

RetinaFusion-XAI: Explainable Multimodal Oculomics for Cardiometabolic Risk Prediction and Remote Monitoring

Dr. Shamik Palit

Post Doctoral Researcher

¹Lincoln University College Malaysia

²Faculty of Natural Sciences, University of Stirling - Ras Al Khaimah Campus, United Arab Emirates

Email: shamik1980@gmail.com

Abstract:

The commonly used traditional methods for evaluating the risk of developing cardiovascular diseases, diabetes and hypertension often overlook early microvascular lesions. These changes which quietly occur in the body's smallest blood vessels are precisely the potential risk signals that fail to be detected in time. By the time traditional methods can identify abnormalities, the optimal window for intervention has usually already passed.

Taking retinal images allows doctors to examine the details of vascular tissues in the eyes without surgery which was originally an excellent way to detect risks in advance. However, most existing artificial intelligence tools used to analyze these images have unavoidable critical flaws: they only operate on the single data type of retinal images; after reaching a conclusion about a disease risk, they cannot explain which specific details they relied on to derive their judgment. What is more, they cannot build corresponding models to track changes in a single patient's vascular health for several years or even more than a decade, failing to accumulate usable data to support continuous monitoring.

This paper proposes RetinaFusion-XAI, a multimodal analysis framework capable of clarifying its reasoning logic. It no longer relies solely on a single type of data and can integrate various routine clinical indicators collected during medical visits, retinal biomarkers extracted from retinal images that reflect the body's state and feature data generated by deep learning models for joint analysis, filling the critical gaps that could not be addressed in previous risk assessments of this type.

Previous analytical tools have long failed to solve two core problems: first, they cannot properly analyze images captured by different imaging devices, resulting in extremely poor device compatibility; second, the conclusions they produce cannot explain the judgment logic to clinicians, making them unusable as reliable references for diagnosis and treatment. In response to these long-standing issues, this new framework incorporates targeted specialized designs: it has a built-in quality control mechanism that filters out qualified data, privacy protection measures for users, the ability to complete risk prediction tasks for several different diseases simultaneously, the function to hierarchically decompose the influence logic of each feature on the final conclusion, and a dedicated remote monitoring panel for medical teams allowing doctors to review all of a patient's monitoring data without needing to be present on-site.

RetinaFusion-XAI establishes a reliable non-invasive architecture that requires no invasive procedures to be performed on patients throughout the entire workflow. It can be used for early screening of related diseases in the general population, as well as for long-term continuous monitoring of a single patient's vascular health changes. Nevertheless, before this system can be formally integrated into hospitals' routine clinical work and deployed in settings such as outpatient departments and physical examination centers, multi-site prospective assessments

must first be completed. Only after tracking and verifying its performance among asymptomatic populations across different sites and confirming its stable and reliable efficacy can it be officially put into clinical use.

Keywords: Retinal oculomics; Cardiometabolic risk; Multimodal learning; Explainable artificial intelligence; Remote patient monitoring

1. Introduction

Chronic diseases such as cardiovascular disease, diabetes and hypertension do not emerge suddenly from the time the root cause of illness takes root in the body to the actual development of the disease, it takes several years of accumulation before the condition gradually manifests. The routine clinical examinations that ordinary people undergo at hospitals are usually conducted at intervals of several months or even longer. As a result, the slow, subtle changes that occur in blood vessels between two medical appointments are very likely to be entirely missed.

Fortunately, the human retina provides doctors with a window to directly observe microvascular health. Through technologies including fundus photography, OCT and OCTA, the tiny blood vessels distributed across the retina can be visualized non-invasively without the need for surgical incisions. To date, multiple sets of data measured from the retina have been correlated with the prognosis of blood pressure conditions, diabetes and cardiovascular disease [1]-[6].

Later, the introduction of deep learning technology expanded the applications of such analyses. These tools can not only predict a person's overall health status and future risk of developing cardiovascular disease [7] but also anticipate the onset of conditions such as chronic kidney disease and type 2 diabetes in advance [8]-[11]. The subsequent emergence of foundation models further optimized the ability of these analytical models to extract and learn information [12]. However, even as the capabilities of these models continue to improve, a host of pressing practical problems must be resolved before they can be truly implemented in clinical diagnosis and treatment in hospitals. These problems include the uneven quality of captured fundus images, the inability to adapt to feature differences among populations of different ages and backgrounds, and the lack of clarity in the basis for the models' conclusions—what the industry commonly refers to as the explainability problem.

