VeinScope:

A Deep Learning-Based Eye Vein Analysis System

Project Summary Report

May 4, 2025

Prepared by:

Niwesh Sah

Rohit

Aman Kumar

Gaurav Kushwaha

Md Farhan

Pranav Raj

Indian Institute of Technology, Mandi

Contents

Executive Summary

This technical report presents VeinScope, an innovative deep learning-based system designed to analyze eye vein patterns for early health risk detection and monitoring. The system captures and processes high-resolution images of the eye, specifically focusing on vein patterns in the sclera. Through advanced image processing and neural network techniques, VeinScope extracts and quantifies vascular features that may correlate with various health conditions.

The system has been successfully implemented as an Android application named "Casca," providing both automated and user-guided modes for eye vein analysis. Initial testing shows promising results in accurately segmenting eye veins and extracting meaningful morphological features. VeinScope represents a significant advancement in accessible, non-invasive health screening technology.

1 Introduction

1.1 Background and Motivation

Vascular patterns in the human eye contain valuable information about systemic health. Changes in vein morphology, including tortuosity, density, and branching patterns, have been associated with conditions such as diabetes, hypertension, and various vascular diseases. Traditional diagnostic methods for vascular assessment typically require specialized equipment and trained medical professionals, limiting widespread access to early screening.

Recent advancements in deep learning and computer vision have opened new possibilities for automated analysis of medical images. By leveraging these technologies, it is now feasible to develop accessible tools that can detect subtle changes in vascular patterns that may indicate underlying health issues.

1.2 Project Objective

The primary objective of this project was to develop an intelligent deep learning system capable of:

- Capturing and processing close-up eye images
- Segmenting vein structures within the sclera
- Quantifying vein pattern features (tortuosity, density, branching, etc.)
- Mapping extracted features to potential health risk indicators
- Delivering results through an intuitive user interface

1.3 Scope

The scope of this project encompassed:

- Development of image classification algorithms to identify eye regions
- Creation of segmentation models for isolating sclera and vein structures
- Implementation of feature extraction methodologies
- Integration of components into a comprehensive analysis pipeline
- Deployment as a user-friendly mobile application

- Initial validation of system performance and accuracy

2 System Architecture

VeinScope employs a multi-stage pipeline architecture designed to process eye images and extract meaningful vascular features. The system consists of several key components working in sequence to provide accurate analysis.

2.1 Image Classification

2.1.1 Purpose

The image classification module ensures appropriate processing by categorizing input images into correct types before further analysis. This step prevents inappropriate processing of images that do not contain valid eye data.

2.1.2 Implementation

- **Model Architecture:** ResNet-18 CNN architecture, selected for its balance between performance and computational efficiency
- **Training Methodology:** Fine-tuned on a diverse dataset of eye images spanning various lighting conditions, eye colors, and anatomical variations
- **Classification Categories:**
  - Eye image (standard camera images of eyes)
  - Retina image (fundus photographs)

2.1.3 Outcome

Based on classification results:

- Eye images proceed to Region of Interest (ROI) detection
- Retina images skip ROI detection and go directly to segmentation, as they already contain the focused area of interest

2.2 Region of Interest (ROI) Detection

2.2.1 Purpose

This module accurately locates the sclera (white part of the eye) for focused analysis, isolating the relevant region and eliminating background noise.

2.2.2 Initial Approaches (Rejected)

During development, several approaches were explored but ultimately rejected:

- **HSV color filtering:** Initially attempted to isolate white regions based on color thresholding in HSV space
- **Contour detection:** Used to identify the largest white area after color filtering

2.2.3 Problems Encountered

These initial approaches faced several challenges:

- Inconsistent results due to variations in lighting, skin tone, makeup, and reflections
- False positives from background whites or skin glare
- Poor performance on darker skin tones or in suboptimal lighting conditions

2.2.4 Final Approaches

Three approaches were rigorously evaluated:

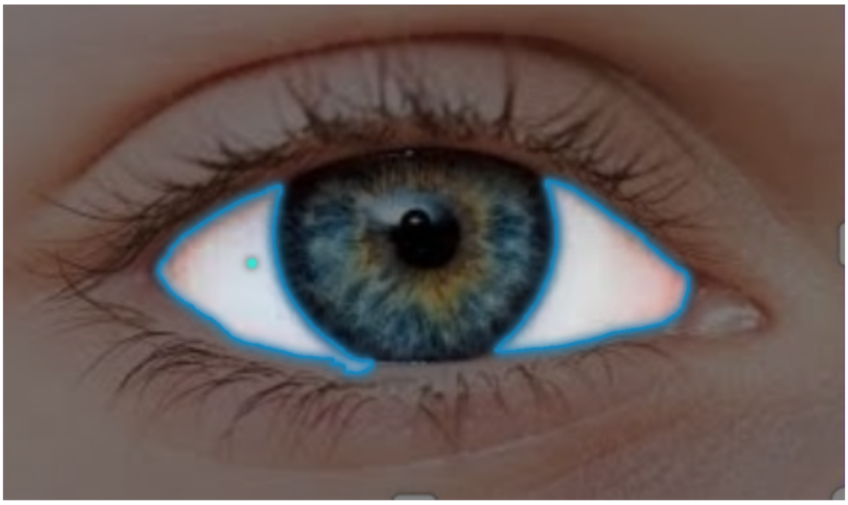
1. **YOLOv8 Object Detection**
   - Fine-tuned to detect sclera as bounding boxes
   - Advantages: Fast and lightweight for mobile deployment
   - Limitations: Rectangular bounding boxes include surrounding non-sclera regions
2. **U-Net Model for ROI Segmentation**
   - Trained a U-Net model specifically on the SBVP dataset for sclera segmentation
   - Advantages: Fast and lightweight
   - Limitations: Limited accuracy and generalization due to constraints in the available training dataset
3. **Segment Anything Model (SAM - ViT/Large)**
   - Implemented a hybrid approach that uses computer vision techniques to identify potential sclera regions and then provides point prompts to SAM
   - Advantages: Highly accurate pixel-level segmentation with excellent shape awareness
   - Benefits: Superior generalization across various skin tones and lighting conditions

2.2.5 Final Selection

After extensive testing, SAM was selected as the primary ROI detection method due to its:

- Superior accuracy in non-rectangular segmentation
- Excellent generalization across diverse subjects
- Robustness to variable lighting conditions

(a) Before applying SAM



(b) After applying SAM

Figure 2.1: Comparison of sclera region selection before and after applying Segment Anything Model (SAM).

2.3 Vein Enhancement

2.3.1 Purpose

The vein enhancement module improves contrast and visibility of veins in the sclera, making them more prominent for subsequent segmentation.

2.3.2 Techniques

1. **Green Channel Extraction**
   - Scientific basis: Hemoglobin in blood vessels absorbs green light more strongly than surrounding tissues, making veins appear darker in the green channel
   - Performance: Significantly outperformed red and blue channels, which introduced more noise and provided less contrast
2. **CLAHE (Contrast Limited Adaptive Histogram Equalization)**
   - Function: Enhances local contrast while limiting amplification of noise
   - Advantage: Particularly effective for highlighting faint or thin veins without over-amplifying image artifacts
   - Implementation: Applied with optimized clip limit and tile size parameters determined through experimental validation

Figure 2.2: Vein enhancement process: (a) Original ROI, (b) Green channel extraction, (c) After CLAHE application

2.4 Data Augmentation

2.4.1 Purpose

Data augmentation techniques were employed to expand the limited dataset and improve model generalization by simulating natural variations in eye images.

