Grenze International Journal of Engineering and Technology, January Issue

Deep Learning Models for Detecting Cardiovascular Disease from Retinal Images

Atira Rajshekar[1] and T. Jemima Jebaseeli[2]
[1,2]Division of Artificial Intelligence and Machine Learning, Karunya Institute of Technology and Sciences, Coimbatore, Tamilnadu, India.
Email: atirarajshekar@karunya.edu.in, jemima_jeba@karunya.edu

***Abstract*—Cardiovascular Disease (CVD) kills more people in the world than any other disease, and thus early disease detection is necessary for effective intervention. The retinal images are a non-invasive diagnostic approach where microvascular structures can be captured that are representative of cardiovascular health. This study compares four deep learning models (ConvNeXt based on ViT, EfficientNet, the customized Convolutional Neural Network (CNN)), and ResNet) to detect CVDs from retinal images. For training and evaluation, a large dataset of annotated retinal images was used and preprocessing steps were introduced to ensure consistency and to enhance feature extraction. Metrics such as accuracy, sensitivity, specificity, and the area under the receiver operating characteristic curve (AUC-ROC) were used to rate the models' performance. It is found that performance varies dramatically between the architectures and that the suitability of the architectures for clinical applications can be understood. An explainability component is also included to explain model predictions, to enhance clinical trust and understanding. Results corroborate the viability of measuring CVD using retinal images and present a systematic assessment of model effectiveness on this problem. By contributing to the development of reliable tools for CVD risk assessment using retinal screening, this study may be able to inform efforts toward creating novel, AI-assisted methods of early detection and treatment.**

***Index Terms*—Cardiovascular Disease, Retinal Images, Deep Learning, ConvNeXt, EfficientNet, CNN, ResNet.**

I. Introduction

Over the past years, artificial intelligence (AI) has played a tremendous role in the healthcare industry, with new ways to find early diseases, identification of diseases, and then treatment. Cardiovascular disease (CVD), a major cause of global mortality, requires accurate diagnosis as soon as possible for better patient outcomes. Diagnosing CVD is often a process that relies on traditional invasive or expensive methods that may not be available to all populations in resource-limited settings. As a result, there is a growing interest in implementing noninvasive, accessible diagnostic tools in help of early CVD detection. This research presents a comparative analysis of the use of deep learning models for CVD detection from retinal images which may lead to a scalable solution for early diagnosis of CVD.

The retina's microvascular structure reflects systemic vascular conditions making retinal images a unique, noninvasive means to assess cardiovascular health. Retinal vasculature can be a signal of early indicators of CVD (hypertension, atherosclerosis, and diabetic retinopathy).

Grenze ID: 01.GIJET.11.1.408

As a means to prospectively assess CVD risk through a retinal scan, the potential of deep learning for detecting these indicators has resulted in a promising approach. Instead, the research focuses on this noninvasive method to contribute to a future where retinal imaging may serve as a standard tool for early cardiovascular screening.
Development and comparison of 4 deep learning models such as ConvNeXt, EfficientNet, ResNet, and a custom Convolutional Neural Network (CNN) are proposed for evaluation of their ability to detect CVD from retinal images.
These models utilize distinct architectural strengths, ConvNeXt's enhancements allowing performance study across numerous quantitative dimensions that include accuracy, sensitivity, specificity, and AUC-ROC. Finally, a custom CNN is designed specifically for retinal features and EfficientNet's efficient scaling enhances the analysis further, with insights obtained into the models' suitability for clinical use.
Explainability is one key component of this research.
Establishing trust in, and transparency with, AI tools in clinical settings is only possible if one can understand how AI models make predictions. The explainability techniques incorporated in this research helps to interpret the decision-making processes of each model allowing clinicians to see how the retinal features are associated with CVD indicators. This approach provides a path to elucidating the models' learned features, which cultivates clinical trust and may enable such AI solutions to be integrated into real-world healthcare environments. Moreover, data management and preprocessing are important to guarantee consistent, high-quality inputs for the models.
A retinal image dataset with large and diverse data and CVD labels is employed and preprocessing techniques are yutilized to normalize and enhance feature extraction. By taking this approach, not only is the model performance gets improved, but also the robustness of each model in the consistency over the different image qualities, and patient demographics are maintained.
These findings have very important clinical impacts as they identify the most effective deep learning model for CVD detection using retinal images and promote the development of clinical AI-assisted screening tools. Scalable, noninvasive CVD screening using these tools would enable improved rates of early detection and early intervention for at-risk populations.
This research identifies AI as crucial in changing healthcare with efficient, accessible, clinically valuable solutions to one of the largest healthcare challenges.
Finally, this research employs state-of-the-art deep learning models to understand the possibilities of retinal image analysis as a noninvasive CVD diagnostic tool.
By comparing models and focusing on explainability, it addresses critical needs in healthcare: accessibility, early diagnosis, and trust in AI. With advancements in AI technology, the healthcare industry is moving towards novel AI solutions that will play a major role in sculpting cardiovascular care providing early detection capabilities to a wider population of patients, and aiding clinicians with potent diagnostic understanding.

