DIABETIC RETINOPATHIC DETECTION USING CNN

A PROJECT REPORT

21CSC305P - MACHINE LEARNING

(2021 Regulation)

III Year/ V Semester

Academic Year: 2024 -2025

Submitted by

DHARSAN L [RA2211003010232]

ALEN GEORGE [RA2211003010235]

MUKIL M S [RA2211003010252]

ANUVIND A [RA2211003010266]

Under the Guidance of

Dr. Anbazhagu U V

Assistant Professor, Department of Computing Technologies

in partial fulfillment of the requirementsfor the degree of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE ENGINEERING

SCHOOL OF COMPUTING

COLLEGE OF ENGINEERING AND TECHNOLOGY

SRM INSTITUTE OF SCIENCE ANDTECHNOLOGY

KATTANKULATHUR- 603 203

NOVEMBER 2024

SRM

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

KATTANKULATHUR – 603 203

BONAFIDE CERTIFICATE

Certified that **21CSC305P - MACHINE LEARNING project** report titled **"DIABETIC RETINOPATHIC DETECTION"** is the Bonafide work of **"DHARSAN L [RA2211003010232], ALEN GEORGE [RA2211003010235], MUKIL M S [RA2211003010252], ANUVIND A [RA2211003010266]"** who carried out the task of completing the project within the allotted time.

12/4/24

SIGNATURE	**SIGNATURE**
Dr. Anbazhagu U V	Dr. G. Niranjana
Course Faculty	**Head of the Department**
Assistant Professor	Professor
Department of Computing Technologies	Department of Computing Technologies
SRM Institute of Science and Technology	SRM Institute of Science and Technology
Kattankulathur	Kattankulathur

Department of Computing Technologies Kattankulathur 603 203 SRM IST

ABSTRACT

Diabetic retinopathy is a leading cause of vision impairment and blindness worldwide, particularly among individuals with diabetes. Early detection and timely intervention are essential to prevent disease progression and irreversible damage. This project aims to develop a machine learning-based system for automated detection and classification of diabetic retinopathy stages using retinal images. By leveraging advanced image processing techniques and deep learning algorithms, the model is trained to identify various stages of diabetic retinopathy with high accuracy, sensitivity, and specificity. The proposed system provides a scalable and efficient tool for assisting healthcare professionals in diagnosing diabetic retinopathy, ultimately contributing to improved patient outcomes and reduced healthcare burdens. This solution demonstrates the potential of artificial intelligence in enhancing diagnostic capabilities and accessibility in ophthalmology, especially in regions with limited access to specialized care.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABREVIATIONS

Acronym	Description
AI	Artificial Intelligence
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
AUC	Area Under Curve
BNN	Bayesian Neural Network
CV	Computer Vision
EM	Expectation maximization
GPU	Graphics Processing Unit
TN	True Negative
TP	True Positive
KNN	K-Means Nearest Neighbor

CHAPTER 1
INTRODUCTION

Diabetic Retinopathy (DR) is a major cause of blindness, particularly in individuals with diabetes. Early detection is crucial to prevent vision loss, but it requires expert analysis of retinal images, which can be time-consuming and complex. As DR progresses without noticeable symptoms in the early stages, many cases go undiagnosed until significant damage occurs.

Machine learning, especially deep learning models like Convolutional Neural Networks (CNNs), has emerged as a promising solution for automating DR detection. By training on large datasets of retinal images, these models can accurately identify signs of DR, such as microaneurysms, hemorrhages, and exudates. This approach reduces the need for manual intervention and speeds up the diagnostic process, making it more accessible in regions with limited access to specialists.

However, challenges remain, including data imbalance, which can lead to poor model performance on rare cases, and the need for model interpretability to gain trust in clinical settings. Additionally, ensuring that models generalize well across different datasets is another hurdle. Despite these challenges, machine learning holds significant potential to improve early DR detection, enhancing the prevention of vision loss and supporting more efficient screening methods.

