2024 전기 졸업과제 중간 보고서

Eye Disease Detection AI Model

Topic: Glaucoma Eye Disease Detection AI Model

Team Name: Octas

Team Number: 41

Team members: Nemekhbayar Nomin, Battulga Bazarsad

1. Project Background

Glaucoma is a severe eye disease caused by increased fluid pressure in the eyes, which damages the optic nerves and can lead to partial or complete vision loss. The World Health Organization (WHO) has identified glaucoma as a major cause of blindness worldwide. It is a progressive and slow-developing condition, often only detectable at later stages, necessitating detailed screening and detection of retinal images to prevent vision forfeiture. [1]

This study aims to detect glaucoma at early stages with the help of deep learning-based feature extraction. Retinal or fundus images are the source of detection of disease in the eyes. The key factor causing glaucoma is a disproportion in the volume of fluid that maintains the eye shape, leading to increased pressure on the optic nerve head (ONH), resulting in damage. By analyzing fundus images, our goal is to develop an AI model capable of detecting glaucoma early, potentially improving patient outcomes and reducing the risk of vision loss.

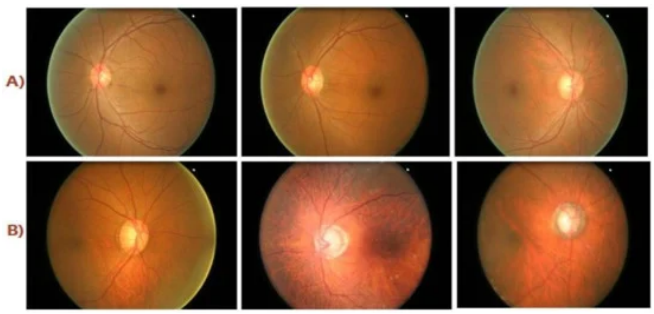


Figure 1. (**A**) Healthy Class Fundus Images; (**B**) Diseased Class Fundus Images.

2. Task Objectives

The primary objectives of this study are as follows:

1. Early Detection of Glaucoma: Develop a deep learning-based system capable of detecting glaucoma at early stages using retinal fundus images.
2. Pre-Processing and Segmentation: Implement pre-processing techniques to enhance image quality and reduce noise, followed by segmentation to extract the region of interest (ROI) on datasets without OC mask.
3. Feature Extraction: Utilize hybrid feature descriptors, including convolutional neural networks (CNN), local binary patterns (LBP), histogram of oriented gradients (HOG), and speeded-up robust features (SURF), to extract detailed features from the optic disc (OD) and optic cup (OC) regions.
4. Feature Selection and Ranking: Employ the MR-MR method for selecting and ranking the most representative features, ensuring efficient and accurate feature representation.
5. Classification: Use multi-class classifiers such as support vector machine (SVM), random forest (RF), and K-nearest neighbor (KNN) to classify the fundus images as healthy or diseased.
6. Performance Evaluation: Conduct various experiments to assess the performance of the proposed system, aiming for high accuracy in early glaucoma detection.

By achieving these objectives, the study seeks to provide a robust, accurate, and efficient method for the early detection of glaucoma, ultimately aiding in better patient outcomes and reducing the risk of vision loss.

3. Requirements Analysis and Realistic Constraints

After the initial report, We have faced few constraints and obstacles as we carried out actual tasks.

3.1 Dataset Analysis

3.1.a Variations in datasets

We are using open datasets including G1020, ORIGA, and REFUGE. The G1020 dataset, in particular, contains the most faulty fundus images and shows the greatest variation across the three datasets, presenting a significant challenge. [2]

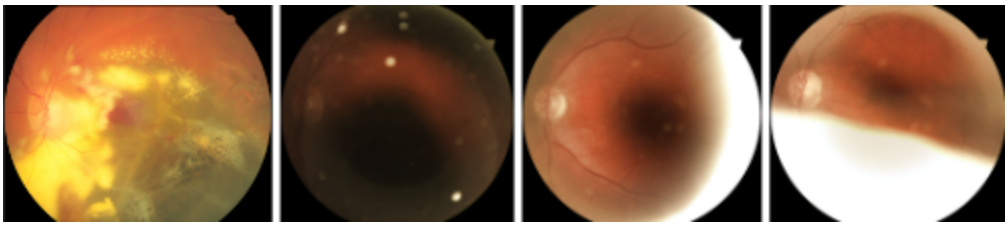


Figure 2. Example of Faulty fundus images in G1020 dataset

3.1.b Faulty results of Extraction of ROI

The extraction of the ROI process was inconsistent across different datasets, and the execution time was high. Our initial algorithm used was the Brightest spot algorithm and ellipse detection to find the brightest spot on the image and extract ellipse parameters from the brightest point. However, due to faulty fundus

images, noise and image variations, our proposed method did not accurately locate the ROI (figure 3).

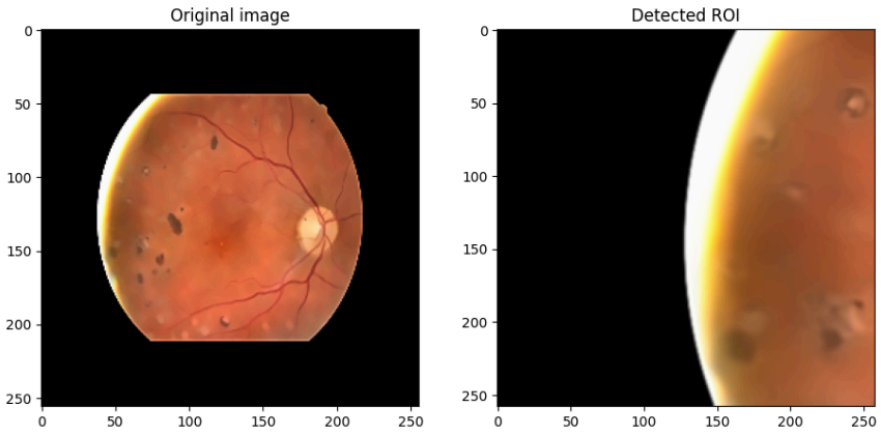


Figure 3. Faulty ROI extraction

3.2 Constraints

- Computational Resources: Training deep learning models requires significant computational power. Ensuring access to high-performance CPU and implementing memory efficient approach.
- Processing time: model must perform classification in relatively short time compared to other similar models.
- Generalization: The model must be generalized well to handle variations in image to achieve high accuracy and precision. This includes dividing the dataset into training and validation.