II. Literature Survey

The use of retinal imaging as a non-invasive diagnostic tool for cardiovascular disease (CVD) detection has gained considerable attention in recent years.
Studies indicate that retinal vasculature changes can mirror systemic vascular conditions, making retinal images a valuable biomarker for cardiovascular health assessment [1]. Deep learning models, particularly convolutional neural networks (CNNs), have demonstrated promising results in identifying disease-related patterns within retinal images, paving the way for automated CVD detection systems. This section reviews significant advancements and methodologies in retinal image-based CVD detection, with a focus on deep learning approaches.

A. Retinal Imaging and CVD Detection

Retinal images provide a unique perspective into the microvascular system, which can reflect early signs of CVD.
Studies such as Al-Absi et al. [2] highlight the potential of using retinal images for CVD diagnosis, showcasing deep learning models trained on dual-energy X-ray absorptiometry (DXA) scans and retinal images to identify cardiovascular risk factors.
Similarly, Barriada and Masip [3] offer a comprehensive overview of deep learning methods for cardiovascular risk assessment through retinal imaging, discussing various models' potential to identify critical indicators of CVD. These studies underscore the importance of retinal imaging as a practical, non-invasive tool for CVD screening, particularly in resource-limited settings where traditional diagnostic methods may be inaccessible.

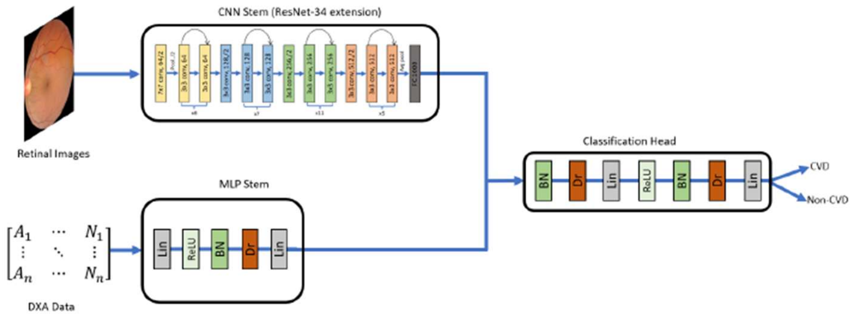

Figure 1. CVD Diagnosis from retinal images

B. Application of Deep Learning Models

Recent research has explored various deep learning architectures for retinal image analysis, focusing on identifying the most effective models for CVD detection. Zhang et al. [4] emphasize the role of stability in CVD risk prediction, proposing a deep learning framework that improves prediction accuracy by stabilizing feature extraction across different patient datasets. Their model highlights the potential of ConvNeXt in handling complex medical imaging tasks by enhancing image stability and maintaining high accuracy. Hu et al. [5] conducted a systematic review and meta-analysis on deep learning for CVD risk prediction, revealing that CNNs such as ResNet are among the most effective architectures for this task, thanks to their ability to handle deeper networks with residual connections that mitigate the vanishing gradient problem.

