discriminative (i.e., had higher feature importance) than those of the right eye. We posit that the impact on the latent vectors associated with the left eye is interconnected with the superior image quality of the images of the left eye within our cohort 8. The UKBB standard operating procedure stipulated that the second eye imaged was consistent with the left eye. This protocol was not randomised. As a result, there may be potential systematic disparities between left and right eye OCTs (for example, left eye scans might consistently exhibit better quality because they are the second scan performed and patients are potentially more adept at following instructions). These collective findings underscore the improved performance achieved by integrating retina OCT imaging and metadata in the classification task.

Figure 4 presents confusion matrices for the seven classifiers investigated. A consistent observation across all our results is that the BE-MTDT-RF classifier misclassified fewer instances in the CVD+ (39 out of 130) group, than all other classifiers, which is consistent with the classification metrics summarised in Table 3. Similarly, the LE-MTDT-RF and RE-MTDT-RF classifiers exhibited good sensitivity by incurring few false negative errors, i.e. 40 cases out of 130 instances in the CVD+ group were incorrectly classified. Although the MTDT-RF classifier incurred marginally fewer false negative errors (56) than classifiers utilising only OCT features (LE-RF with 62 and RE-RF with 61), the former misclassified a significantly higher number of CVD- cases, i.e. incurred a significantly higher number of false positive errors than the latter.

Subsequently, we conducted a rigorous comparative analysis between the best classifier identified from the previous experiments, namely BE-MTDT-RF, and the QRISK3 algorithm, the current clinical standard for assessing patients at risk of stroke or acute myocardial infarction. As elucidated in Table 3, our model showed superior classification performance in terms of accuracy, sensitivity, and AUC. Specifically, for the QRISK3 model, performance metrics were as follows: accuracy (0.553), sensitivity (0.6), specificity (0.545), and AUC (0.597). In contrast, BE-MTDT-RF achieved the following metrics: accuracy (0.69), sensitivity (0.71), specificity (0.68), and AUC (0.75).

Classification metrics				
Modality	Accuracy	Sensitivity	Specificity	AUC-Value
BE-MTDT-RF	**0.69**	**0.71**	**0.68**	**0.75**
QRISK	0.553	0.6	0.545	0.597

Table 3: Comparison of classification metric results between our model employing both ocular data and metadata (BE-MTDT) and the QRISK algorithm.

This discrepancy in performance at separating the CVD+ and CVD- patient groups underscores the improved discriminative capacity provided by our model which combines retinal OCT imaging data with basic patient information (see Table 1), relative to using an extensive set of demographic and clinical variables (refer to Table xx in the Appendix for more information on the variables of the patients required to calculate the QRISK3 scores), as in the case of the widely used QRISK3 algorithm.

3.2. Model Explainability

In order to provide global explanations for the behavior of all classifiers investigated in this study, we analysed the most important features identified by each model (refer to Figure 5) for distinguishing between the CVD+ and CVD- groups. Important features identified for the best performing classifier, namely, BE-MTDT-RF in particular, provided some noteworthy insights. As highlighted in Figure 5(a), we found that a latent variable learned from the left-eye OCT image, denoted zl066, had the most influence on the classifier's ability to separate CVD+ and CVD- patient groups. Additionally, among the top 10 most important features identified for the BE-MTDT-RF classifier, 9 of the features pertained to latent variables learned from the left-eye OCT image. BMI was the only feature from the basic set of patient metadata used to train the classifier, that was identified to have a significant influence on the classifier's predictions. Furthermore, looking at the global explanations summarised in Figure 5(b), (e) and (f), we observe that latent variable zl066 consistently ranks among the top two most important features for the LE-MTDT-RF, BE-RF and LE-RF classifiers, respectively. This indicates that the retinal features encoded by zl066 are consistently considered to be relevant across all four classifiers presented in Figure 5(a), (b), (e) and (f). Among the classifiers which combined retinal OCT-image derived features with patient metadata, namely, BE-MTDT-RF, LE-MTDT-RF and RE-MTDT-RF, we observed that only two features, namely, BMI and HbA1c, ranked among the top 10 most important features for the classification task. Both features are known and established cardiovascular risk factors, and importantly, we infer from these results that the latent variables learned from the retinal OCT images, had a greater influence on the classifiers' predictions than the patient metadata variables. As previously highlighted, we hypothesize that the significant importance of the latent variables corresponding to the left eye can be attributed to the superior image quality of left-eye OCT images within our cohort (as illustrated in Figure 8 in the Appendix). As a result, corresponding latent vectors **z** effectively capturing image features that potentially enhance predictive capabilities.