- Variability in Data: Retinal images can vary widely due to factors such as lighting conditions, camera quality, and patient demographics. The model must be robust enough to handle these variations.
- Deployment: The model must be deployable in real-world settings, which requires considerations for scalability, ease of use, and integration with existing medical systems.

4. Details of design and changes

Our modified proposal:

Algorithm : Pseudocode for the modified System.

Input: $Images = \{I1, I2, I3, \ldots\ldots\ldots Ik\}$

Output: *Classified Images*

Start:

record(i) ← 1….*k*

While(record(i)!= eof){

Perform Pre-processing over the images *Ik, clean the csv metadata table and split the dataset into training and testing dataset for further compilation*

Roi Extraction ← BS with sliding window algorithm

CNNF ← 2DCNN Perform Feature Computation

HOGF ← Feature Computation using HOG

LBPF ← Feature Extraction using LBP

SIFT ← Feature Extraction using SIFT

Fea_Vec ← (CNN, LBP, HOG, SIFT)

Fea_Comb ← (CNN with HOG, SURF, CNN with HOG and LBP, CNN with LBP, HOG, and SIFT)

Selected ← MR-MR()

Labels ← Annotations (Diseased, Healthy)

Class ← (SVM (Selected, CL, testSet) KNN (Selected, CL, testSet) RF (Selected, CL, testSet))

j = 1;

While (*j*≤n)

{

IF(Class(*j*)) ← Diseased

{Output “Glaucoma Detected”

ELSE IF(Class(*j*)) ← Healthy

Output “Healthy Eye”

}

END

4.1 Preprocessing and ROI

a. Preprocessing:

- Format Conversion: Images are converted to the TIFF format to preserve the image quality.

- Downscaling: The images are downscaled using a bilinear technique to reduce computational complexity and noise.

- Enhancement and Contrast: Images are enhanced and added more contrast using the CLAHE (Contrast Limited Adaptive Histogram Equalization) method for enhanced quality.

b. ROI Extraction

In case of fundus images, the portion including and around the disc is considered to be the ROI of a fundus image. As a characteristic of retinal fundus image, generally the pixel with highest intensity value lies in the optic cup. With the help of this fact, we propose a more robust, automatic method for ROI extraction using the brightest spot algorithm and ellipse method. Brightest spot algorithm [3] alone did not show promising results, however, by combining the brightest algorithm with iteration and score function, we achieved 98.31% accuracy in ORIGA dataset (1.69% mismatch), 95.92% accuracy in REFUGE dataset (4.08% mismatch), and 83% accuracy in G1020 dataset (16.86% mismatch) respectively. (figure 4).

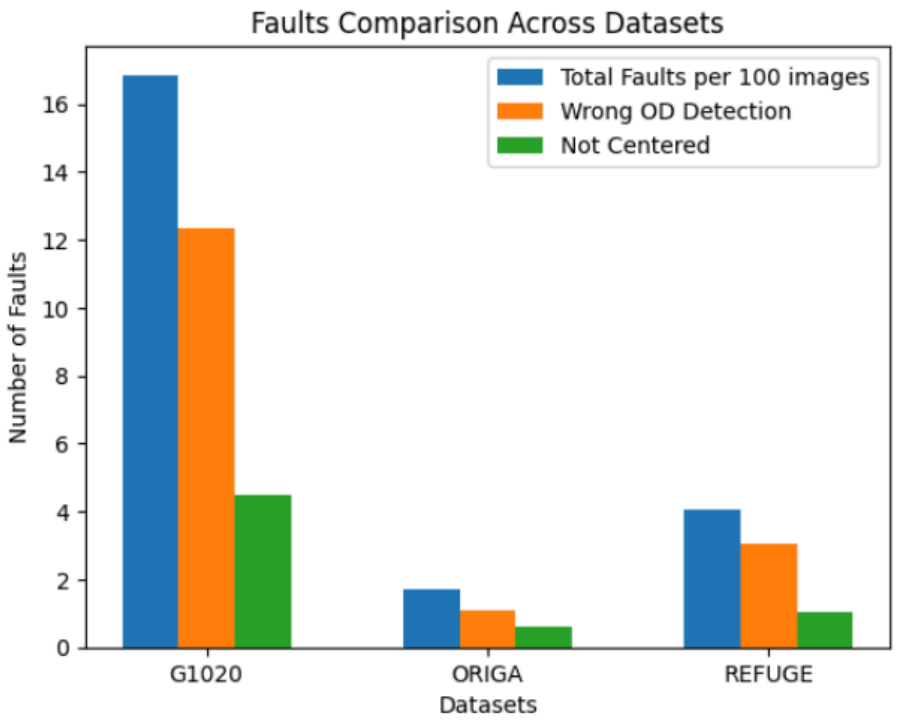


Figure 4. Faults comparison in G1020, ORIGA, and REFUGE datasets

Our proposed algorithm iterates through the fundus image with height / 2 x width / 2 window with step size of height / 4. The algorithm de-noises the image with a gaussian filter and then calculates the brightest spot in the window, then, thresholding the window with brightest value - 20 using binary thresholding. (figure 5)