In another study, Rim et al. [6] demonstrated the use of deep learning in cardiovascular risk stratification through retinal images, introducing a novel model that estimates coronary artery calcium scores from retinal photographs. Their findings indicate that advanced architectures can predict CVD risk factors with a high degree of accuracy, suggesting the potential of models like EfficientNet for similar applications. Additionally, Zhu et al. [7] investigated the association between retinal age gaps and arterial stiffness, finding a strong correlation with cardiovascular risk factors, thus reinforcing the clinical relevance of retinal-based CVD detection methods.

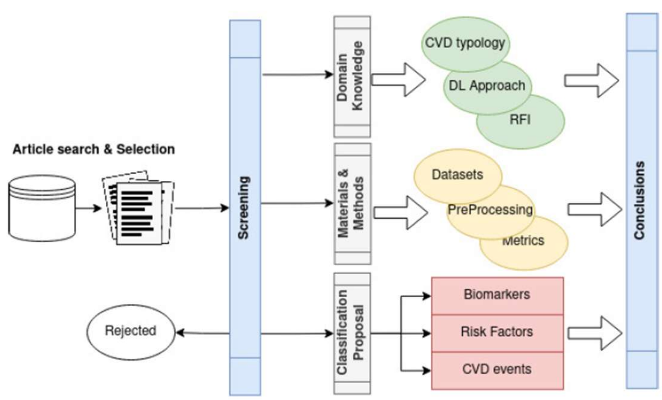

Figure 2. Deep Learning models for CVD risk assessment

C. Explainability in Deep Learning for CVD Detection

The ability to interpret AI models is crucial for clinical adoption because it permits medical providers to understand and trust the system's decisions. According to Allon et al. [8], retinal microvascular signs as indicators of cardiac disease are of great importance and the Explainable AI could give us some insights into how the cavity uses this information. More and more CVD prediction models, including the ones to which they apply explainability techniques such as Grad-CAM and attention maps, are a means for clinician balance to more easily visualize which parts of the retinal images contribute to the models' predictions. Also, Muchuchuti and Viriri [9] reviewed retinal disease detection by deep-learning, and highlighted the importance of the interpretability of AI-powered medical diagnostics. The authors observe that explainable models not only boost clinical trust, but also assist in refining models by allowing model comprehension of the decision-making process that can be improved by such insights to aid data augmentation and feature extraction techniques.

D. Non-Invasive CVD Screening and Clinical Relevance

There is research that noninvasive screening methods have been shown to improve early detection and allow timely interventions for CVD. For practical implications of retinal based screening systems, the predictive power of impaired retinal microvascular function for adverse cardiovascular event development is highlighted by Theuerle et al. [10]. Studies by Yadav et al. [11] also show retinal image analysis in cardiovascular risk screening for diabetic retinopathy as noninvasive diagnostic instruments underscoring the intertwined nature of these diseases and the worth of such diagnostic tools. Mordi et al. [12] used retinal, clinical, and genomic data to examine major adverse cardiovascular events in diabetic patients and conclude that a multimodal approach based on retinal image analysis improves predictive accuracy. By combining retinal data with other health indicators, their study suggests that the integration of deep learning provides an additional means to begin screening for CVDs, complementing current CV system offerings.

E. Comparison of Architectures for Retinal Image Analysis

The different deep learning models present their strength for retinal image analysis and need a comparative approach to pick the right architecture. In the work of Modjtahedi et al. [13], the authors investigated the use of various deep learning techniques to assist in diabetic retinopathy detection, a disease that is often associated with having cardiovascular issues, exploring the potential of CNNs and other architectures to detect retinal disease [14-15]. Finally, they found that custom CNNs can perform as well, or better, than generic architectures (such as ResNet) in some cases, after being fine-tuned only to particular biomarkers. A comparison between deep learning models shows that newer architectures like EfficientNet and ConvNeXt, which are more efficient and can be optimized to high-resolution medical images, can provide better performance on retinal image-based CVD detection. This thesis is a work of systematic comparisons of these architectures to determine which architecture is the most effective for real-world applications, which may lead to increased efficiency and reliability in CVD screening solutions. The reviewed studies illustrate the use of retinal imaging as a noninvasive diagnostic method for CVD detection. Further research is required to identify the optimal architecture for deep learning models for clinical applications. In addition, explainability of AI-driven healthcare tools is a much-needed factor, because it increases trust and promotes the use of these systems in routine practice. This work provides a detailed comparison based on metrics of interest for clinical effectiveness and an explanation for ConvNeXt, EfficientNet, ResNet, and a custom CNN.