Using the insights gained from analysing the global explanations of classifier behaviour summarised in Figure 5, we propose a novel approach based on latent space traversals to translate the former into local explanations that provide insights to regions of the OCT image that contain relevant information for correctly identifying patients at risk of cardiovascular disease. Specifically, having identified latent variable zl066, derived from left-eye OCT images, as being the most important feature for classification, our local explainability approach (refer to section 2.4) begins by perturbing the values of the latent variable for any image in the CVD+ group, reconstructs the OCT image using the perturbed latent representation (using the pretrained VAE) and then estimates optical flow maps between the original and perturbed OCT image reconstructions, seeking to pinpoint the specific image regions that change as a result of the perturbation. We conducted a qualitative analysis involving estimation of optical flow maps between the original reconstructed B-scan OCT images and their perturbed counterparts. This analysis was specifically focused on representative B-scan examples from 5 patients correctly diagnosed with CVD, where each patient corresponds to a row in Figure 7. We considered 3 B-scans per patient, including the 1st one (upper row), 64th (middle row) and 128th (bottom row). Finally, by overlaying the estimated optical flow maps onto the original OCT images, we obtained visual interpretations of the retinal image features encoded by the latent variable of interest.
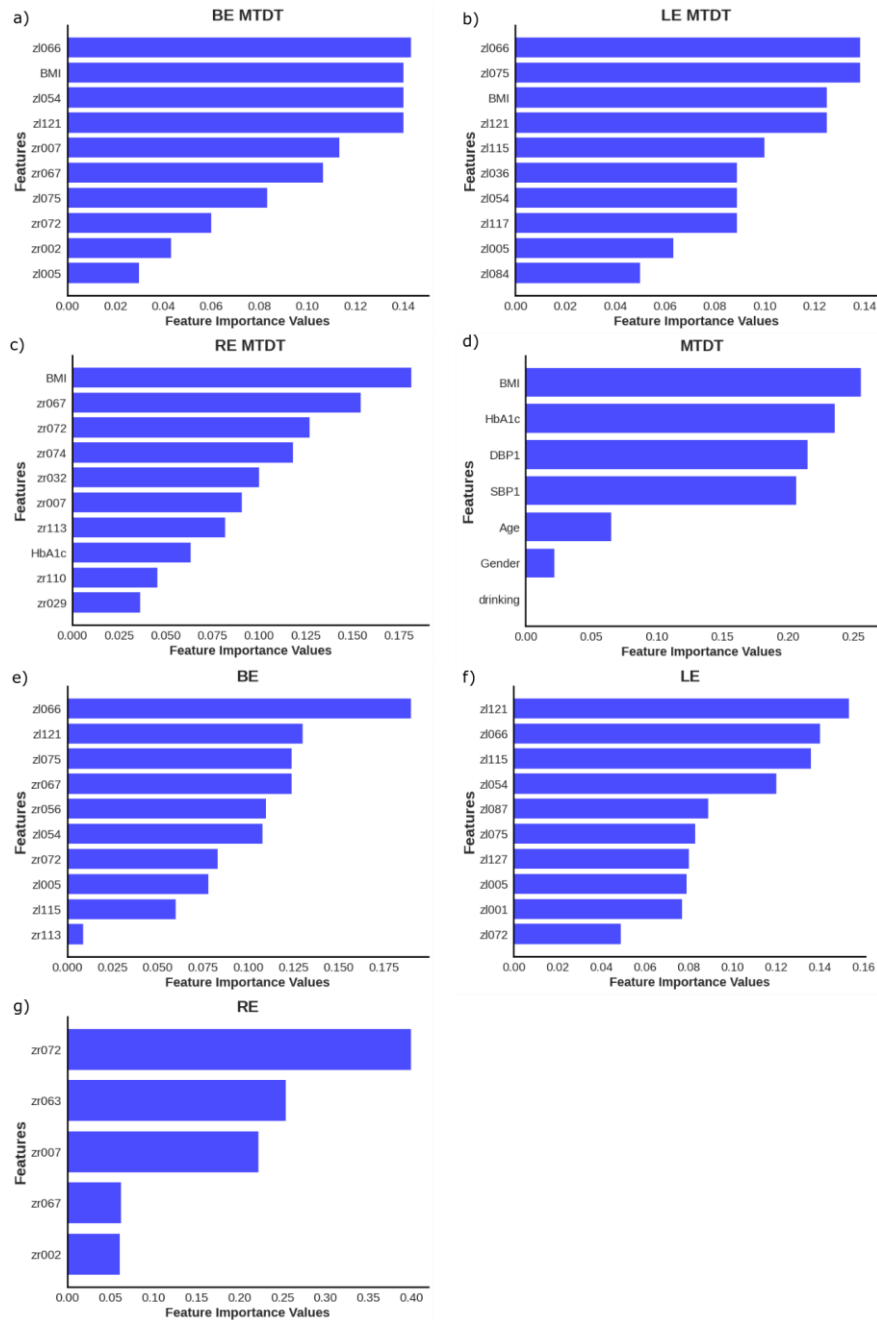
Figure 5: Calculation of feature importance magnitudes for the seven different classifiers investigated, where each classifier uses different combinations of data data channels/modalities.