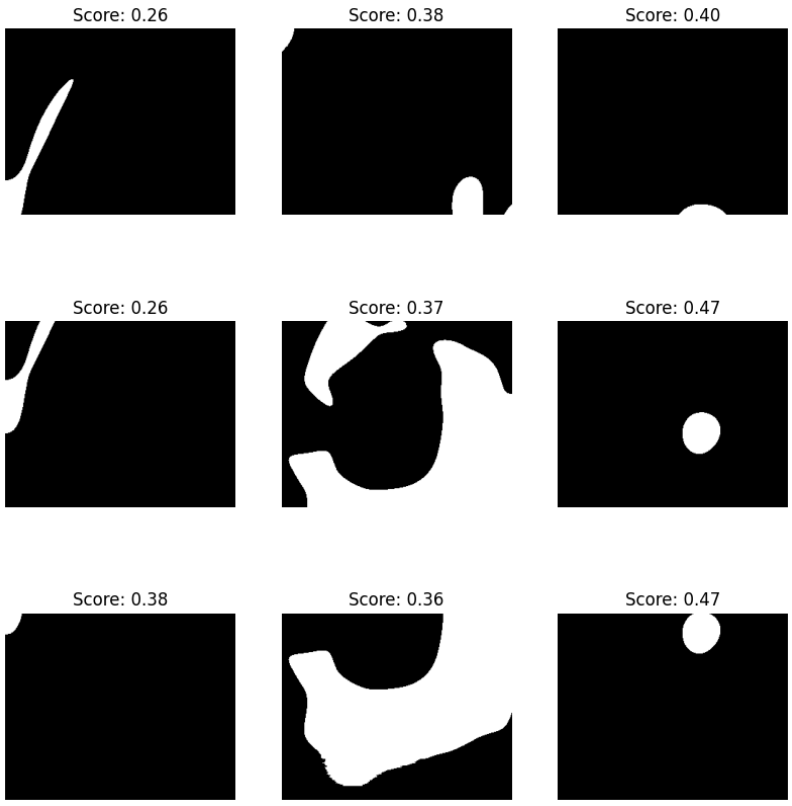


Figure 5. Thresholded windows

After thresholding, contouring is applied to fit an ellipse into the thresholded window. If it finds an ellipse, the scoring function scores each ellipse with weighted mean brightness of the area inside the ellipse, standard deviation, and aspect ratio. (figure 6)

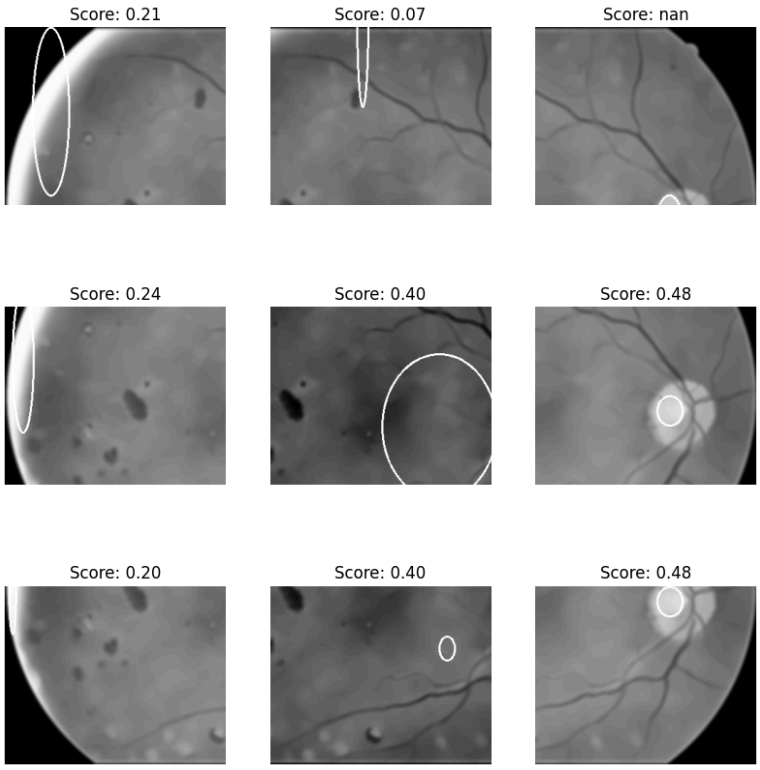


Figure 6. Scored ellipses in each window

The algorithm finally chooses an ellipse with the highest score, saves the cropped image around the ellipse center to the directory storing ROIs.

Example output:

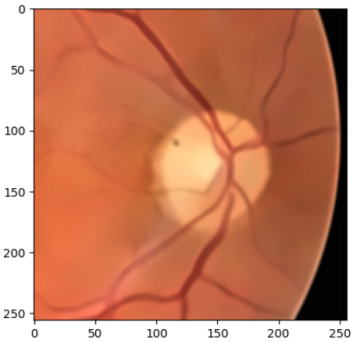


FIgure 7. Extracted ROI

c. Segmentation

Our proposed method detects the brightest point in the fundus image. The brightest point is then used to locate the center of the OD and crop the surrounding area, which produces the ROI. [3]

We choose an active-contour algorithm for segmentation of the OD in the ROI. [4] d on o

Baseur testing, compared to blue and green channels, the red channel provided the edges more clearly and distinctly, used to segment OD from the ROI. (figure 8).

