Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 14, Issue. 2 : 2023
ISSN : **1906-9685**

Deep Learning Models for Heart Attack Risk Prediction Using Retinal Imaging Data

ALANKARAM SHOBITHA LAKSHMI1, PIDUGU BHARATH TEJA2

#1Assistant Professor, Department of CSE, PBR Visvodaya Institute of Technology and Science, Kavali

#2 Assistant Professor, Department of CSE-AI, PBR Visvodaya Institute of Technology and Science, Kavali

ABSTRACT_

The development of reliable prediction models for early identification of heart disease is crucial, since it is a top cause of death worldwide. Retinal imaging is one non-invasive alternative to traditional technologies that have recently gained popularity. Using retinal scans and machine learning methods, this study aims to develop a dependable model for predicting the incidence of heart disease. In order to forecast the occurrence of cardiovascular illness, earlier studies have used a variety of techniques, including logistic regression, k Nearest Neighbor (KNN), and naive Bayes. Although these techniques provide a variety of approaches, they could not be as accurate or have a smaller dataset. Furthermore, classification techniques based on machine learning have been investigated; nevertheless, dealing with imbalanced classes and features is a difficulty.

Using image processing, the proposed research would examine retinal fundus pictures for characteristics that may indicate the likelihood of heart illness, such as the size of blood vessels and the amount of light coming from the backdrop. Improving prediction accuracy is the goal of the model, which makes use of Random Forest Classifier (RFC) and Support Vector Machine (SVM) methods. A new approach for predicting the occurrence of heart disease is the end result of this project's data preparation, picture segmentation, and outcome analysis phases. The new model outperforms prior methods in terms of efficiency and accuracy in predicting the incidence of heart disease from retinal pictures, according to comparative research. The model improves its performance and allows for prompt interventions for at-risk people by incorporating modern methods such as Recurrent Neural Networks (RNNs) and Fuzzy C-Means clustering. This initiative highlights the possibility of transforming cardiovascular health monitoring via the merger of modern machine learning and medical imaging.

1.INTRODUCTION

In the center of the body's circulatory system, which also includes the lungs, is a muscular organ called the heart, which pumps blood throughout the body.A system of blood vessels, including veins, arteries, and capillaries, is also part of the cardiovascular system.Blood is distributed throughout the body by these vessels.Cardiovascular diseases (CVD) include a wide range of conditions characterized by irregularities in the heart's normal blood flow.The leading causes of death on a global scale are heart diseases.One7.5 million people die every year as a result of cardiovascular diseases, according to a report by the World Health Organization (WHO).In low- and middle-income nations, cardiovascular disease accounts for more than 75% of all deaths.Further, heart attacks and stroke account for 80% of all deaths caused by CVDs.Tools for the prediction of heart diseases and the ability to predict cardiac abnormalities at an early stage can save lives by assisting doctors in developing effective treatment plans that ultimately reduce mortality rates due to cardiovascular diseases.

Big Data in Electronic Health Records Systems and other modern healthcare innovations have made previously inaccessible patient records usable for the purpose of developing prediction models for cardiovascular illnesses. Discoveries may be made by evaluating massive data from several perspectives using data mining or machine learning.Extracting implicit, previously unknown, and potentially usable information from data is the goal of data mining, which involves encapsulating it into useable information.These days, healthcare industries produce massive amounts of data related to patient information, disease diagnoses, and other related topics.Data mining offers a variety of methods for unearthing previously unseen patterns or similarities in large datasets.

As a result, this paper proposes a machine learning algorithm for a heart disease prediction system that has been validated on two open-access datasets.Data mining is the practice of mining enormous databases for valuable information using computational methods.The non-trivial information found in massive amounts of evidence is what makes data mining so useful for exploratory analyses. When it comes to clinical information, medical data mining offers a lot of untapped potential for uncovering hidden patterns.

It is possible to use these patterns for healthcare diagnosis.Having said that, the raw medical data that is currently available is vast, varied, and dispersed.It is necessary to gather this data in a structured manner.A medical information system may be formed by integrating the acquired data.Data mining offers a user-centric method for discovering and utilizing hidden patterns in data.Data mining methods are helpful for healthcare professionals in forecasting the occurrence of different illnesses and addressing business queries.Data mining plays a crucial role in disease prediction.The prediction of heart disease using classification algorithms is examined in this paper.These patterns in healthcare data can be used for health diagnosis. When dealing with indefinite patterns in data, data mining technology provides an efficient method.Better services can be obtained by healthcare administrators using the identified information.In nations like India and the United States, heart disease was the leading cause of death.Here, we're using classification algorithms to foretell cases of heart disease.Classification algorithms like DNN Classifications and logistic regression are examples of machine learning approaches that may be used to investigate many types of heart-related issues.

