UNIVERSITY SCHOOL OF AUTOMATION AND ROBOTICS

COMPUTER VISION BASED HEALTHCARE SYSTEM FOR IDENTIFICATION OF DIABETES & ITS TYPES USING AI

Avinash Sharma a , K.D.V. Prasad b,c , Sadashiva V. Chakrasali d , Dankan Gowda Ve, Chanakya Kumarf , Abhay Chaturvedi g , A. Azhagu Jaisudhan Pazhani h

By Eshaan Gupta
14519051622
AI&ML-2026

OVERVIEW

- Introduction
- Problem Statement & Objectives
- Overview of Diabetic Retinopathy
- Dataset and Preprocessing Steps
- Image Augmentation
- Model Architectures

1. SVM Classifier
2. Fuzzy C-Means
3. Random Forest
4. CNN

- Model Evaluation and Comparison
- Key Findings
- Conclusion

INTRODUCTION

- Diabetes Prevalence: The International Diabetes Federation (IDF) predicts that by 2030, 550 million people will have diabetes.

- Need for AI & ML: Early detection of diabetes is crucial for managing complications. AI and ML can assist in automated diagnosis.

- Research Focus: How CNN-based deep learning models can help in detecting diabetes and its types efficiently.

What is diabetic retinopathy ?

- Diabetic Retinopathy (DR) is a complication of diabetes that affects the retina, the light-sensitive tissue at the back of the eye.
- It occurs due to prolonged high blood sugar, which damages the blood vessels in the retina.
- Damaged vessels can leak fluid or blood, causing vision problems and eventually vision loss.

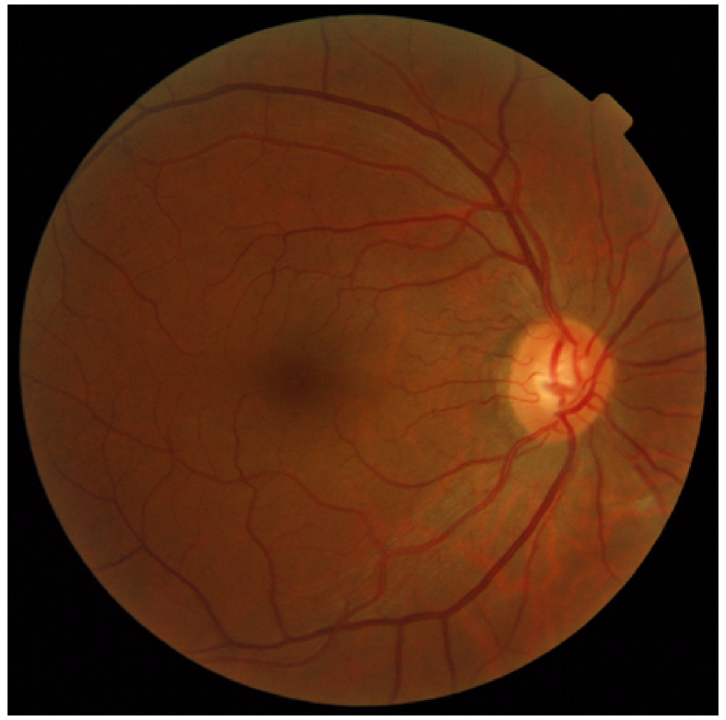
4 stages of diabetic retinopathy

DR progresses through four stages:

- Mild NPDR – Microaneurysms
- Moderate NPDR – Vessel swelling and leakage
- Severe NPDR – Blocked blood flow
- Proliferative DR – New abnormal vessels grow

In the next slide, we will see example for each stage.

4 stages of diabetic retinopathy


No diabetic retinopathy


Mild NPDR


Moderate NPDR


Severe NPDR


Proliferative diabetic retinopathy

- It is evident from the previous slide that raw retinal images alone are not sufficient to effectively train a model to predict the stage of diabetic retinopathy.
- Therefore, image preprocessing is essential before developing a machine learning model to extract meaningful features and maximize relevant information from the images.

KEY FEATURES

- **Vessel Growth**

Neovascularization indicating advanced stages

- **Exudates**

Bright lipid deposits—early signs of leakage

- **Hemorrhages**

Dark/red blood spots—caused by ruptured vessels

- **Fluid Deposits**

Signs of retinal swelling or macular edema

These features are crucial indicators used to identify and classify the stage of Diabetic Retinopathy.

PRE PROCESSING STEPS

- Image loading and resizing
- Gaussian Blurring
- CLAHE [Contrast Limited Adaptive Histogram Equalization] on the green channel
- Canny Edge Detection and Morphological Closing
- Lession Highlighting using Morphological and colour techniques
- Visualization of results

STEP 1

IMAGE LOADING AND RESIZING

- The dataset was sourced from Kaggle, containing a total of 1,744 retinal images.
- A CSV file accompanied the dataset, mapping each image to its corresponding class label.
- To streamline the classification process, images were organized into separate directories based on their labels Class_0, Class_1, Class_2, Class_3, and Class_4
- This structure transformed the data into a more model-friendly format, enabling efficient pre-processing and training.
- Resizing the images will help to maintain uniformity for future processes. The images were resized to height and width of 512 and 512 pixels respectively.

STEP 2

GAUSSIAN BLURRING

- **Purpose:** To reduce high-frequency noise while retaining important anatomical features of the retina.

- **How it works:** Applies a low-pass filter that smooths the image by averaging pixel values using a Gaussian kernel.