1.1 Background on Diabetic Retinopathy

Diabetic Retinopathy (DR) is a serious complication of diabetes that impacts the blood vessels in the retina, the light-sensitive tissue at the back of the eye responsible for transmitting visual information to the brain. Prolonged high blood sugar levels in diabetic patients can damage these retinal blood vessels, causing them to leak fluid or bleed, leading to various forms of retinal damage, including swelling, scar tissue formation, and, in severe cases, retinal detachment. DR progresses through several stages, from mild non-proliferative abnormalities to more advanced proliferative retinopathy, where abnormal blood vessels form and pose a significant risk of vision loss.

Early-stage DR is typically asymptomatic, which means patients may not notice any changes in vision until the disease has significantly progressed. This makes regular screening essential, particularly for individuals with diabetes. Conventional diagnosis relies on trained ophthalmologists analyzing high-resolution retinal images, such as those from fundus photography, to identify signs of DR. However, in many regions, especially in low-resource settings, there is a shortage of eye care specialists, which limits access to timely diagnosis. This shortage highlights the need for automated and accessible screening solutions that can support early diagnosis and treatment, reducing the risk of blindness in diabetic populations.

1.2 Role of Machine Learning in DR Detection

Machine learning, and more specifically deep learning, offers a promising solution for automating the detection and classification of diabetic retinopathy. Deep learning models, particularly Convolutional Neural Networks (CNNs), have been highly effective in image analysis tasks because they can automatically learn complex patterns and features from large datasets. In the context of DR detection, CNNs are trained on vast datasets of retinal images, allowing the models to recognize subtle visual markers, such as microaneurysms, hemorrhages, exudates, and neovascularization, which are characteristic of different stages of DR. With enough high-quality labeled data, these models can achieve diagnostic accuracy comparable to human experts.

The potential of machine learning for DR detection is significant, as it can make screening more accessible and scalable, even in areas with limited medical resources. Despite the promise, there are several challenges to address. For instance, data imbalance can occur if certain stages of DR are underrepresented in the dataset, leading to a model that may struggle with rare cases. Furthermore, model interpretability is crucial in healthcare, as clinicians need to understand and trust the predictions made by machine learning models. Another key challenge is generalization—ensuring the model performs well across different populations, camera types, and imaging conditions.

Ongoing research is focused on overcoming these challenges by refining models, improving dataset diversity, and incorporating techniques like attention mechanisms to improve interpretability. By addressing these issues, machine learning has the potential to greatly enhance diabetic retinopathy detection, supporting earlier interventions and ultimately reducing the burden of vision impairment related to diabetes.

1.3 Software Requirement Specification

Category	Details
Operating System	Windows 10/11 or Linux (Ubuntu 18.04 or higher)
Programming Language	Python 3.8 or higher
Libraries/Frameworks	- TensorFlow 2.x (for deep learning models)
	- Keras (high-level API for TensorFlow)
	- OpenCV (image preprocessing/augmentation)
	- NumPy (numerical operations)
	- Pandas (data handling)
	- Matplotlib (visualization)
	- Scikit-learn (evaluation metrics)
Development Environment	Visual Studio Code (VSCode) or PyCharm; Jupyter Notebook for testing
Version Control	Git (version management)
Additional Tools	Docker (containerization)

Table 1: Software Requirements Table

CHAPTER 2

LITERATURE SURVEY

1. An Advanced Mask R-CNN Network for Cattle Individual Recognition

Jianping Wang

School of Computer Science and Technology, Henan Institute of Science and Technology, Xinxiang, China

Xueyan Zhang

School of Computer Science and Technology, Henan Institute of Science and Technology, Xinxiang, China

Guohong Gao

School of Computer Science and Technology, Henan Institute of Science and Technology, Xinxiang, China

Yingying Lv

School of Computer Science and Technology, Henan Institute of Science and Technology, Xinxiang, China

This paper presents OP Mask R-CNN, a novel network designed for the individual identification of cattle to enhance large-scale farm management and promote digital advancements in animal husbandry. The proposed method merges Open Pose with the Mask R-CNN framework and incorporates three main strategies to improve identification accuracy. First, the authors optimize the Mask R-CNN's backbone, ResNet101, by refining its convolutional layers. Second, they introduce a bovine skeleton feature extraction method using Open Pose to enhance data representation. Finally, they develop a fusion mechanism that integrates the attention module (CBAM), Open Pose, and ResNet101 for improved performance. This approach balances precision and computational efficiency, offering a lightweight solution for effective cattle identification in real-time monitoring systems.