2.LITERATURE SURVEY

[1] "Can deep learning on retinal images augment known risk factors for cardiovascular disease prediction in diabetes? A prospective cohort studyfrom the national screening programme in Scotland" Paul M. McKeigue, Helen M. Colhoun, Amos J. Storkey, Luke Blackbourn, Alan Fleming, Alan Fleming, Stuart J. McGurnaghan, Luke Blackbourn, and Caroline Styles The purpose of this research is to examine how deep learning (DL) applied to retinal pictures could improve diabetes-related CVD prediction. To predict the risk and causes of CVD, DL models were trained using data from Scotland's diabetic retinopathy screening program. Although DL ratings did not significantly increase prediction accuracy compared to models using clinical risk variables alone, they were shown to be independently linked with incident CVD in both type 1 and type 2 diabetes cohorts. The research implies that further methods, including serial image analysis, could be necessary for substantial therapeutic use of retinal pictures, which do include predictive information for CVD in diabetes. There is a high correlation between DL scores and CVD, however adding them to prediction models improved performance metrics just little. These results provide insight on the need for more clinical testing and improvement of DL's use of retinal imaging for CVD risk assessment in diabetic populations, which helps shed light on the current state of the field.

[2] "Retinal age gap as a predictive biomarker for mortality risk" by Zhuoting Zhu, Danli Shi, Peng Guankai, Zachary Tan, XianwenShang,Wenyi Hu, Huan Liao,Xueli

Zhang,Yu Huang,Honghua Yu, Wei Meng,Wei Wang ,Zongyuan Ge ,Xiaohong Yang,Mingguang

He Zhu et al.'s research presents a novel method for predicting mortality risk by comparing chronological age with retinal age (as calculated by deep learning models). A non-invasive and cost-effective way to evaluate biological aging is offered by this unique technology that leverages fundus pictures. The researchers achieved impressive accuracy in age prediction from fundus photos by training deep learning models on a massive dataset from the UK Biobank. A predictive biomarker for mortality risk was the "retinal age gap," which is the difference between anticipated retinal age and chronological age.

In particular, the results show that mortality risk is higher for non-cardiovascular and non-cancer-related reasons when the age difference between the retinas is larger. Accordingly, fundus pictures of the aged retina may represent systemic alterations that point to an increased risk of death. The research highlights the potential of retinal imaging to help with healthcare risk assessment and intervention delivery on an individual level. Retinal imaging has the potential to transform techniques for early identification and focused intervention in age-related health outcomes due to its accessibility and simplicity, particularly with increasing technology such as smartphone-based cameras. Research like this paves the way for further studies on the effects of retinal age assessment on healthy aging and lifespan, as well as its potential therapeutic applications.

3.PROPOSED SYSTEM

This section provides a high-level overview of the proposed system and shows how all the parts, methods, and tools were put together to build it.If we want to build a heart disease prediction system that is both smart and easy to use, we need a software tool that can train numerous machine learning algorithms on large datasets.Once the robust algorithm with the best accuracy and performance measures is chosen, work will be finished on the smartphone app to detect and predict the risk level of heart disease.A variety of retinal diseases may be detected using retinal fundus pictures. This paper demonstrates the feasibility of using deep-learning models trained on external pictures of the eyes to identify diabetic macular oedema, diabetic retinopathy (DR), and inadequate management of blood glucose. Using images of diabetic patients' eyes taken from 301 DR screening sites, we trained the models, and then tested them on four tasks and four validation datasets using data from 198 more screening locations. Predictions generalized to patients with dilated pupils, to patients from a different DR screening program, and to a general eye care program that included diabetics and non-diabetics; furthermore, the deep-learning models outperformed logistic regression models that relied on self-reported demographic and medical history data for all four tasks. Additionally, we investigated the possibility of using the deep-learning models to identify cases with increased lipid levels. Using photos from various cameras and patient demographics may further confirm the efficacy of external eye photographs for illness diagnosis and treatment.