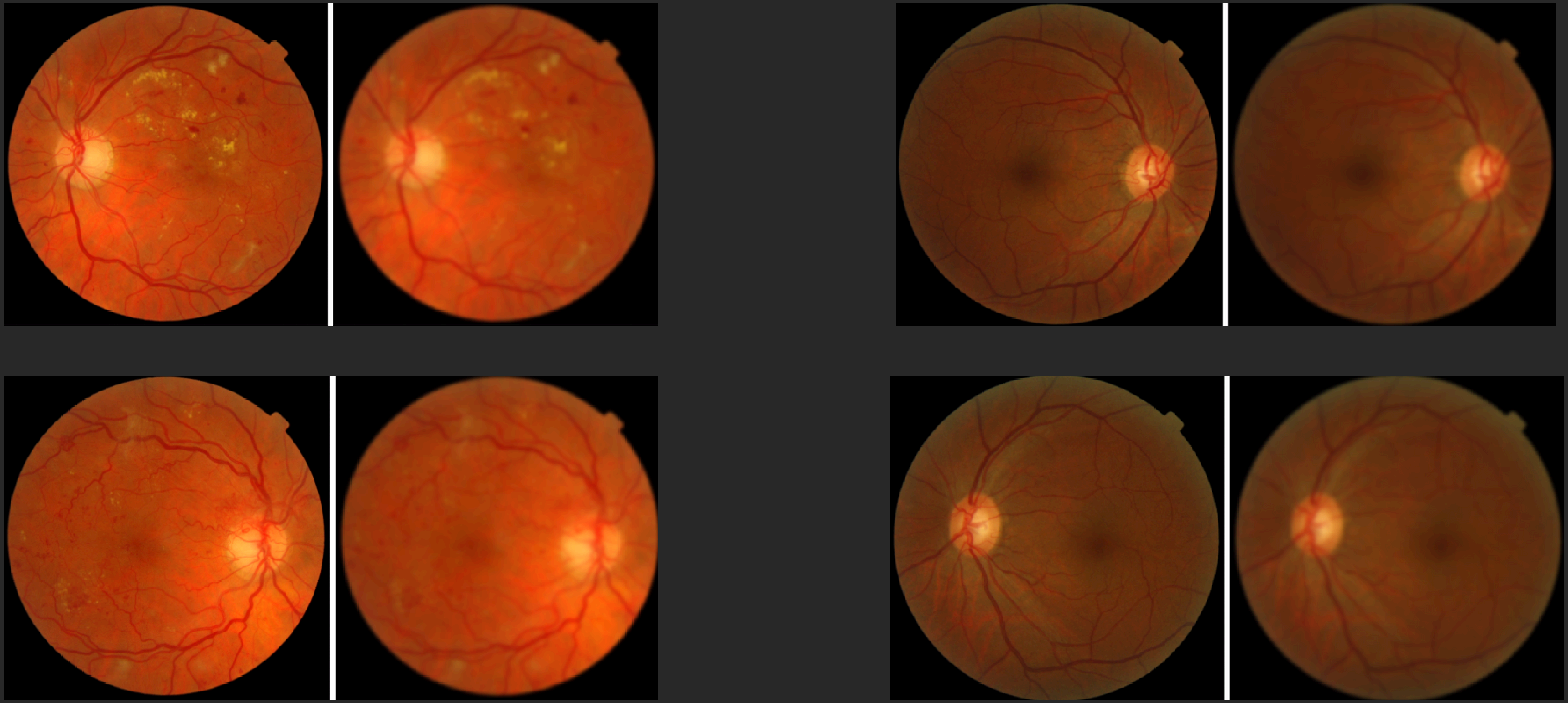
- **Preserves:** The general structure and major features such as blood vessels, exudates, and hemorrhages.

- A kernel size of 15×15 was used, which provides a strong smoothing effect — ideal for minimizing background noise and enhancing the visibility of clinically relevant features.

- This step ensures that subtle retinal abnormalities are not masked by random noise in the raw image.

STEP 2

Some examples of images after applying gaussian blur



STEP 3

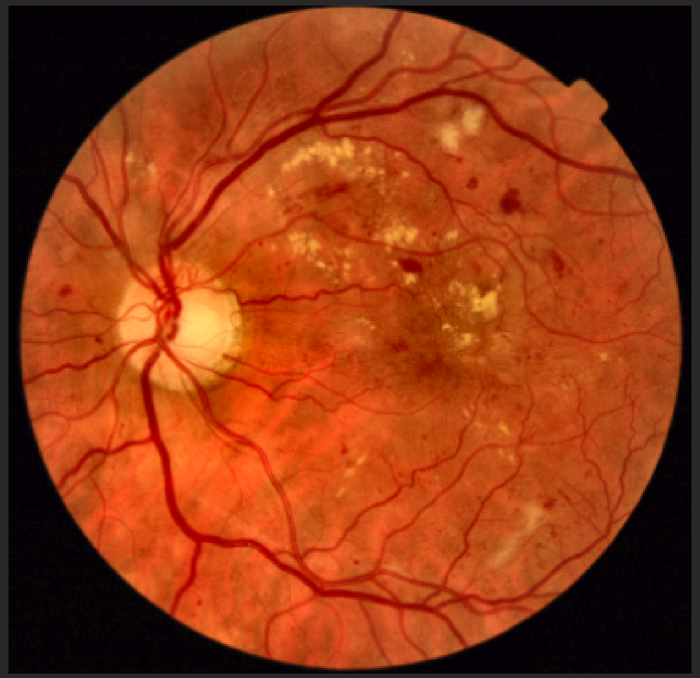
CONTRAST ENHANCING USING CLAHE

- CLAHE stands for **Contrast Limited Adaptive Histogram Equalization.**
- Enhances local contrast by working on small regions of the image.
- Avoids amplifying noise by clipping histogram values.
- Especially effective in medical images, where subtle texture differences (like lesions) are important.

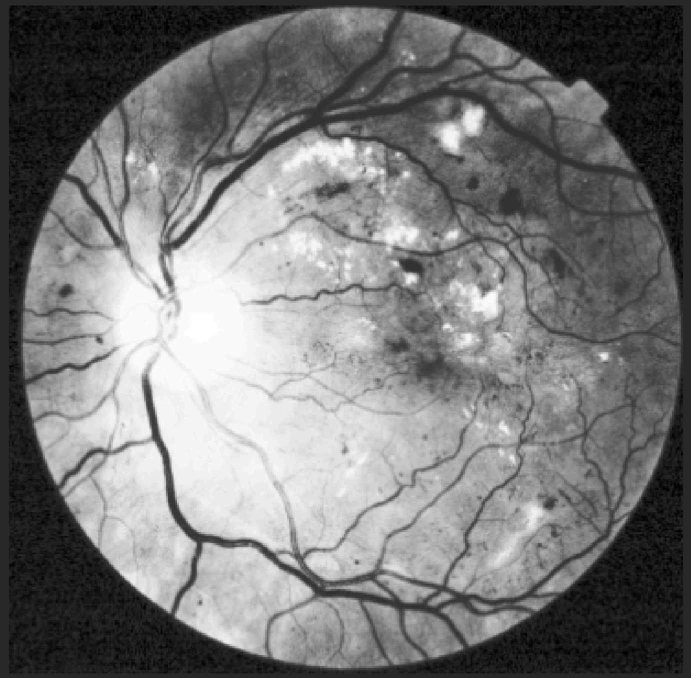
STEP 3

Why CLAHE ?

