Machine Learning Techniques For Diabetes Detection Using Iris And Conjunctival Images

PRESENTED BY:

MR. EHTESHAM SANA

Department of Computer Engineering,
Z.H.C.E.T, A.M.U.

SUPERVISOR:

DR. NADEEM AKHTAR

Department of Computer Engineering,
Z.H.C.E.T, A.M.U.

CO-SUPERVISOR:

DR. HAMID ASHRAF

Rajiv Gandhi Centre for Diabetes and
Endocrinology, J.N.M.C.H

OVERVIEW

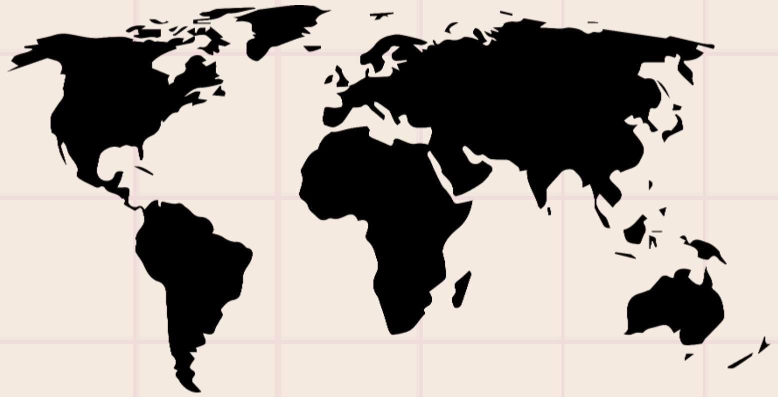
1. Introduction
2. Literature Review
3. Problem Statement
4. Proposed Solution
5. Experiments
   5.1 Experiment 1 – Utilizing Iris Features
   5.2 Experiment 2 – Utilizing Conjunctival Features
   5.3 Experiment 3 – Utilizing Iris-conjunctival Features
6. Conclusion & Future Work
7. References

1. Introduction

1. INTRODUCTION

537 million adults are living with diabetes.[1]

240 million people living with undiagnosed diabetes

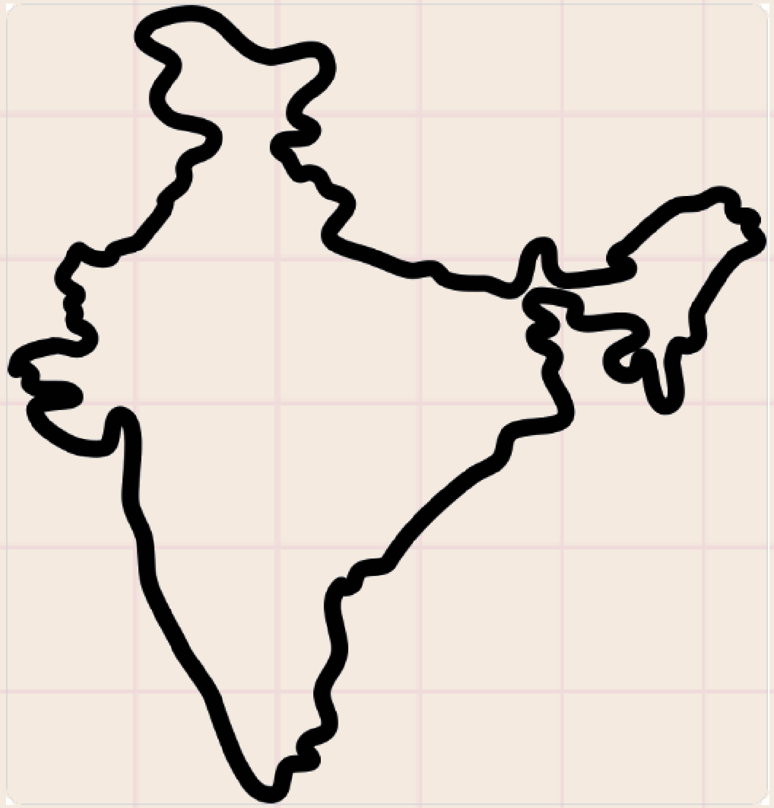


1. INTRODUCTION

- India is the '***diabetes capital of the world***', According to a study by the Indian Council of Medical Research. [2]

101 million diabetic patient

50% unaware of their diabetic status



1. INTRODUCTION

Methods For Diabetes Detection:

1. **Traditional Methods (Commercially Available)**
2. **Non-Traditional Methods (Under Research)**

1. INTRODUCTION

1. Traditional Methods For Diabetes Detection

Traditional Method	Details
1.Hemoglobin A1c Test (HbA1c):	• Reflects average blood sugar level for the past two to three months. • Diabetes: A1c level of 48 mmol/mol (6.5%) or higher on two separate tests. [3]
2. Fasting Plasma Glucose Test (FPG):	• Measures blood sugar after an overnight fast. • Diabetes: 126 mg/dL or higher on two separate tests. [4]
3.Oral Glucose Tolerance Test:	• Measures blood glucose before and 2 hours after consuming a glucose-containing beverage. • Diabetes: 200 mg/dL or higher after two hours. [5]
4. Random Plasma Glucose Test:	• Measures blood sugar without regard to last meal timing. • Diabetes: 200 mg/dL or higher. [6]

1. INTRODUCTION

1. Traditional Methods For Diabetes Detection

Traditional Method	Details
5. Glycosuria Test:	• Detects the presence of glucose in the urine. • Positive results typically require confirmation with more specific blood glucose tests. [7]
6. Semi-Invasive Glucose Monitoring Patches	• **Easy Tracking:** These patches use small sensors under the skin to measure sugar levels, offering immediate results. • **Better Health Choices:** They allow constant monitoring of sugar levels, aiding in better decision-making for meals, physical activities, and medicine use.

1. INTRODUCTION

2. Non-Traditional Methods For Diabetes Detection [8]

Method	*Definition*	*Requirement*
1. Breath Analysis:	Detecting higher acetone levels in diabetic patients' breath.	Under research
2. Saliva Testing:	Researching the correlation between saliva and blood glucose levels.	Under research
3. Tear Analysis:	Developing devices to monitor tear glucose levels.	Under research
4. Sweat Monitoring:	Track glucose levels continuously via sweat analysis using biosensors.	Under research
5. Spectroscopy:	Using light waves to measure glucose levels.	Under research
6. Iris & Conjunctival Analysis:	It involves the examination and processing of high-resolution images of the iris and conjunctiva to identify early microvascular changes indicative of diabetes.	Can use existing Hardware

1. INTRODUCTION

Why Iris & Conjunctival Analysis a top non-invasive method?

1. **Universality:** Eye provides a consistent sample unaffected by environment.
2. **Use of Existing Infrastructure:** Does not require specialized hardware.
3. **Broad Accessibility :** High-quality cameras, including those on smartphones, can be used for image capturing.
4. **Automated Analysis:** Machine learning can provides a fast, objective, and automated analysis of iris and conjunctival images.
5. **Patient Comfort:** Quick, painless, non-intrusive image capturing.
6. **Telemedicine Compatibility:** Ideal for remote screening, especially in underserved areas.

2. LITERATURE REVIEW

2. LITERATURE REVIEW

Iris Method:

Study	Year	Method Used	Dataset Description	Accuracy
Önal et al. [9]	2023	Deep learning	30 diabetic patients and 38 healthy	80%
Aminah et al. [10]	2019	Machine Learning	16 non-diabetic and 11 diabetic subjects	84.6%
Parsa et al. [11]	2018	Machine Learning	138 diabetic patients and 138 healthy	91.8%
Lesmana et al. [12]	2011	Deep learning	30 DM images and 20 healthy images	83.3%

2. LITERATURE REVIEW

Diving deeper into Iris Method:

1. **Iridology:** The study of iris patterns and colors for systemic health detection, including diabetes.

2. **Rubeosis Indication:** Iris imaging can reveal rubeosis, or neovascularization, an early sign of diabetes.

3. **Machine Learning:** Algorithms effectively spot minute iris changes, improving detection accuracy.

4. **Early, Non-invasive Diagnosis:** Through iris changes, iridology allows potential early diabetes detection in a painless, accessible way.

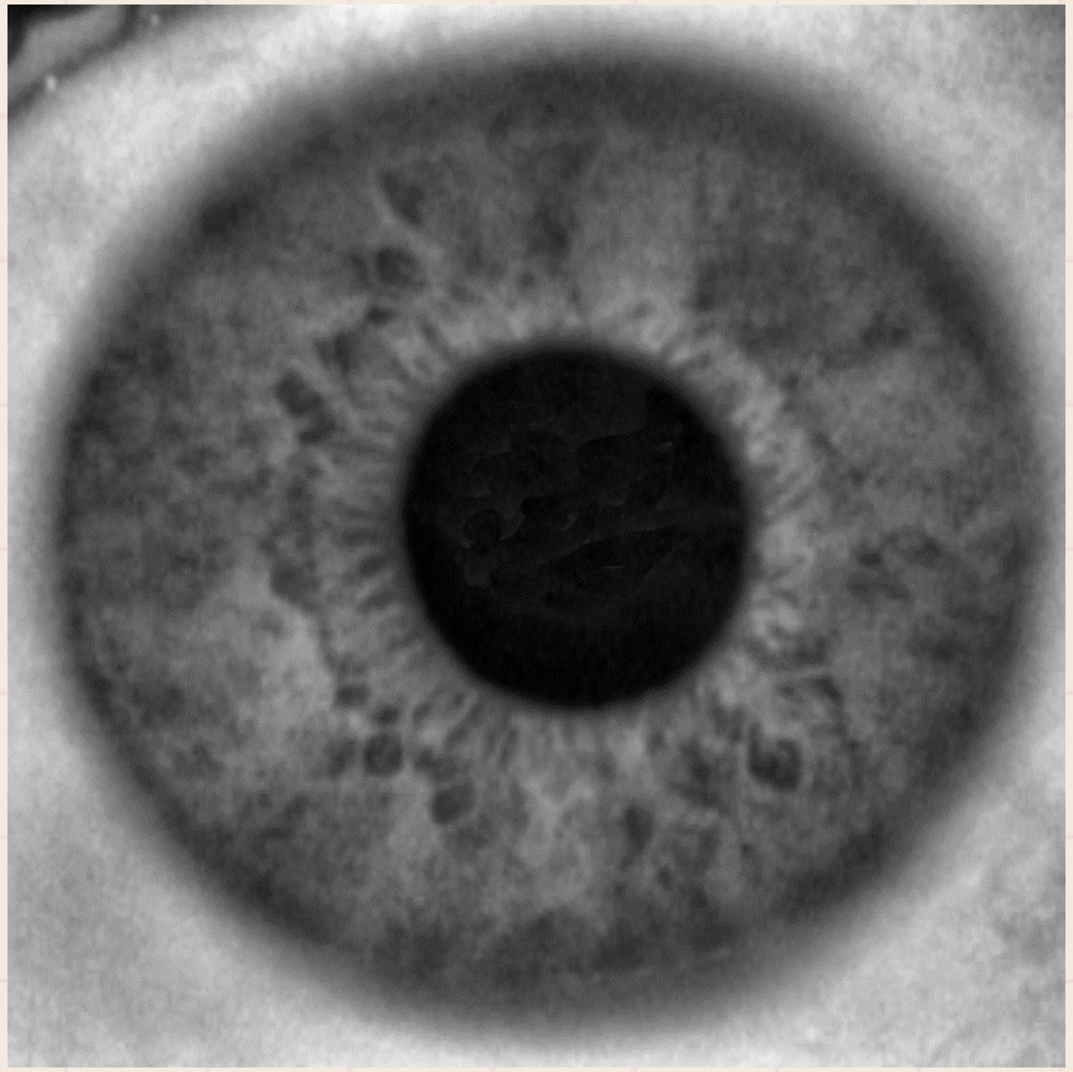


Fig 1: Iris Image

2. LITERATURE REVIEW

Conjunctival Method:

Study	Year	Key Findings
Zhang et al. [13]	2022	Employed deep learning method for diabetes detection using conjunctival images. The study achieved 75% accuracy with a dataset comprising 405 images from 68 diabetic patients and 206 from 62 healthy individuals.
Maziyar et al. [14]	2016	Demonstrated a correlation between various stages of diabetic retinopathy and conjunctival microvasculature images, providing insights into potential early detection methods.
Wilson et al. [15]	2011	Successfully established a correlation between conjunctival microangiopathy and retinopathy in Type-2 Diabetes Mellitus patients, highlighting the potential for early pathological detection in the retina via the bulbar conjunctiva.