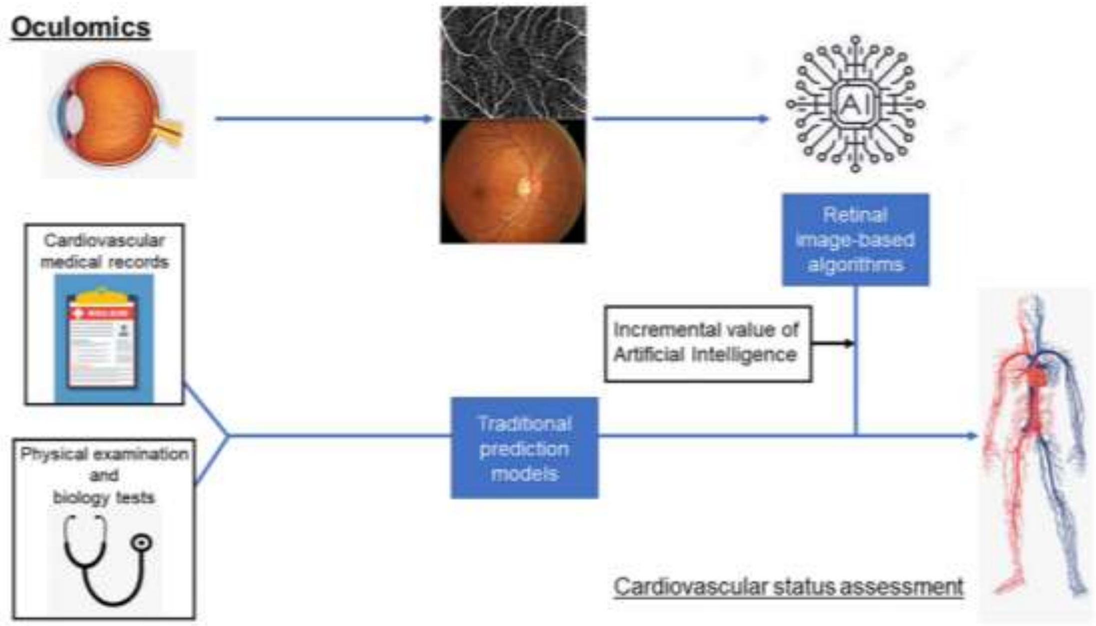

Fig 1: Architecture Diagram

3.1 IMPLEMENTATION

3.1.1 DataPre-processing:

To process all incoming documents and texts, this file provides all the necessary pre-processing routines.We began by reading the training, testing, and validation data files, and then we did some preprocessing, such as tokenizing and sorting, etc.Response variable distribution and data quality checks (such as missing values and null values) are examples of exploratory data analysis.

3.1.2 Feature Extraction:

Extraction Using the sci-kit-learn python packages, we have extracted features and selected them in this file.We have utilized methods such as simple bag-of-words and n-grams, as well as term frequency techniques like TF-TDF weighting, for features selection.Although POStagging and word2vecha are not being used at this stage in the project, we have also used ord2vec and POStagging to extract the features.

3.1.3 Classification:

All the classifiers for heart attack disease detection have been established here.Various classifiers are fed the extracted features.Neural Network Classifications, Stochastic Gradient Decomposition, Linear Support Vector Machines, Fuzzy C-Means Clustering, and Naive Bayes are some of the classifiers that we have used from sklearn.All of the classifiers utilized each of the retrieved characteristics.We checked the confusion matrix and compared the f1 score after fitting the model.

Fuzzy C-Means Clustering and RNN were chosen as the two best-performing models after fitting all the classifiers for heart disease classification.We have selected the best-performing parameters for the classifier and executed parameter tuning on the candidate models using the GridSearchCV method.

With the probability of truth, the final model was used for heart disease detection.Also, to find out which terms are most essential in each class, we have retrieved the top 50 characteristics from our term-frequency vectorizer.

To examine how the training and test sets perform as we increase the quantity of data in our classifiers, we have also employed precision-recall and learning curves.

3.1.4 Prediction:

We saved the algorithm with the name final_model.sav after it was our best-performing and final classifier.You will be able to use this model to classify heart diseases in the prediction.py file after you close this repository. The model takes a news story as input and uses it to provide final categorization output, along with the likelihood of truth, which is then given to the user.