Original



CLAHE Equalized Image



Histogram Equalized image

- Standard histogram equalization can make images overly bright, leading to the loss of important features.
- CLAHE (Adaptive Histogram Equalization) improves contrast tile-by-tile, enhancing local details without overexposing.
- It helps in preserving critical features like blood vessels, exudates, and hemorrhages.

STEP 3

Enhancing RGB and LAB channels using CLAHE

Original L Channel R Channel G Channel B Channel

- L: Lightness (from black to white)
- A: Green–Red component
- B: Blue–Yellow component

Combining G channel CLAHE enhancement with L channel helps to capture the exudates, blood vessels and hemorrhages clearly while maintaining the colour of the image.

STEP 4

GAMMA CORRECTION

Gamma correction is a nonlinear operation that adjusts the luminance of an image. It changes how light or dark the image appears to the human eye.

Why is it Used ?

1. Enhance Image Visibility
2. Helps improve visibility in dark or bright areas without losing details.
3. Pre-processing for Computer Vision
4. Adjusts lighting to normalize input before feeding into models or further algorithms.

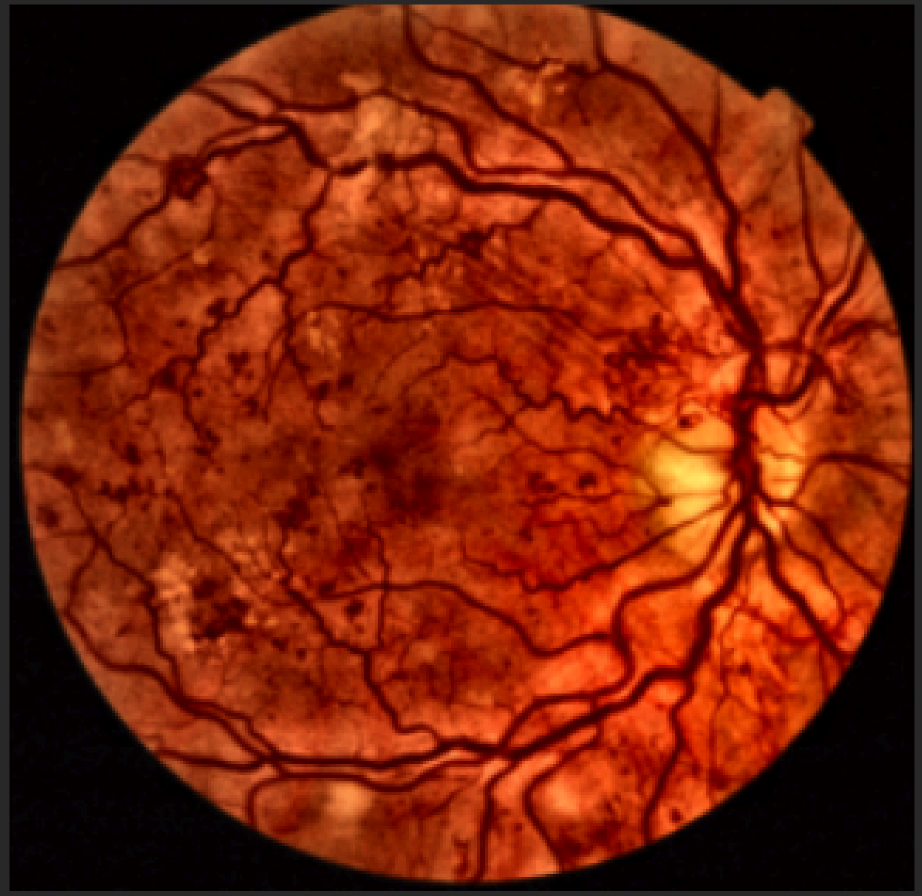
STEP 4

Original

- Vessels not visible
- Exudates not clearly visible
- Hemorrhages not visible

Gamma corrected image with value 1.3

- Vessels clearly visible
- Exudates clearly highlighted
- Hemorrhages clearly highlighted



Gamma corrected image with value 1.6

- Vessels clearly visible
- Exudates clearly visible
- Some hemorrhages not clearly visible

STEP 5

The pre processing for step 5 was approached in two different ways:

- Images were converted to grayscale and Canny Edge Detection was implemented to highlight the edges
- Lesion highlighting using morphological and color techniques

STEP 5

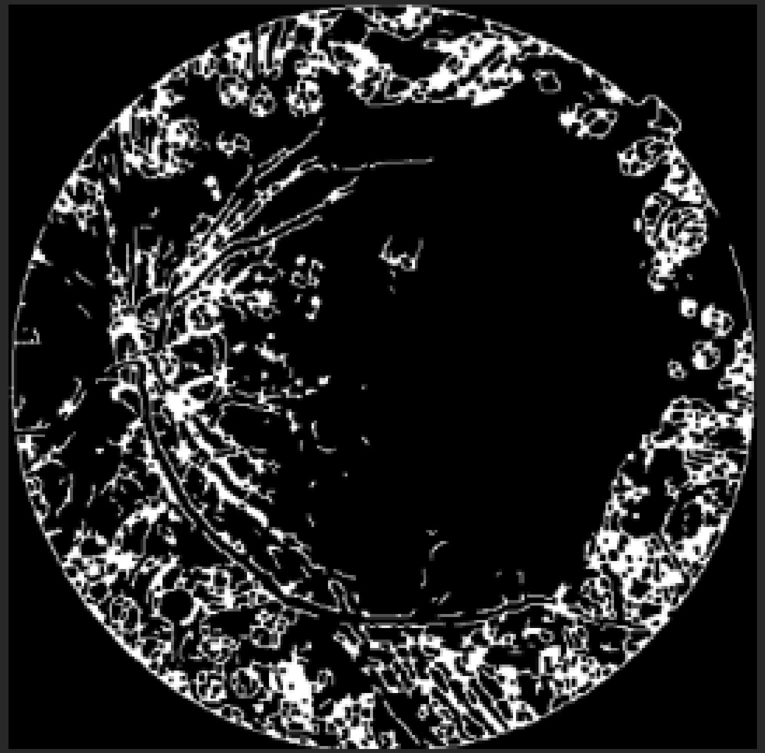
CANNY EDGE DETECTION

Detects sharp edges in an image — often used in object detection and computer vision.

