

Machine Learning-Driven Heart Disease Risk Prediction and Stratification from Retinal Images

Mrs. Smita S Wagh¹, Mr. Tejas Khatkale², Ms. Devshri Khairnar³, Mr. Yash Mali⁴, Ms. Sanika Najan⁵
^{1,2,3,4,5} Department of Computer Engineering, JSPM's Jayawantrao Sawant College of Engineering Pune, India

Abstract- Heart attacks and hypertension are major healthcare issues which affect microvascular structure and function adversely. Visualizing early pathological changes in retinal blood vessels becomes possible with non-catheter-based retinal fundus imaging. The study develops an AI-driven system using machine learning methods to diagnose cardiovascular diseases in preclinical stages by evaluating retinal vessel pictures. The research approach begins with retinal image collection followed by vascular segmentation methods which clean and enhance meaningful vascular features of the retina without superfluous structures. The trained deep learning models use separated features to identify situations of heart attacks and hypertension. The system intends to assist in prompt medical detection through its capacity to spot minimal vessel irregularities that usually escape human notice especially within younger groups. Through this system the correct detection and prediction of cardiovascular risks utilizes retinal image processing combined with morphological analysis and artificial intelligence. This medical approach demonstrates dual value in diagnostic precision within ophthalmology as well as cardiology and generates superior therapeutic designs and established clinical outcome measures.

Keywords: Cardiovascular Disease, Retinal Vessel Segmentation, Machine Learning, Hypertension Detection, Deep Learning in Healthcare, Retinal Image Analysis

I. INTRODUCTION

Heart attacks [1] and hypertension collectively form the main causes of death and illness which affect individuals worldwide. Early detection along with prevention strategies present essential problems during clinical practice since symptoms typically escape younger patients until extensive damage develops. Non-invasive imaging technologies have discovered retinal fundus imaging as a novel technique for performing early diagnosis. Due to its enhanced observation of vascular system health the retina functions as the perfect site for identifying vascular

abnormalities which could signal cardiovascular problems. [2][3]

The medical diagnostic value of retinal blood vessel narrowing together with vessel tortuosity and branching irregularities helps identify hypertension and heart disease markers. Manual clinical observation proves challenging to spot such alterations although they exist in early stages of cardiovascular changes. The detection accuracy and efficiency of healthcare professionals requires automated intelligent diagnostic systems due to these limitations.[4][5]

This research demonstrates an artificial intelligence system with deep learning and machine learning functionalities which analyzes retinal images to identify initial cardiovascular disease indicators. The system utilizes contemporary vessel segmentation tools that separate retinal blood vessels to obtain cardiovascular disease diagnosis-specific morphological characteristics. The predictive models use these features to determine whether heart attack or hypertension indicators exist.[6][7]

This study concentrates on preclinical detection to build an early intervention tool which benefits people who do not show obvious symptoms. The application of artificial intelligence in retinal imaging serves both ophthalmological needs and cardiological purposes along with improving medical decisions regarding patient treatments. This interdisciplinary method shows excellent potential to connect early vision-related warning signs with prompt medical responses for cardiovascular care systems.[8][9].

II LITERATURE SURVEY

Author demonstrated the training of a deep learning algorithm that forecast multiple cardiovascular risk aspects through assessments of retinal fundus images. The research utilized information from UK Biobank and EyePACS to have their model forecast both

demographics and cardiovascular risks as well as systolic blood pressure values alongside smoking patterns and heart disease prognosis for five years ahead. Through its training process CNNs developed expertise in identifying distinct vascular patterns within retinal images thereby confirming retina scans provide an reliable non-invasive method to forecast systemic vascular health. This study served as a strong base for AI cardiovascular risk assessment by showing how retinal images could indicate early diseases.[10][11]

Srinivasan et al. (2020) conducted research about precise blood vessel segmentation in retinal images through the application of U-Net-based deep learning architecture. The research tackled the requirement for exact vessel segmentation because this procedural step functions as preprocessing foundation for automated diagnostic systems. The model processing publicly accessible DRIVE and STARE datasets enabled the authors to obtain accurate results in major and minor vessel detection. Better extraction of important morphological indicators like vessel diameter along with branching patterns and tortuosity became possible through this model since it surpassed traditional processing approaches in effectiveness.