To address these pain points, this study proposes RetinaFusion-XAI, a multimodal framework that allows clinicians to understand the judgment logic. It integrates three types of data: first, biomarker data with clear definitions that can directly reflect the body's state; second, deep feature data extracted from retinal images; and third, historical clinical diagnosis and treatment data left by patients at hospitals. This approach pursues two goals: one is to improve the accuracy of cardiometabolic risk stratification, to accurately distinguish patients with different levels of disease risk; the other is to support long-term remote monitoring of patients, enabling them to track changes in their condition without frequent visits to the hospital.

2. Related Work

The review covers retinal microvascular and OCT/OCTA biomarkers, deep retinal oculomics, foundation models, multimodal AI, explainability and privacy-aware deployment. Literature searches spanned PubMed/MEDLINE, Scopus, Web of Science, IEEE Xplore and Google Scholar to capture comprehensive biomedical, technical and multidisciplinary evidence.

2.1 Eligibility Criteria

We have established several unified entry criteria for the studies to be included in our synthesis, all of which must be met for a study to be selected. First, all observation subjects in the study must be adults. Second, the retinal images of these subjects, including fundus, OCT or OCTA must be of sufficient quality to be analyzed and graded, and cannot be blurred unusable waste images. Third, the study must have fully collected the subjects' demographic information and clinical data and studies with missing or omitted items are not eligible. Fourth, the study must clearly mark the cardiometabolic-related information associated with the research subjects. For longitudinal studies with long-term follow-up, there are additional requirements: multiple repeated assessments must be conducted on the same patient throughout the study's follow-up period, and studies with only one assessment do not meet the requirements.

There are also clear rules and if anyone is met, the relevant study will not be included in our synthesis. Specifically, these situations include: the retinal images used in the study have severe interference artifacts that make them completely unanalyzable; in models used to predict disease risk, the patient had already developed the outcome that the model aims to predict before they were enrolled in the study. The research subjects have severe eye diseases or have received relevant eye treatments that would affect the measurement of retinal-related indicators. The clinical data collected by the study is incomplete and fails to meet the entry requirements and repeated images of the same patient are mistakenly used as data from different independent patients.

2.2 Evidence Synthesis

In research on systemic diseases related to the retina, the first approach adopted by researchers was to use classic retinal biomarkers. Studies of this type could associate retinal vascular morphology with systemic diseases relying only on a small number of manually extracted features [1]-[6]. Later, deep retinal oculomics emerged which eliminates the need for manual feature annotation and can directly excavate signals related to systemic diseases from retinal images but the entire derivation process cannot be made transparent, no one can clarify which image details the model used to calculate these associations [7]-[11]. After that OCT and OCTA technologies for capturing retinal images were developed, which can obtain fine-structured data of the retina with far richer details than traditional images. However, the hardware parameters and scanning processes of different devices cannot be unified, and the data generated by different studies have always been inconsistent, making it impossible to pool the data for combined use [4], [5]. In recent years, foundation models adapted for retinal analysis have emerged, with improved transferability across study cohorts, eliminating the need to retrain the model every time the patient cohort changes. But when applied to routine clinical scenarios, doctors still struggle to understand the model's judgment logic and dare not directly use its conclusions as a reference for diagnosis and treatment [12]. The academic community has proposed addressing this gap by using explainability frameworks and uncertainty assessment. These methods can indeed improve medical staff's trust in the model, but both types of methods still require clinician-centered verification, which aligns the conclusions output by the model with existing professional concepts in the retinal field, before they can be truly implemented for practical use [13], [14].

Table 1. Comparison of representative related work

Study	Input	Systemic target	Explicit biomarkers	Deep representation	Multimodal / longitudinal	Explanation / uncertainty
McGeechan et al. [1]	Fundus	Coronary heart disease	Vessel calibre	No	No	No
Ong et al. [2]	Fundus	Stroke mortality	Fractal dimension	No	Longitudinal outcome	Biomarker association
Mordi et al. [6]	Fundus + clinical + genomic	Major cardiovascular events	Geometry	Limited	Yes	Feature-level
Poplin et al. [7]	Fundus	Risk factors	No	CNN	No	Attention maps
Cheung et al. [8]	Fundus	Cardiovascular risk	Vessel calibre	CNN	No	Biomarker-linked
Zhou et al. [12]	Fundus + OCT pretraining	Ocular/systemic tasks	No	Foundation model	Multi-dataset	Limited
Hu et al. [15]	Fundus	10-year CVD risk	No	CNN	External longitudinal validation	Risk report
RetinaFusion-XAI	Fundus + OCT/OCTA + clinical	Diabetes, hypertension, CVD risk	Yes	Foundation embeddings	Multimodal + longitudinal	Concept attribution + calibrated uncertainty