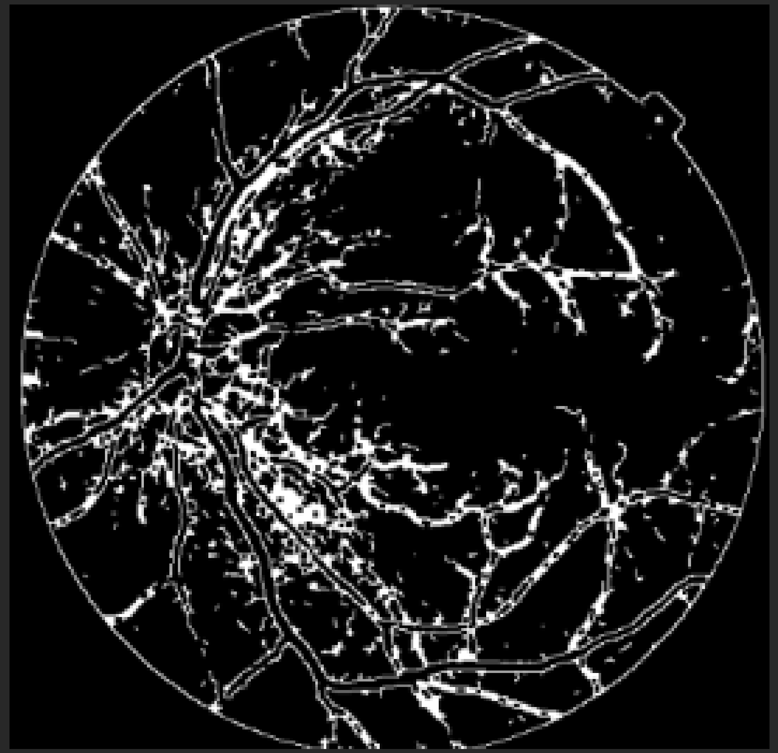
MORPHOLOGICAL CLOSING

Fills small holes or gaps in binary images — helps connect disjointed parts of an object.