2. LITERATURE REVIEW

Diving deeper into Conjunctival Method:

1. **Vascular Changes:** Conjunctival images capture diabetes-induced alterations in eye blood vessels.
2. **Diabetes Indicators:** Visible abnormalities in conjunctiva, such as microaneurysms and hemorrhages, suggest diabetes.
3. **Machine Learning:** Algorithms effectively identify subtle diabetes signs in conjunctival images.
4. **Early Detection:** Conjunctival imaging, aided by machine learning, identifies early, subtle vascular changes caused by elevated glucose levels, enabling potential early diabetes diagnosis.

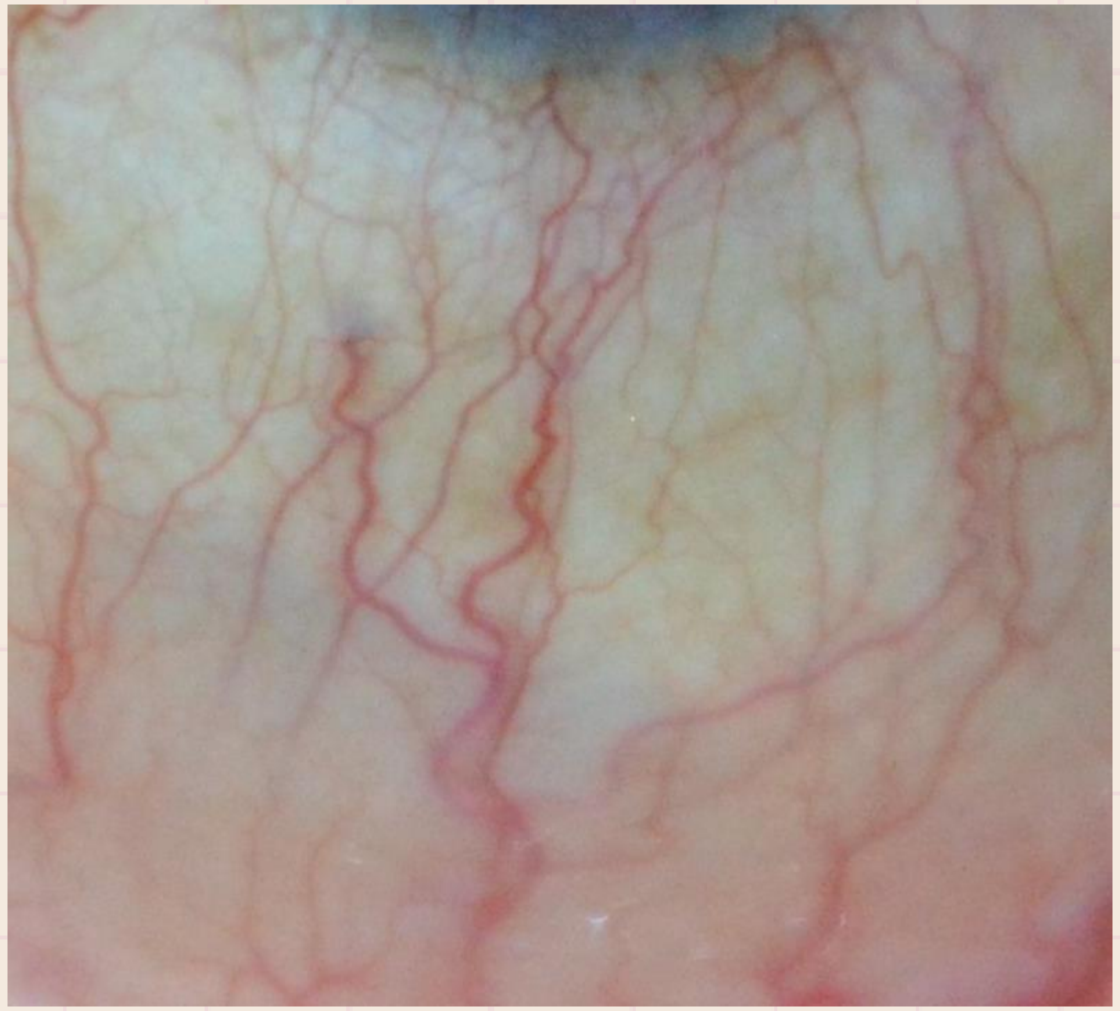


Fig 2: Conjunctival Image

3. PROBLEM STATEMNT

3. PROBLEM STATEMENT

1. Current gap: Despite medical advances, 50% of diabetes remains undetected.

2. Existing methods: Current tests are invasive or lacking in accuracy and early detection.

3. Under Utilization of Existing Infrastructure: Current non-invasive methods demand hardware which are under research.

4. ML vs. DL: Deep learning shows promise but needs extensive resources and lacks interpretability. Machine learning techniques provide computationally cheaper and interpretable alternative.

5. Dataset limitation: The lack of publicly available and geographically diverse (particularly for the Indian demographic) iris and conjunctival image datasets with diabetic and non-diabetic labels hampers the development of effective machine learning models.

6. Fusion of Eye Features: Machine learning's potential to merge iris and conjunctival image features for diabetes detection remains largely unexplored.

4. PROPOSED SOLUTION

4. PROPOSED SOLUTION

- **Non-Invasive Techniques**: Utilizing iris and conjunctival images for diabetes detection, offering a pain-free, risk-free approach.
- **Machine Learning**: Employing advanced algorithms to analyze the above images, enhancing accuracy in early diabetes detection.
- **Iris Images Experiment**: Dedicated analysis using machine learning on iris images, validating their significance in diabetes detection.
- **Conjunctival Images Study**: Exploration of under-utilized conjunctival images, connecting their distinct features with diabetes indicators.
- **Feature Fusion**: Combining the power of iris and conjunctival features for comprehensive, more accurate detection outcomes.

4. PROPOSED SOLUTION

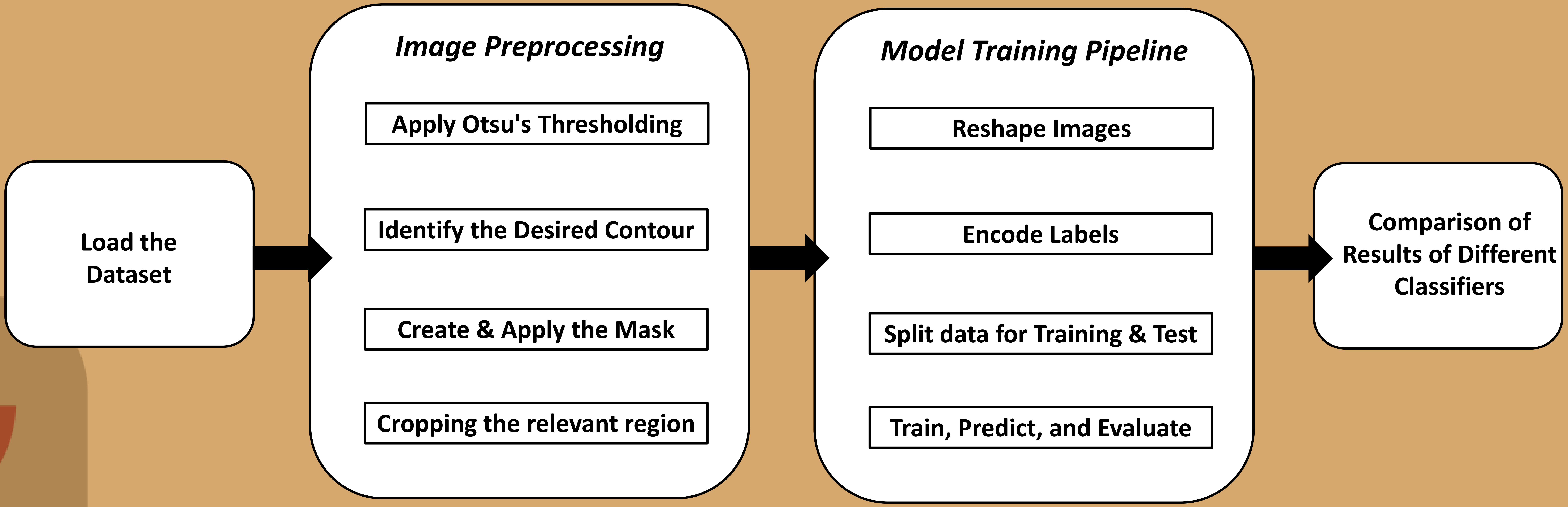


Fig 3: Different Stages of Proposed Methodology

4. PROPOSED SOLUTION

Bridging Data Gaps:

A. Iris Dataset

1. **Issue with Existing Data**: The only publicly existing dataset with iris images contained 236 images, inadequately labeled for diabetic and non-diabetic categories.

2. **Need for Comprehensive Data**: Due to the data scarcity, it was essential to gather a more extensive dataset for a thorough analysis.

3. **Collection of New Data**: Collected 800 new iris images, meticulously labeled as diabetic and non-diabetic.

4. **Formation of Larger Dataset**: This new collection led to the formation of a larger, more comprehensive dataset for enhanced investigation.

4. PROPOSED SOLUTION

Bridging Data Gaps:

B. Conjunctival Dataset

1. **Absence of Public Data**: No publicly accessible dataset was available for conjunctival images for diabetes research.

2. **Need for Specialized Data**: The absence called for the creation of a specific dataset for better exploration and understanding.

3. **Creation of New Data**: Gathered a new set of 572 conjunctival images, diligently categorized as diabetic and non-diabetic.

4. **Enriched Data Resource**: This initiative resulted in a valuable, tailor-made dataset for advanced study in the field of diabetes detection.