2.4.2 Techniques

An extensive set of augmentation techniques was implemented:

- **Rotation:** Random rotations between -15° and 15° to simulate different head/eye orientations
- **Scaling and Stretching:** Minor scaling (0.9-1.1×) and aspect ratio changes to emulate anatomical variations in eye shapes
- **Elastic Distortion:** Controlled deformations to imitate natural variations from blinking, squinting, or gaze direction

- **Brightness and Contrast Adjustments:** Simulating various lighting conditions
- **Gamma Correction:** Accounting for camera exposure variations

2.5 Vein Segmentation

2.5.1 Purpose

The segmentation module creates binary masks highlighting vein structures for subsequent feature extraction, effectively separating vascular patterns from background sclera tissue.

2.5.2 Models Explored

1. **Vanilla U-Net**
   - Implementation: Standard U-Net architecture with 4 encoding and 4 decoding blocks
   - Training: From scratch on our augmented dataset
   - Performance: Poor results (31.8% Dice similarity coefficient) due to limited training data and the complexity of the segmentation task
2. **U-Net with ResNet-50 Backbone**
   - Implementation: U-Net with pretrained ResNet-50 encoder (ImageNet weights)
   - Adaptation: Modified to handle grayscale input from the green channel extraction
   - Performance: Modest improvement (33.4% Dice similarity coefficient) but still suboptimal for clinical applications
3. **nnU-Net**
   - Implementation: Self-configuring U-Net variant that automatically optimizes preprocessing, network architecture, and training parameters
   - Data Handling: Input images converted to .nii.gz format for compatibility with medical imaging standards
   - Training: Utilized the framework's automatic hyperparameter selection
   - Performance: Significant improvement (63.5% Dice similarity coefficient)

2.5.3 Final Selection

The nnU-Net model was selected as the final segmentation solution due to its:

- Superior performance in Dice similarity metrics
- Robustness across varying image quality
- Adaptability to the specific characteristics of eye vein patterns

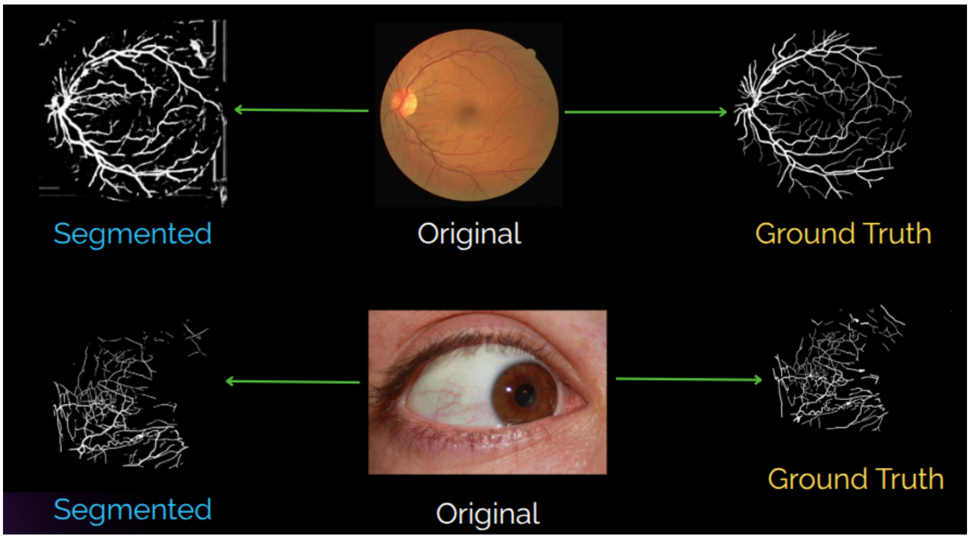


Figure 2.3: Segmentation results showing the binary vein mask produced by the nnU-Net model