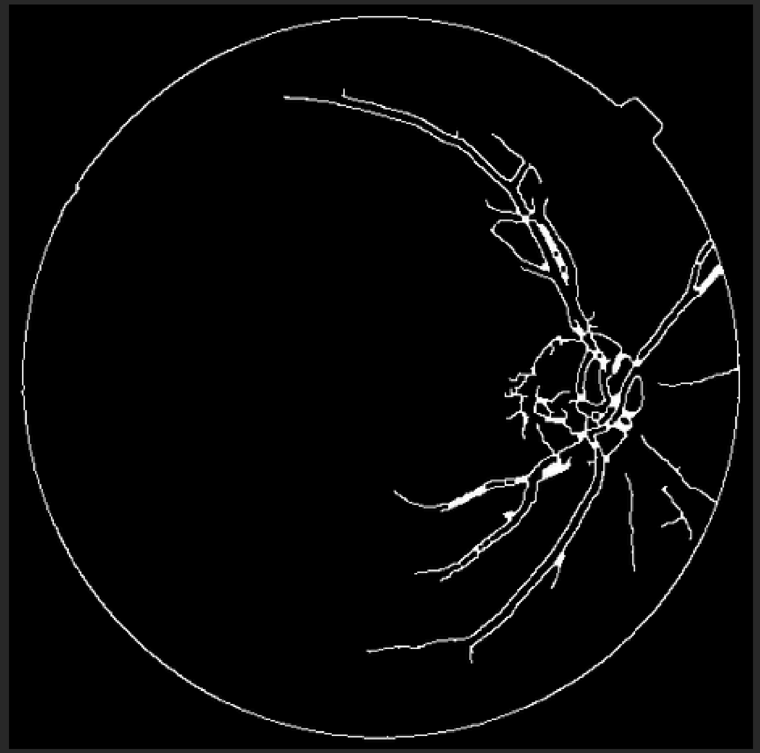
STEP 5



Class 4



Class 2

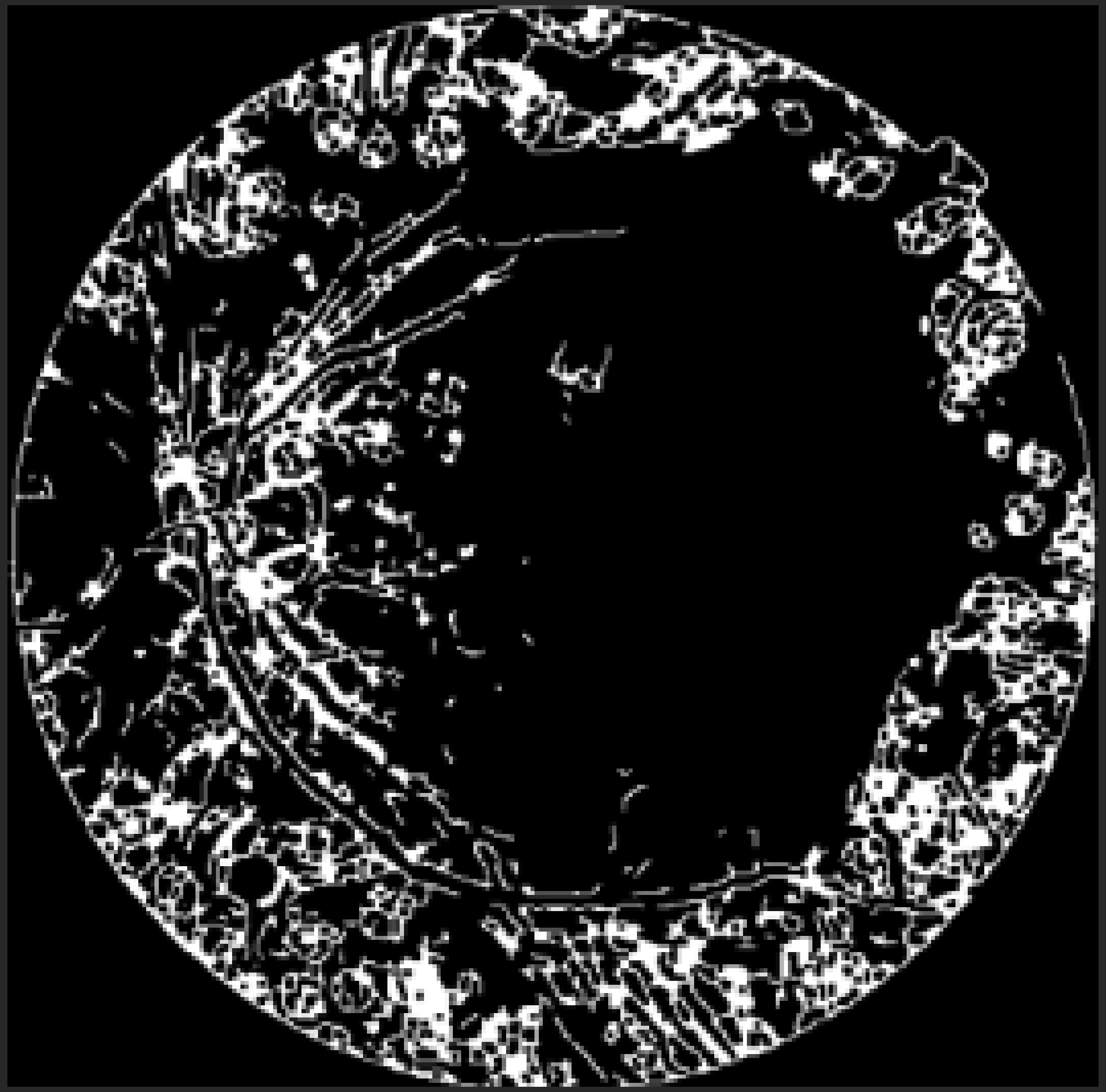


Class 0

STEP 5

Need for another approach

- Missing chaotic vessel structure
- Exudates and hemorrhages not distinguished
- No clear distinction between classes



Class 4

STEP 5

LESION HIGHLIGHTING USING MORPHOLOGICAL AND COLOR TECHNIQUES

Detects sharp edges in an image — often used in object detection and computer vision.

Morphological operations used are White top hat and Black hat

1. White top hat highlights small white regions on a dark background which is useful for finding exudates.
2. Black hat highlights small dark regions on a bright background which is useful for finding hemorrhages.

STEP 5

LESION HIGHLIGHTING USING MORPHOLOGICAL AND COLOR TECHNIQUES

- After applying white top hat and black hat operations, another hsv yellow kernel was applied on the image to catch any remaining yellow pigmentation on the retinal scans.
- This ensures that maximum information is extracted from the images.

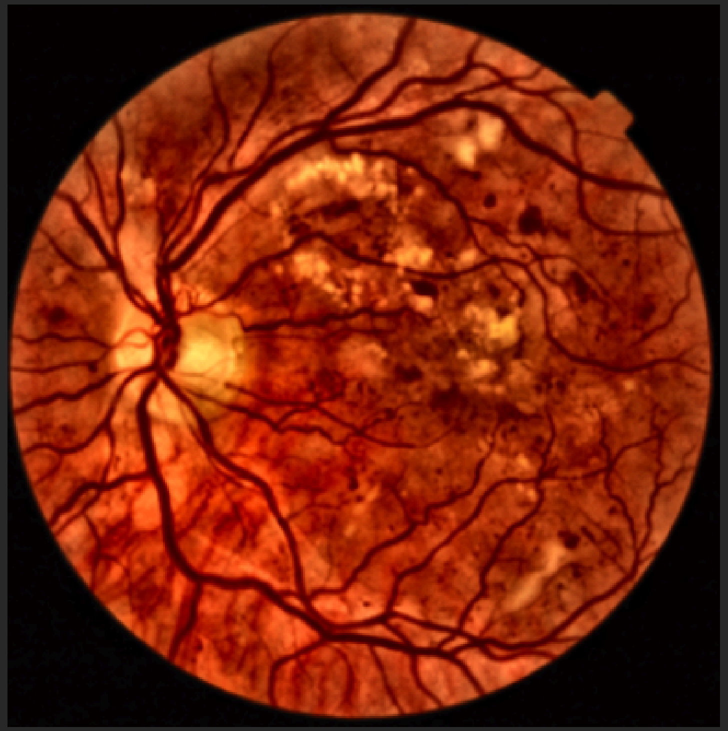
STEP 5

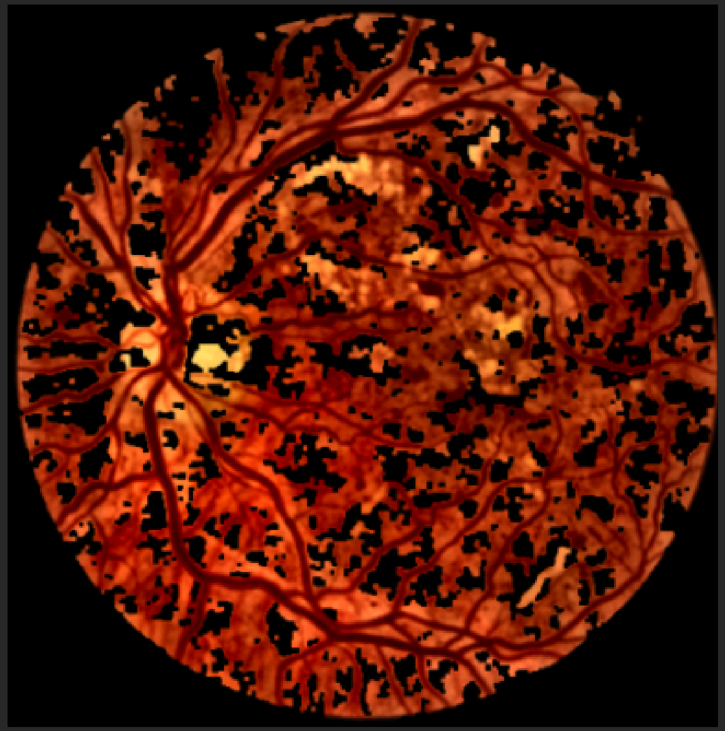
LESION HIGHLIGHTING USING MORPHOLOGICAL AND COLOR TECHNIQUES

After applying the morphological operations, the red regions in the image were effectively suppressed, turning them black, while important structures such as blood vessels, exudates, and hemorrhages were preserved.