The optical flow maps generated by our model predominantly accentuated the choroidal layer in the majority of B-scans, with additional identification of layers adjacent to the choroid, including the retinal pigment epithelium (RPE). Additionally, the optical flow maps highlighted the regions in the inner retinal layers, likely corresponding to the retinal nerve fiber layer (RNFL) and ganglion cell layer (GCL) [37]. Although some other layers received comparatively less emphasis, the main focus remained on the choroidal layer and the innermost layers. The optical
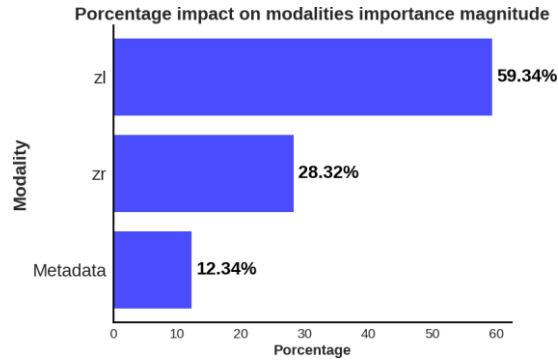
Figure 6: Global explanations of features from different data modalities/channels which were considered important by the predictive model for separating the CVD+ and CVD- groups. Bar plot summarises the relative importance of latent variables from left (zl) and right (zr) eye OCT images and patient metadata, as percentages.

flow maps provide precise localisation of the image regions modelled by latent variable zl066 (visualised as landmarks, as shown in Figure 7), thereby providing local explanations for the most discriminative regions within the OCT, and providing insights to which retinal layers may contain relevant information for predicting risk of CVD in patients. In particular, these local explanations highlighted the relevance of information contained within the choroidal layer of the retina, for distinguishing between CVD+ and CVD- patient groups. Such findings have significant implications for understanding the potential association between changes to the choroidal layer in the retina and the onset and progression of cardiovascular diseaes.

4. Discussion

Our findings indicate that the use of retinal OCT images in conjunction with VAE and multimodal RF classification has potential to improve the ability to identify patients at risk of CVD (within a five-year interval), relative to the use of the current clinical standard, that is, the QRISK3 score. This approach can serve as a complementary strategy to existing clinical procedures used for the prevention of primary CVD [3]. Our investigation included the deployment of a self-supervised VAE coupled with an RF classifier framework, which incorporates B-scan OCT images and metadata as distinct modalities. This integration allowed our model to discern the specific attributes within the OCT images that contribute significantly to the prediction of CVD. Importantly, our study distinguishes itself by interpreting the particular OCT image features (at both the global i.e. class/category, and local i.e. instance, levels), which are relevant to the classification task and thereby provide insights to the key regions of the retinal image that are most discriminative. To the best of our knowledge, some studies have ventured into the application of OCT within a primary care framework for CVD. However, these studies were limited in their explanatory capacity regarding the effects of including images from both eyes and different types of patient data, and used a small portion of the OCT B scans, limiting the information from the entire volume. Nevertheless, the performance results show promising outcomes for OCT as a modality in the primary care of CVDs [15, 37].