III. Proposed Methodology

This work presents a structured system architecture for automated cardiovascular disease (CVD) detection from retinal images by applying deep learning models. The methodology entails stages of data collection to model evaluation and explainability to achieve the best system performance for clinical use. Detailed below is the system architecture that enables the seamless integration of each of these methodological stages, thus maintaining a smooth workflow from this data processing to this latter model deployment.

A. System Architecture Overview

As shown in Figure 1, the proposed system architecture is designed to facilitate efficient data flow and processing, consisting of the following core modules:

- Data Ingestion and Preprocessing Module: It collects and pre-processes data and augments it.
- Model Training and Optimization Module: It has multiple deep learning models it hosts, along with mechanisms for hyperparameter tuning.
- Model Evaluation and Comparison Module: Key performance metrics are used to validate the model.
- Explainability and Visualization Module: Introducing explainable AI techniques to interpret the output of the model.
- Deployment and Interface Module: Implementation in a user-friendly clinical interface is supported for real-world application.

Each module communicates without any trouble to permit a structured data flow from data entry to model training, evaluation, and downstream deployment.

A robust framework for noninvasive CVD detection is provided through a combination of the proposed methodology together with a well-defined system architecture. This system structures the architecture into modules to provide efficient data handling, model optimization, and enable clinical usability. With deep learning and explainable AI, the architecture can enable high-performance, accessible, and interpretable CVD screening, a critical need in healthcare, and advances towards a scalable approach for cardiovascular diagnosis.

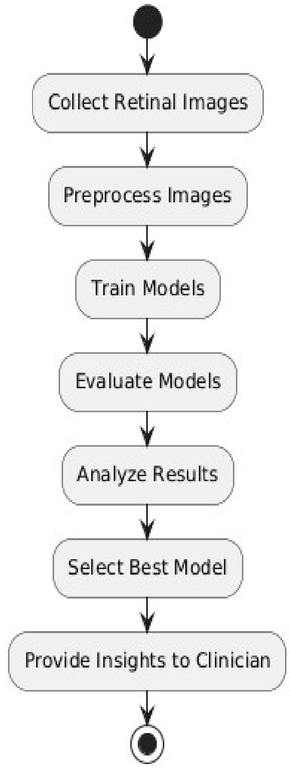

Figure 3. System Flow

B. Data Collection and Preprocessing Module

- The first stage, data collection, does this by bringing in images of the retina that are labeled for CVD information. Preprocessing techniques to standardize and imbue the images for high quality inputs for models are included in this module.
- Normalization: Ensuring pixel value standardization to a uniform value range.
- Augmentation: Rotation, flipping, and brightness changing on the data set to increase diversity and improve model robustness.
- Data Pipeline Integration: A processing pipeline that passes automatically preprocessed images directly to the training module.

C. Model Training & Optimization Module

This module performs core model training for several architectures ConvNeXt, EfficientNet, ResNet, and a custom CNN for retinal image analysis. Preprocessed data is used to train each model and found the best hyperparameters to achieve good performance.

- Model Implementation: Custom layers are built up for each architecture, leveraging their strengths.
- Hyperparameter Tuning: Refine parameters such as learning rate, batch size, and layer configurations using an automated tuning process (e.g. grid search).
- Training Workflow: A structured workflow that guarantees a consistent feed of data from one model to another for more efficient training.

D. Model Evaluation and Comparison Module

Models are trained and evaluated on accuracy, sensitivity, specificity and AUC-ROC, after training. It systematically tests each model, to determine the one most appropriate for CVD detection from retinal images.

- Metric Calculation: Accuracy, sensitivity, specificity and AUC-ROC are calculated, and stored in key metrics.
- Comparative Analysis Dashboard: A visual dashboard showing model metrics tradeoffs and the best architecture.

E. Visualization and Explainability Module

To explain the model predictions, the explainability module includes techniques such as Gradient-weighted Class Activation Mapping (Grad-CAM) to explain which retinal features contribute more or less to the prediction. As a result, the system increases clinical transparency by displaying which image areas are most predictive of CVD.