The statistical research of Cheung et al. (2017) evaluated the hypertension connection to retinal vascular characteristics in a large population sample. Investigators analyzed retinal photographs to obtain quantitative values from vessels including their measurements and calculated ratios and levels of vessel curvature which produced robust correlations with blood pressure elevation patterns. Doctors found evidence in their research that supports how retinal imaging displays hypertension-induced microvascular changes which can act as non-invasive measures to detect early clinical cardiovascular risks. Researchers used the study to verify the link between vascular conditions affecting the body and eye health.

Li et al. (2021) presented a machine learning system integration of traditional image processing with deep learning techniques for the identification of cardiovascular disease indicators in retinal pictures. The technique started with deep neural networks processing vessels then extracting vessel width data along with density data before analyzing bifurcation angles. A support vector machine (SVM) together with other ML classifiers analyzed the extracted features to detect hypertension and cardiovascular risks. The

technique demonstrated strong performance in disease classification and also enabled medical teams to understand the prediction process thus ensuring its potential for real-world clinical use.

III. PROBLEMS IN EXISTING SYSTEMS

Current systems using retinal images for heart disease prediction have overcome many obstacles yet still face numerous important restrictions. Models face their greatest challenge because they need top-quality retinal images yet obtaining these clear images proves challenging when working in real clinical environments. The predictive models exhibit high susceptibility to image noise which distorts vital retinal features thus leading to prediction inaccuracies. The lack of sufficient large retinal image datasets with proper labeling acts as a substantial barrier because it creates difficulties for machine learning models to achieve effective generalization across different patient demographic groups. The computational requirements of deep learning architectures mostly affect convolutional neural networks (CNNs) since they need significant processing resources and sustained training duration to operate effectively.[12][13] Challenging deployment conditions arise because of this requirement. The risk of overfitting occurs frequently in models created from minimally sized and structurally unbalanced datasets since they develop reduced robustness during practical utilizations.[14] The high prediction accuracy from deep learning approaches comes with a major drawback because their lack of interpretability hinders medical professionals from understanding the basis of model decision-making. Intelligent cardiovascular diagnostics that use retinal imaging require improved efficient and interpretable solutions to overcome existing challenges.[15][16]

IV. SYSTEM DESIGN

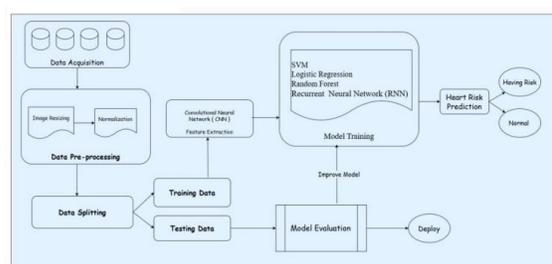


Figure 1: System Architecture

The designed framework follows multiple stages to detect risks associated with heart conditions through retinal fundus image analysis. Data acquisition serves as the first step before retinal images collect from public datasets or medical practitioners to secure both reliable and diverse input information. Before the processing starts the system performs standardization through two steps: first it resizes the images to maintain uniformity and second it applies pixel value normalization to support model training. The processed dataset shifts into training and testing components to support both modeling procedures and evaluation operations.