3. Key Contribution

- RetinaFusion-XAI has not launched a set of fully clinically validated medical models that can be directly deployed for practical use. What it does is to connect and assess all types of existing research evidence in the field and integrate the previously scattered retinal and ocular omics research into a coherent research framework. Its core components include:
- A multimodal input module that supports the reception of three types of data from different sources: fundus photographs, OCT/OCTA images, and clinical data. It can process data normally without issues even if one type of data is missing.
- A dual knowledge flow module that integrates the biomarker information required for clinical practice and the embedding vector information generated by the base model for combined use.
- A concept-guided cross attention mechanism that can associate the deep abstract features extracted by the deep learning model with specific ocular anatomical concepts, so as to complete multi-task prediction work.
- A joint risk assessment function that can simultaneously assess the onset risk of three diseases: diabetes, hypertension and cardiovascular disease. During the risk calculation process, it considers both the inherent uncertainty of the conditions and clinical referral rules.
- A patient-specific longitudinal modeling function that integrates multiple consecutive scan results from the same patient converts them into a trajectory of the patient's disease risk changing over time, to support remote disease monitoring.
- A complete validation protocol that separately assesses the tool's generalization ability, performance across different populations, privacy protection capability and clinical utility to comprehensively test the tool's actual performance.

4. Method, Experiments and Results

4.1 RetinaFusion-XAI Architecture

The8- Stage Workflow of the Proposed Approach in Figure 1 Retinal anatomy is segmented or localized post quality and privacy checks. Explicit biomarkers and deep embeddings are extracted and merged with clinical variables. Multi-task prediction forms conditional risks, and interpretation and uncertainty module provides evidence as a human-understandable output. Lastly, a dashboard for risk, confidence and time process is displayed.

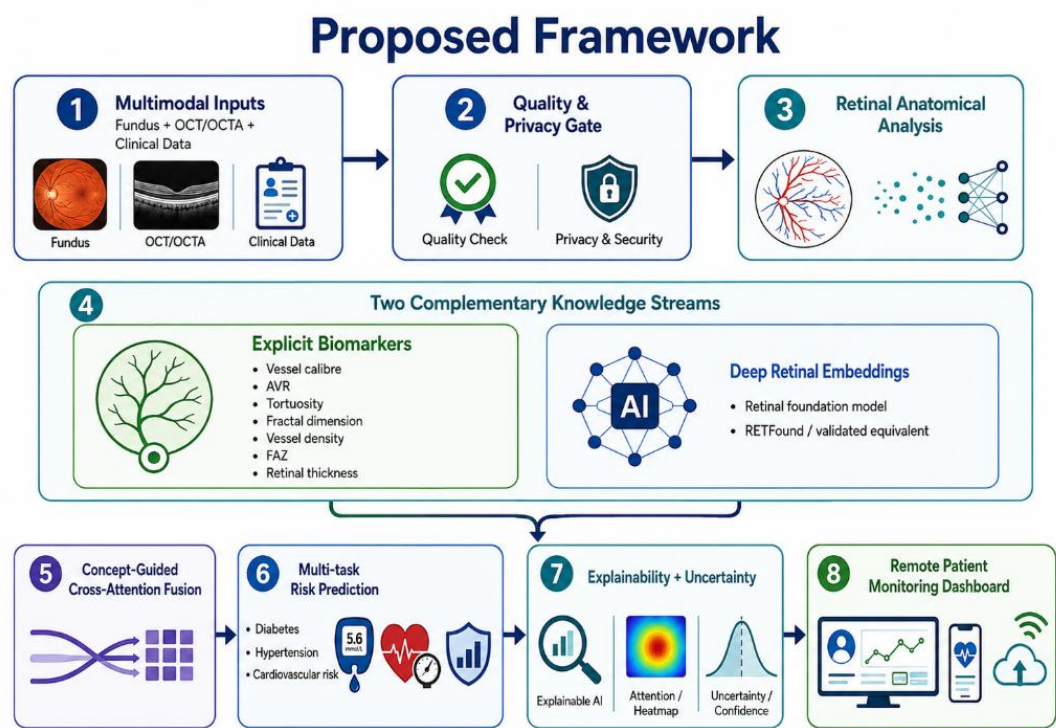


Figure 1. Proposed RetinaFusion-XAI multimodal biomarker-fusion framework.