Interestingly, our results suggest that choroidal morphology is a predictor of identifying patients at risk of CVD, which is consistent with previous studies [1] that have reported significant associations between choroidal characteristics and the risk of stroke and acute myocardial infarction. Given that the choroid has the highest flow per perfused volume of any human tissue and that there is growing evidence that changes in the choroid microvasculature can be indicative of systemic vascular pathology [13], our findings offer clinical interpretability to the predictions of our classifier. Currently, UK Biobank images are captured using a spectral domain (SD) OCT [17]. SD-OCT images suffer significant light scattering at the choroid, which limits the resolution of this layer. Despite this limited resolution, it is encouraging to observe that the proposed approach focused on features within the choroidal layer to identify patients at risk of stroke or myocardial infarction. To provide a greater context to the key findings reported in this study, it appears probable that once image modalities with deeper tissue penetration, such as swept source OCT, become available at scale in population imaging initiatives (such as UK Biobank), the predictive performance of learning-based
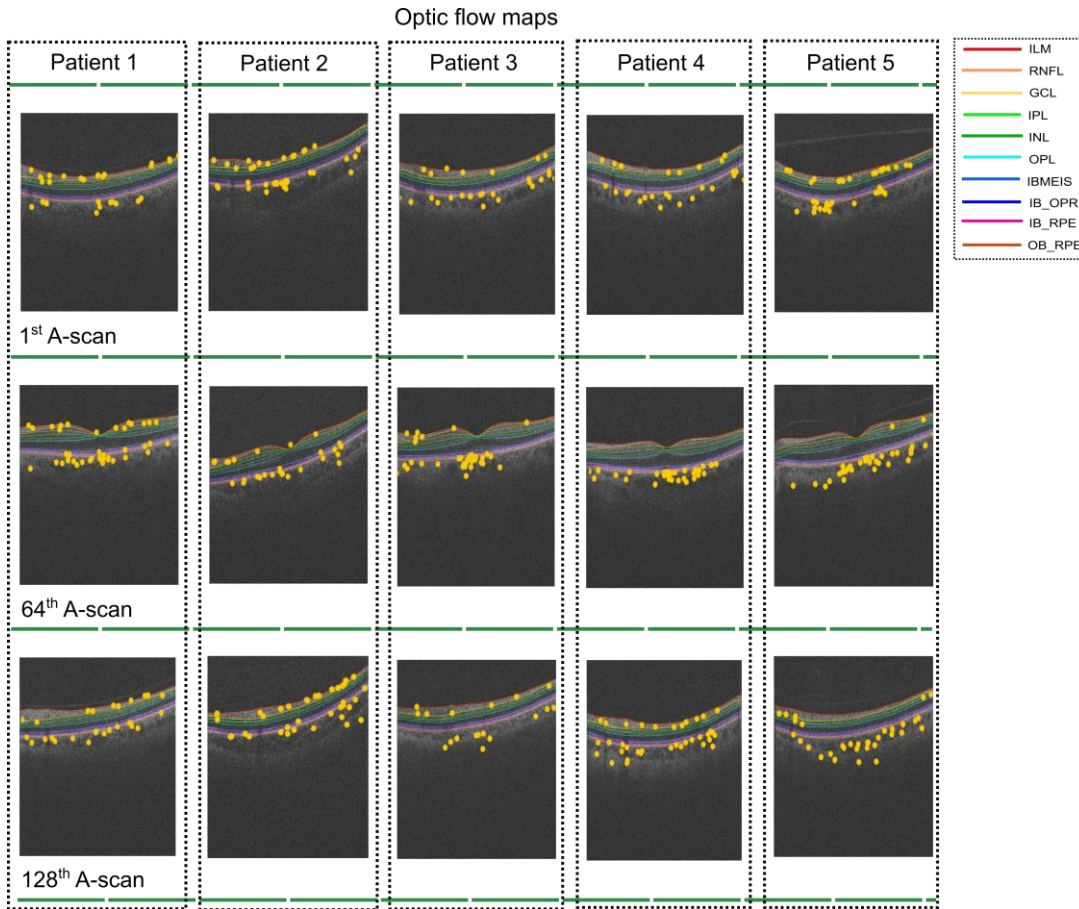
Figure 7: Optical flow maps presented for three differents B-scans for the left eye. The top row corresponds to the 1st B-scan, the middle image for the 64th B-scan while the bottom row depicts the final B-scan. The yellow circles represent the regions that the optical flow maps highlight when modifying the latent variable zl066. The legend includes the names of layer boundaries in the optical coherence tomography images, as follows: Internal Limiting Membrane; RNFL: Retinal Nerve Fiber Layer; GCL: Ganglion Cell Layer; IPL: Inner Plexiform Layer; INL: Inner Nuclear Layer; OPL: Outer Plexiform Layer; BMEIS: Boundary of Myoid and Ellipsoid of Inner Segments; IB_OPR: Inner Boundary of Outer Segment Retinal Pigment Epithelium Complex; IB_RPE: Inner Boundary of Retinal Pigment Epithelium; OB_RPE: Outer Boundary of Retinal Pigment Epithelium

systems such as ours will improve [8]. In addition to the choroidal characteristics, our results indicated that the inner retinal layers contributed to the classification, including the retinal nerve fibre layer (RNFL), the ganglion cell layer (GCL), and the inner plexiform layer (IPL). These aspects of the neurosensory retina consist of retinal ganglion cells, their synapses with bipolar cells, and their axons. Regarding the choroid, thinning and defects in these layers have received extensive study in relation to established CVD, but their role as predictors of future disease has received limited attention [6, 25]. Mechanisms that may underpin the role of neurosensory retina morphology as a predictor of CVD are yet to be elucidated, although it could be hypothesised that subclinical ocular circulatory pathology could explain morphological changes through local ischaemic damage [6], or neuronal degeneration could even occur through silent / subclinical cerebral ischaemic / vascular changes manifesting in the inner retina through transneuronal retrograde degeneration [21, 28].

5. Conclusion

This study presents a predictive framework comprising a self-supervised representation learning approach based on a VAE and a Random Forest classifier, which effectively integrates multi-modal imaging (OCT imaging) and non-imaging (e.g. patient demographic and clinical variables) data, to identify patients at risk of stroke or acute myocardial infarction. Although the focus of this study was on spectral domain OCT imaging, future improvements to the presented work could include the use of more informative retinal imaging modalities such as swept source OCT imaging or wide-field OCT angiography (OCTA) imaging. We hypothesise that learning representations from multi-modal retinal imaging (e.g. fundus photographs, OCT, OCTA) may improve the classification performance of the proposed approach. Additionally, a key benefit of the proposed approach is that it lends itself to explaining model predictions in both a global (across all instances) and local (instance-specific) sense, and thereby, provides insight into which retinal layers contain the most relevant information to identify risk of CVD. In general, this investigation has supported the utility and prospective predictive value of OCT imaging to identify people at risk of stroke or acute myocardial infarction. As OCT imaging is a cost-effective and noninvasive imaging modality, the results presented in this study make a compelling case for future exploration of OCT imaging as a tool for screening individuals at risk of cardiovascular disease.