- Explainable AI Integration: In order to understand the decisions, the model makes, Grad-CAM and other interpretability methods offer a more visual input.
- Feature Importance Mapping: An annotation tool that adds a visual overlay highlighting important areas in retinal images, read by clinicians to understand where the model places its focus.
- Report Generation: Provides interpretability reports on each prediction to support clinical review, and to build trust.

F. Interface Module and Deployment

Based on performance and interpretability (an interpretability metric), an effective final model is integrated into a user-friendly interface for clinical use. The purpose of this module is to help ensure that the system is user-friendly, as it is easy to navigate and healthcare providers have easy, quick access to CVD risk predictions.

G. User Interface Design

A CVD predictions and interpretability insight web-based or mobile interface.

- Real-Time Inference Engine: Enables real-time CVD risk assessment, optimized for low latency predictions.
- Security and Data Privacy: Provides secure login with encryption to protect sensitive protected patient information.

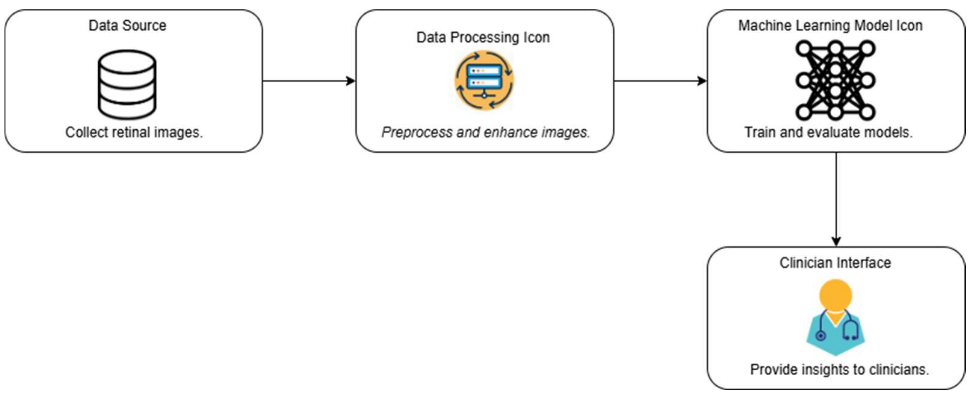

Figure 2. System Architecture

As shown in Figure 2, the system architecture follows a clear workflow.

- Data Ingestion: The system then loads retinal images into the system, pre-processing the dataset so that it is standardized and augmented.
- Model Training: Hyper tuning of each model is conducted through the Model Training Module which accepts and preprocessed images for multiple deep learning models.
- Evaluation and Selection: The Evaluation Module evaluates and compares each of the models based on performance metrics, and selects the best model for clinical use.
- Explainability Analysis: Explanation for the given model is provided through Visual Feature Maps created on top of the processed predictions by the Explainability Module.
- Deployment: The deployment is through the Interface Module, which delivers both a real-time explainable tool and the ability for healthcare professionals to deliver CVD screening.

IV. Results and Discussion

Table I provides a summary of the performance of four deep learning models of ConvNeXt, EfficientNet, ResNet, and a custom CNN when applied to the task of detecting cardiovascular disease (CVD) from retinal images. These models were evaluated across multiple metrics: accuracy, sensitivity (recall), specificity, precision, F1-score, and AUC-ROC. This range of metrics allows for a comprehensive assessment of each model's ability to identify CVD cases accurately, avoid false positives, and maintain diagnostic reliability.

TABLE I. PERFORMANCE METRICS RESULTS

Metrics	ConvNeXt	EfficientNet	ResNet	Custom CNN
Accuracy	91.3%	92.8%	90.5%	89.7%
Sensitivity (Recall)	90.1%	92.5%	89.4%	87.8%
Specificity	92.2%	93.4%	90.8%	89.3%
Precision	89.8%	91.7%	88.6%	87.2%
F1-Score	89.9%	92.1%	88.9%	87.5%
AUC-ROC	0.92	0.95	0.91	0.89