4. PROPOSED SOLUTION

Bridging Data Gaps:

MATERIAL AND METHODS

- This hospital-based study was conducted **on patients both with Diabetes and Non-diabetic history.**
- Valid **consent was taken from each of the patients in writing** after explaining the procedure prior to entering the study.

Period of Study

- 6 Months

Place of Study

- Rajiv Gandhi Centre for Diabetes and Endocrinology, J.N. Medical College and Hospital, A.M.U. Aligarh.

4. PROPOSED SOLUTION

- **Inclusion Criteria**
  - Those who have given their consent to participate in the study.
  - Those who have verifiable Diabetic or Non-Diabetic condition.

- **Exclusion Criteria**
  - Those who have not give their consent to participate in the study or suffer from Conjunctival lesions.

- **Procedure**
  - An Iris scanner was used to capture the Iris image and a DSLR camera was used for Conjunctival images of the Patients.

Fig. 4 Iris Scanner

Fig. 5 DSLR Camera

5. EXPERIMENTS

5. EXPERIMENTS

General Experimental Setting:

Purpose:

- The experiments aim to **evaluate the efficacy of machine learning algorithms** in detecting diabetes through the analysis of iris and conjunctival images.
- The goal is to understand **how effectively different models can differentiate between diabetic and non-diabetic individuals** based on these images.
- The experiments will also investigate the **impact of the size of the dataset on model performance** by comparing results from datasets of different sizes.

Platform & Tool: The experiments will be conducted on a Windows operating system.
Jupyter Notebook, a widely used tool in the data science and machine learning communities, will be the main tool for these experiments. It allows for:

- Efficient manipulation and preprocessing of datasets.
- Implementation of various machine learning models.
- Evaluation of the performance of these models.
- Presentation of results in an understandable and reproducible manner.

5. EXPERIMENTS

Experiment I: Diabetes Detection from Iris Images using Machine Learning

- This experiment leverages machine learning for diabetes detection using 276 iris images from an existing dataset and 800 images from a new JNMCH dataset.
- Evaluate the performance of machine learning models on the Parsa et al. [11] dataset.
- Assess the impact of a larger dataset on model accuracy.

Experiment II: Diabetes Detection from Conjunctival Images using Machine Learning

- This experiment extends the iris image analysis to 572 conjunctival images collected from JNMCH.
- Evaluate the performance of machine learning models on conjunctival images.
- Compare the results with those obtained from iris images.

Experiment III: Combined Iris and Conjunctival Image Analysis for Diabetes Detection using Machine Learning

- This experiment utilizes 572 images each from both iris and conjunctival datasets.
- Develop models utilizing combined features from both iris and conjunctival images.
- Compare the performance of these combined feature models with those trained on individual datasets.

5. EXPERIMENTS

Dataset Description

1. Parsa et al. [11] Iris Dataset
- **Diabetic Images:** 138
- **Non-Diabetic Images:** 138
- **Total Images:** 272

2. JNMCH Iris Dataset
- **Diabetic Images:** 400
- **Non-Diabetic Images:** 400
- **Total Images:** 800

3. JNMCH Conjunctival Dataset
- **Diabetic Images:** 286
- **Non-Diabetic Images:** 286
- **Total Images:** 572

DATA DESCRIPTION

2. JNMCH Iris Dataset	
3.1 Diabetic Data (Iris) • Mean: 43.81 • Median: 44.0 • Range: 47 • Min Age: 18 • Max Age: 65 • Variance (Sigma Squared): 140.71 • Standard Deviation: 11.86	3.2 Non-Diabetic Data (Iris) • Mean: 30.92 • Median: 26.0 • Range: 47 • Min Age: 18 • Max Age: 65 • Variance (Sigma Squared): 122.73 • Standard Deviation: 11.07

DATA DESCRIPTION

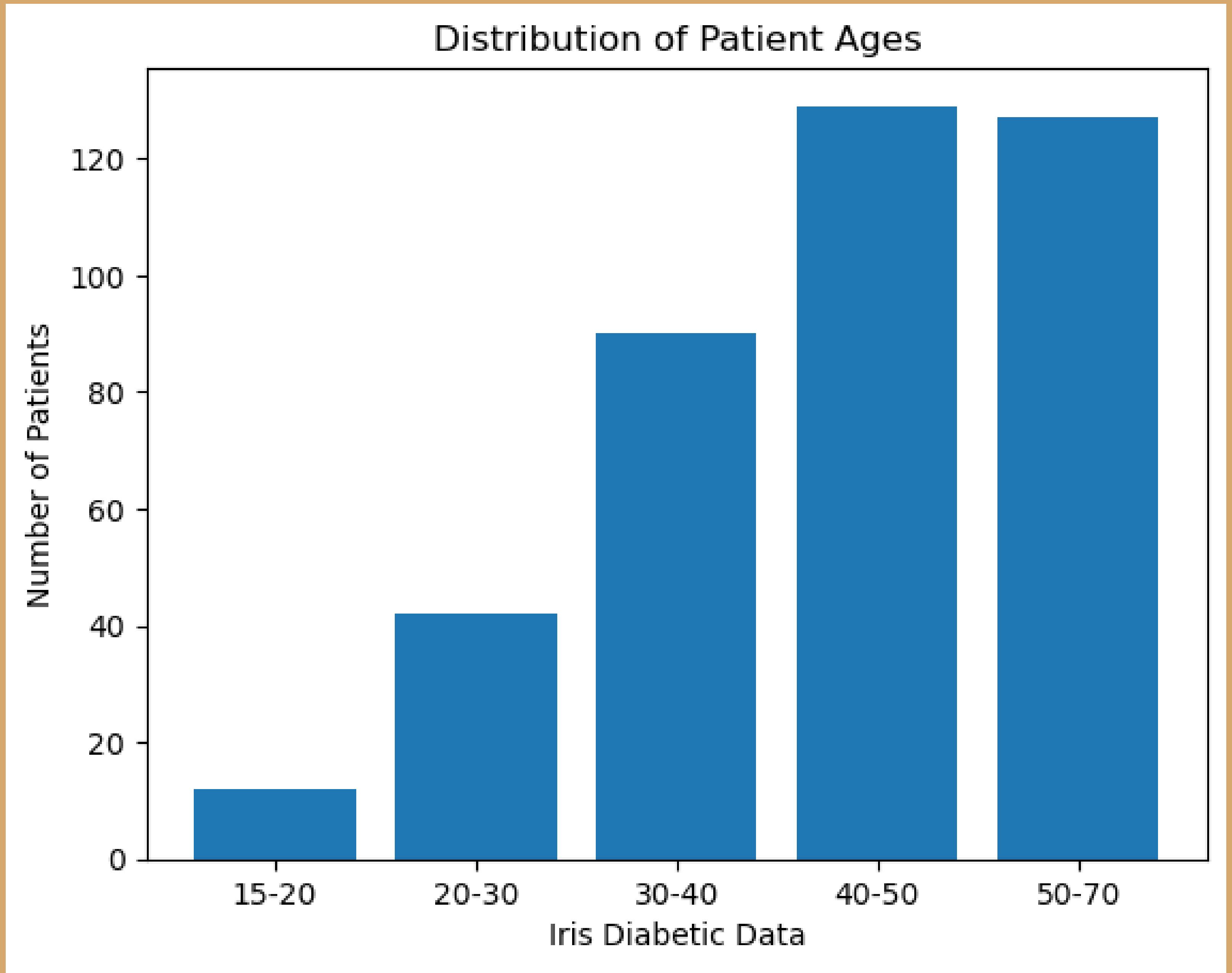


Fig 6: JNMCH Iris Data - Diabetic Patient's Age Distribution

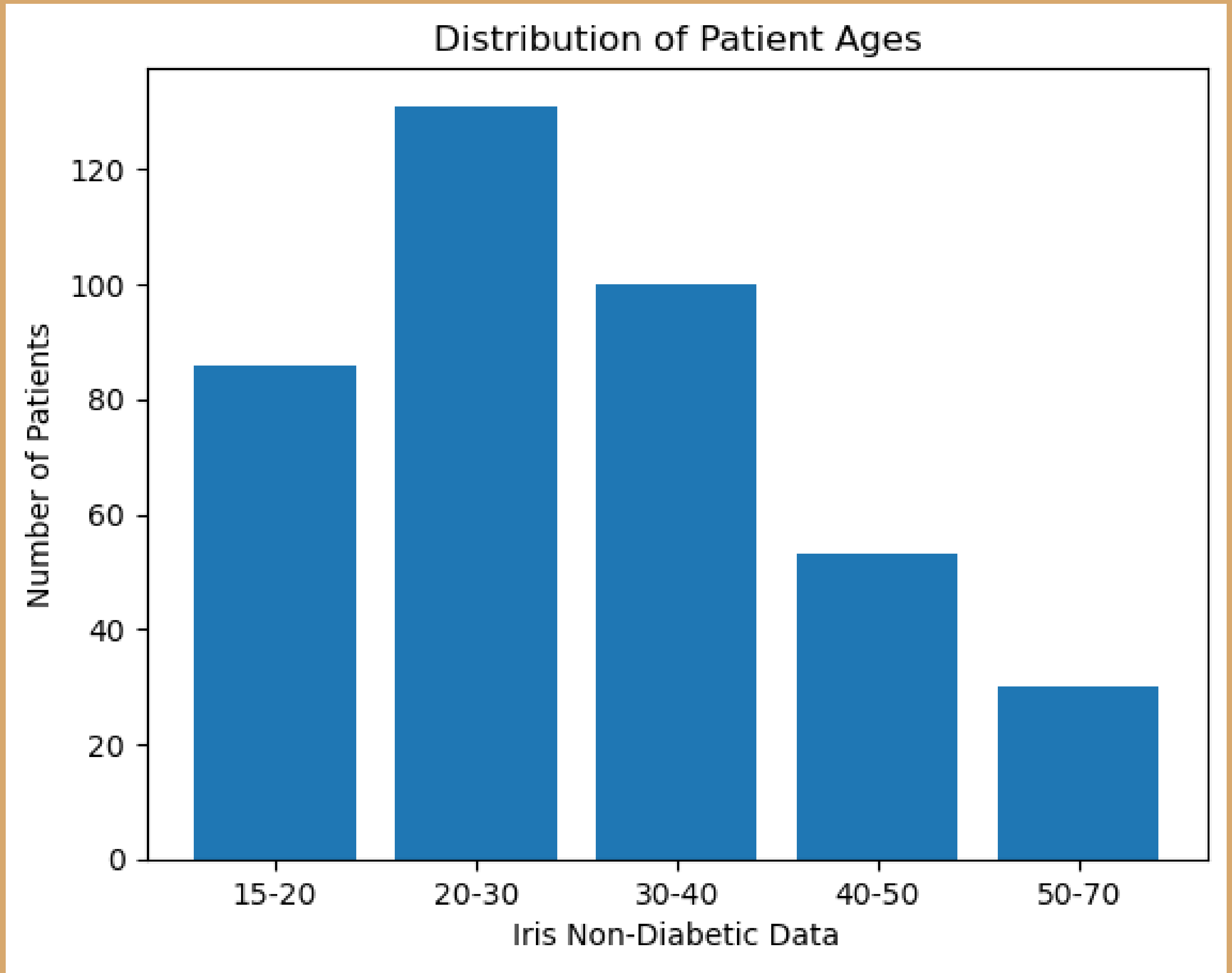


Fig 7: JNMCH Iris Data – Non-Diabetic Participant's Age Distribution

DATA DESCRIPTION

3. JNMCH Conjunctival Dataset	
3.1 Conjunctival Diabetic Data • Mean: 44.8 • Median: 45.0 • Range: 47 • Min Age: 18 • Max Age: 65 • Variance (Sigma Squared): 132.35 • Standard Deviation: 11.50	3.2 Conjunctival Non-Diabetic Data • Mean: 30.45 • Median: 25.0 • Range: 37 • Min Age: 18 • Max Age: 55 • Variance (Sigma Squared): 112.84 • Standard Deviation: 10.62