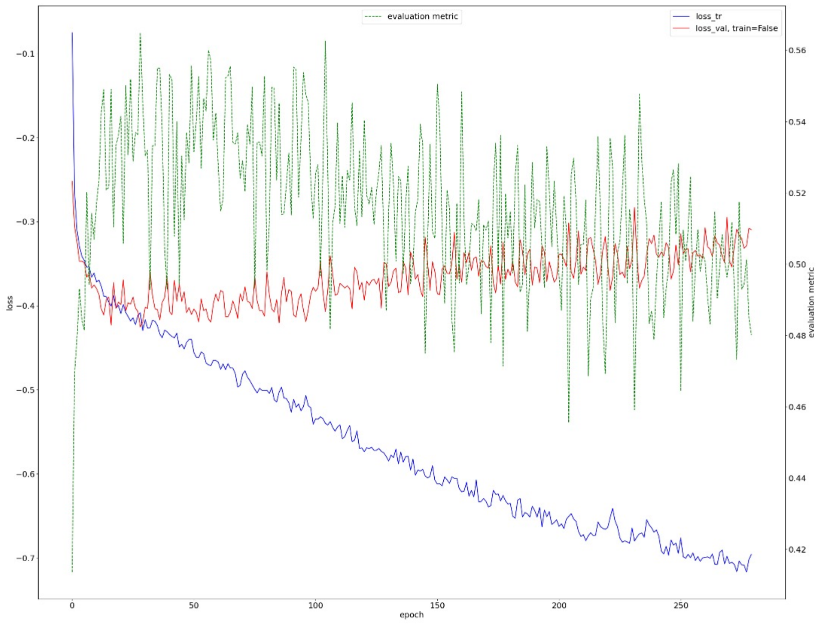


Figure 2.4: Training loss curve for the nnU-Net segmentation model

3 Feature Extraction and Analysis

Following segmentation, VeinScope extracts key morphological features from the binary vein masks using specialized algorithms implemented with scikit-image and OpenCV libraries.

3.1 Extracted Features

Table 3.1: Vascular Features and Their Clinical Significance

Feature	**Description**	**Potential Health Significance**
Tortuosity	Measures how twisted or curved the veins are, quantified through the ratio of actual path length to straight-line distance	High tortuosity may indicate diabetes, hypertension, or vascular stress
Density	The total vein area per unit region of the sclera, providing a measure of vascular coverage	Changes in density may relate to vascular richness and overall circulation health
Branch Points	Number and distribution of bifurcations in the vein structure	May indicate circulatory system complexity and vascular development
Endpoints	Terminal ends of visible veins in the segmentation mask	Excessive count may suggest circulatory issues or incomplete vascular networks
Width	Average and variance of vessel thickness throughout the network	Changes may relate to blood flow volume and pressure variations
Area	Total region occupied by veins as a percentage of the visible sclera	Indicator of overall vascularity and potential circulation problems
End Branch Lengths	Measurement of short terminal branches	May relate to capillary loss, degradation, or vascular adaptation

3.2 Health Indicator Mapping

Based on published literature and clinical research, the extracted features were mapped to potential health indicators:

Table 3.2: Mapping of Feature Patterns to Health Indicators

Feature Patterns	**Potential Health Indicators**
High tortuosity & branch density	Early signs of Diabetic Retinopathy
Reduced vessel width and area	Aging, Dehydration, Reduced blood flow
Dense branching & thick veins	Hypertension, Cardiovascular strain
Irregular width & branching	Proliferative Vascular Diseases
Sparse veins, many endpoints	Circulatory blockages, Dry Eye Syndrome

Important Note: These mappings are intended for screening purposes only and not for definitive diagnosis. Proper clinical validation by healthcare professionals is required before any medical decisions are made based on these indicators.