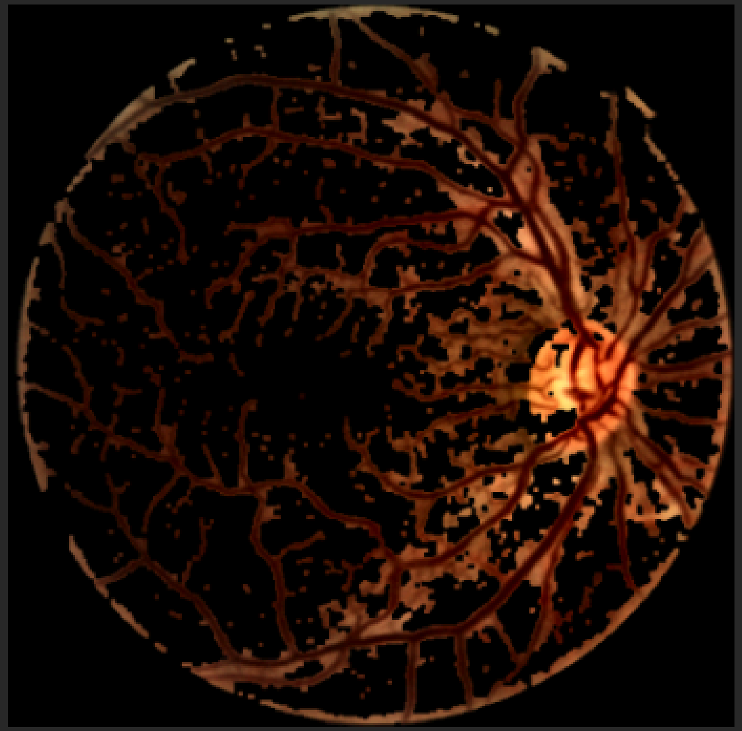
This technique allows the system to focus solely on clinically relevant areas by removing unnecessary distractions and enhancing the visibility of abnormalities.

STEP 5



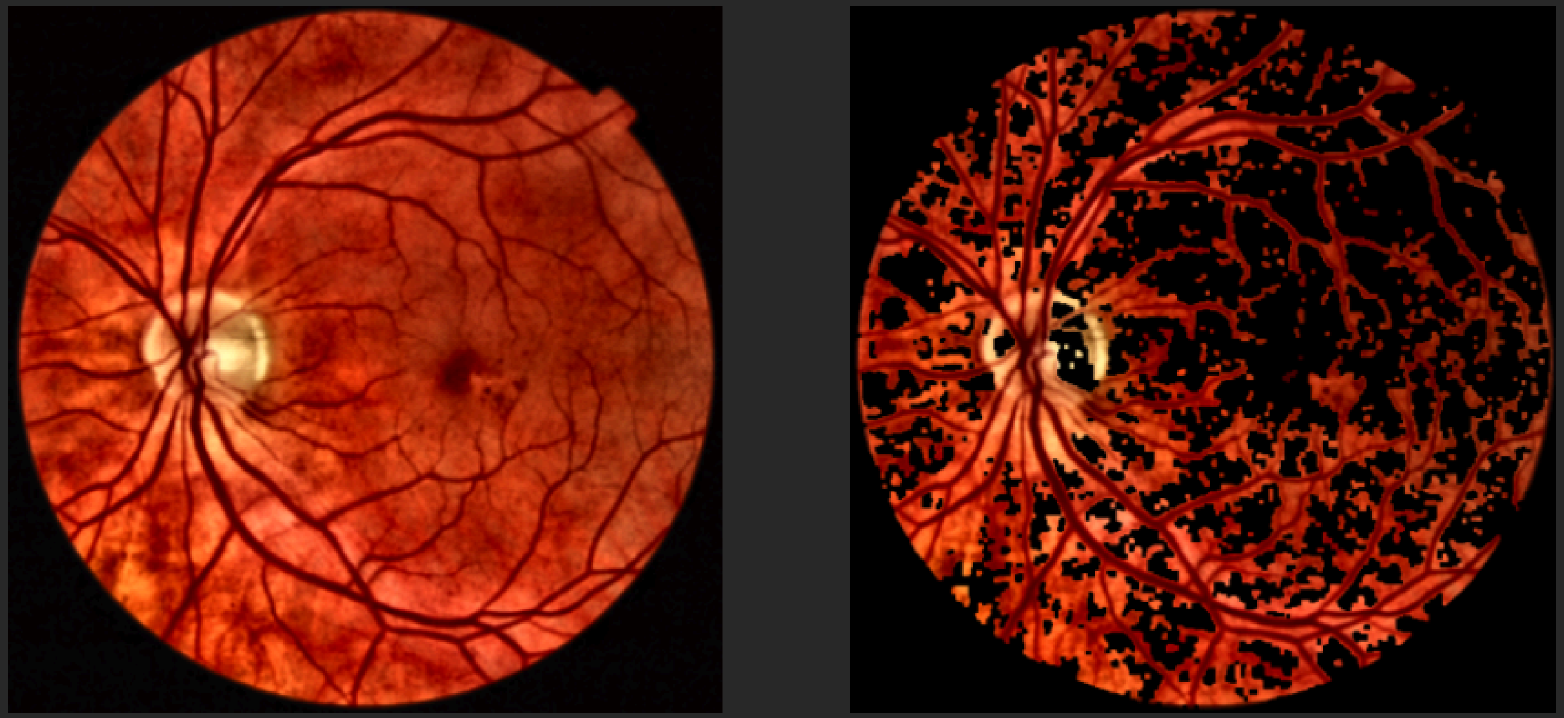


- Chaotic blood vessel structure perfectly captured (Class 4).
- Exudates clearly highlighted.
- Hemorrhages clearly visible as dark red spots.