DATA DESCRIPTION

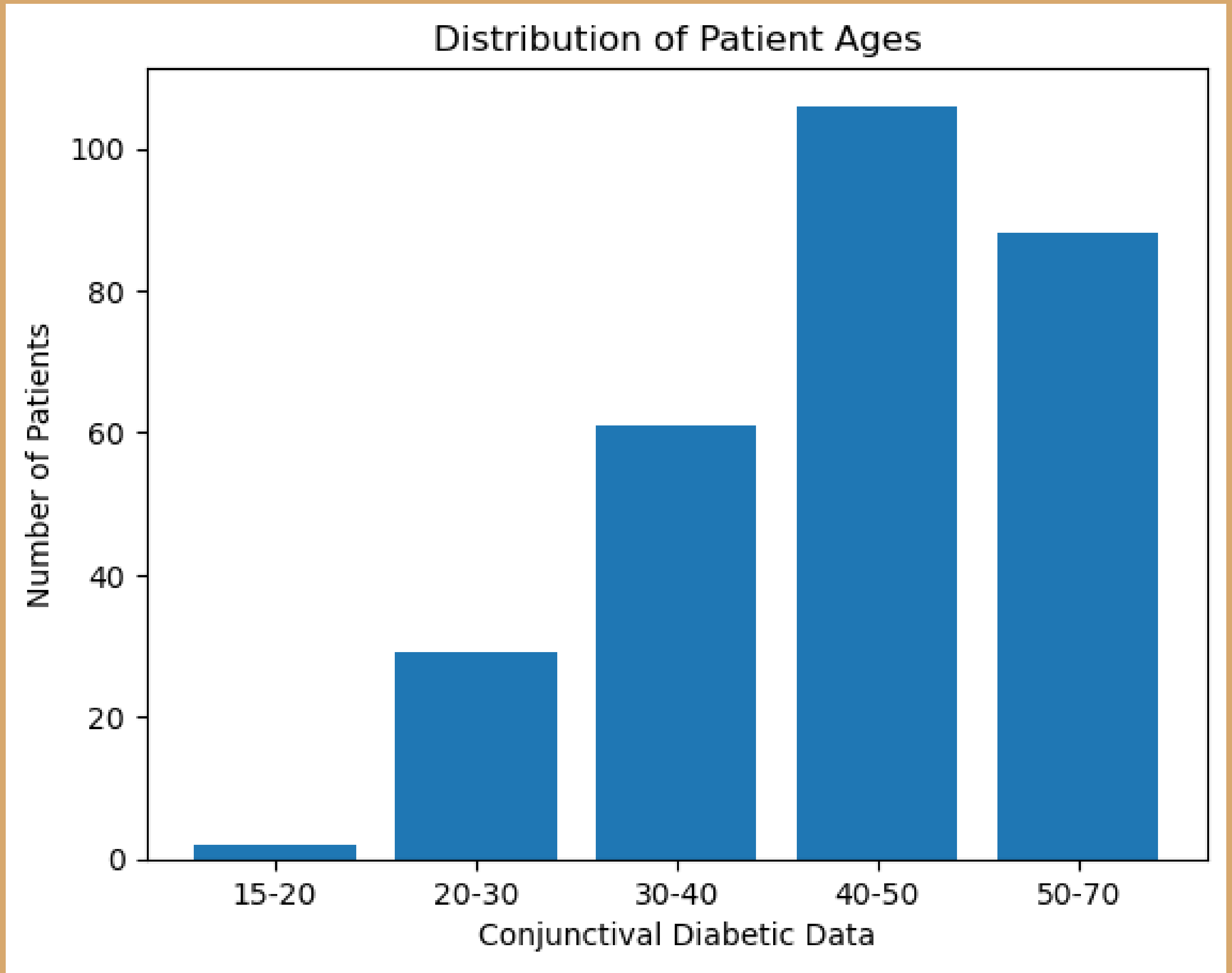


Fig 8: JNMCH Conjunctival Data - Diabetic Patient's Age Distribution

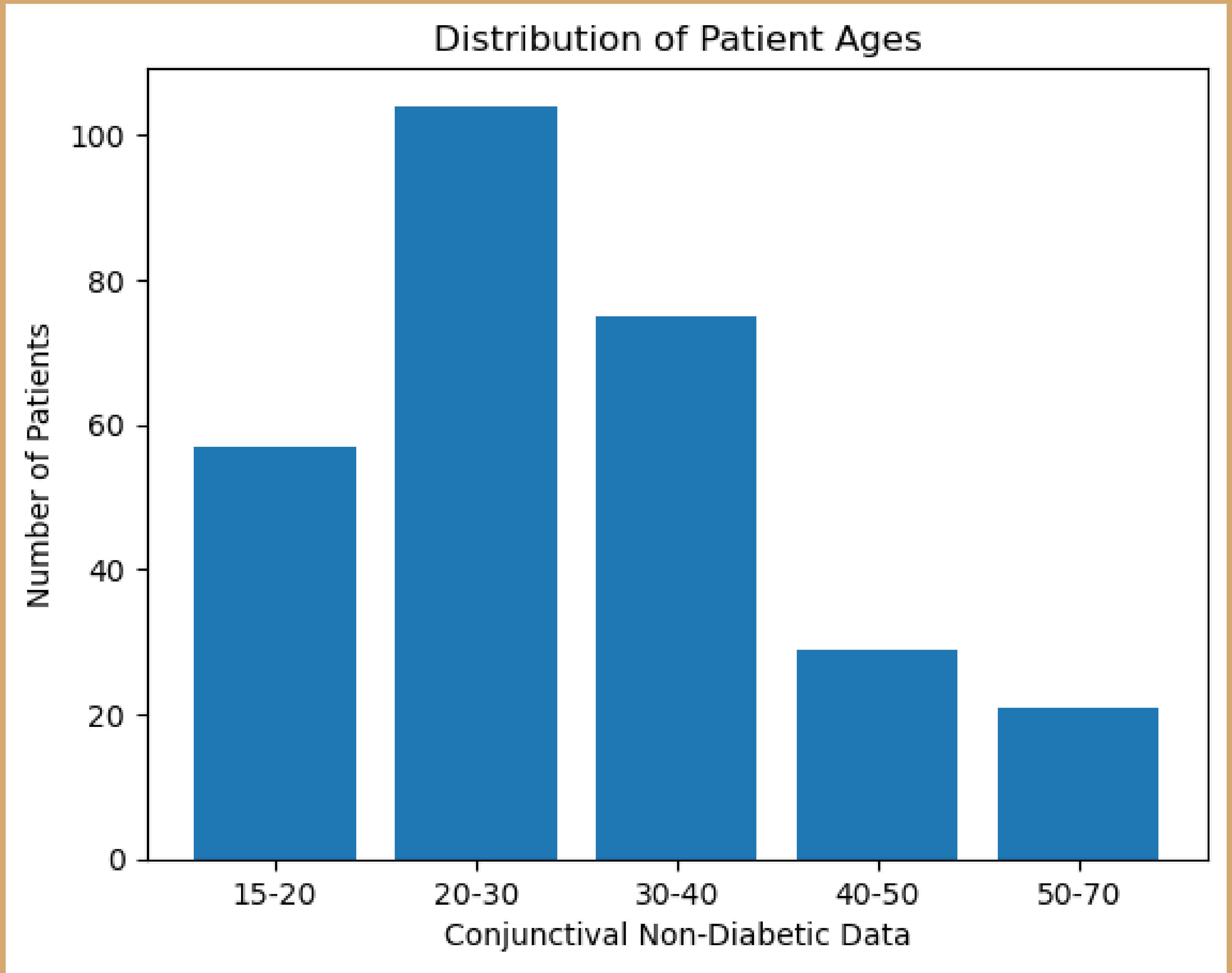


Fig 9: JNMCH Conjunctival Data – Non-Diabetic Participant's Age Distribution

EXPERIMENT – I

(IRIS DATASET)

EXPERIMENT – I

Experiment I: Diabetes Detection from Iris Images using Machine Learning

This experiment will explore the use of machine learning for the detection of diabetes using iris images. Two datasets will be used: an Parsa et. al [11] dataset of 276 images and a larger, newly collected dataset of 800 images from JNMCH.

Aims:

1. Evaluate and compare the performance of machine learning models for diabetes detection using the existing iris image dataset.
2. Investigate the impact of using a larger dataset on the model's performance by applying the same machine learning methods on the newly collected dataset.

EXPERIMENTAL SETUP

1. **Programming Language**: Python - leveraged for its robust libraries and compatibility with machine learning and image processing tasks.

2. **Platform**: Jupyter Notebook - chosen for its interactive coding environment, ideal for data analysis and modeling.

3. **Operating System**: Windows - served as the base system for deploying the Jupyter Notebook and running Python scripts.

4. **Image Processing Library**: OpenCV - used for reading, resizing, and preprocessing the input iris images.

5. **Input Data Specifications**: Images of dimension 150x150 pixels - chosen to keep the data manageable while maintaining sufficient detail for accurate model training.

6. **Machine Learning Algorithms**: Decision Trees, SVM, and KNN - deployed for their individual strengths in classification tasks, and their results compared to identify the optimal algorithm.

Hyperparameter Tuning

- Utilized the Grid Search method extensively during our machine learning model development process.
- Grid Search is a hyperparameter tuning technique.
- It exhaustively tries every combination of hyperparameter values for optimal model performance.

Presented below are the optimized hyperparameters for each model:

1. Decision Tree Classifier:

- Optimized max_depth: 100

2. K-Nearest Neighbors Classifier:

Used default settings

3. Support Vector Machine Classifier:

- Optimized C: 3.0
- Tried different kernels: 'linear', 'poly', 'rbf'

4. Random Forest Classifier:

- Optimized n_estimators: 100

5. Naive Bayes Classifier & Gradient Boosting Classifier:

- Used default settings, fewer hyperparameters available for tuning.

EXPERIMENT I – Pre Processing

1. Load the Image: We first load the image using the function in grayscale mode.

2. Apply Otsu's Thresholding: This step separates the iris (which we assume to be a darker region) from the rest of the image.

3. Morphological Operations: We perform a series of erosions and dilations on the thresholded image. These operations help to remove any small blobs of noise.

4. Find Contours: We find the contours in the thresholded image. A contour is a curve joining all the continuous points along a boundary that have the same color or intensity.

5. **Identify the Iris Contour**: We identify the largest contour as the iris.