Acknowledgements

References

[1] 2020. Choroidal thickness in patients with cardiovascular disease: A review. *Survey of Ophthalmology* 65(4). 473–486.

[2] Anderson, Todd J., Akimi Uehata, Marie D. Gerhard, Ian T. Meredith, S Knab, Danielle Delagrange, Eric H. Lieberman, Peter Ganz, Mark A. Creager & Alan C Yeung. 1995. Close relation of endothelial function in the human coronary and peripheral circulations. *Journal of the American College of Cardiology* 26 5. 1235–41.

[3] Badawy, Mohammed Darwesh, Lin Naing, Sofian Johar, Sok King Ong, Hanif Abdul Rahman, Dayangku Siti Nur Ashikin Pengiran Tengah, Chean Lin Chong & Nik Ani Afiqah Tuah. 2022. Evaluation of cardiovascular diseases risk calculators for cvds prevention and management: scoping review. *BMC Public Health* 22.

[4] Breiman, L. 2001. Random forests. *Machine Learning* 45. 5–32.

[5] Brown, Alex, Nenad Tomaev, Jana von Freyberg, Yuan Liu, Alan Karthikesalingam & Jessica Schrouff. 2022. Detecting shortcut learning for fair medical ai using shortcut testing. *Nature communications* 14 1. 4314.

[6] Chen, Yanping, Yixiong Yuan, Shiran Zhang, Shaopeng Yang, Junyao Zhang, Xiao lei Guo, Wenyong Huang, Zhuoting Zhu, Mingguang He & Wen Wang. 2023. Retinal nerve fiber layer thinning as a novel fingerprint for cardiovascular events: results from the prospective cohorts in uk and china. *BMC Medicine* 21.

[7] lui Cheung, Carol Yim, Dejiang Xu, Ching-Yu Cheng, Charumathi Sabanayagam, Yih-Chung Tham, Marco Yu, Tyler Hyungtaek Rim, Chew Yian Chai, Bamini Gopinath, Paul Mitchell, Richie Poulton, Terrie E. Moffitt, Avshalom Caspi, Jason C S Yam, Clement C. Tham, Jost Bruno Jonas, Ya Xing Wang, Su Jeong Song, Louise M. Burrell, Omar Farouque, Lingjun Li, Gavin S. W. Tan, Daniel Shu Wei Ting, Wynne Hsu, Mong Li Lee & Tien Yin Wong. 2020. A deep-learning system for the assessment of cardiovascular disease risk via the measurement of retinal-vessel calibre. *Nature biomedical engineering* .

[8] Copete, Sergio, Ignacio Flores-Moreno, Javier Antonio Montero, Jay S. Duker & Jose´ M. Ruiz-Moreno. 2013. Direct comparison of spectral-domain and swept-source oct in the measurement of choroidal thickness in normal eyes. *British Journal of Ophthalmology* 98. 334 – 338.

[9] Diaz-Pinto, Andrés, Nishant Ravikumar, Rahman Attar, Avan Suinesiaputra, Yitian Zhao, Eylem Levelt, E Dall"armellina, Marco Lorenzi, Qingyu Chen, Tiarnán D. L. Keenan, Elvira Agrón, Emily Y. Chew, Zhi gang Lu, Chris P. Gale, Richard Gale, Sven Plein & Alejandro F. Frangi. 2022. Predicting myocardial infarction through retinal scans and minimal personal information. *Nature Machine Intelligence* 4. 55–61.

[10] Donkor, Eric S. 2018. Stroke in the 21st century: A snapshot of the burden, epidemiology, and quality of life. *Stroke Research and Treatment* 2018. https://api.semanticscholar.org/CorpusID:57193376.

[11] Farrah, Tariq E., Baljean Dhillon, Pearse A. Keane, David John Webb & Neeraj Dhaun. 2020. The eye, the kidney, and cardiovascular disease: old concepts, better tools, and new horizons. *Kidney International* 98. 323 – 342.

[12] Farrah, Tariq E., David John Webb & Neeraj Dhaun. 2019. Retinal fingerprints for precision profiling of cardiovascular risk. *Nature Reviews Cardiology* 16. 379–381.

[13] Ferrara, Mariantonia, Gaia Lugano, Maria Teresa Sandinha, Victoria R. Kearns, Brendan Geraghty & David H. W. Steel. 2021. Biomechanical properties of retina and choroid: a comprehensive review of techniques and translational relevance. *Eye* 35. 1818 – 1832.

[14] Flammer, Josef, Katarzyna Konieczka, Rosa Maria Bruno, Agostino Virdis, Andreas J. Flammer & Stefano Taddei. 2013. The eye and the heart. *European Heart Journal* 34. 1270 – 1278. https://api.semanticscholar.org/CorpusID:2384003.

[15] Garc′ıa, Cynthia Maldonado, Rodrigo Bonazzola, Nishant Ravikumar & Alejandro F. Frangi. 2022. Predicting myocardial infarction using retinal oct imaging. In *Annual conference on medical image understanding and analysis*, .