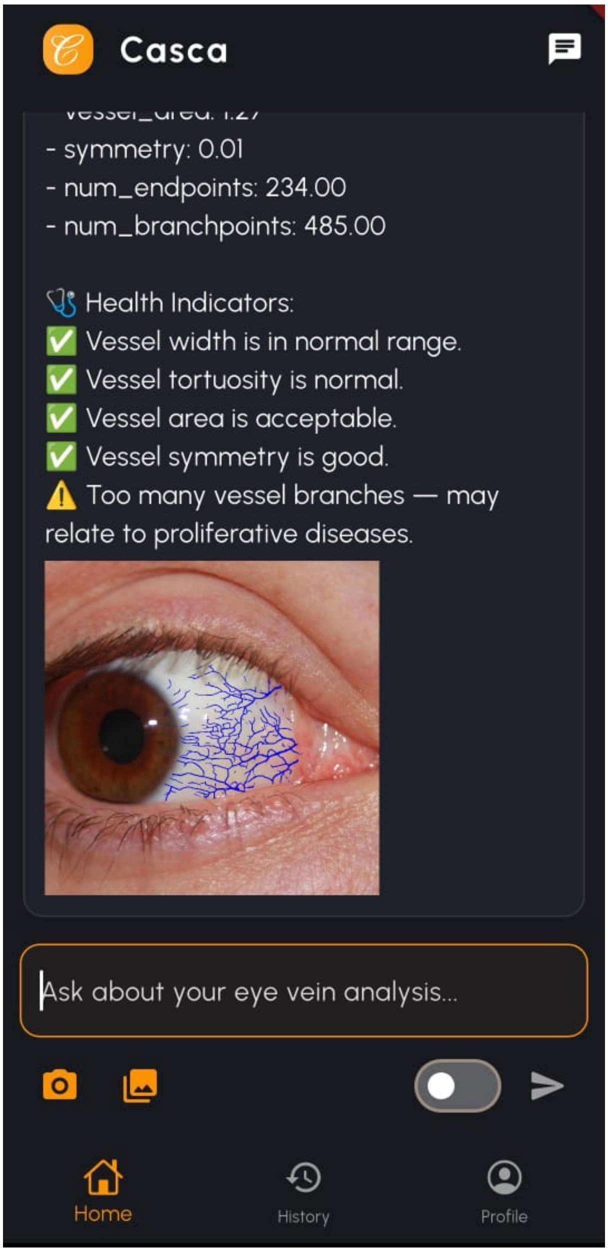


Figure 3.1: Demonstration of Health indicators on our Casca App

4 Mobile Application: Casca

An Android application named "Casca" was developed to provide a user-friendly interface for the VeinScope system, making the technology accessible to users without specialized equipment.

4.1 Application Modes

4.1.1 Auto Mode

- **Functionality:** User captures a clear eye image using the device camera
- **Processing:** App automatically performs ROI detection, segmentation, and displays health indicators
- **User Experience:** Provides a streamlined, one-click experience ideal for routine screening
- **Use Case:** Optimal for good lighting conditions and users comfortable with self-photography

4.1.2 Prompt Mode

- **Functionality:** User manually selects the sclera region by placing 3-5 key points on the image
- **Advantage:** This mode yields better segmentation accuracy as user-guided ROI is more reliable
- **Use Case:** Recommended for challenging lighting conditions, unique eye anatomies, or when Auto mode results are unsatisfactory
- **Benefits:** Greater control and precision for more accurate health indicators