2. **Food Image Recognition and Calorie Prediction Using Faster R-CNN and Mask R-CNN**

Elena Denny Cherpanath

dept. Electronics and Communication Engineering (KTU), Rajagiri School of Engineering and Technology (KTU), Ernakulam, India

P R Fathima Nasreen

dept. Electronics and Communication Engineering (KTU), Rajagiri School of Engineering and Technology (KTU), Ernakulam, India

Karthika Pradeep

dept. Electronics and Communication Engineering (KTU), Rajagiri School of Engineering and Technology (KTU), Ernakulam, India

Malavika Menon

dept. Electronics and Communication Engineering (KTU), Rajagiri School of Engineering and Technology (KTU), Ernakulam, India

V S Jayanthi

dept. Electronics and Communication Engineering (KTU), Rajagiri School of Engineering and Technology (KTU), Ernakulam, India

In response to the rising health concerns associated with obesity and overeating, there is increasing public awareness around maintaining proper dietary habits to avoid conditions like hypertension, diabetes, and heart disease. The World Health Organization (WHO) reports that over 2.8 million deaths annually are related to being overweight or obese. To help individuals monitor and manage their dietary intake, a project has been proposed that leverages deep learning techniques for calorie estimation through image analysis. This approach involves a structured, layer-based process: capturing images of food items, classifying the food types, and predicting their calorie content. This innovative use of deep learning could empower users to make informed decisions about their diet and contribute to better health management.

3. Visible Image Recognition of Power Transformer Equipment Based on Mask R-CNN

Anfeng Jiang
State Grid Shanghai Energy, Interconnection Research Institute, Shanghai
Nannan Yan
State Grid Shanghai Energy, Interconnection Research Institute, Shanghai
Fa Wang
State Grid Shanghai Energy, Interconnection Research Institute, Shanghai
Hao Huang
State Grid Shanghai Energy, Interconnection Research Institute, Shanghai, China
Huangru Zhu
State Grid Shanghai Energy, Interconnection Research Institute, Shanghai
Bengang Wei
State Grid Shanghai Energy, Interconnection Research Institute, Shanghai

This paper addresses the importance of visible image recognition of power substation equipment for intelligent operation, fault location, and infrared defect diagnosis. Previous image recognition methods have faced challenges such as limited detail capture, susceptibility to environmental interference, and low recognition accuracy. To overcome these issues, the paper introduces a visible image recognition model for substation equipment based on Mask R-CNN and compares it to a model using Faster R-CNN. Both models are trained on 7,000 labeled images of power substation equipment, and their performance is evaluated using 2,000 additional images, focusing on the recognition of 11 common substation components, including transformers, current transformers, and bushings.

The study reveals that the Mask R-CNN-based model can achieve pixel-level recognition, making it highly effective for precise identification of substation equipment. However, the Faster R-CNN model, while lacking pixel-level detail, performs better in terms of speed and accuracy for target box detection. Despite these findings, both models require further optimization to enhance their effectiveness in real-world power substation environments.

4. Impact of Computer Vision with Deep Learning Approach in Medical Imaging Diagnosis

Charleen

Computer Science, Binus University, Jakarta, Indonesia

Cheryl Angelica

Computer Science, Binus University, Jakarta, Indonesia

Hendrik Purnama

Computer Science, Binus University, Jakarta, Indonesia

Fredy Purnomo

Computer Science, Binus University, Jakarta, Indonesia

This research investigates the impact and challenges associated with the use of computer vision in medical imaging and its potential effects on patient safety. Medical experts traditionally interpret medical images, but their evaluations can be influenced by subjectivity and the inherent complexity of the images. The purpose of this research is to determine whether the integration of computer vision in healthcare could negatively impact patients and to assess the challenges faced during its implementation. A systematic literature review on the application of computer vision in medical imaging is conducted to explore how deep learning algorithms compare to healthcare professionals in disease classification. The findings indicate that deep learning approaches in computer vision can significantly assist physicians by enhancing diagnostic accuracy. These technologies have proven to be safe and effective for aiding medical professionals in interpreting medical images. The research concludes that, despite the challenges in implementing computer vision in healthcare, it serves as a reliable tool that supports doctors in improving diagnostic precision without posing harm to patients.