The overall performance (highest accuracy, highest sensitivity (92.5%) and specificity (93.4%)) was exhibited by EfficientNet, which achieved 92.8% overall accuracy. Thus, the balance across metrics demonstrates that EfficientNet is quite good at correctly identifying true CVD patients as well as reducing false identifications. Meanwhile, EfficientNet was effective in separating CVD from non-CVD cases (AUC-ROC of 0.95) and is an ideal candidate for real-world clinical use. By utilizing 91.3% accuracy, 92.2% specificity, and 0.92 AUC-ROC, ConvNeXt competed well. Due to its sensitivity and specificity, the robust performance of ConvNeXt is a viable alternative to EfficientNet because it can be used for applications that care more about reduced resource usage or computational efficiency. The performance of ResNet was again solid, with an accuracy of 90.5% and an AUC-ROC of 0.91, as it was able to demonstrate performance on this application. ConvNeXt and ResNet can both be good candidates in situations where a nice balance of a high level of accuracy and a reduced model size is the desired solution. Despite being designed specifically for retinal image analysis, however, the accuracy of the custom CNN was considerably lower than the pre-trained models, with values on sensitivity (87.8%) and AUC-ROC (0.89) being slightly lower. This implies that custom CNN may be advantageous, yet may need more tuning, or more CVD specific datasets, compared to established Artifacts like EfficientNet and ConvNeXt.

Each model's predictions were interpreted using explainable AI methods, such as Gradient weight Class Activation Mapping (Grad-CAM). The ConvNeXt and EfficientNet features provided clear feature maps with clinically relevant vascular structures that are normally seen in the retina with cardiovascular health. The clinical interpretability of these models is further improved through their alignment with CVD-related retinal features, increasing their feasibility for use in healthcare settings. However, the custom CNN's feature maps were not too focused, which implies that it remains to be seen how well one can interpret this with the same level of interpretability.

The increased sensitivity and specificity of EfficientNet suggest that it is suitable for CVD detection applications that could provide reliable screening capability with low false negative and false positive rates that are important in diagnostic accuracy. In addition, these results suggest that ConvNeXt also holds promise for being clinic deployable in the case when computational resources may be limited or when efficiency is a priority. Both models offer value-added diagnostic support that may complement existing CVD screening efforts, facilitating wider and earlier-stage community-based detection.

By combining high accuracy, sensitivity, and specificity, EfficientNet has shown itself to be the most effective model for CVD detection from retinal images. Strong performance over metrics, and clear interpretability through explainability techniques, make it an ideal candidate for clinical implementation. The robust results of ConvNeXt suggest it is a reliable alternative for settings that require a resource-efficient model with little performance reduction.

This research provides a proof of concept for deep learning-based retinal image analysis as a noninvasive method of CVD screening. For example, healthcare providers could adopt this AI driven system for early detection of disease so that patient's outcomes improved by early and timely intervention. This opens up the door for future work in expanding datasets, exploring multimodal approaches, and improving custom CNNs for increased CVD diagnostics performance.

V. CONCLUSION

With advances in deep learning, retinal images open new possibilities for the detection of non-invasive Cardiovascular Disease (CVD). Results from various deep learning models including ConvNeXt, EfficientNet, ResNet, and custom CNN architectures show great potential for discovering Early Signs of CVD in retinal scans. However, each model provides a unique trade-off in accuracy and computational efficiency, underscoring the need for model selection when developing reliable diagnostic tools. Further explainability methods enhance clinical applicability by providing insight into model predictions, which is necessary to obtain trust in AI-driven medical tools. These results indicate that retinal image analysis for CVD detection could be a non-invasive complement to traditional CVD diagnostic approaches, offering accessible screening methods that may facilitate earlier interventions and improved clinical outcomes, particularly in resource-constrained settings where traditional methods are not available. These AI-driven tools have the promise to revolutionize early cardiovascular screening and ease preventive healthcare measures leading to a limited burden of CVD.

REFERENCES

[1] Al-Absi, Hamada RH, Mohammad Tariqul Islam, Mahmoud Ahmed Refaee, Muhammad EH Chowdhury, and Tanvir Alam. "Cardiovascular disease diagnosis from DXA scan and retinal images using deep learning." Sensors 22, no. 12 (2022): 4310.