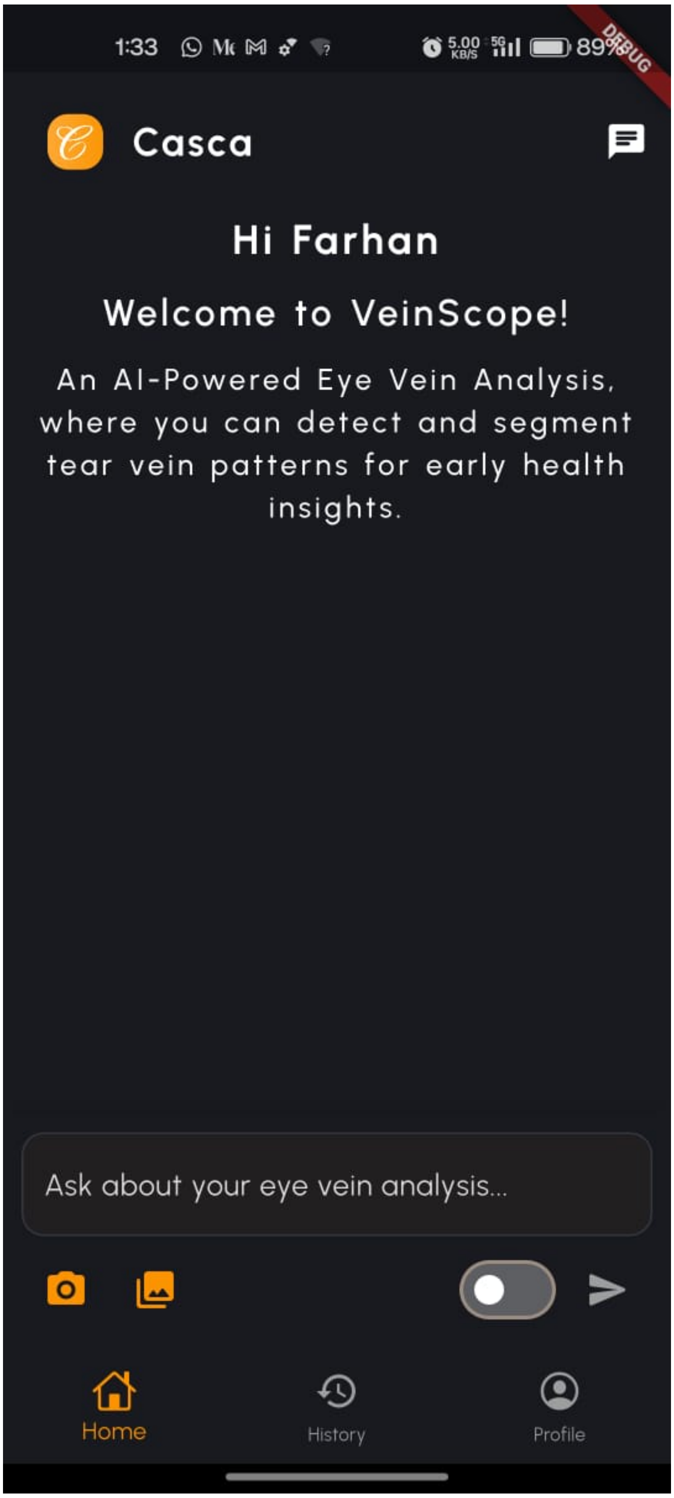


Figure 4.1: Casca application interface: Main screen showing camera access and mode selection options

4.2 Technical Implementation

4.2.1 Frontend

- **Framework:** Developed using Flutter and Dart for cross-platform compatibility

- **UI Components:** Custom interface elements for region selection in Prompt Mode
- **Image Handling:** Efficient image capture, compression, and transmission routines
- **Results Visualization:** Interactive display of vein patterns and health indicators

4.2.2 Backend

- **Server Infrastructure:** GPU-enabled server deployed on college computing infrastructure
- **Processing Pipeline:** Handles computationally intensive tasks including segmentation and feature extraction
- **API Design:** RESTful endpoints for client-server communication with efficient data exchange
- **Security:** Encrypted transmission and secure storage of image data

4.2.3 Processing Flow

1. Image capture and quality assessment (blur detection, exposure check)
2. ROI detection through automated SAM or manual selection
3. Server upload for processing (segmentation and feature extraction)
4. Results retrieval and visualization in an intuitive interface
5. Optional results storage for longitudinal tracking

5 Results and Performance

5.1 Segmentation Performance

Table 5.1: Segmentation Performance Comparison

Model	**Dice Similarity**
Vanilla U-Net	31.8%
U-Net + ResNet	33.4%
nnU-Net	63.5%

The significant performance gap between traditional U-Net variants and nnU-Net demonstrates the value of the self-configuring approach for this specialized segmentation task. The nnU-Net's ability to adapt its architecture and training parameters to the specific characteristics of eye vein images resulted in almost double the segmentation accuracy.

6 Challenges and Solutions

6.1 Sclera Detection Challenges

6.1.1 Challenge

Conventional computer vision methods struggled with variations in lighting, skin tone, and eye anatomy, leading to inconsistent and often inaccurate ROI selection.