5. Research on Water Conservancy Project Monitoring and Intelligent Identification Method Based on Computer Vision Technology

Jin Zhu

Zhejiang Province Water Conservancy and Hydropower Technology Consultation Center, Hangzhou, Zhejiang, China

Feng Yang

Zhejiang Zhongshui Engineering Technology Co., Ltd., Hangzhou, China

Monitoring and maintenance of water resources projects is essential to ensure project safety and environmental protection. Traditional monitoring methods often rely on manual inspections and sensor data, but these methods suffer from high cost, low efficiency and limited monitoring range. With the rapid development of computer vision technology, its potential for application in water conservancy engineering has gradually emerged. The purpose of this paper is to study the water conservancy project monitoring and intelligent identification methods based on computer vision technology. We first outline the basic concepts of computer vision and its application scenarios in water conservancy engineering, including structural health monitoring, environmental monitoring and disaster management. Then, the specific applications of image processing techniques, machine learning and deep learning in intelligent recognition are described in detail, especially how these techniques can be utilized to achieve an efficient and accurate real-time monitoring system. Through several case studies, this paper demonstrates the effects and advantages of the application of computer vision technology in actual water conservancy projects. Finally, it summarizes the main findings of the current study and looks forward to the future technological development and application prospects, pointing out the direction of further research. The findings of this paper show that computer vision-based monitoring and intelligent identification methods not only improve the accuracy and efficiency of monitoring, but also provide new technical means for the management and maintenance of water conservancy projects.

6. Traffic Sign Recognition for Computer Vision Project-Based Learning

David Gerónimo

Computer Vision Center and the Department of Computer Science, Universitat Autònoma de Barcelona, Bellaterra, Spain

Joan Serrat

Computer Vision Center and the Department of Computer Science, Universitat Autònoma de Barcelona, Bellaterra, Spain

Antonio M. López

Computer Vision Center and the Department of Computer Science, Universitat Autònoma de Barcelona, Bellaterra, Spain

Ramon Baldrich

Computer Vision Center and the Department of Computer Science, Universitat Autònoma de Barcelona, Bellaterra, Spain

This paper outlines a graduate course project focused on computer vision, specifically the detection and recognition of traffic signs in video sequences captured by an on-board vehicle camera. Traffic sign recognition poses significant challenges for driving assistance systems, making it an ambitious and engaging problem for students to tackle. The project allows students to confront the complexities of real-world computer vision tasks and evaluate the effectiveness of contemporary vision and pattern recognition methods introduced during the course.

The paper details the learning objectives of the course and discusses the design constraints, including the varied backgrounds of students and the time commitment required from both students and instructors. The course content, schedule, and implementation of project-based learning are described to illustrate how theoretical knowledge is applied practically. The outcomes of the course are analyzed, highlighting student performance and feedback. This comprehensive approach helps students gain hands-on experience, better understand the challenges of computer vision applications, and appreciate the potential and limitations of the techniques covered in the curriculum.

7. X-Ray Chest Image Classification by A Small-Sized Convolutional Neural Network

Ege Kesim
Department of Computer Engineering, Koc University, Istanbul, Turkey
Zumray Dokur
Department of Electronics and Communication Engineering, Istanbul Technical University, Istanbul, Turkey
Tamer Olmez
Department of Electronics and Communication Engineering, Istanbul Technical University, Istanbul, Turkey

This study examines the application of Convolutional Neural Networks (CNNs) for classifying X-ray chest images, addressing common challenges associated with using pre-trained, large-scale deep learning models. While these models deliver strong performance on training data, they face practical limitations such as reduced generalization when exposed to different imaging modalities and unsuitability for real-time applications due to their large size.