4.RESULTS AND DISCUSSION

Cluster Screen

You can see how the data points are organized into clusters according to their commonalities on the clusters screen. It has a big visual representation of the clusters, with various colors for each cluster. Each cluster is color-coded and described in the screen's legend. Users may pan and zoom to see more details inside each cluster, and the display is fully interactive. When analyzing data, the clusters screen is a lifesaver since it lets users see trends and patterns at a glance. The data is easily navigable and interpreted because to its user-friendly interface and intuitive design, which provides significant insights for decision-making.

Fig – 4.1: Cluster Screen

Cluster Output

The clusters output is an interactive and aesthetically pleasing representation of the data points that have been clustered according to their commonalities. Because it allows users to swiftly see trends and patterns in the data, the clusters output is an effective tool for data analysis. The data is easily navigable and interpreted because to its user-friendly interface and intuitive design, which provides significant insights for decision-making. Users are able to have a better knowledge of their patterns and trends via the clusters' output, which provides a complete and fascinating picture of the data.

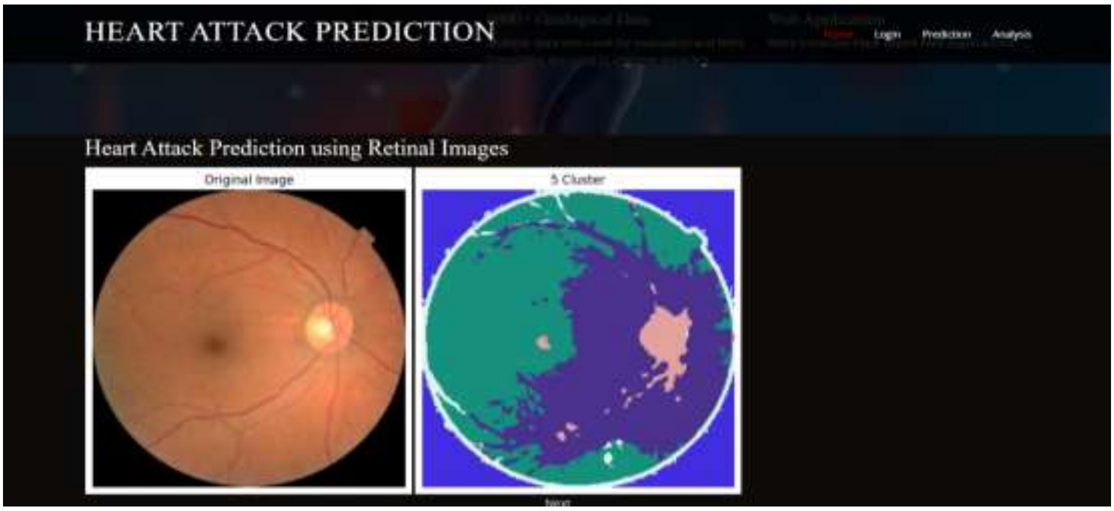

Fig - 4.2: Cluster Output

Results

For each input variable, such age, there is an associated "attribution score" that indicates the relative importance of that factor in determining the value difference between the patient's estimated CVD risk and the total risk in the source dataset. According to the American College of Cardiology, age, systolic blood pressure (SBP), diastolic blood pressure (DBP), body mass index (BMI), hemoglobin A1c (HbA1c), and risk of heart attack are the most important variables in determining an individual's cardiovascular risk.

Fig - 4.3: Results

Features caused Heart Attack

The given graph is associated with a model for predicting heart attacks using retinal pictures. The chart provides a comprehensive list of heart attack-related characteristics, together with their respective values. Age, Diastolic Blood Pressure (DBP), Body Mass Index (BMI), Systolic Blood Pressure (SBP), and Hemoglobin are the characteristics that are given.Some of the values occur more than once, and the range is from 100 to 180. It is the hope of the chart's creators that these characteristics correlate with an increased risk of heart attack.