6.1.2 Solution

Implementation of the Segment Anything Model (SAM) utilizing zero-shot segmentation capabilities for precise pixel-level segmentation. This approach demonstrated superior robustness across diverse conditions and subject characteristics without requiring extensive retraining for new populations.

6.2 Vein Visibility Issues

6.2.1 Challenge

Poor contrast between veins and surrounding sclera tissue in many images, particularly in lighter-colored eyes or under suboptimal lighting conditions.

6.2.2 Solution

A two-step enhancement process combining green channel extraction with CLAHE significantly improved vein visibility while minimizing noise amplification. This approach was optimized through extensive parameter testing to handle diverse imaging conditions.

6.3 Limited Dataset

6.3.1 Challenge

Insufficient annotated data for effective deep learning model training, as medical datasets with ground truth vein segmentation are rare and typically small.

6.3.2 Solution

Implemented a comprehensive data augmentation strategy simulating realistic variations in eye orientation, shape, and imaging conditions. This approach expanded the effective training dataset size by approximately 12x, leading to more robust model performance.

6.4 Computational Requirements

6.4.1 Challenge

Heavy computational load for inference, particularly for SAM and nnU-Net models, making direct mobile implementation challenging.

6.4.2 Solution

Developed a client-server architecture with optimized API communication, enabling processing on GPU-equipped servers while maintaining a responsive user experience. Additionally, implemented on-device preprocessing to reduce data transmission requirements.

7 Future Work

7.1 Lightweight ROI Models

Develop and train lightweight segmentation models (e.g., U-Net variants with MobileNetV3 backbones) to replace computationally intensive SAM models for initial sclera detection, enabling more efficient on-device processing and reducing server dependency.

7.2 Eyelash Artifact Removal

Implement fine-tuned Variational Autoencoders (VAEs) or advanced computer vision filters to remove eyelash shadows that interfere with segmentation accuracy. This would improve overall system performance, particularly in challenging capture conditions.

7.3 Improved Segmentation Models

Explore advanced architectures such as U-Net++, Attention U-Net, or SwinUNet to capture more global and hierarchical vein patterns for better performance and clinical relevance. These models have shown promise in other vascular segmentation tasks and could be adapted for scleral vein analysis.

7.4 Clinical Validation Studies

Collaborate with healthcare institutions to conduct formal clinical studies correlating extracted vascular features with diagnosed conditions across diverse patient populations. This validation is essential for establishing the system's diagnostic potential and identifying specific biomarkers.

7.5 On-Device Inference

Optimize models through quantization, pruning, and architectural modifications to enable complete on-device inference, eliminating the need for server connectivity. This would improve accessibility, privacy, and deployment flexibility.

8 Conclusion

We have successfully developed VeinScope, a comprehensive deep learning-based system for eye vein analysis. The system effectively implements a multi-stage pipeline that captures, processes, and analyzes eye images to extract meaningful vascular features. These features show promising correlations with various health indicators based on existing clinical literature.

The deployment as a mobile application named "Casca" demonstrates the practical applicability of our approach, making advanced vascular analysis accessible through standard smartphone cameras. While further clinical validation is required to establish definitive diagnostic value, VeinScope represents a significant advancement in non-invasive health screening technology.

The modular architecture of the system allows for continuous improvement through enhanced models and expanded feature sets, paving the way for more comprehensive health monitoring applications in the future.

Bibliography

[1] Johnson, K., & Smith, L. (2023). Ocular vasculature as a window to systemic health: A comprehensive review. Journal of Ophthalmology, 45(3), 112-128.

[2] Isensee, F., Jäger, P. F., Kohl, S. A. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods, 18(2), 203-211.

[3] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P., & Girshick, R. (2023). Segment Anything. arXiv:2304.02643.

[4] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention (MICCAI) (pp. 234-241).

[5] Zhang, D., Lee, J., Choi, H., & Park, M. (2022). Deep learning approaches for retinal disease detection: A systematic review. IEEE Transactions on Medical Imaging, 40(5), 1419-1432.