[16] Hanssen, Henner, Lukas Streese & Walthard Vilser. 2022. Retinal vessel diameters and function in cardiovascular risk and disease. *Progress in Retinal and Eye Research* 91. https://api.semanticscholar.org/CorpusID:249995065.

[17] Keane, Pearse A., Carlota M. Grossi, Paul J. Foster, Qi Yang, Charles A Reisman, Kinpui Chan, Tunde Peto, Dhanes Thomas & Praveen J. Patel. 2016. Optical coherence tomography in the uk biobank study – rapid automated analysis of retinal thickness for large population-based studies. *PLoS ONE* 11.

[18] Khozeimeh, Fahime, Danial Sharifrazi, Navid Hoseini Izadi, Javad Hassannataj Joloudari, Afshin Shoeibi, Roohallah Alizadehsani, Mehrzad Tartibi, Sadiq Hussain, Zahra Alizadeh Sani, Marjane Khodatars et al. 2022. Rf-cnn-f: random forest with convolutional neural network features for coronary artery disease diagnosis based on cardiac magnetic resonance. *Scientific Reports* 12(1). 11178.

[19] Kingma, Diederik P. & Max Welling. 2014. Auto-encoding variational bayes. *CoRR* abs/1312.6114.

[20] Kirshner, Howard S, Jose Biller & Alfred S. Callahan. 2005. Long-term therapy to prevent stroke. *The Journal of the American Board of Family Practice* 18 6. 528–40.

[21] Langner, Sönke, Jan Henrik Terheyden, Clara F Geerling, Christine Kindler, Vera C W Keil, Christopher A Turski, Gabrielle N. Turski, Charlotte Behning, Maximilian W. M. Wintergerst, Gabor C. Petzold & Robert P. Finger. 2022. Structural retinal changes in cerebral small vessel disease. *Scientific Reports* 12.

[22] Li, Yan, Matthew Sperrin & Tjeerd Pieter van Staa. 2019. R package "qrisk3": an unofficial research purposed implementation of clinrisk's qrisk3 algorithm into r. *F1000Research* https://api.semanticscholar.org/CorpusID:219492836.

[23] Lindstrom, Megan et al. 2023. Deaths from cardiovascular disease surged 6030 years: Report. *World Heart Federation* .

[24] Lucas, Bruce D. & Takeo Kanade. 1981. An iterative image registration technique with an application to stereo vision. In *International joint conference on artificial intelligence*, .

[25] Matuleviiūtė, Indrė, Agnė Sidaraitė, Vacis Tatarunas, Audronė Veikutienė, Olivija Dobilienė & Dalia Zaliuniene. 2022. Retinal and choroidal thinning—a predictor of coronary artery occlusion? *Diagnostics* 12.

[26] Muraoka, Yuki, Akitaka Tsujikawa, Kyoko Kumagai, Masahiro Akiba, Ken Ogino, Tomoaki Murakami, Yumiko Akagi-Kurashige, Kazuaki Miyamoto & Nagahisa Yoshimura. 2013. Age- and hypertension-dependent changes in retinal vessel diameter and wall thickness: an optical coherence tomography study. *American journal of ophthalmology* 156 4. 706–14.

[27] Nusinovici, Simon, Tyler Hyungtaek Rim, Marco Yu, Geunyoung Lee, Yih-Chung Tham, Ning Cheung, Crystal Chong, Zhi Da Soh, Sahil Thakur, Chan Joo Lee, Charumathi Sabanayagam, Byoung Kwon Lee, Sungha Park, Sung Soo Kim, Hyeon Chang Kim, Tien Yin Wong & Ching-Yu Cheng. 2022. Retinal photograph-based deep learning predicts biological age, and stratifies morbidity and mortality risk. *Age and Ageing* 51.

[28] Park, Hae-Young Lopilly, Young Gun Park, A Hyun Cho & Chan Kee Park. 2013. Transneuronal retrograde degeneration of the retinal ganglion cells in patients with cerebral infarction. *Ophthalmology* 120 6. 1292–9.

[29] Poplin, Ryan, Avinash V. Varadarajan, Katy Blumer, Yun Liu, Michael V. McConnell, Greg S Corrado, Lily H. Peng & Dale R. Webster. 2018. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nature Biomedical Engineering* 2. 158–164.

[30] Rudnicka, Alicja Regina, Roshan Alex Welikala, Sarah Ann Barman, Paul J. Foster, Robert N. Luben, Shabina A Hayat, Kay-Tee Khaw, Peter H. Whincup, David P. Strachan & Chris G. Owen. 2022. Artificial intelligence-enabled retinal vasculometry for prediction of circulatory mortality, myocardial infarction and stroke. *The British Journal of Ophthalmology* 106. 1722 – 1729.

[31] Samsa, Gregory P, Stuart James Cohen, Larry B. Goldstein, Arthur J. Bonito, Pamela W. Duncan, Cam E. Enarson, Gordon H. DeFriese, Ronnie D Horner & David Bruce Matchar. 1997. Knowledge of risk among patients at increased risk for stroke. *Stroke* 28 5. 916–21.