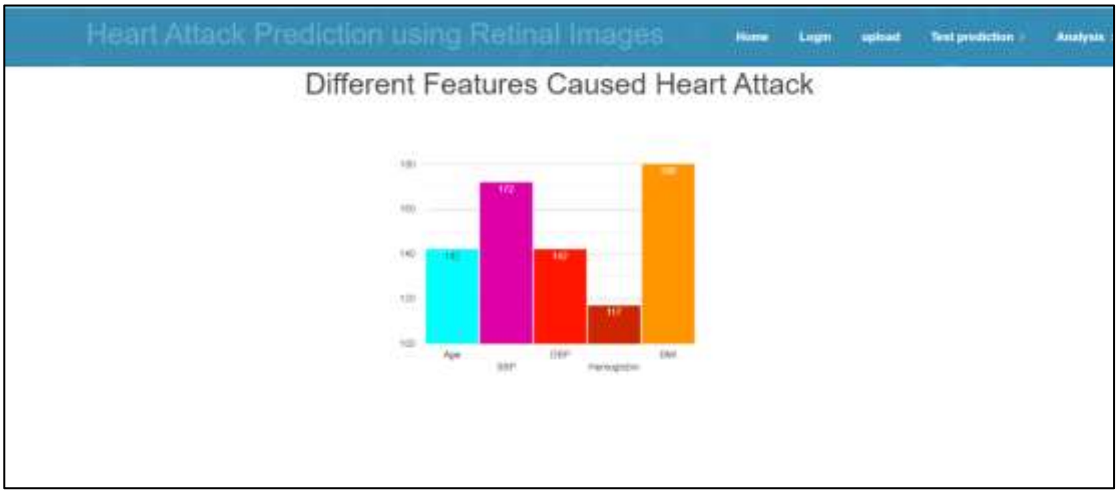

Fig - 4.4: Features caused Heart Attack

Accuracy Plot

The accuracy plot is a visual depiction of a model's performance as measured by its accuracy, which is the percentage of accurate predictions relative to the total number of predictions.In most cases, a parameter like the learning rate, model complexity, or the number of iterations is plotted against the accuracy of the model in the accuracy plot. Parameter values are shown on the x-axis and matching model accuracy on the y-axis of the graphic.

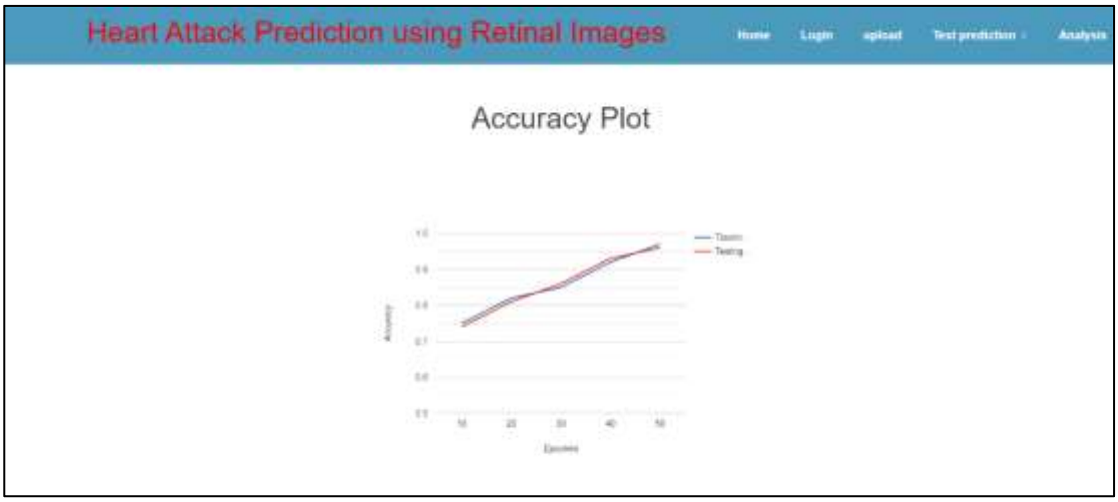

Fig - 4.5: Accuracy Plot

5.CONCLUSION

The "Heart Attack Risk Prediction Using Retinal Eye Images" initiative introduces a fresh method for evaluating the risk of cardiovascular events by combining state-of-the-art retinal imaging technology with sophisticated machine learning algorithms. The results of this study show that non-invasive retinal imaging may provide useful information for estimating the likelihood of a heart attack, which might lead to early identification and treatment for those who are at high risk.

The system accurately analyzes retinal pictures and identifies important patterns suggestive of cardiovascular illness by combining Recurrent Neural Networks (RNNs) with Fuzzy C-Means clustering. Incorporating age, blood pressure, body mass index (BMI), and hemoglobin A1c levels into the prediction model further strengthens its dependability and resilience.

An intuitive online interface facilitates easy system interaction, data entry, and real-time risk prediction for healthcare professionals, further highlighting the project's focus on user-friendly design. To guarantee the system's efficacy in clinical situations, the implementation approach has prioritized scalability, performance, and security.