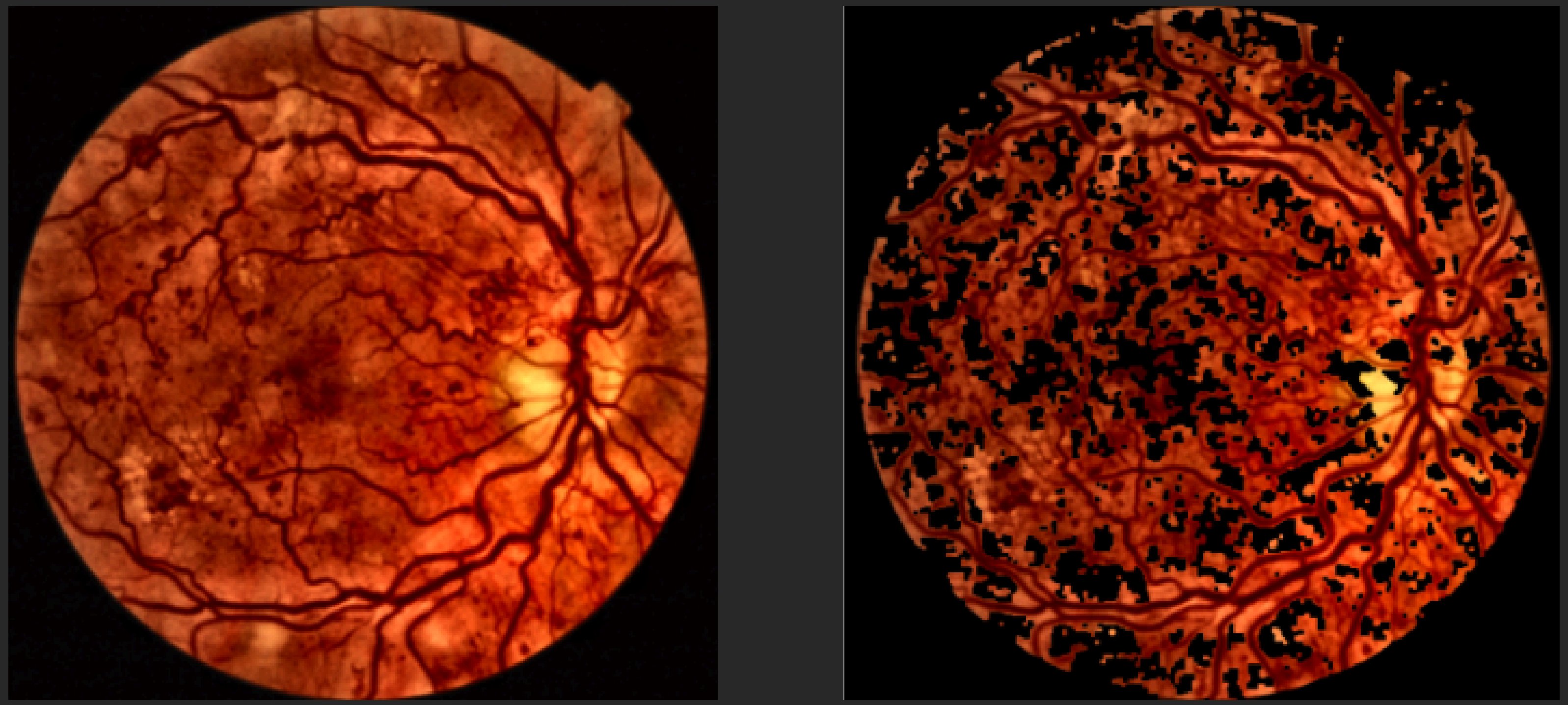


- Normal blood vessels structure as expected (Class 0).
- No exudates.
- No hemorrhages.

STEP 5



Class 2



Class 3

STEP 6

FINAL RESULTS OF PRE PROCESSED IMAGES

After successfully testing these operations on a sample of images from different classes, they were applied to the entire dataset to preprocess and prepare the images for training a machine learning model.