4.2 Multimodal Inputs and Data Organization

All analyses are carried out on a per-patient unit, from data organization to all subsequent studies; all processing work is advanced by treating each patient as an independent basic unit. Each visit record contains three types of images: fundus, OCT, OCTA as well as clinical data and demographic statistics apart from the clinical information directly related to the patient's condition and the demographic information recording the patient's basic personal situation. These three types of medical images are also uniformly included in each patient's visit data. Multiple sequential visit records of the same patient are linked together by an anonymous identifier. This identifier will not disclose the patient's real identity and is only used to connect all the visit data of the same person from different time points, preventing the mixing of information from other patients. The three groups of patients used for model

development, validation and testing will never overlap. The same patient can only appear in one group and no patient will appear in two or more groups, ensuring that the three sets of data are completely separated. If conducting an experiment on the risk of disease onset, patients who already have the target condition at the start of the experiment will be excluded. After all, such experiments are used to assess the risk of subsequent onset of the condition, and patients who already have the disease before the experiment starts do not meet the experiment's inclusion criteria.

Table 2. Input streams and derived information

Input stream	Preprocessing	Derived information	Quality control
Fundus photography	Field normalization, illumination correction, laterality check	Calibre, AVR, tortuosity, fractal dimension, deep embedding	Gradability, blur, field definition, artefacts
OCT	Layer alignment, scan resampling, segmentation review	Layer thickness and structural representation	Signal strength, motion, segmentation failure
OCTA	Projection and motion correction, slab harmonization	Vessel density, FAZ and capillary geometry	Signal strength, decorrelation artefacts, device protocol
Clinical data	Unit harmonization, temporal alignment, missingness flags	Conventional risk context and labels	Range, consistency and timestamp checks
Repeated encounters	Patient-level ordering and interval encoding	Change features and risk trajectory	Minimum interval and duplicate detection

4.3 Quality and Privacy Gate

This workflow first sets up two core checkpoints. The first is the quality checkpoint, which assigns an analyzability score to each image to be processed. Any image that fails to meet the quality requirements for analysis will be directly rejected. After rejection, there are three compliant processing paths: collect a new image that meets the requirements, use another available examination modality at hand as a substitute, or directly abandon the analysis of this image, without forcing results to be produced from unqualified data.

Next is the privacy checkpoint, which is responsible for completing four tasks related to data security and privacy. First, all personally identifiable information is erased. Then, all the data retained is encrypted and protected. Next, access permissions are set for personnel who can access this data. Finally, all records of data access and operations are retained to form traceable audit logs.

If this workflow is used for multi-site collaboration, and local regulations prohibit the centralized transmission and processing of data from all regions at a single location, a federated learning scheme can be adopted [16].

4.4 Retinal Anatomical Analysis and Explicit Biomarkers

When conducting anatomical analysis of the retina, the key structures on the retina as well as each layer of OCT and OCTA are segmented one by one. In simple terms, this process clearly separates different tissues and layers that were originally mixed in the same medical image, turning each into an independent unit from which data can be extracted separately. After these basic structures are demarcated, a dedicated explicit calculation process is

initiated to derive a set of clear biomarkers from these segmented regions, including vascular equivalent values, AVR, tortuosity, density and thickness. Each calculated indicator is attached with a corresponding quality score, which verifies that the measurement result of this indicator is reliable, and not randomly calculated from unclear image data. All these sorted features can be presented in a format that clinicians can understand. Doctors can directly compare these indicators with the confirmed disease-related retinal changes and verify the accuracy of the disease risk estimate generated by the system by matching it with known pathological patterns.

4.5 Deep Retinal Embeddings

This deep computing workflow leverages pre-trained foundational large models such as RETFound and converts all input data from fundus photographs and OCT/OCTA into the same set of universal feature encodings, which are stored in a unified latent space [12]. During the model's training process, a modality dropout mechanism is intentionally added to the framework: during training, a portion of the examination data from certain modalities is randomly hidden, allowing the model to gradually learn to cope with scenarios where certain examination data is missing in real-world use. Metadata such as the site where each piece of data was collected and the device used is only retained exclusively for robustness analysis and is never allowed to participate in the model's core prediction process. This prevents these additional pieces of information from becoming uncontrolled shortcuts, that is, it avoids the model generating results solely relying on these condition-irrelevant pieces of information which would prevent it from learning the true pathological features that need to be extracted from fundus images and disrupt the model's judgment.