Ultimately, this initiative adds to the continuing drive to enhance cardiovascular health monitoring via the provision of a cutting-edge instrument that may facilitate the early identification and mitigation of heart attacks. It is possible that the system may become an invaluable resource for healthcare practitioners and patients when it is further developed and tested with bigger datasets. Exploring alternative machine learning approaches to further improve prediction accuracy, expanding its applicability to other medical disorders, and integrating with current healthcare systems are all potential future avenues for this research.

This project showcases the power of multidisciplinary cooperation in healthcare technology. It combines medical imaging, machine learning, and user-friendly features to improve patient outcomes and alleviate the strain on healthcare systems.

REFERENCES

[1] "Can deep learning on retinal images augment known risk factors for cardiovascular disease prediction in diabetes? A prospective cohort studyfrom the national screening programme in Scotland"by Joseph Mellor, Wenhua Jiang, Alan Fleming, Stuart J. McGurnaghan, Luke Blackbourn, Caroline Styles, Amos J. Storkey, Paul M. McKeigue, Helen M. Colhoun

[2] "Retinal age gap as a predictive biomarker for mortality risk" by Zhuoting Zhu, Danli Shi, Peng Guankai, Zachary Tan, XianwenShang,Wenyi Hu, Huan Liao,Xueli Zhang,Yu Huang,Honghua Yu, Wei Meng,Wei Wang ,Zongyuan Ge ,Xiaohong Yang,Mingguang He

[3] "Heart Attack Risk Prediction Using Retinal EyeImages Based On Machine Learning And ImageProcessing" by Mrs. Sushma V, Sindhu K U, Sonupriya

[4] "Retinal photograph-based deep learning predictsbiological age, and stratifies morbidity andmortality risk" by Simon Nusinovici, Tyler Hyungtaek Rim, Marco Yu, Geunyoung Lee, Yih-Chung Tham, Ning Cheung, Crystal Chun Yuen Chong, Zhi Da Soh,Sahil Thakur, Chan Joo Lee, CharumathiSabanayagam,Byoung Kwon Lee, Sungha Park,Sung Soo Kim, Hyeon Chang Kim, Tien-Yin Wong, Ching-Yu Cheng

[5] "Heart, Eye, and Artificial Intelligence: A Review" by Kush Guptaa, Sarath Reddy

[6] "Insights into Systemic Disease through RetinalImaging-Based Oculomics" by Siegfried K. Wagner, Dun Jack Fu, Livia Faes,Xiaoxuan Liu, Josef Huemer,Hagar Khalid, Daniel Ferraz, Edward Korot, Christopher Kelly,Konstantinos Balaskas, Alastair K. Denniston, Pearse A. Keane

[7] "Deep-learning-based cardiovascular risk stratification using coronary artery calcium scores predicted from retinal photographs" by Tyler Hyungtaek Rim, Chan Joo Lee, Yih-Chung Tham, Ning Cheung, Marco Yu, GeunyoungLee, Youngnam Kim, Daniel S W Ting, Crystal Chun Yuen Chong, Yoon Seong Choi, Tae KeunYoo, IkHeeRyu, Su Jung

Baik, Young AhKim, Sung Kyu Kim, Sang-Hak Lee, Byoung Kwon Lee, Seok-Min Kang, Edmund YickMun Wong, Hyeon Chang Kim, Sung Soo Kim,Sungha Park, Ching-Yu Cheng, Tien Yin Wong

[8] "A Review: Prediction of Multiple AdverseHealth Conditions from Retinal Images" by Lakshmi Kala Pampana,Manjula Sri Rayudu

[9] "Using Artificial Intelligence to Analyse the RetinalVascular Network: The Future of Cardiovascular RiskAssessment Based on Oculomics? A Narrative Review" by Louis Arnould, Fabrice Meriaudeau, Charles Guenancia, Cle´mentGermanese,Ce´cileDelcourt, Ryo Kawasaki, Carol Y. Cheung, Catherine Creuzot-Garcher, Andrzej Grzybowski

[10] "A foundation model for generalizable disease detection from retinal images" by Yukun Zhou, Mark A. Chia, Siegfried K. Wagner, Murat S. Ayhan, Dominic J. Williamson,Robbert R. Struyven, Timing Liu, Moucheng Xu, Mateo G. Lozano, Peter Woodward-Court, Yuka Kihara, UK Biobank Eye & VisionConsortium, Andre Altmann, Aaron Y. Lee, Eric J. Topol, Alastair K. Denniston, Daniel C. Alexander,Pearse A. Keane