and is u

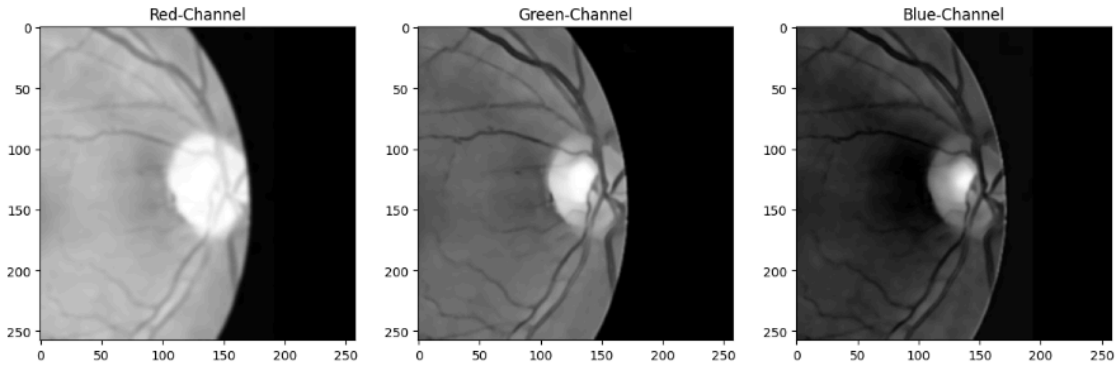


Figure 8. Difference in color channel

It first converts the image to grayscale, applies a Gaussian filter to smooth it, and then uses an active contour (snake) algorithm to detect the optic disc's boundary. The detected contour is used to create a mask (figure 9), which is then smoothed and applied to the original image (figure 9).

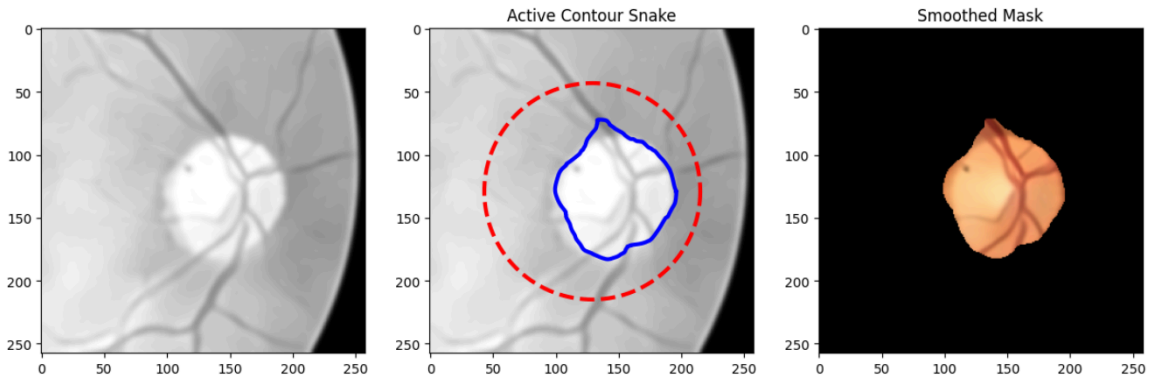


Figure 9. Active contour method

Example output:

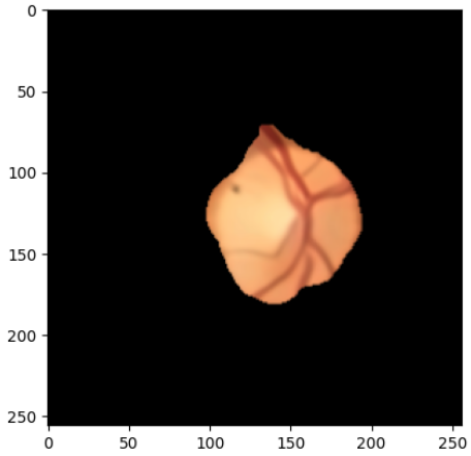


Figure 10. Segmented Optic Disk

4.2 Feature Extraction

Feature Extraction

Extract features from the ROI using hybrid DCNN approach.

We will aim to extract both high- and low-level features from fundus images using combined feature descriptors of CNN, SIFT, LBP and HOG to potentially improve the overall performance. [5]

CNN for High-Level Feature Extraction:

- The CNN architecture begins with a convolutional layer with a filter size of 20 and a stride of 1.
- The subsequent layer is a pooling layer with a size of 2 × 2 and a stride of 1.
- Another convolutional layer with a filter size of 32 and stride of 1 follows, alternating with the pooling layer for the first six layers.
- The seventh layer is an activation layer using the Rectified Linear Unit (ReLU), followed by an eighth convolutional layer with a filter size of 40 (4 × 4 × 32).
- The final layer is a Softmax function layer. The input image, resized to 100 × 100 × 1, is transformed to 1 × 1 × 2 via forward propagation through all layers.

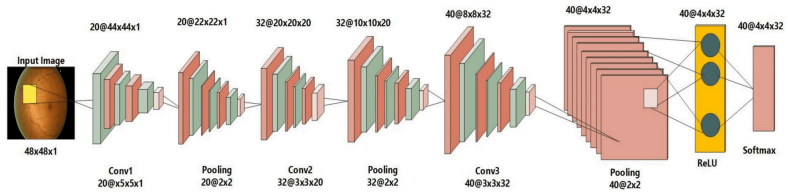


Fig 11. Deep CNN architecture for an input image.

HOG Descriptors:

- These capture gradient-based texture information. Images are divided into contiguous blocks of sizes ranging from 28 × 28 to 6 × 6, with each block measuring 2 × 2 and a stride of 4.
- A total of 9 bins are created, with 7 × 7 pixels per cell, chosen for their accuracy.