4.6 Concept-Guided Cross-Attention Fusion

The task of this fusion module is to integrate two completely distinct types of features. One type is the explicit biomarkers that can be directly listed, that is, the various indicators that clinicians can understand briefly. The other type is the deep feature encodings generated by the deep learning model. To integrate these two types of features well, the cross-attention mechanism is relied upon. In addition to the basic fusion logic, a concept alignment loss term is added to the module to keep the model's attention focused ensuring that wherever the model places its attention, it always aligns with the corresponding anatomical region and does not randomly focus on areas irrelevant to the task. The final generated fused feature integrates all information from the entire medical image, while retaining a clear traceable path that links the final analysis result to each individual tissue detection dataset, preventing it from becoming an unintelligible black box.

4.7 Multi-task Risk Prediction

This Multi-task Risk Prediction workflow is carried out in two steps. First, the universal shared layer in the model extracts common cardiometabolic features across all cases which are shared patterns applicable to all three diseases. After extracting these general pieces of information, the system passes these features to independent modules dedicated to each task. Each dedicated module uses probability or survival functions to calculate the risk of developing three types of diseases: diabetes, hypertension, and cardiovascular disease, respectively. This system outputs the risk results for each disease separately, rather than blending all results into a vague comprehensive score. Each result is presented independently so that anyone can understand the level of risk for each disease avoiding the situation where it is unclear which risk factor is problematic within a single score.

4.8 Explainability, Calibration and Uncertainty

The model can provide three types of XAI content. The first is a spatial map that can mark which regions on the image have the greatest impact on the final prediction result. The second is the contribution of each biomarker to the prediction result. The third is the contribution of clinical features to the prediction result. Calibration capability testing is completed using industry-standard metrics, specifically including Brier scores and expected calibration error. The model also calculates an uncertainty value for each prediction and sets an automatic analysis abandonment threshold based on this value. If any one of three conditions triggered excessively high prediction uncertainty, poor image quality or inconsistent results across different modalities. The system will transfer this case to a clinician for manual review [13], [14].

4.9 Longitudinal Remote Patient Monitoring

When conducting multiple follow-up assessments for the same patient, this framework tracks and records changes in the patient's various indicators and deep feature encodings throughout the entire process. The time series model equipped in the system calculates the trend of changes in the patient's condition risk step by step. Even if the patient cannot come to the hospital for examinations at fixed intervals and the intervals between hospital visits are completely irregular. This model can still complete its calculation normally and will not be affected by disrupted timelines.

The system also generates a visual operation panel, which clearly lists the current risk estimate value for the patient, the corresponding confidence interval for this estimate, the historical values from the patient's previous examinations, and the specific features that influenced the current risk calculation. All this displayed information is only used to assist doctors in long-term monitoring of the patient's condition changes and will never replace clinicians in making any clinical decisions.

Any system alert can only be triggered after meeting preset fixed thresholds, and after an alert is generated, it must undergo manual review before it can be further processed and circulated.

4.10 Experimental Design

The efficacy evaluation we conducted for the RetinaFusion-XAI framework proceeds through four sequential stages. The first is internal validation: we first run the model with a set of data in our possession to thoroughly test its basic capabilities. The second is temporal validation: we test it again with a set of data collected at different times to check whether the model's performance changes over time and if it remains stable. The third is external validation: we collaborate with completely independent research institutions to test the model with data they have collected that covers different population characteristics, to verify whether the model can adapt to unfamiliar data sources. The last is long-term follow-up validation, which is specifically used to predict whether new, previously undiagnosed health events will occur in the future, to test the model's ability to predict long-term risks.

All research reports related to this evaluation must clearly state several key pieces of information. The full process flow of all study subjects from initial enrollment to the completion of the entire study must be documented who withdrew from the study midway and for what reason and who successfully completed all tests must all be recorded. The overall quality of all data used in the study must also be stated along with the number and proportion of the various health outcomes we tested for that occurred in the study population, and the demographic composition of each subgroup among all participants.