[32] Stein, Daniel M, Hiroshi Ishikawa, Roopa Hariprasad, Gadi Wollstein, Robert J. Noecker, James G. Fujimoto & Joel S. Schuman. 2006. A new quality assessment parameter for optical coherence tomography. *British Journal of Ophthalmology* 90. 186 – 190.

[33] Wagner, Siegfried Karl, Dun Jack Fu, Livia Faes, Xiaoxuan Liu, Josef Huemer, Hagar Khalid, Daniel Araújo Ferraz, Edward Korot, Christopher J. Kelly, Konstantinos Balaskas, Alastair K. O. Denniston & Pearse A. Keane. 2020. Insights into systemic disease through retinal imaging-based oculomics. *Translational Vision Science & Technology* 9.

[34] Wei, Wen Bin, Liang Xu, Jost Bruno Jonas, Lei Shao, Kuifang Du, Shuang Wang, Chang Xi Chen, Jie Xu, Ya Xing Wang, Jin Qiong Zhou & Qi Sheng You. 2013. Subfoveal choroidal thickness: the beijing eye study. *Ophthalmology* 120 1. 175–80.

[35] Wong, Dragon Y.L., Mary C. Lam, An Ran Ran & Carol Yim lui Cheung. 2022. Artificial intelligence in retinal imaging for cardiovascular disease prediction: current trends and future directions. *Current Opinion in Ophthalmology* 33. 440 – 446.

[36] Yang, Li, Haibin Wu, Xiaoqing Jin, Pinpin Zheng, Shiyun Hu, Xiaoling Xu, Wei Yu & Jing Yan. 2020. Study of cardiovascular disease prediction model based on random forest in eastern china. *Scientific reports* 10(1). 5245.

[37] Zhou, Yukun, Mark A Chia, Siegfried K. Wagner, Murat S. Ayhan, Dominic J Williamson, Robbert R. Struyven, Timing Liu, Moucheng Xu, Mateo G. Lozano, Peter Woodward-Court, Yuka Kihara, Naomi John E. J. Thomas Tariq Paul Graeme Panagiotis Den Allen Gallacher Littlejohns Aslam Bishop Black Ser, Naomi Allen, John E. J. Gallacher, Thomas Littlejohns, Tariq Aslam, Paul Bishop, Graeme Black, Panagiotis I. Sergouniotis, Denize Atan, Andrew D. Dick, Cathy Williams, Sarah Ann Barman, Jenny H. Barrett, Sarah Mackie, Tasanee Braithwaite, Roxana O. Carare, Sarah Ennis, Jane Gibson, Andrew J. Lotery, Jay Self, Usha Chakravarthy, Ruth E. Hogg, Euan N. Paterson, Jayne Woodside, Tunde Peto, Gareth J. McKay, Bernadette Mcguinness, Paul J. Foster, Konstantinos Balaskas, Anthony P. Khawaja, Nikolas Pontikos, Jugnoo S Rahi, Gerassimos Lascaratos, Praveen J. Patel, Michelle Chan, Sharon Y. L. Chua, Alexander Day, Parul Desai, Catherine Egan, Marcus Fruttiger, David F. Garway-Heath, Alison Hardcastle, Sir Peng Tee Khaw, Tony Moore, Sobha Sivaprasad, Nicholas G. Strouthidis, Dhanes Thomas, Adnan Tufail, Ananth C. Viswanathan, Bal Dhillon, Tom Macgillivray, Cathie Sudlow, Veronique Vitart, Alexander Doney, Emanuele Trucco, Jeremy A. Guggeinheim, James Morgan, Chris J. Hammond, Katie M. Williams, Pirro G. Hysi, Simon Harding, Yalin Zheng, Robert Luben, Phil Luthert, Zihan Sun, Martin Mckibbin, Eoin O'Sullivan, Richard Oram, Mike Weedon, Chris G. Owen, Alicja Rudnicka, Naveed Sattar, David Steel, Irene Stratton, Robyn Tapp, Max M. Yates, Axel Petzold, Savita Madhusudhan, André´ Altmann, Aaron Y.

Lee, Eric J. Topol, Alastair Keith Denniston, Daniel C. Alexander & Pearse Andrew Keane. 2023. A foundation model for generalizable disease detection from retinal images. *Nature* 622. 156 – 163. https://api.semanticscholar.org/CorpusID:264168236.