The equation of pixel (*x*, *y* gradient)'s magnitude M and direction and the direction of the pixel is calculated as below:

$$M\ (x,y) = \sqrt[2]{Ix^2 + Iy^2} \qquad \vartheta = \tan^{-1}\frac{I_y}{I_x}$$

Fig 12. HOG equations

LBP for Texture Classification:

- Local Binary Patterns (LBP) create dual feature vectors (e.g., 2^k, where K indicates the number of adjacent pixels).
- For 16 pixels, this results in 65,536 feature vectors, with the number increasing as the size of surrounding pixels increases.

$$LBP = \sum_{i=0}^{1} k\,(p-c).2^i,$$

$$k\,(x) = \begin{cases} 1 & if\ x \geq 0 \\ 0 & otherwise \end{cases}$$

Fig 13. The LBP's mathematical equation.

SIFT for Key Point and Descriptor Extraction:

The Scale-Invariant Feature Transform (SIFT) method is employed to extract key points and descriptors from the images.

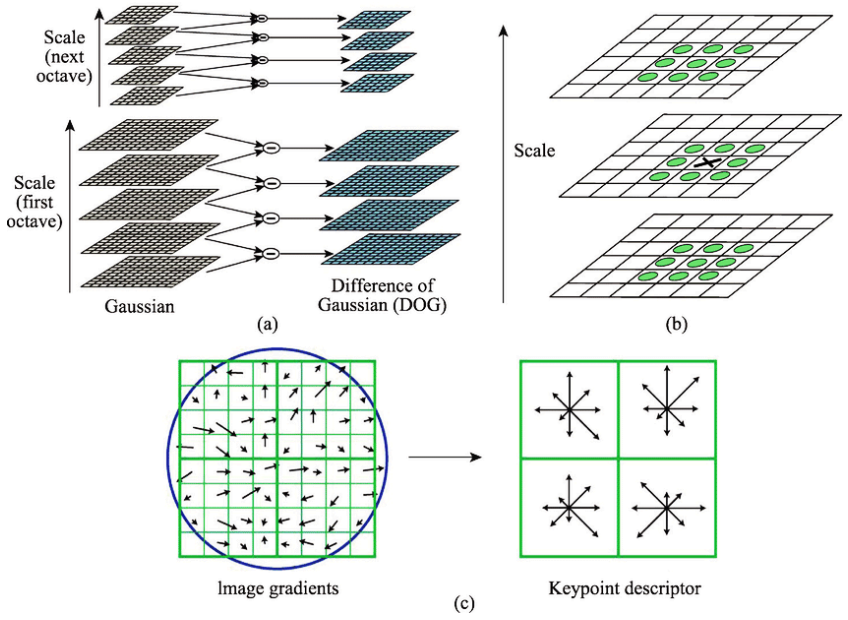


Fig 14. SIFT descriptor framework. (a) DoG, (b) Keypoint localization, (c) Orientation assignment and keypoint description. [5]

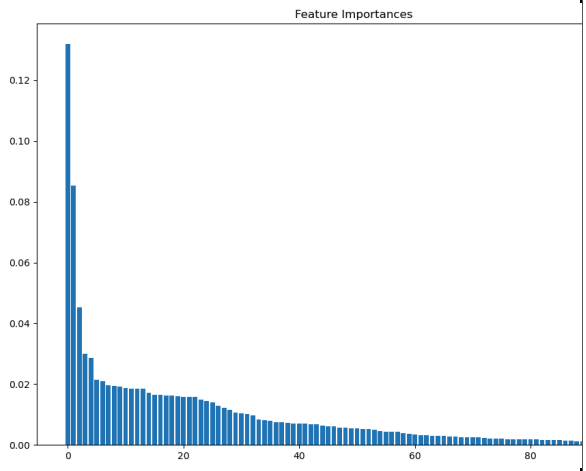
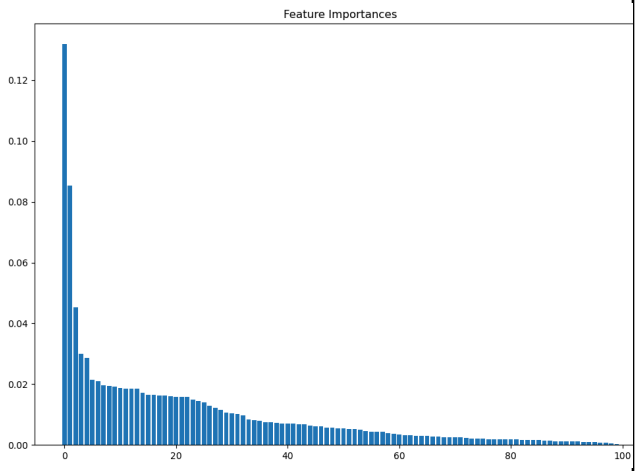
4.3 Feature Selection

We utilized the Minimum Redundancy Maximum Relevance (MRMR) algorithm for feature selection. Two variations were tested:

a) A `mrmr_selection` method referenced from GitHub. [6]

b) The `mrmr_classif` function with "num_features_to_select = 100", tested on a dataset of 686 training and 458 testing images labeled as ‘Glaucoma’ and ‘Non-Glaucoma’.

a)

Training_dataset	Testing_dataset
 **Fig 15** a. training_selected_features (k=100)	 **Fig 15** b. testing_selected_features (k=100)

b)