The analysis utilizes a Convolutional Neural Network to learn essential heart disease-related patterns which identify characteristic cardiovascular characteristics. Extracted feature data points from multiple machine learning classifiers including Support Vector Machine (SVM) and Logistic Regression together with Random Forest and Recurrent Neural Networks (RNN) receive training inputs to categorize images as either “Heart Risk” or “Normal”. The model evaluation phase depends on accuracy measures together with precision and recall metrics and F1-score to measure performance before doing any required iterative model improvement. A web-based platform enables medical practitioners to submit new retinal images which results in immediate predictions from the optimal model. The system produces two distinct categories as a final output which determines whether patients have heart risks or not helping physicians make preclinical diagnosis and initiate timely interventions.

V. PROPOSED SYTEM

The system development utilizes retinal fundus images together with deep learning and conventional machine learning algorithms to detect cardiovascular risk factors with emphasis on heart disease early diagnosis. The system builds an efficient modular pipeline which runs on Python infrastructure through open-source frameworks while employing libraries. The method extracts vascular patterns from retinal images before employing morphological features for low and high cardiovascular condition risk identification.

A web-based Flask application enables medical professionals to upload images so they can conveniently submit retinal fundus images. After

uploading the images into the system they receive preprocessing that includes conversion to 224×224 pixels and grayscale transformation alongside pixel normalization for learning consistency. The system applies specific CNN architecture to extract vascular features automatically yet uses data from flatten layer output

to train different machine learning classifiers. The prediction system integrates SVM alongside KNN and Random Forest together with Decision Tree as various models that assess heart disease likelihood.

The system enables GPU acceleration of model development through Google Colab training capabilities. The system features modular design where the CNN training occurs in `cnn_model.py` and results in traditional ML classification through `ml_classifiers.py` and visualization is handled in `visualization.py`. Accuracy plots and confusion matrices attached to visual performance reports assist both model performance assessments and improvement processes.

Through the web interface users can achieve an easy-to-use experience because uploaded images get processed via the backend before receiving classification results as either “Low Risk” or “High Risk.” Feature extraction with classification takes place in the backend section which the frontend shows as easy-to-understand predictive results. The deployable model for retinal heart disease detection exists as `retinal_heart_disease_model.h5` format. During data loading the binary-labeled folders (`class_0` for low risk and `class_1` for high risk) allow automatic labeling. The system incorporates an extensive arrangement that enables precise and scalable diagnostic help access from clinical areas as well as remote environments.

VI RESULTS ANALYSIS

The system performance analysis utilizes a bar chart to display accuracy results of different traditional machine learning models when training with features extracted from CNN. The Support Vector Machine (SVM) model exhibited the highest accuracy in the Accuracy Comparison (ML Models) figure followed by Random Forest. The performance achievements of K-Nearest Neighbors (KNN) and Decision Tree fell behind SVM but they produced spot-on results. Support Vector Machine functions as the most optimal

classification method in this framework for retinal-based cardiovascular risk evaluation. The prediction system demonstrates dependable forecasting capabilities because all models reach accuracy levels above 65% which demonstrates the strong combination of CNN-based feature analysis and the effectiveness of retinal imaging for non-invasive diagnosis.

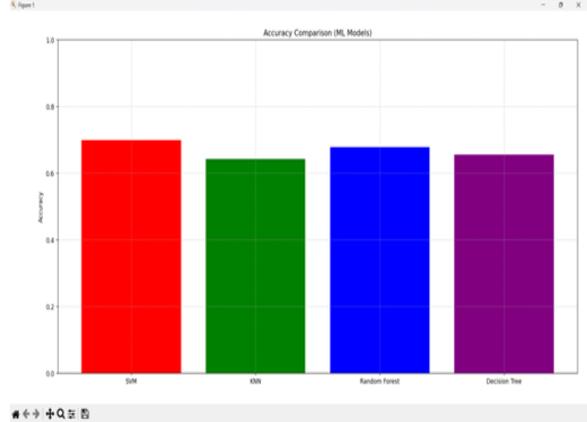


Figure 2: Model Accuracy Comparison

The Support Vector Machine (SVM) model achieved its classification results for retinal images using the data presented in the confusion matrix of Figure X. The model detected 408 non-risk cases correctly together with 169 at-risk cases yet it showed 237 false negative results that revealed difficulties in identifying high-risk patients. The SVM displayed strong precision in its results but its significant false negative rate makes it less functional than CNN-based models for essential health screening purposes due to the better accuracy retrieved from our CNN implementation.