6. **Create a Mask**: We create a mask (an array of the same size as the original image), and fill in the contour of the iris.

7. Apply the Mask to the Original Image: We apply the mask to the original image, using a bitwise-and operation. This results in an image with only the iris.

Cropped Image

Fig 10: Preprocessing

EXPERIMENT I – Pre Processing

Fig 11. Input Image

Fig 12. Otsu Threshold

Fig 13. Cropped Region

EXPERIMENT I – ML Classifier

8. Load the Cropped Image: We first load the cropped image.

9. Reshape Images: Convert the 3D matrix of each image dataset to a 2D matrix for machine learning model input.

10. Encode Labels: Transform categorical labels to numerical values using .

11. Setup k-Fold Cross-Validation: Implement a 10-fold cross-validation for model evaluation.

12. Train, Predict, and Evaluate with ML: In each fold, train the ML model, make predictions, and compute accuracy. Also generate classification report and confusion matrix for each fold.

13. Calculate Average Accuracy: After all folds, compute and print the average accuracy as the final model performance metric.

Fig 14: ML Classifier

EXPERIMENT I – RESULTS

Model	Accuracy
SVM (Poly)	95%
SVM (rbf)	92.5%
SVM (linear)	91%
Decision Tree	83.9%
KNN	84%
Random Forest	94.16%
Gradient Boosting	91%
Naïve Baiyes	81%

Table 1: JNMCH Iris Dataset

Model	Accuracy
SVM (Poly)	93%
SVM (rbf)	91%
SVM (linear)	90%
Decision Tree	83%
KNN	89%
Random Forest	94%
Gradient Boosting	91%
Naïve Baiyes	78%

Table 2: Parsa et al. [11] Iris Dataset

EXPERIMENT I – DISCUSSION

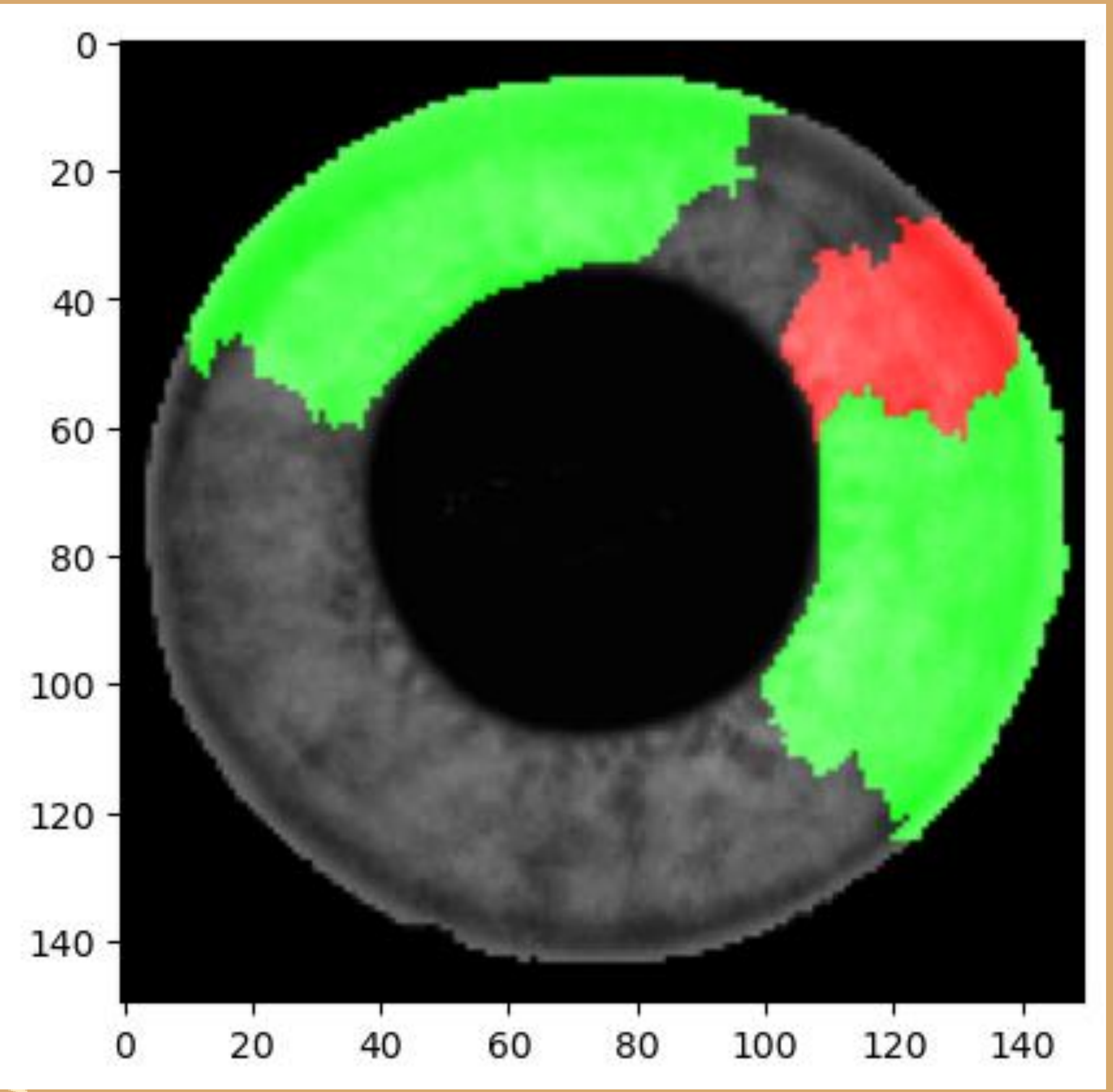


Fig 15: Region of Interest SVM (poly)

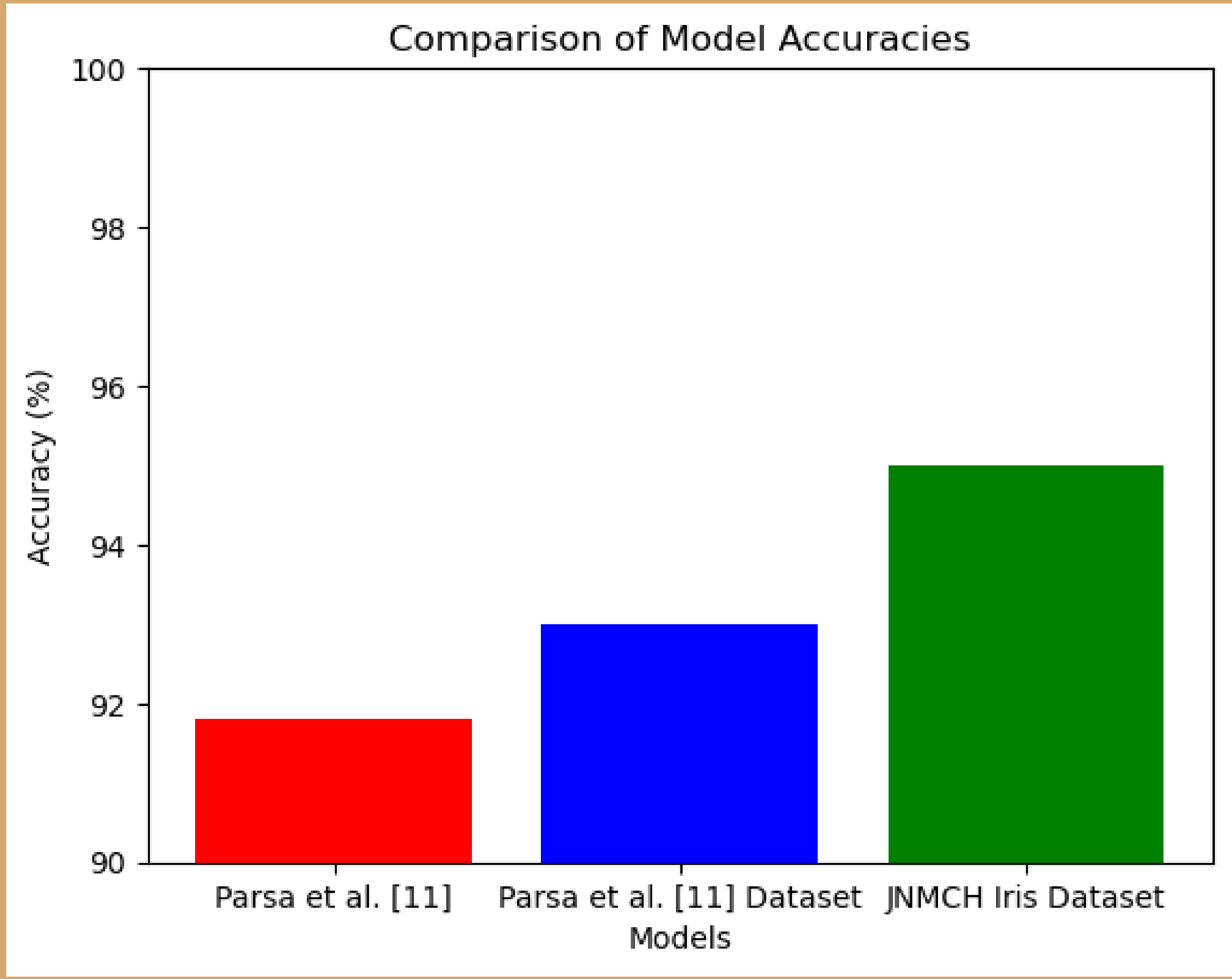


Fig 16: Experiment 1 Comparison Bar graph

EXPERIMENT– II

(CONJUNCTIVAL DATASET)

EXPERIMENT II

Experiment II: Diabetes Detection from Conjunctival Images using Machine Learning

- Building upon Experiment I, this experiment will expand the analysis to conjunctival images.
- A dataset of 572 conjunctival images (both diabetic and non-diabetic) collected from JNMCH will be used.

Aims:

1. Apply the machine learning techniques to conjunctival images for diabetes detection.
2. Evaluate the performance of the models on the conjunctival image dataset.

EXPERIMENTAL SETUP

1. **Programming Language**: Python - leveraged for its robust libraries and compatibility with machine learning and image processing tasks.

2. **Platform**: Jupyter Notebook - chosen for its interactive coding environment, ideal for data analysis and modeling.

3. **Operating System**: Windows - served as the base system for deploying the Jupyter Notebook and running Python scripts.

4. **Image Processing Library**: OpenCV - used for reading, resizing, and preprocessing the input conjunctival images.

5. **Input Data Specifications**: Images of dimension 150x150 pixels - chosen to keep the data manageable while maintaining sufficient detail for accurate model training.

6. **Machine Learning Algorithms**: Decision Trees, SVM, and KNN - deployed for their individual strengths in classification tasks, and their results compared to identify the optimal algorithm.

Hyperparameter Tuning

- Utilized the Grid Search method extensively during our machine learning model development process.
- Grid Search is a hyperparameter tuning technique.
- It exhaustively tries every combination of hyperparameter values for optimal model performance.