Training_dataset	Testing_dataset

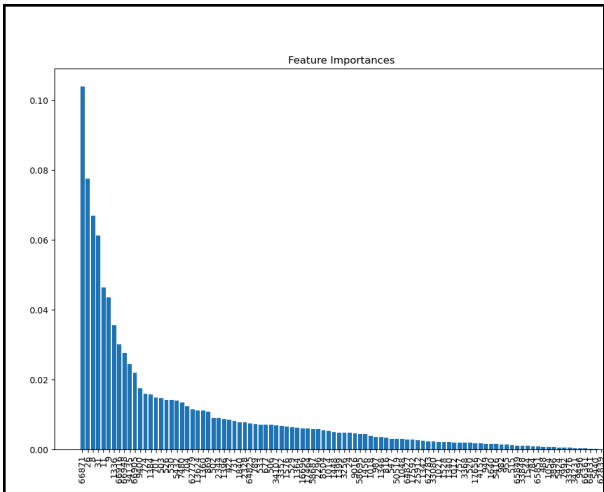


Fig 16 a. training_selected_features (k=100)

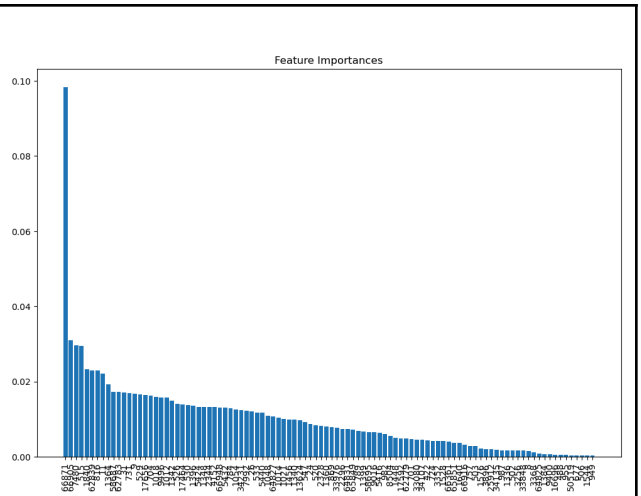


Fig 16 b. testing_selected_features (k=100)

From the plots of Feature Importances from two kinds of MRMR feature selection methods compared can be concluded that b) proceeds with better results showing more gradual decline, indicating a broader selection of moderately important features, which could capture more nuances in data. By observing, it shows that the algorithms prioritize different aspects of the data, suggesting differences in handling the features.

4.4 Classification

After the feature extraction and selection phase, the most important step is to classify the selected features. For the detection of glaucoma, selected features have been passed to our trained classifiers for the categorization of images into diseased and healthy images based on the ROI. The ROI for the detection of glaucoma is an enlarged optic disc (OD) comprising of optic cup (OC). For the training of the classifiers, we have employed train and test sets of fundus images split as 60% and 40% respectively. We employed the following classifiers [7]:

Support Vector Machine (SVM):

- A memory-efficient supervised learning method used to classify the images into healthy and unhealthy categories.

Random Forest (RF):

- This algorithm uses multiple de-correlated decision trees and is particularly effective on large datasets.

K-Nearest Neighbors (KNN):

- The simplest classifier, it uses a voting method to classify data. For instance, when k = 1, the object is assigned to the closest neighbor's class. [8]

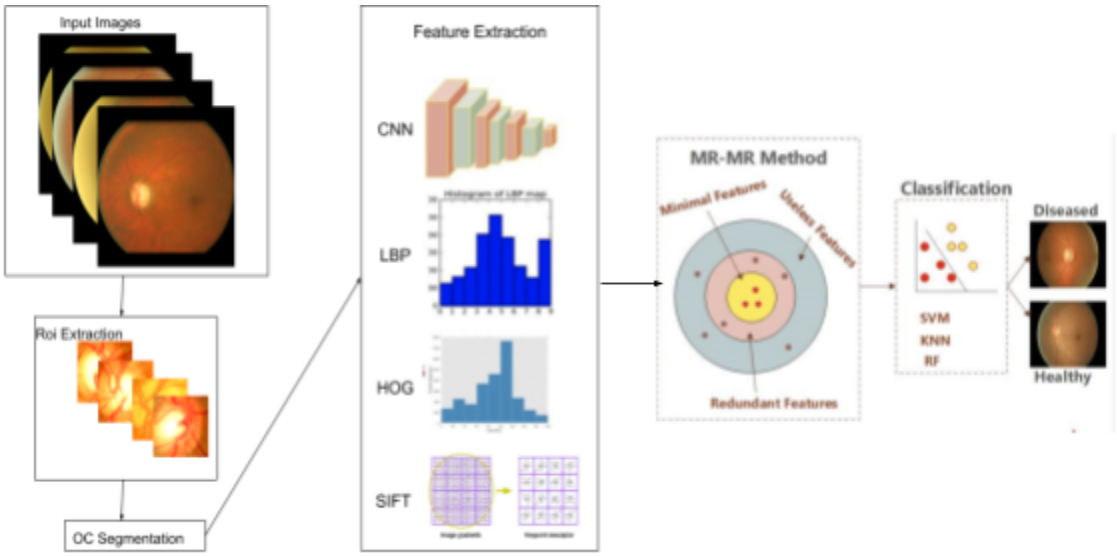


Fig 17. Proposed model

5. Experimental Evaluation

5.1 Experimental run

During the training phase, features have been extracted using a combination of CNN with HOG, LBP, and SURF, respectively. In CNN, the total number of convolutional layers was 20 and the pooling layer's size was 2. Moreover, both types of layers were alternatively present three times in a network. The ReLU layer was employed for the activation. For the classification purpose, the Softmax function was utilized at the final layer.

We employed three classification algorithms, i.e., SVM, KNN, and RF, to classify the images into two classes, i.e., healthy, and unhealthy. Initially, an SVM algorithm was employed and trained over samples of positive and negative classes. Then, KNN and RF are trained for all combinations of feature descriptors.