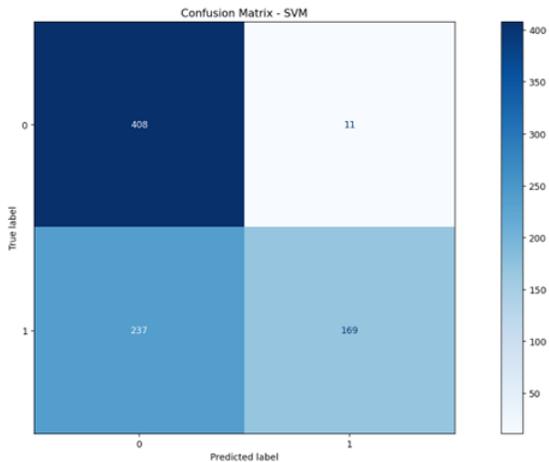


Figure 3: ConfusionMatrix

VII RESULTS



Figure 3: Result 1



Figure 4: Result 2

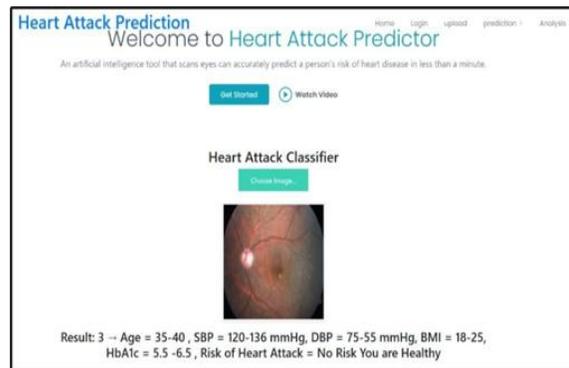


Figure 5: Result 3

The designed user interface of the heart attack prediction system enables easy navigation for both patients and clinicians through its intuitive and interactive layout. A user who accesses the first screen can begin their workflow by selecting either of two options which lead to project initiation or video viewing. The prediction solution accepts cardiovascular health indicators including age and

systolic and diastolic blood pressure combined with body mass index (BMI) and HbA1c results. Users activate the retinal image classification by pressing the 'Choose Image' button in order to start the machine learning-powered backend classification procedure. The system evaluates retinal scans after upload by combining vascular features retrieval with health parameter evaluation to generate complete predictions. On the third screen the results show both age range and blood pressure measurements including systolic and diastolic and also display BMI results as well as HbA1c range with a direct statement about heart attack risk. This particular output reveals that the system positions the user under the category "No Risk – You are Healthy." The system's combination of AI analytics with current health metrics produces prompt correct feedback that facilitates preventive health services. The real-time prediction feature proves the clinical effectiveness of this system when applied to clinical and telemedicine environments.

VII CONCLUSION

The Retinal Image-Based Heart Disease Prediction and Risk Management System reveals the strong relationship between machine learning and medical imaging for diagnosing healthcare conditions. Medical staff benefits from the system through retinal scans combined with advanced AI models to identify heart disease indicators at an early stage which leads to prompt medical care. Generative AI algorithms increase model performance when they supplement insufficient datasets which makes the systems both reliable and applicable to expanded purposes. Improved cardiovascular diagnostic capabilities and enhanced patient results are expected through ongoing toolkit advancement combined with bigger real-world data collection and live deployment of this innovative solution.

REFERENCES

[1] A. A. Almazroi, E. A. Aldahri, S. Bashir and S. Ashfaq, "A Clinical Decision Support System for Heart Disease Prediction Using Deep Learning," *IEEE Access*, vol. 11, pp. 61646–61659, 2023, doi: 10.1109/ACCESS.2023.3285247.