Four types of models are used for comparison. We test the complete RetinaFusion-XAI model and these four types of models under the same scenario. The four types are: a model using only clinical information, a model based on explicit biomarkers, a model that only performs deep analysis of fundus images and an unguided multimodal model. To ensure that the comparison standards for all models are completely unified, and that differences in data partitioning do not affect the fairness of the results, all participating models are evaluated with the same data grouping.

Finally, when conducting statistical analysis, we use the bootstrap resampling method to calculate the corresponding confidence intervals for all performance metrics. When evaluating the model's performance, we do not draw conclusions solely from independent significance tests. Instead, we focus on effect sizes and quantified uncertainty to make a more comprehensive judgment of the model's actual capabilities.

Table 3. Planned ablation experiments

Experiment	Clinical	Explicit biomarkers	Deep embeddings	Fundus + OCT/OCTA	Concept fusion	Longitudinal
A: Clinical baseline	Yes	No	No	No	No	No
B: Biomarker model	Yes	Yes	No	No	No	Optional
C: Fundus deep model	Yes	No	Yes	Fundus only	No	Optional
D: Multimodal deep	Yes	No	Yes	Yes	No	Optional
E: RetinaFusion-XAI	Yes	Yes	Yes	Yes	Yes	Yes

4.11 Outcome Measures

For binary outcomes, evaluation includes area under the receiver-operating characteristic curve, area under the precision-recall curve, sensitivity, specificity, F1-score and positive and negative predictive values at clinically justified thresholds. For continuous risk factors, mean absolute error, root-mean-square error and correlation are reported. Survival models use concordance, time-dependent discrimination and calibration. Clinical value is examined through decision-curve analysis and net benefit. Robustness analysis stratifies performance by sex, age group, ethnicity, site, device, image quality and modality availability.

4.12 Current Results: Structured Evidence Synthesis

The current work yields a framework-level result rather than a clinical performance estimate. The structured synthesis identifies five consistent findings. First, explicit retinal vascular geometry has established associations with systemic vascular status and cardiovascular outcomes [1]-[6]. Second, deep learning extracts cardiometabolic information beyond a small set of handcrafted measures [7]-[11]. Third, OCT and OCTA add structural and capillary information, although device and software variability remain important [4], [5]. Fourth, retinal foundation models improve reusable representation learning but do not by themselves solve interpretability or calibration [12]. Fifth, real-world and external evaluations show promise while revealing subgroup, acquisition and generalization limitations [15].

These findings support the architecture shown in Figure 1 and the comparative design in Table 3. They do not establish that RetinaFusion-XAI improves clinical outcomes. That claim requires prospective multi-site data, pre-

specified endpoints and comparison with accepted clinical risk tools. The proposed experiments directly test whether explicit biomarkers add value beyond clinical variables, whether foundation embeddings add value beyond biomarkers, whether OCT/OCTA improve robustness and whether concept guidance improves interpretability without unacceptable loss of predictive performance.

5. Discussion

RetinaFusion-XAI integrates auditable clinical biomarkers with broad foundation embeddings via concept-guided fusion. Remote monitoring enables patient-specific surveillance, requiring the system to distinguish true biological changes from technical variations using quality metrics, metadata, and uncertainty measures. Successful clinical translation demands accurate calibration, equitable subgroup performance, external validation, clinically assessable explanations, and uncertainty-aware referral pathways.

6. Limitations and Future Work

This paper proposes a framework rather than a completed study. Limitations include restricted dataset availability, variations in clinical definitions and devices, disease related gradability issues, foundation model biases, and explanation fidelity. Future work should implement this architecture on multi-site data, compare it with standard clinical tools, assess its performance in the case of missing modalities, and evaluate clinical dashboards. Ultimately, a prospective study must test whether the system can improve patient management without increasing avoidable workload.

7. Conclusion

Retinal imaging provides non-invasive insights into cardiometabolic health. RetinaFusion-XAI addresses current research gaps in multimodal fusion, interpretability, and longitudinal assessment by integrating fundus, OCT/OCTA, and clinical data through complementary biomarker and embedding streams. This framework defines a reproducible, safe, and interpretable path for testing clinical value, though prospective multi-site validation remains essential before deployment.

Declarations

Conflict of interest: There is no conflict of interest in writing the research article.

Funding: No external funding for this conceptual framework.