Experimental classification run on (686, 458) training and testing images which is labeled 'Glaucoma' and 'Non-Glaucoma', which we have successfully extracted ROI and segmented the OC, for the time initial run. We plan to expand our dataset after perfecting OC segmentation for more accurate results, experiment with increased images, improving our classification accuracy result.

a) mrmr_selection

```
Results for SVM:
Accuracy: 0.6739
Precision: 0.6986
Recall: 0.6739
F1-Score: 0.6760
              precision    recall  f1-score   support

Non-Glaucoma       0.78      0.62      0.69        81
    Glaucoma       0.58      0.75      0.66        57

    accuracy                           0.67       138
   macro avg       0.68      0.69      0.67       138
weighted avg       0.70      0.67      0.68       138

--------------------------------------------------
Results for Random Forest:
Accuracy: 0.7174
Precision: 0.7275
Recall: 0.7174
F1-Score: 0.7194
              precision    recall  f1-score   support

Non-Glaucoma       0.79      0.70      0.75        81
    Glaucoma       0.64      0.74      0.68        57

    accuracy                           0.72       138
   macro avg       0.71      0.72      0.71       138
weighted avg       0.73      0.72      0.72       138

--------------------------------------------------
Results for KNN:
Accuracy: 0.6377
Precision: 0.7019
Recall: 0.6377
F1-Score: 0.6328
              precision    recall  f1-score   support

Non-Glaucoma       0.82      0.49      0.62        81
    Glaucoma       0.54      0.84      0.66        57

    accuracy                           0.64       138
   macro avg       0.68      0.67      0.64       138
weighted avg       0.70      0.64      0.63       138

--------------------------------------------------
```

Fig 18 a. Classification report

b) mrmr_classif

```
Results for SVM:
Accuracy: 0.8718
Precision: 0.8822
Recall: 0.8718
F1-Score: 0.8718
              precision    recall  f1-score   support

Non-Glaucoma       0.81      0.94      0.87        18
    Glaucoma       0.94      0.81      0.87        21

    accuracy                           0.87        39
   macro avg       0.88      0.88      0.87        39
weighted avg       0.88      0.87      0.87        39

--------------------------------------------------
Results for Random Forest:
Accuracy: 0.7949
Precision: 0.7996
Recall: 0.7949
F1-Score: 0.7951
              precision    recall  f1-score   support

Non-Glaucoma       0.75      0.83      0.79        18
    Glaucoma       0.84      0.76      0.80        21

    accuracy                           0.79        39
   macro avg       0.80      0.80      0.79        39
weighted avg       0.80      0.79      0.80        39

--------------------------------------------------
Results for KNN:
Accuracy: 0.9231
Precision: 0.9239
Recall: 0.9231
F1-Score: 0.9229
              precision    recall  f1-score   support

Non-Glaucoma       0.94      0.89      0.91        18
    Glaucoma       0.91      0.95      0.93        21

    accuracy                           0.92        39
   macro avg       0.93      0.92      0.92        39
weighted avg       0.92      0.92      0.92        39

--------------------------------------------------
```

Fig 18 b. Classification report

5.2 Evaluation Metrics

Various metrics have been used for the performance evaluation of the proposed model.

Accuracy: Accuracy is calculated to examine the performance of the proposed system over the specific data. The equation of accuracy is shown below.

$$Accuracy = (TP + TN) / (TP + TN + FP + FN),$$

where true positive (*TP*) denotes the total images that were classified correctly as non-healthy, whereas false positive (*FP*) represents the images that were incorrectly classified as positive class such as diseased. The false negative (*FN*) represents the percentage of images that our proposed method was not able to detect as non-healthy. Moreover, true negative (*TN*) represents the images that are healthy and classified as healthy by our proposed system.

Recall: Recall *R* refers to the percentage of the images that were diseased and recalled by the system.

$$R = TP/(TP + FN) * 100,$$

Precision: Precision *P* refers to the ratio of the frames that are correctly classified through the proposed method. The equation of precision is presented below:

$$P = TP/(TP + FP) * 100,$$

F1 Score: It is also known as the F Score denoted by F. The equation is given below:

$$R = 2.(P * R)/(P + R),$$

5.3) Future Work

Moving forward, we plan to experiment with different combinations of feature extraction methods and classifiers. The goal is to identify the combination that yields the best performance using the equations and methods described above. [9]

The combined feature extraction equation of CNN and LBP is presented below:

$$CNN_{LBP} = \sum_{s=1}^{l} \sum_{t=1}^{w} fst.I_{x+s-1,y+t-1} + \sum_{i=0}^{1} LBP_{p,r}$$

$$LBP_{p,r} = \frac{(I * fst)_{x,y} - (I * fst)_{x,y}}{fst_{xc-x+1,yc-y+1}}$$

$$LBP_{p,r} = \sum_{i=0}^{1} k(p_i - c_i).2^i = I(xc, yc) \;\&\; k(x) = \begin{cases} 1, \; if \; x \geq 0 \; ; \; 0, \; otherwise \; , \end{cases}$$

The combined equation for CNN and HOG became as below:

$$(I_{M,\vartheta} * K)_{x,y} = \sum_{s=1}^{l} \sum_{t=1}^{w} fst.I_{x+s-1,y+t-1} \; . \; e^{s\vartheta(x+s-1,y+t-1)}$$