[2] K. U. Kumar, K. U. Singh and M. Kumar, "A Clinical Data Analysis Based Diagnostic Systems for Heart Disease Prediction Using Ensemble Method,"

Big Data Mining and Analytics, vol. 6, no. 4, pp. 513–525, Dec. 2023, doi:10.26599/BDMA.2022.9020052.

[3] M. S. A. Reshan et al., "A Robust Heart Disease Prediction System Using Hybrid Deep Neural Networks," *IEEE Access*, vol. 11, pp. 121574–121591, 2023, doi:10.1109/ACCESS.2023.3328909.

[4] S. Mondal et al., "An Efficient Computational Risk Prediction Model of Heart Diseases Based on Dual-Stage Stacked Machine Learning Approaches," *IEEE Access*, vol. 12, pp. 7255–7270, 2024, doi: 10.1109/ACCESS.2024.3350996.

[5] G. N. Ahmad et al., "Efficient Medical Diagnosis of Human Heart Diseases Using Machine Learning Techniques with and Without GridSearchCV," *IEEE Access*, vol.10, pp. 80151–80173, 2022, doi: 10.1109/ACCESS.2022.3165792.

[6] A. Noor et al., "Heart Disease Prediction Using Stacking Model with Balancing Techniques and Dimensionality Reduction," *IEEE Access*, vol. 11, pp. 116026–116045, 2023, doi:10.1109/ACCESS.2023.3325681.

[7] S. Ghorashi et al., "Leveraging Regression Analysis to Predict Overlapping Symptoms of Cardiovascular Diseases," *IEEE Access*, vol. 11, pp. 60254–60266, 2023, doi:10.1109/ACCESS.2023.3286311.

[8] M. C. Badescu et al., "Cardiovascular risk factors in patients with congenital hemophilia: A focus on hypertension," *Diagnostics*, vol. 12, no. 12, p. 2937, Nov.2022.

[9] C. Xiao, Y. Li and Y. Jiang, "Heart Coronary Artery Segmentation and Disease Risk Warning Based on a Deep Learning Algorithm," *IEEE Access*, vol. 8, pp. 140108–140121, 2020, doi: 10.1109/ACCESS.2020.3010800.

[10] J. P. Li et al., "Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare," *IEEE Access*, vol. 8, pp. 107562–107582, 2020, doi:10.1109/ACCESS.2020.30011.

[11] Punam V. Maitri, Dattatray S. Waghole, Vivek S. Deshpande "Low latency for file encryption and decryption using Byte Rotation Algorithm", *Proceedings of IEEE International conference on International Conference on Pervasive Computing*, 2015.

[12] Vivek S Deshpande, Dattatray S Waghole, "Performance analysis of FMAC in wireless sensor networks", pp. 1-5, *IEEE, Eleventh International*

Conference on Wireless and Optical Communications Networks (WOCN), 2014

[13] Dattatray Waghole, Vivek Deshpande, Divya Midhunchakkaravarthy, Makarand Jadhav, “Position aware congestion control (PACC) algorithm for disaster management system using WSN to improve QoS”, Design Engineering, pp. 11470-11478, 2021

[14] Dattatray S Waghole, Vivek S Deshpande, Punam V Maitri,” pp. 1-5, IEEE, International Conference on Pervasive Computing (ICPC) 2015.

[15] Dattatray S Waghole, Vivek S Deshpande,” Analyzing the QoS using CSMA and TDMA protocols for wireless sensor networks”, pp. 1-5, IEEE, International Conference for Convergence for Technology-2014.

[16] Omkar Udawant, Nikhil Thombare, Devanand Chauhan, Akash Hadke, Dattatray Waghole, “Smart ambulance system using IoT”,pp 171-176, IEEE, International conference on big data, IoT and data science (BID),2017