To overcome these issues, the research explores new network architectures and optimizes input image sizes for efficient classification. The proposed network successfully classifies chest images into twelve distinct classes with an accuracy of approximately 86%. The lightweight structure of the network enables rapid training and ensures it is viable for real-time use. The system's performance was tested on an embedded platform equipped with a camera, and it was observed that classification results were generated in under one second, demonstrating the model's practicality for real-time medical image analysis.

8. **Image classification based on image pixel value**

Mukesh Chandra Arya
GBPEC Pauri Garhwal, India
Ashish Semwal
Computer Science & Engineering Department, GBPEC Pauri Garhwal, India

This paper focuses on image classification based on the analysis of pixel values. By examining pixel intensity, it is possible to determine whether two images are identical. The proposed method aims to classify images into different categories based on pixel intensity levels. The central approach emphasizes comparing pixel intensities and using these comparisons to classify images into various groups, where each class represents images with similar pixel intensity levels.

The method leverages pixel-level analysis to enhance image classification accuracy and streamline the categorization process. By grouping images with matching pixel intensity distributions, this approach provides a straightforward yet effective way to classify images. The results demonstrate that this pixel-intensity-based method can accurately and efficiently sort images into distinct classes, highlighting its potential for applications where precise image comparison and classification are required.

9. **Projected pattern on three-dimensional objects for image feature classification and recognition**

Gridsada Phanomchoeng
Department of Mechanical Engineering, Chulalongkorn University, Bangkok, Thailand
Ratchatin Chanchareon
Department of Mechanical Engineering, Chulalongkorn University, Bangkok, Thailand

This paper introduces a novel technique aimed at enhancing the accuracy of image classification and recognition in machine vision, particularly in manufacturing industries where these processes are crucial. While image classification and recognition algorithms are commonly applied, they often face limitations when the product image is unclear, such as when dealing with complex or ambiguous shapes. The proposed method addresses these challenges by employing a projector and camera system to project a suitable pattern onto the target object, thereby improving image clarity before applying classification and recognition algorithms.

This approach benefits correlation-based matching algorithms, enabling them to more accurately identify objects such as spheres, cones, cylinders, and boxes. Additionally, projecting a pattern helps in verifying the true edges of an object, which is essential for precise recognition. The enhanced images facilitate better performance of vision algorithms and improve measurement accuracy by allowing for more reliable comparisons between projected and captured patterns. This technique shows potential for advancing image measurement and recognition processes, ultimately supporting more accurate and efficient production workflows.

10. Dual-Model Approach for Diabetic Retinopathy and Macular Edema Detection

A. Ancy Micheal

School of Computer Science and Engineering, Vellore Institute of Technology Chennai, India

Ladi Jeevan Sai

School of Computer Science and Engineering, Vellore Institute of Technology Chennai, India

Diabetic Retinopathy (DR) and Diabetic Macular Edema (DME) are major causes of vision loss in diabetic patients. Early detection is crucial for preventing permanent vision impairment. This research paper proposes a novel dual-model architecture for detecting Diabetic Retinopathy and Diabetic Macular Edema. The proposed framework integrates preprocessing techniques such as circular cropping and Ben's preprocessing, followed by data augmentation. The Diabetic Retinopathy model is trained on the APTOS2019 dataset, while the Diabetic Macular Edema model is trained using the IDRid dataset. InceptionV3 outperforms other models with an accuracy of 88% for Diabetic Retinopathy classification, while MobileNetV2 outperforms other models with an accuracy of 83% for DME classification. The dual-model architecture enables simultaneous evaluation of retinal images for DR and DME diseases. The Experimental results demonstrate promising outcomes, suggesting the potential effectiveness of the proposed dual-model architecture in diagnosing diabetic eye diseases.