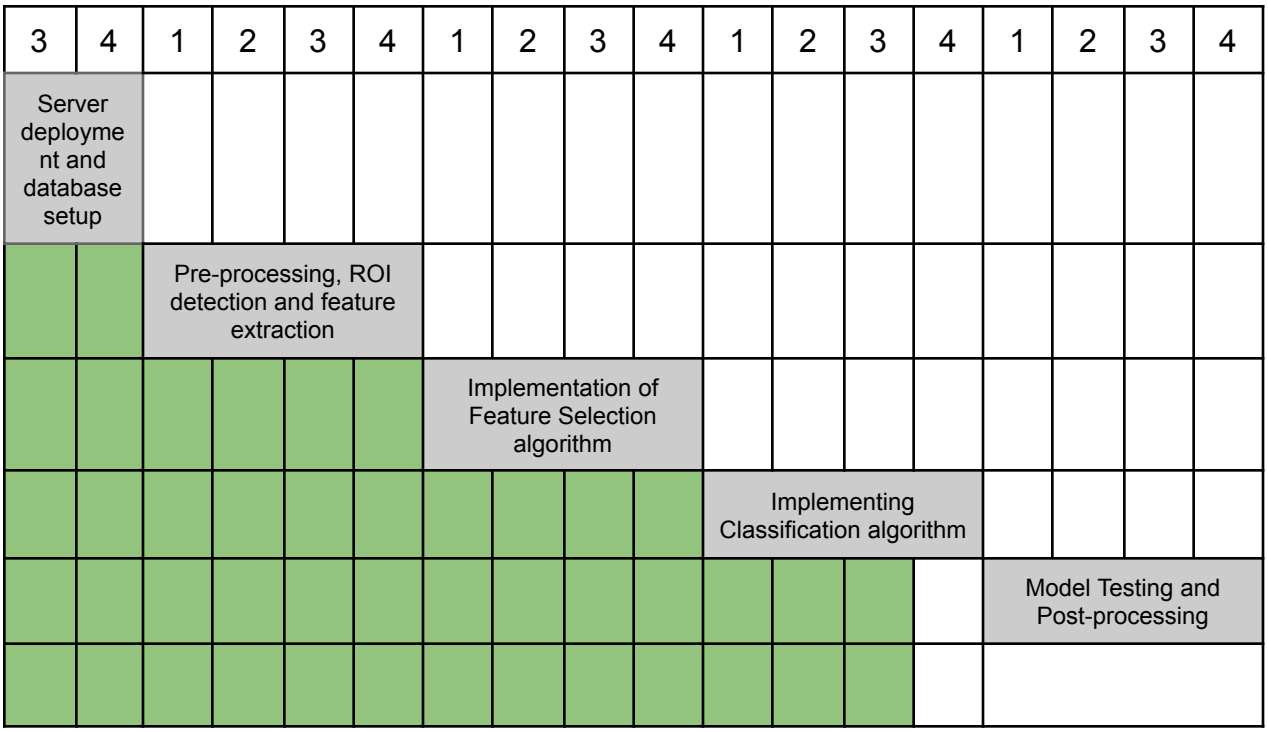
Classifier Name
SVM + LBP
SVM + HOG
SVM + CNN
SVM + SURF
SVM + HOG + CNN
SVM + CNN + LBP
SVM + CNN + SURF
SVM + HOG + CNN + LBP
SVM + HOG + CNN + LBP + SURF
RF + HOG
RF + CNN
RF + SURF
RF + LBP
RF + CNN + LBP
RF + CNN + SURF
RF + HOG + CNN + LBP
RF + HOG + CNN + LBP + SURF
RF + HOG + CNN
KNN + HOG
KNN + CNN
KNN + LBP
KNN + SURF
KNN + CNN + HOG
KNN + HOG + CNN
KNN + HOG + CNN + LBP
KNN + HOG + CNN + LBP + SURF

Fig 19. Possible Combinations

6. Development schedule and role division

6.1 Schedule

May	Jun	Jul	Aug	Sep



6.2 Roles

Name	Tasks
Nomin	- Data collection and cleaning - Implementation of Feature Extraction algorithm - Implementation of Feature Selection algorithm
Bazarsad	- Image Preprocessing - ROI Detection algorithm - OC Segmentation
Common	- Deep learning model development - Testing and improving - Presentation - Implementation of Classification algorithm

7. References

[1] Bhat, S.H.; Kumar, P. Segmentation of Optic Disc by Localized Active Contour Model in Retinal Fundus Image. In *Smart Innovations in Communication and Computational Sciences*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 35–44. (*Google Scholar*)

[2] Fundus images and OD/OC masks from ORIGA, REFUGE, and G1020 datasets. (Kaggle)

[3] Rutuja Shinde, Glaucoma detection in retinal fundus images using U-Net and supervised machine learning algorithms, Intelligence-Based Medicine, Volume 5, 2021, p4. (ScienceDirect)

[4] Mahum, R.; Rehman, S.U.; Okon, O.D.; Alabrah, A.; Meraj, T.; Rauf, H.T. A Novel Hybrid Approach Based on Deep CNN to Detect Glaucoma Using Fundus Imaging. *Electronics* **2022**, p6. (MDPI)

[5] Modified Bag of Visual Words Model for Image Classification Article's Information Abstract - Scientific Figure on ResearchGate. (Research Gate)

[6]Mahum, R.; Rehman, S.U.; Okon, O.D.; Alabrah, A.; Meraj, T.; Rauf, H.T. A Novel Hybrid Approach Based on Deep CNN to Detect Glaucoma Using Fundus Imaging. *Electronics* **2022**, pp. 6-8. (MDPI)

[7] Smazzanti, mRMR feature selection Implementation. **2022**, p9.. (Github)

[8] Mahum, R.; Rehman, S.U.; Okon, O.D.; Alabrah, A.; Meraj, T.; Rauf, H.T. A Novel Hybrid Approach Based on Deep CNN to Detect Glaucoma Using Fundus Imaging. *Electronics* **2022**, p. 9. (MDPI)

[9] Mahum, R.; Rehman, S.U.; Okon, O.D.; Alabrah, A.; Meraj, T.; Rauf, H.T. A Novel Hybrid Approach Based on Deep CNN to Detect Glaucoma Using Fundus Imaging. *Electronics* **2022**, p. 11. (MDPI)