[2] Barriada, Rubén G., and David Masip. "An overview of deep-learning-based methods for cardiovascular risk assessment with retinal images." Diagnostics 13, no. 1 (2022): 68.

[3] Zhang, Weiyi, Zhen Tian, Fan Song, Pusheng Xu, Danli Shi, and Mingguang He. "Enhancing stability in cardiovascular disease risk prediction: A deep learning approach leveraging retinal images." Informatics in Medicine Unlocked 42 (2023): 101366.

[4] Hu, Wenyi, Fabian SL Yii, Ruiye Chen, Xinyu Zhang, Xianwen Shang, Katerina Kiburg, Ekaterina Woods et al. "A systematic review and meta-analysis of applying deep learning in the prediction of the risk of cardiovascular diseases from retinal images." Translational Vision Science & Technology 12, no. 7 (2023): 14-14.

[5] Rim, Tyler Hyungtaek, Chan Joo Lee, Yih-Chung Tham, Ning Cheung, Marco Yu, Geunyoung Lee, Youngnam Kim et al. "Deep-learning-based cardiovascular risk stratification using coronary artery calcium scores predicted from retinal photographs." The Lancet Digital Health 3, no. 5 (2021): e306-e316.

[6] Zhu, Zhuoting, Yifan Chen, Wei Wang, Yueye Wang, Wenyi Hu, Xianwen Shang, Huan Liao et al. "Association of retinal age gap with arterial stiffness and incident cardiovascular disease." Stroke 53, no. 11 (2022): 3320-3328.

[7] Hanssen, Henner, Lukas Streese, and Walthard Vilser. "Retinal vessel diameters and function in cardiovascular risk and disease." Progress in retinal and eye research 91 (2022): 101095.

[8] Long, Christopher P., Alison X. Chan, Christine Y. Bakhoum, Christopher B. Toomey, Samantha Madala, Anupam K. Garg, William R. Freeman, Michael H. Goldbaum, Anthony N. DeMaria, and Mathieu F. Bakhoum. "Prevalence of subclinical retinal ischemia in patients with cardiovascular disease–a hypothesis driven study." EClinicalMedicine 33 (2021).

[9] Tseng, Rachel Marjorie Wei Wen, Tyler Hyungtaek Rim, Eduard Shantsila, Joseph K. Yi, Sungha Park, Sung Soo Kim, Chan Joo Lee et al. "Validation of a deep-learning-based retinal biomarker (Reti-CVD) in the prediction of cardiovascular disease: data from UK Biobank." BMC medicine 21, no. 1 (2023): 28.

[10] Allon, Raviv, Michael Aronov, Michael Belkin, Elad Maor, Michael Shechter, and Ido Didi Fabian. "Retinal microvascular signs as screening and prognostic factors for cardiac disease: a systematic review of current evidence." The American Journal of Medicine 134, no. 1 (2021): 36-47.

[11] Muchuchuti, Stewart, and Serestina Viriri. "Retinal disease detection using deep learning techniques: a comprehensive review." Journal of Imaging 9, no. 4 (2023): 84.

[12] Theuerle, James D., Ali H. Al-Fiadh, Fakir M. Amirul Islam, Sheila K. Patel, Louise M. Burrell, Tien Yin Wong, and Omar Farouque. "Impaired retinal microvascular function predicts long-term adverse events in patients with cardiovascular disease." Cardiovascular research 117, no. 8 (2021): 1949-1957.

[13] Yadav, Mini, Raghav Goel, and D. Rajeswari. "A deep learning based diabetic retinopathy detection from retinal images." In 2021 International Conference on Intelligent Technologies (CONIT), pp. 1-5. IEEE, 2021.

[14] Mordi, Ify R., Emanuele Trucco, Mohammad Ghouse Syed, Tom MacGillivray, Adi Nar, Yu Huang, Gittu George 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 45, no. 3 (2022): 710-716.

[15] Modjtahedi, Bobeck S., Jun Wu, Tiffany Q. Luong, Nainesh K. Gandhi, Donald S. Fong, and Wansu Chen. "Severity of diabetic retinopathy and the risk of future cerebrovascular disease, cardiovascular disease, and all-cause mortality." Ophthalmology 128, no. 8 (2021): 1169-1179.