Presented below are the optimized hyperparameters for each model:

1. Decision Tree Classifier:

- Optimized max_depth: 100

2. K-Nearest Neighbors Classifier:

Used default settings

3. Support Vector Machine Classifier:

- Optimized C: 10.0
- Tried different kernels: 'linear', 'poly', 'rbf'

4. Random Forest Classifier:

- Optimized n_estimators: 100

5. Naive Bayes Classifier & Gradient Boosting Classifier:

- Used default settings, fewer hyperparameters available for tuning.

EXPERIMENT II – Pre Processing

1. Load the Image: We first load the image using the function in grayscale mode.

2. Apply Otsu's Thresholding: This step separates the iris (which we assume to be a darker region) from the rest of the image.

3. Morphological Operations: We perform a series of erosions and dilations on the thresholded image. These operations help to remove any small blobs of noise.

4. Find Contours: We find the contours in the thresholded image. A contour is a curve joining all the continuous points along a boundary that have the same color or intensity.

5. **Identify the Iris Contour**: We identify the largest contour as the iris.

6. **Create a Mask**: We create a mask (an array of the same size as the original image), and fill in the contour of the iris.

7. Apply the Mask to the Original Image: We apply the mask to the original image, using a bitwise-and operation. This results in an image with only the iris.

Cropped Image

Fig 17: Preprocessing

EXPERIMENT II – Pre Processing

Fig. 18 Input Image

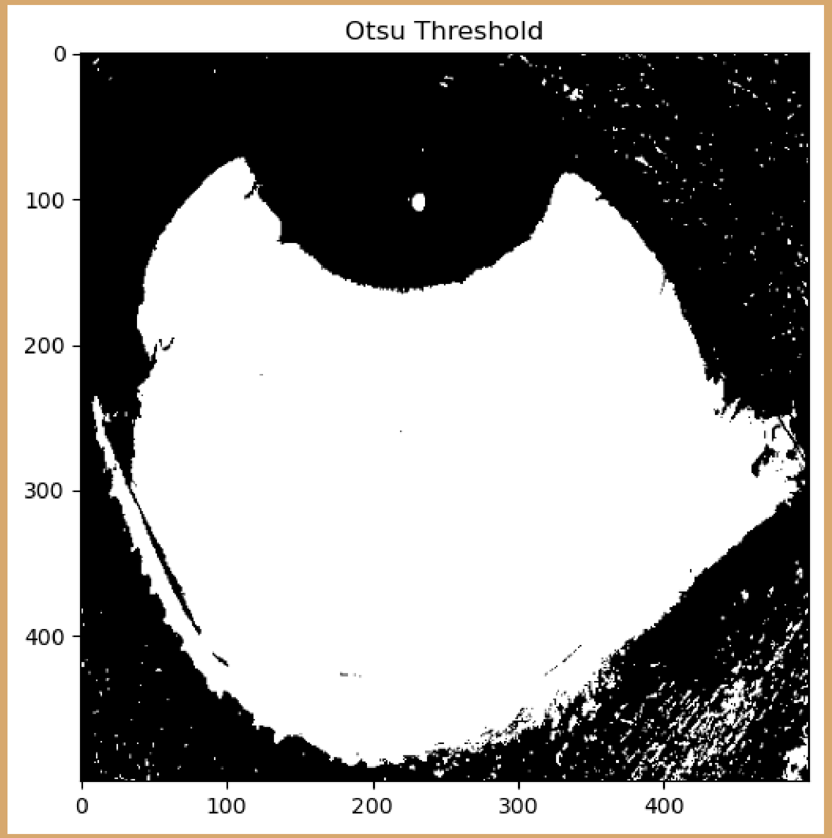


Fig 19. Otsu Threshold

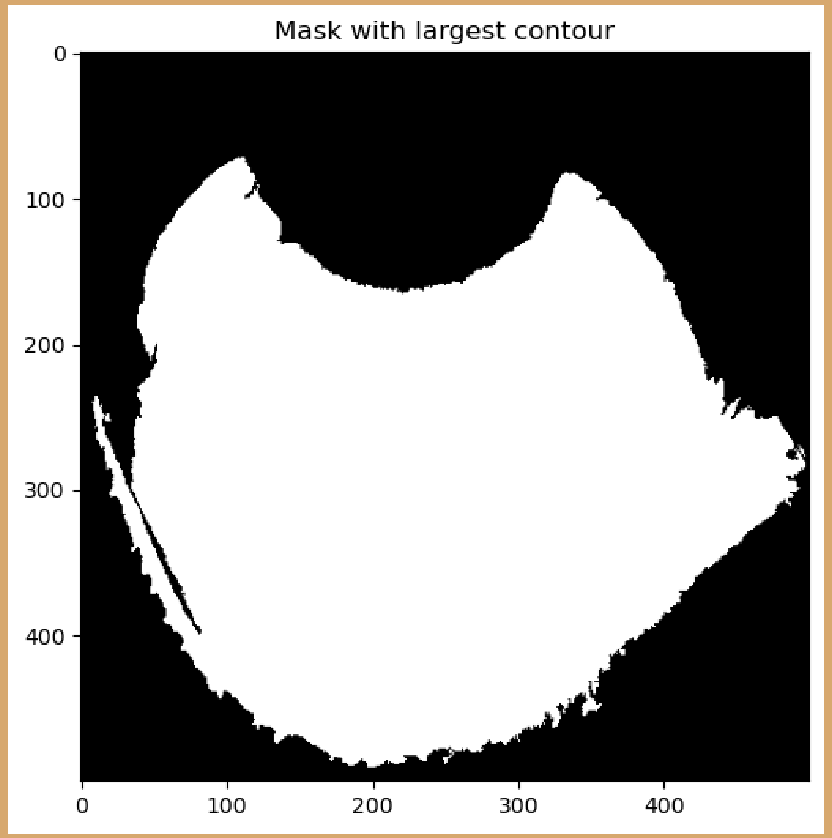


Fig. 20
Mask with largest Contour

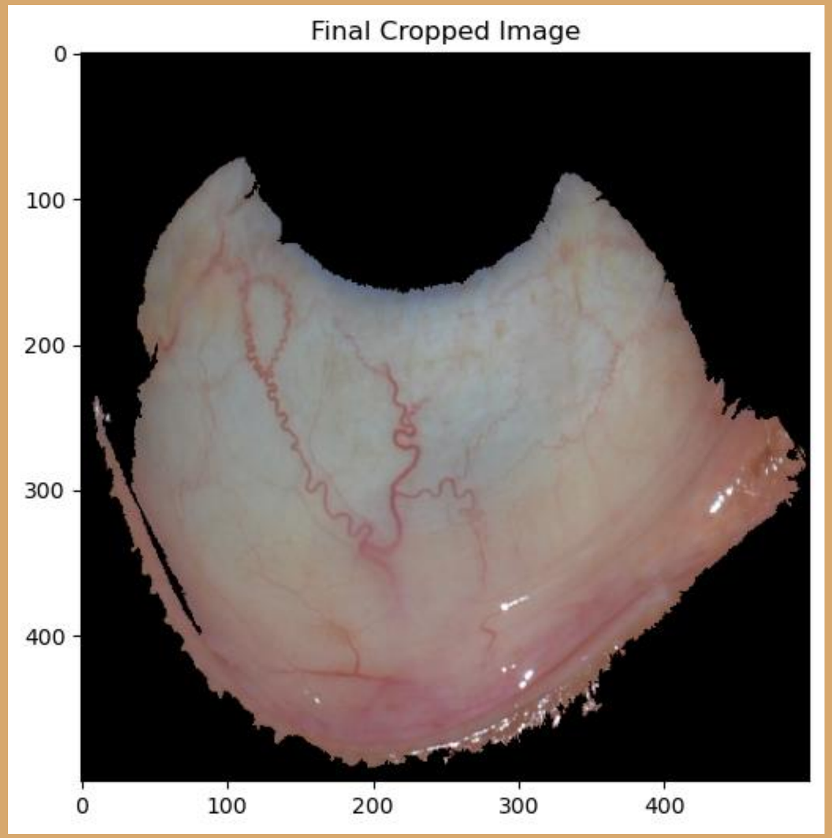


Fig. 21 Cropped Region

EXPERIMENT II – ML Classifier

8. Load the Cropped Image: We first load the cropped image.

9. Reshape Images: Convert the 3D matrix of each image dataset to a 2D matrix for machine learning model input.

10. Encode Labels: Transform categorical labels to numerical values using .

11. Setup k-Fold Cross-Validation: Implement a 10-fold cross-validation for model evaluation.

12. Train, Predict, and Evaluate with ML: In each fold, train the ML model, make predictions, and compute accuracy. Also generate classification report and confusion matrix for each fold.

13. Calculate Average Accuracy: After all folds, compute and print the average accuracy as the final model performance metric.

Fig 22: ML Classifier

EXPERIMENT II – RESULTS

Model	Accuracy
SVM (Poly)	93.3%
SVM (rbf)	93 %
SVM (linear)	90%
Decision Tree	84%
KNN	85%
Random Forest	93 %
Gradient Boosting	93%
Naïve Baiyes	79%