Appendix
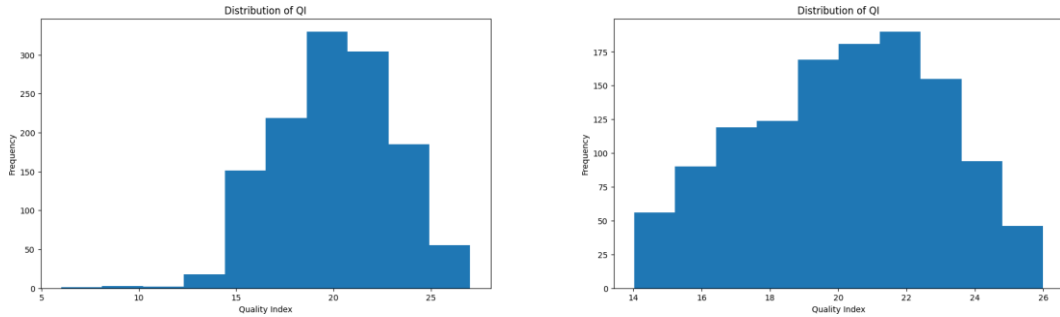
(a)   Distribution of QI of the left eye   (b) Distribution of QI of the right eye

Figure 8: Comparation of the distribution of the Quality Index from both eyes

The Quality Index is calculated as the product of two terms referred to as Intensity Ratio (IR) and Tissue Signal Ratio (TSR). The IR is akin to the signal-to-noise ratio (SNR), but rather than considering the maximum SNR value among all A-scans, it encompasses the entire image. Meanwhile, TSR represents the ratio of highly reflective pixels to those with lower reflectivity. Further details regarding the formula are provided in [32].
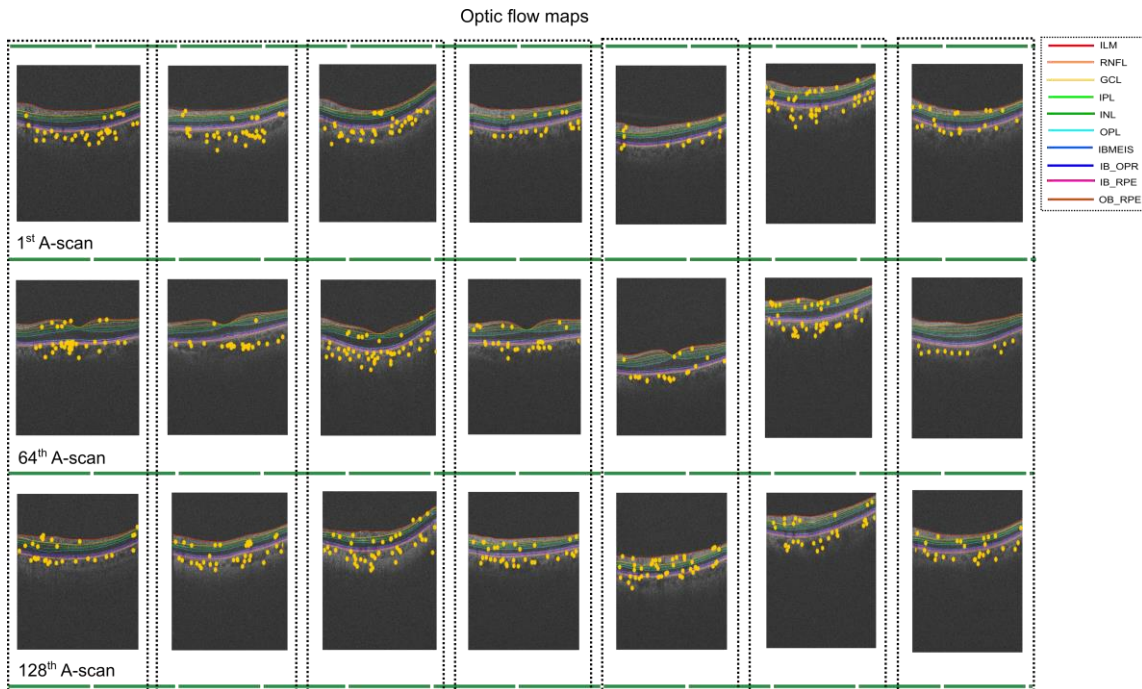

Figure 9: Optical flow maps present three different B-scans for the left eye. The top row corresponds to the 1st B-scan, the middle image to the 64th B-scan, and the bottom row depicts the final B-scan. The yellow circles highlight regions of the optical flow maps when modifying the latent variable zl066. The legend includes the names of layer boundaries in the optical coherence tomography images: Internal Limiting Membrane, Retinal Nerve Fiber Layer (RNFL), Ganglion Cell Layer (GCL), Inner Plexiform Layer (IPL), Inner Nuclear Layer (INL), Outer Plexiform Layer (OPL), Boundary of Myoid and Ellipsoid of Inner Segments (BMEIS), Inner Boundary of Outer Segment Retinal Pigment Epithelium Complex (IB_OPR), Inner Boundary of Retinal Pigment Epithelium (IB_RPE), and Outer Boundary of Retinal Pigment Epithelium (OB_RPE).

Variables used in QRISK3
Gender
Age
Atrial fibrillation
Atypical antipsy
Regular steroid tablets
Erectile disfunction
Migraine
Rheumatoid arthritis
Chronic kidney disease
Severe mental illness
Systemic lupus erythematosis
Blood pressure treatment
Diabetes1
Diabetes2
Weight
Height
Ethnicity
Heart attack relative
Cholesterol HDL ratio
Systolic blood pressure
Std systolic blood pressure
Smoke
Townsend

Table 4: List of Variables used to calculate QRISK3