Original Smooth CLAHE Gamma Morphological operations

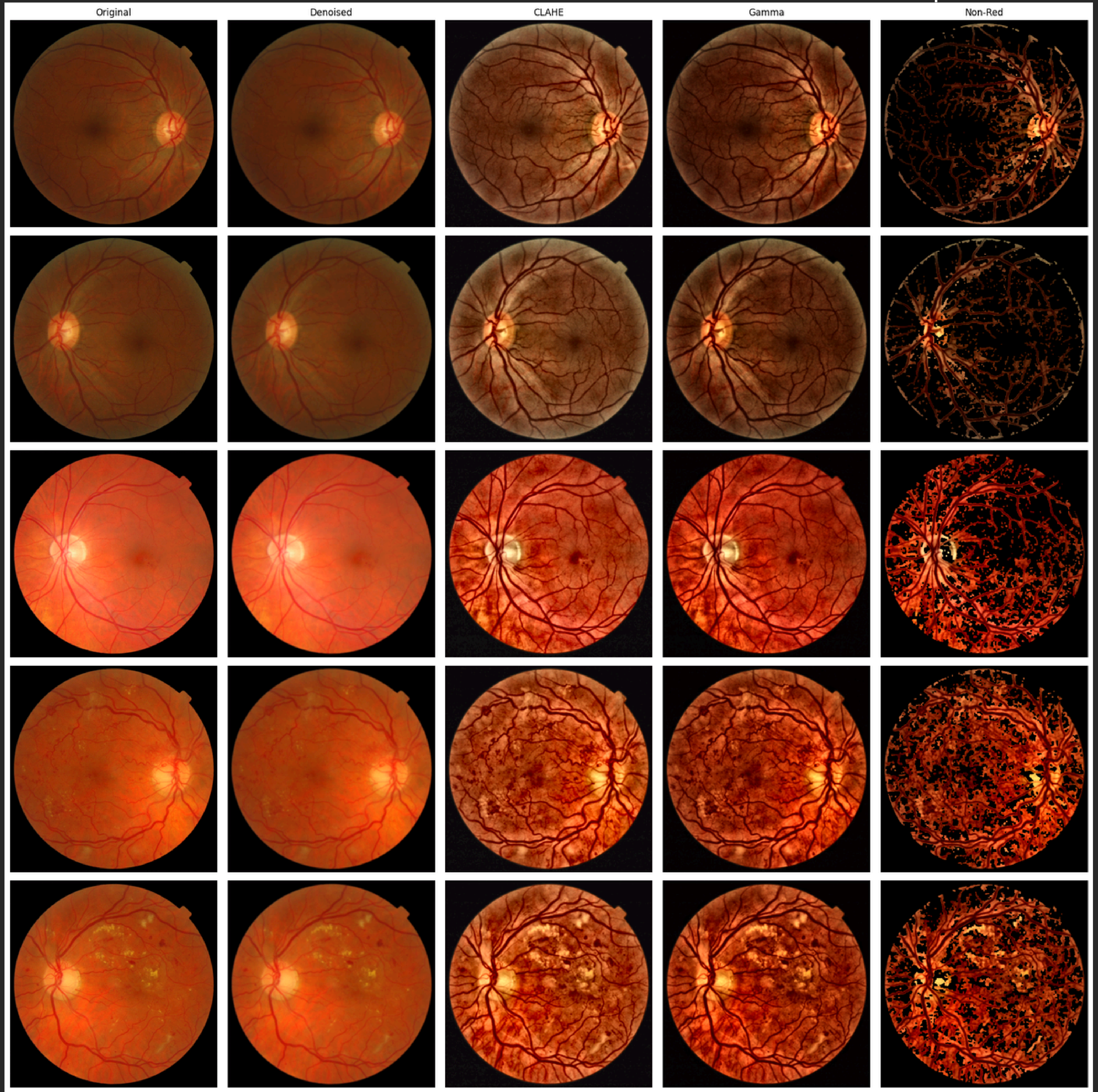
Class 0

Class 1

Class 2

Class 3

Class 4



FINAL STEP

IMAGE AUGMENTATION

To enhance model training and improve efficiency, all images were augmented through 90-degree and 180-degree rotations in both clockwise and counter-clockwise directions. This not only introduced variability but also helped in balancing the dataset. Augmentations were repeated as needed until each class contained exactly 2,000 images, ensuring uniform representation across all categories.

MODEL TRAINING

To evaluate the effectiveness of the preprocessing pipeline, I experimented with three different models: Fuzzy C-Means Clustering, Random Forest Classifier, Support Vector Machine, and a Convolutional Neural Network (CNN). These models were selected to represent both classical machine learning and deep learning approaches.

Due to hardware and resource constraints, I was unable to train the models on the full dataset of 10,000 images. Instead, I limited the training set to 1,250 images, selecting 250 images per class. While this reduced dataset may impact the final accuracy scores, it still provides a reasonable estimate of how well the preprocessing pipeline enhances feature quality and class separability.

SVM CLASSIFIER

THE HISTOGRAM OF ORIENTED GRADIENTS (HOG)

The Histogram of Oriented Gradients (HOG) is a feature extraction method used in computer vision for object detection and classification. It works by converting an image to grayscale, dividing it into small cells (e.g., 16x16 pixels), and computing gradient directions and magnitudes for each pixel.

These gradients are grouped into orientation histograms per cell, typically using 9 bins. To improve robustness to lighting and contrast, histograms are normalized across overlapping blocks. The resulting feature vector effectively captures the image's structural information for use in classifiers.

SVM CLASSIFIER

SUPPORT VECTOR MACHINE

Support Vector Machine (SVM) is a supervised classification algorithm that finds the optimal boundary (hyperplane) to separate classes by maximizing the margin between them. If the data isn't linearly separable, it uses the "kernel trick"—in this case, the RBF kernel—to project data into a higher-dimensional space where separation is possible.

The SVM is trained on HOG feature vectors from labeled images, learning to classify based on patterns in the data. Its accuracy is then evaluated on unseen images to ensure it generalizes well.

FUZZY C MEANS

Fuzzy C-Means (FCM) is a clustering algorithm where each data point can belong to multiple clusters with varying degrees of membership, unlike traditional methods like K-Means where each point belongs to only one cluster. This soft clustering approach is more flexible and better at capturing complex patterns in data.

Key Characteristics:

- Soft Assignment: Instead of hard labels, each sample has a membership score for every cluster.
- Cluster Centers: The algorithm iteratively updates cluster centers based on weighted membership values.
- Membership Matrix: Determines how strongly each data point belongs to each cluster.

FUZZY C MEANS

Parameters Used:

- Number of Clusters (c): Set to 5, matching the number of classes in the dataset.
- Fuzziness Parameter (m = 2): Controls the level of cluster overlap—higher values allow more overlap.
- Error Threshold (0.005): Defines the convergence criteria for stopping.
- Max Iterations (1000): Limits how long the algorithm runs.

Application Insight:

- Helps explore hidden patterns in data without relying on labels.
- Useful for evaluating the quality of feature representation, especially when ground truth may be limited.

RANDOM FOREST

Random Forest is an ensemble learning method that builds multiple decision trees and merges their results to improve prediction accuracy and control overfitting.

How It Works:

- Multiple decision trees are trained on different subsets of the data and features.
- Each tree independently predicts a class.
- Final prediction is made by majority voting among trees.

RANDOM FOREST

Advantages:

- Robust to overfitting on large datasets.
- Handles high-dimensional data well.
- Works well with non-linear decision boundaries.
- Provides feature importance ranking.

Parameters Used:

- Number of Estimators: 100 trees were used for better generalization.
- Random State: Ensures reproducibility of results.

CONVOLUTIONAL NN

CNN is a deep learning architecture designed to process visual inputs by automatically learning spatial hierarchies of features.

Architecture Highlights:

- Convolution Layers: Extract local features from the image using learnable filters.
- Max Pooling Layers: Reduce spatial dimensions and retain important information.
- Flatten Layer: Converts the 2D feature maps into a 1D vector for the dense layers.
- Dense (Fully Connected) Layers: Perform classification based on extracted features.

Final Layer:

- Uses Softmax Activation to output probabilities across 5 classes.

CONVOLUTIONAL NN

Training Setup:

- Optimizer: Adam – adaptive learning rate optimization.
- Loss: Categorical Crossentropy – suitable for multi-class classification.
- Data normalized to [0, 1] and split into training, validation, and test sets.

The categorical cross-entropy formula is expressed as:

$$L(y, \hat{y}) = -\sum_{i=1}^{C} y_i \log(\hat{y}_i)$$

RESULTS OF MODELS

Model	Type	Feature Input	Accuracy
SVM	Classical	HOG	83%
Fuzzy C-Means	Unsupervised	Raw/HOG	28%
Random Forest	Classical	Flattened pixels	82%
CNN	Deep Learning	Raw images	83%

CONCLUSION

- The project demonstrates the importance of image preprocessing in enhancing the visibility of key retinal features such as exudates, hemorrhages, and vessel structure.
- Among the models tested, SVM and CNN achieved the highest accuracy (83%), suggesting that both classical and deep learning methods can perform well when paired with high-quality preprocessing.
- Despite being limited to only 1,250 images due to resource constraints, the models provided a reasonable estimation of how well preprocessed images aid in classification.
- The results validate the potential of AI-based systems for early detection of diabetic retinopathy using retinal scans.

THANK YOU