Table 3: JNMCH Conjunctival Dataset Model Results

EXPERIMENT II – DISCUSSION

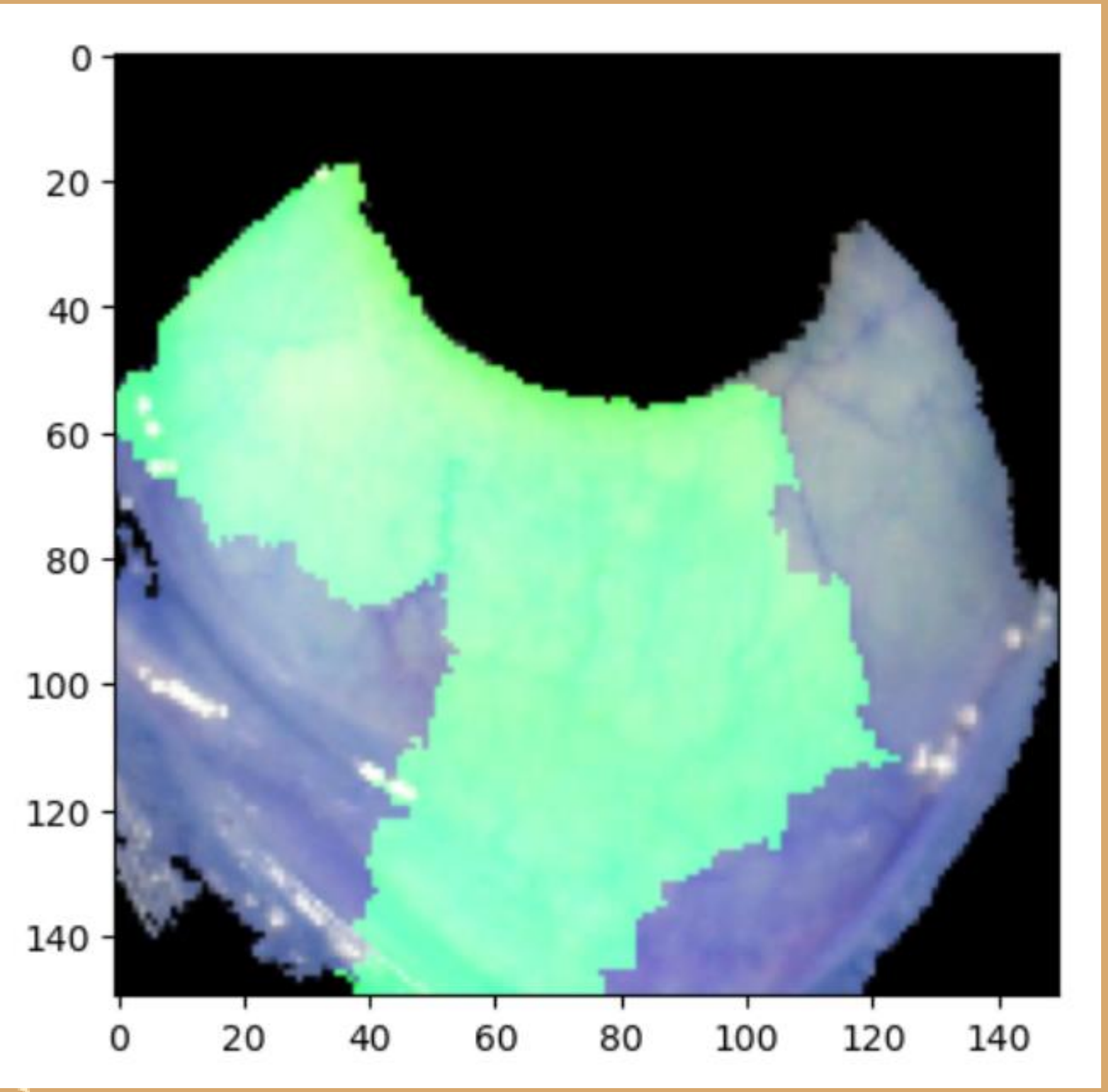


Fig 23: Region of Interest SVM (poly)

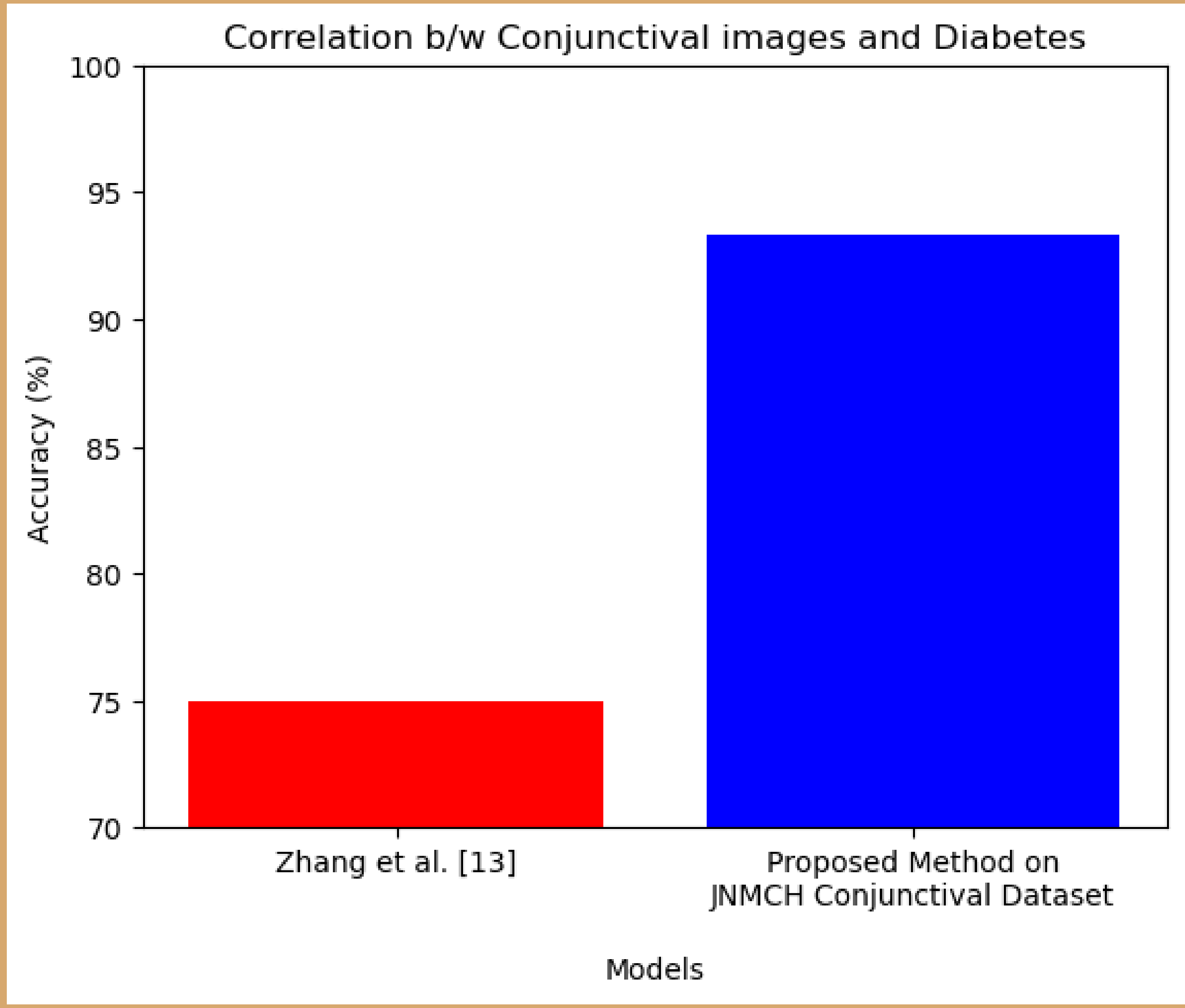


Fig 24: Experiment II Comparison Bar graph

EXPERIMENT III

(IRIS+CONJ DATASET)

EXPERIMENT III

Experiment III: Combined Iris and Conjunctival Image Analysis for Diabetes Detection using Machine Learning

This experiment aims to integrate the features from both iris and conjunctival images to improve the detection of diabetes. Equal-sized datasets of 572 images each from both iris and conjunctival datasets will be used.

Aims:

1. Develop method that can effectively use combined features from both types of images to enhance the accuracy of diabetes detection.
2. Compare the performance of these models with those trained on individual datasets to determine the benefits of the integrated approach.

EXPERIMENTAL SETUP

1. **Programming Language**: Python - leveraged for its robust libraries and compatibility with machine learning and image processing tasks.

2. **Platform**: Jupyter Notebook - chosen for its interactive coding environment, ideal for data analysis and modeling.

3. **Operating System**: Windows - served as the base system for deploying the Jupyter Notebook and running Python scripts.

4. **Image Processing Library**: OpenCV - used for reading, resizing, and preprocessing the input conjunctival images.

5. **Input Data Specifications**: Images of dimension 150x150 pixels - chosen to keep the data manageable while maintaining sufficient detail for accurate model training.

6. **Machine Learning Algorithms**: Decision Trees, SVM, and KNN - deployed for their individual strengths in classification tasks, and their results compared to identify the optimal algorithm.

Hyperparameter Tuning

- Utilized the Grid Search method extensively during our machine learning model development process.
- Grid Search is a hyperparameter tuning technique.
- It exhaustively tries every combination of hyperparameter values for optimal model performance.

Presented below are the optimized hyperparameters for each model:

1. Decision Tree Classifier:

- Optimized max_depth: 100

2. K-Nearest Neighbors Classifier:

Used default settings

3. Support Vector Machine Classifier:

- Optimized C: 3.0
- Tried different kernels: 'linear', 'poly', 'rbf'

4. Random Forest Classifier:

- Optimized n_estimators: 100

5. Naive Bayes Classifier & Gradient Boosting Classifier:

- Used default settings, fewer hyperparameters available for tuning.

EXPERIMENT III – Pre Processing

1. Load the Image: We first load the image using the function in grayscale mode.

↓

2. Apply Otsu's Thresholding: This step separates the iris (which we assume to be a darker region) from the rest of the image.

↓

3. Morphological Operations: We perform a series of erosions and dilations on the thresholded image. These operations help to remove any small blobs of noise.

↓

4. Find Contours: We find the contours in the thresholded image. A contour is a curve joining all the continuous points along a boundary that have the same color or intensity.

↓

5. **Identify the Iris Contour**: We identify the largest contour as the relevant region.

↓

6. **Create a Mask**: We create a mask (an array of the same size as the original image), and fill in the contour of the iris.

↓

7. Apply the Mask to the Original Image: We apply the mask to the original image, using a bitwise-and operation. This results in an image with only the relevant region

↓

Cropped Image

Fig 25: Preprocessing

EXPERIMENT III – ML Classifier

1. Load the Cropped Image: We first load the cropped image.

2. Reshape Images: Convert the 3D matrix of each image dataset to a 2D matrix for machine learning model input.

3. Encode Labels: Transform categorical labels to numerical values using .

4. Combine the Features: Concatenate the features of both image datasets into a single dataset.

5. Setup k-Fold Cross-Validation: Implement a 10-fold cross-validation for model evaluation.

6. Train, Predict, and Evaluate with ML: In each fold, train the ML model, make predictions, and compute accuracy. Also generate classification report and confusion matrix for each fold.

7. Calculate Average Accuracy: After all folds, compute and print the average accuracy as the final model performance metric.

Fig 26: ML Classifier

EXPERIMENT III – RESULTS

ML Model	Accuracy
SVM (Poly)	96.78%
SVM (rbf)	92%
SVM (linear)	90%
Decision Tree	80%
KNN	91%
Random Forest	88%
Gradient Boosting	93%
Naïve Baiyes	86%

Table 4: ML Model Results

DL Model	Accuracy
ResNet101	88%
ResNet50	90%
DenseNet121	86%
MobileNet	90%
XceptionNet	87%
VGG16	89%
VGG19	84%