Ethics: This article presents a framework and literature synthesis and does not report new experiments involving human participants. Any future clinical study must obtain institutional ethical approval and informed consent or an approved waiver, as applicable.

References

- [1] K. McGeechan et al., "Meta-analysis: Retinal vessel caliber and risk for coronary heart disease," *Ann. Intern. Med.*, vol. 151, no. 6, pp. 404-413, 2009, <https://doi.org/10.7326/0003-4819-151-6-200909150-00005>
- [2] Y. T. Ong et al., "Retinal vascular fractal dimension and its relationship with cardiovascular and ocular risk factors," *Am. J. Ophthalmol.*, vol. 153, no. 4, pp. 663-674, 2012, <https://doi.org/10.1016/j.ajo.2012.04.016>
- [3] C. Y. Cheung et al., "Retinal vascular tortuosity, blood pressure, and cardiovascular risk factors," *Ophthalmology*, vol. 118, no. 5, pp. 812-818, 2011, <https://doi.org/10.1016/j.ophtha.2010.08.045>
- [4] G. C. M. Cheung et al., "Impact of cardiometabolic factors on retinal vasculature: A 3 × 3, 6 × 6 and 8 × 8-mm optical coherence tomography angiography study," *Eye*, vol. 35, pp. 2269-2279, 2021, <https://doi.org/10.1111/ceo.13913>
- [5] C. Chua et al., "Systemic hypertension associated retinal microvascular changes can be detected with optical coherence tomography angiography," *Sci. Rep.*, vol. 10, art. 9580, 2020, <https://doi.org/10.1038/s41598-020-66736-w>
- [6] I. R. Mordi et al., "Prediction of major adverse cardiovascular events from retinal, clinical, and genomic data in individuals with type 2 diabetes: A population cohort study," *Diabetes Care*, vol. 45, no. 3, pp. 710-716, 2022, <https://doi.org/10.2337/dc21-1124>
- [7] R. Poplin et al., "Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning," *Nat. Biomed. Eng.*, vol. 2, pp. 158-164, 2018, <https://www.nature.com/articles/s41551-018-0195-0>
- [8] C. Y. Cheung et al., "A deep-learning system for the assessment of cardiovascular disease risk via the measurement of retinal-vessel calibre," *Nat. Biomed. Eng.*, vol. 5, pp. 498-508, 2021, <https://doi.org/10.1038/s41551-020-00626-4>
- [9] A. R. Rudnicka et al., "Artificial intelligence-enabled retinal vasculometry for prediction of circulatory mortality, myocardial infarction and stroke," *Br. J. Ophthalmol.*, vol. 106, no. 12, pp. 1722-1729, 2022, <https://doi.org/10.1136/bjo-2022-321842>
- [10] K. Zhang et al., "Deep-learning models for the detection and incidence prediction of chronic kidney disease and type 2 diabetes from retinal fundus images," *Nat. Biomed. Eng.*, vol. 5, pp. 533-545, 2021, <https://doi.org/10.1038/s41551-021-00745-6>
- [11] D. Squirrell et al., "Blood pressure predicted from artificial intelligence analysis of retinal images correlates with future cardiovascular events," *JACC Adv.*, vol. 3, no. 12, art. 101410, 2024, <https://doi.org/10.1016/j.jacadv.2024.101410>
- [12] Y. Zhou et al., "A foundation model for generalizable disease detection from retinal images," *Nature*, vol. 622, pp. 156-163, 2023, <http://nature.com/articles/s41586-023-06555-x>

- [13] M. Filos et al., "A systematic comparison of Bayesian deep learning robustness in diabetic retinopathy tasks," arXiv:1912.10481, 2019. <https://doi.org/10.48550/arXiv.1912.10481>
- [14] J. A. Leibig et al., "Leveraging uncertainty information from deep neural networks for disease detection," Sci. Rep., vol. 7, art. 17816, 2017, <https://doi.org/10.1038/s41598-017-17876-z>
- [15] W. Hu et al., "Real-world feasibility, accuracy and acceptability of automated retinal photography and AI-based cardiovascular disease risk assessment in Australian primary care settings: A pragmatic trial," npj Digit. Med., vol. 8, art. 84, 2025, <https://doi.org/10.1038/s41746-025-01436-1>
- [16] S. Gholami et al., "Distributed training of foundation models for ophthalmic diagnosis," Commun. Eng., vol. 4, art. 60, 2025, <https://doi.org/10.1038/s44172-025-00341-5>