Fundus Image Segmentation for Clinical Application – Quantitative Analysis of Retinal Structures and Pathologies

Alicja Krzemińska-Ściga

Supervisors: Prof. Dr. Raphael Sznitman, Dr. Sandro De Zanet
Institutions: ARTORG Center for Biomedical Engineering Research, University of Bern
Ikerian AG (formerly RetinAI Medical AG)

Introduction

Fundus imaging is a non-invasive technique widely used for diagnosing and monitoring retinal diseases. Among these, Diabetic Retinopathy (DR) is a vision-threatening condition affecting over 100 million people worldwide. Early detection and progression monitoring are essential for effective disease management, driving the development of automated analysis methods. Deep Learning approaches have achieved high accuracy in determining patient referral, but typically provide no insight into the underlying pathology. This work proposes a solution that enables explainable screening by combining segmentation that detects relevant lesions and rule-based classification, as visualized in Fig. 1.

Materials and Methods

The segmentation model was trained and optimized using an internal dataset from RetinAI. A U-Net architecture with an EfficientNet encoder backbone was employed. The segmentation output includes anatomical structures and DR–related lesions: microaneurysms (MA), hemorrhages (H), hard exudates (HE), and cotton wool spots (CWS). Lesion counts were extracted from the segmentation output and used as biomarkers, serving as features for the classification stage. Several classifiers were evaluated, including a rule-based approach based on the International Clinical Diabetic Retinopathy (ICDR) severity scale, a Decision Tree, and a Random Forest. The proposed approach was validated on public datasets, including DDR (for segmentation) and APTOS, Messidor-2, and EyePACS (for classification).

Results

The segmentation model achieves AUC-PR 55.64 (HE), 32.79 (H), 30.14 (CWS), and 12.05 (MA) on the DDR dataset. Classification performance, evaluated using quadratic weighted kappa (QWK), is reported in Table 1.

Dataset	ICDR	Decision Tree	Random Forest
APTOS	0.23 ± 0.02	0.83 ± 0.03	0.84 ± 0.02
Messidor-2	0.22 ± 0.02	0.59 ± 0.05	0.62 ± 0.06
EyePACS	0.13 ± 0.01	0.31 ± 0.01	0.33 ± 0.03

Table 1. QWK by dataset and classifier

While these scores are lower than direct deep learning classifiers, the interpretability gained through lesion-based reasoning provides added clinical value.

Discussion

In this study, automatic segmentation is combined with DR stage classification. Among the tested classifiers, Decision Tree and Random Forest outperform the rule-based ICDR approach, suggesting that clinical grading does not strictly follow the guidelines, while learned classifiers achieve higher agreement by adapting to the observed patterns. Beyond classification, the segmentation output enables disease progression tracking over time, opening opportunities for longitudinal monitoring and objective assessment of treatment efficacy in clinical trials.

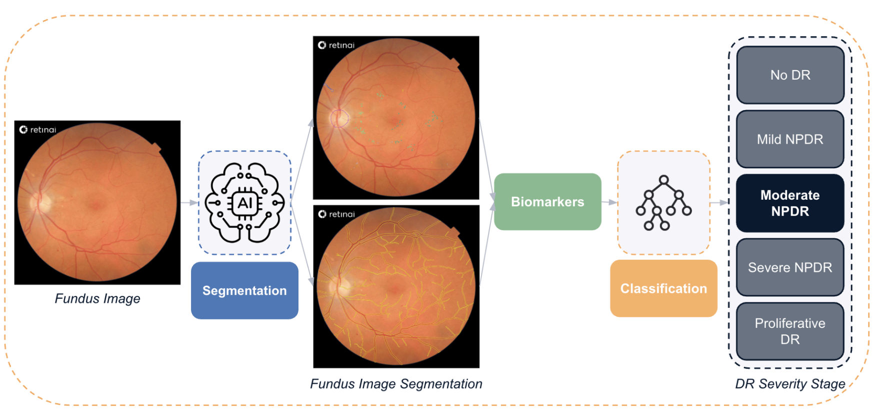

Fig. 1. Architecture of the segmentation-based Diabetic Retinopathy classification solution