Table 5: DL Model Results

EXPERIMENT III – DISCUSSION

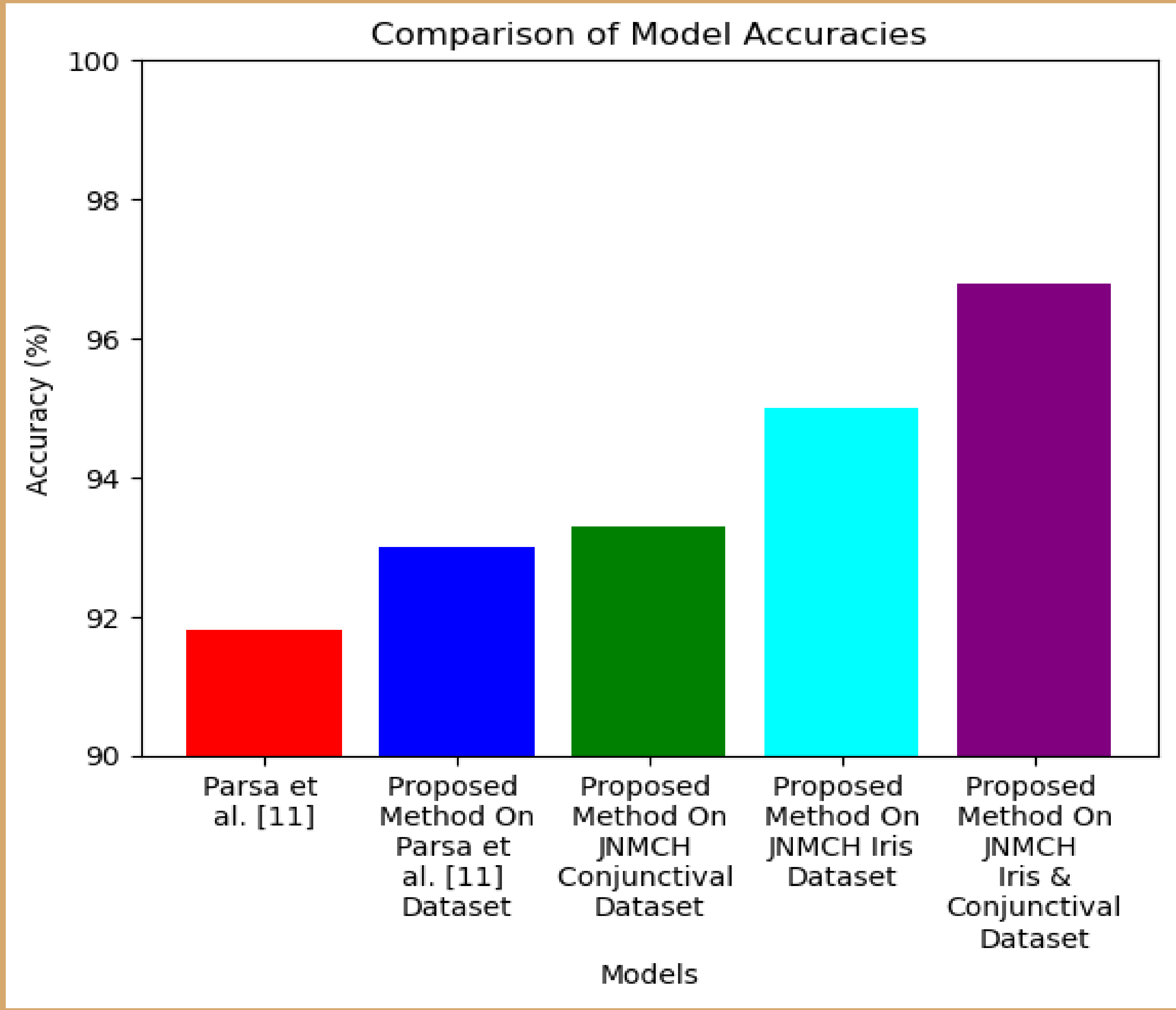


Fig. 27: Experiment III Comparison Bar graph

6. CONCLUSION

CONCLUSION

Key Takeaways from Experiment I: Detecting Diabetes from Iris Images using Machine Learning

Outstanding Model Accuracy on JNMCH Iris Dataset:

- SVM (Poly) model excelled with 95% accuracy.
- High Performance across Multiple Models: SVM (RBF), SVM (Linear), Gradient Boosting, and Random Forest all achieved over 90% accuracy.

Benchmarking on Previous Data:

- Utilized the Parsa et. al [11] 276-image dataset for benchmarking.
- Attained a top accuracy of 93%, outdoing the previous record of 91.8% mentioned in Parsa et al.'s study.

Affirmation of Improved Model Performance:

- Highlighted machine learning's crucial contribution to early, precise diabetes detection via iris images.
- The successful experiments and comprehensive data gathering signify substantial progress in non-invasive diabetes detection.

CONCLUSION

Key Insights from Experiment II: Detecting Diabetes from Conjunctival Images using Machine Learning

- SVM (Poly) model excelled with 93% accuracy.

The Value of Conjunctival Images:
- Demonstrated the effectiveness of conjunctival images for non-invasive diabetes detection.
- High accuracy outcomes validate the potential of conjunctival images to improve detection methods.

Consistent Model Performance:
- Verified SVM (Poly) model's robustness and versatility across both iris and conjunctival images.
- Consistent high accuracies support the use of machine learning in diabetes detection.

Filling the Research Void:
- One of the few studies using conjunctival images, providing new insights and filling a significant research gap.
- Encourages further exploration and usage of conjunctival images in the medical field.
- The significant results expand the scope for non-invasive diabetes detection.

CONCLUSION

Key Insights from Experiment III: Detecting Diabetes by Combining Iris and Conjunctival Features

1. **Excellence of Feature Fusion**:
- The combined iris and conjunctival feature dataset scored an outstanding accuracy of 96.78% using the SVM (Poly) model, thus establishing the efficiency of a blended approach in diabetic detection.

2. **Strong Performance Across Different Models**:
- Apart from SVM (Poly), multiple models like SVM (Linear), SVM (RBF), KNN, and Gradient Boosting also showcased strong results with accuracy scores over 90%.
- This indicates the robustness and versatility of the combined feature set.

3. **Effective Utilization of Pre-Trained Models**:
- The employment of pre-trained models like ResNet50, ResNet101, DenseNet, MobileNet, XceptionNet, and VGG16 provided a combined average accuracy of 88%.
- This underlines the potential of transfer learning methods in enhancing the detection of diabetes.

CONCLUSION

Final Conclusion

1. **Harnessing Non-invasive Techniques:** The study innovatively used non-invasive iris and conjunctival imaging, marking a pioneering step towards early diabetes detection, thereby confronting the global issue of latent diagnoses.

2. **Leveraging Existing Infrastructure:** By using existing medical imaging equipment, the study negated the need for additional hardware or invasive procedures, reshaping the landscape of diabetes detection techniques.

3. **Machine Learning: A Reliable Ally:** The research endorsed machine learning techniques as not just accurate, but also interpretative and efficient, thereby presenting a promising alternative to the often resource-intensive deep learning methods.

4. **Overcoming Data Scarcity:** The research amassed a diverse collection of 800 iris and 572 conjunctival images to overcome the data availability challenges, a contribution that offers significant potential for future research, especially within the Indian demographic context.

CONCLUSION

5. **Experimental Success:** A series of experiments highlighted the potent efficacy of machine learning models, with remarkable accuracy rates of 95%, 93%, and 96.78% for iris, conjunctival, and combined features respectively.

6. **Correlation & Model Efficacy:** The research demonstrated a strong correlation between iris and conjunctival image features and diabetes detection. Significant number of models exceeding the 90% accuracy threshold, the study underscored the credibility of machine learning in early diabetes detection.

7. **Revolutionizing Early Detection:** The convergence of machine learning and non-invasive imaging in this study sets a precedent for highly efficient, accurate early detection systems, thus changing the outlook for future diabetes screening tools.

8. **Implications for Future Research:** Ultimately, this study serves as a robust foundation for future exploration in this field, paving the way for advancements that could transform prognosis and patient experiences globally.

FUTURE WORK

- **Advanced Models Exploration**: Utilize more sophisticated machine learning and deep learning models to potentially enhance results.
- **Multi-Disease Framework**: Expand the developed framework to detect multiple diseases, including other ocular and systemic conditions.
- **Clinical Integration**: Investigate practical implications and benefits of embedding these models in clinical practice for early detection of diabetes.
- **Expansive Diverse Datasets**: Collect larger and more diverse datasets, capturing data from various regions, ages, and disease stages to boost model's universality.
- **Real-time Detection Systems**: Develop real-time disease detection systems using portable devices, moving towards point-of-care diagnostics.
- These future avenues will further bolster the substantial groundwork laid by this research, potentially transforming the domain of non-invasive diabetes detection.

REFERENCES

1. https://idf.org/about-diabetes/facts-figures/ (Retrieved on: 20 Dec 2022)
2. https://economictimes.indiatimes.com/magazines/panache/india-has-over-100-mn-diabetics-136-mn-pre-diabetics-says-new-icmr-study-goa-tops-the-list-up-records-lowest-prevalence/articleshow/100866686.cms?from=mdr (Retrieved on: 20 Dec 2022)
3. https://www.niddk.nih.gov/health-information/diagnostic-tests/a1c-test (Retrieved on: 20 Dec 2022)
4. https://www.cdc.gov/diabetes/basics/getting-tested.html#:~:text=Fasting%20Blood%20Sugar%20Test,higher%20indicates%20you%20have%20diabetes. (Retrieved on: 20 Dec 2022)
5. https://www.ncbi.nlm.nih.gov/books/NBK532915/#:~:text=The%20results%20of%20the%20OGTT,200%20mg%2FdL%20indicates%20diabetes (Retrieved on: 20 Dec 2022)
6. https://www.cdc.gov/diabetes/basics/getting-tested.html#:~:text=Random%20Blood%20Sugar%20Test,higher%20indicates%20you%20have%20diabetes.&text=*Results%20for%20gestational%20diabetes%20can%20differ. (Retrieved on: 25 July 2023)
7. https://medlineplus.gov/lab-tests/glucose-in-urine-test/ (Retrieved on: 25 July 2023)

REFERENCES

8. S. Lekha and S. M, "Recent Advancements and Future Prospects on E-Nose Sensors Technology and Machine Learning Approaches for Non-Invasive Diabetes Diagnosis: A Review," in IEEE Reviews in Biomedical Engineering, vol. 14, pp. 127-138, 2021, doi: 10.1109/RBME.2020.2993591.

9. Önal, M.N., Güraksin, G.E. & Duman, R. Convolutional neural network-based diabetes diagnostic system via iridology technique. Multimed Tools Appl 82, 173–194 (2023). https://doi.org/10.1007/s11042-022-13291-3

10. R. Aminah and A. H. Saputro, "Diabetes Prediction System Based on Iridology Using Machine Learning," 2019 6th International Conference on Information Technology, Computer and Electrical Engineering (ICITACEE), Semarang, Indonesia, 2019, pp. 1-6, doi: 10.1109/ICITACEE.2019.8904125.

11. P. Moradi, N. Nazer, A. K. Ahmadi, H. Mohammadzade and H. K. Jafari, "Discovering Informative Regions in Iris Images to Predict Diabetes," 2018 25th National and 3rd International Iranian Conference on Biomedical Engineering (ICBME), Qom, Iran, 2018, pp. 1-6, doi: 10.1109/ICBME.2018.8703564.

12. I. P. D. Lesmana, I. K. E. Purnama and M. H. Purnomo, "Abnormal condition detection of pancreatic Beta-cells as the cause of Diabetes Mellitus based on iris image," 2011 2nd International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering, Bandung, Indonesia, 2011, pp. 150-155, doi: 10.1109/ICICI-BME.2011.6108614.

REFERENCES

13. ***Li, X., Xia, C., Li, X. et al. Identifying diabetes from conjunctival images using a novel hierarchical multi-task network. Sci Rep 12, 264 (2022). https://doi.org/10.1038/s41598-021-04006-z***

14. *Khansari, Maziyar M., et al. "Automated fine structure image analysis method for discrimination of diabetic retinopathy stage using conjunctival microvasculature images." Biomedical optics express 7.7 (2016): 2597-2606.*

15. *To, Wilson J., et al. "Correlation of conjunctival microangiopathy with retinopathy in type-2 diabetes mellitus (T2DM) patients." Clinical hemorheology and microcirculation 47.2 (2011): 131-141.*

Thank You