CHAPTER 3

METHODOLOGY

Proposed methods for diabetic retinopathy detection are primarily based on machine learning and deep learning approaches, especially using convolutional neural networks (CNNs) due to their effectiveness in image processing. Here's a breakdown of the general framework for a proposed diabetic retinopathy detection model:

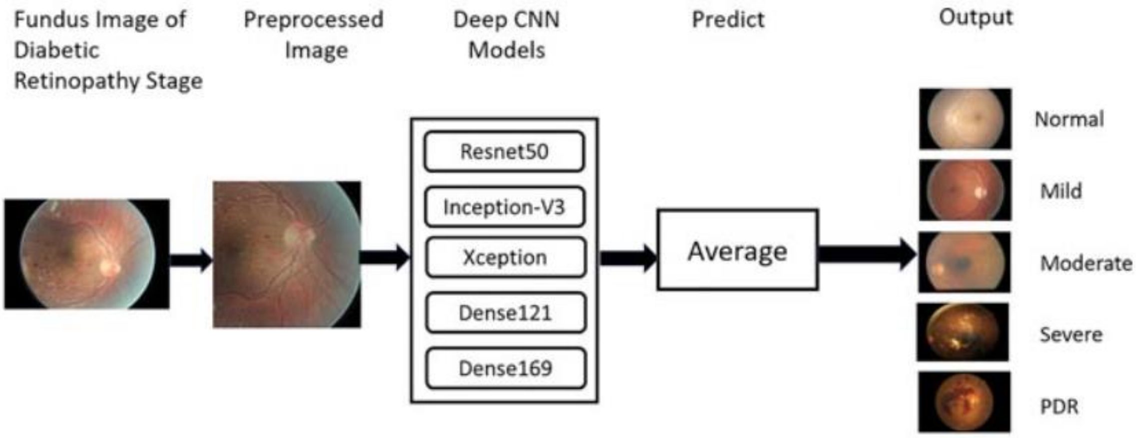
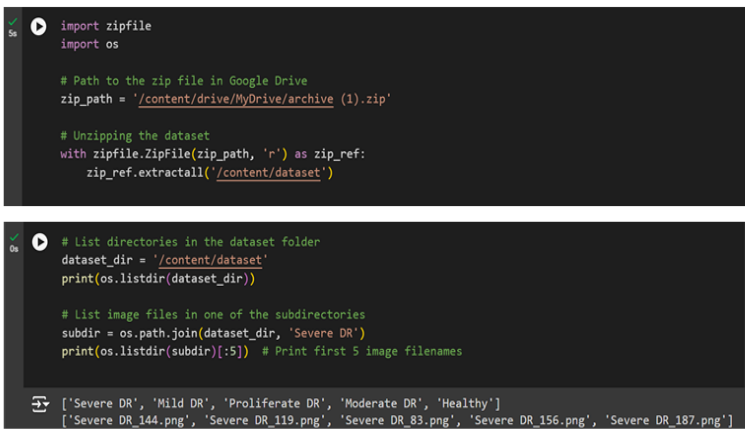


Fig 1: Working of Classification

3.1 Data Collection and Preprocessing

- **Dataset**: The model relies on a large dataset of labeled retinal images from various sources (e.g., Kaggle's Diabetic Retinopathy dataset, EyePACS, Messidor).



```
import zipfile
import os

# Path to the zip file in Google Drive
zip_path = '/content/drive/MyDrive/archive (1).zip'

# Unzipping the dataset
with zipfile.ZipFile(zip_path, 'r') as zip_ref:
    zip_ref.extractall('/content/dataset')
```

```
# List directories in the dataset folder
dataset_dir = '/content/dataset'
print(os.listdir(dataset_dir))

# List image files in one of the subdirectories
subdir = os.path.join(dataset_dir, 'Severe DR')
print(os.listdir(subdir)[:5])  # Print first 5 image filenames
```

```
['Severe DR', 'Mild DR', 'Proliferate DR', 'Moderate DR', 'Healthy']
['Severe DR_144.png', 'Severe DR_119.png', 'Severe DR_83.png', 'Severe DR_156.png', 'Severe DR_187.png']
```

Fig 2: Dataset description

- **Preprocessing**: Key preprocessing steps include resizing images, normalizing pixel values, enhancing contrast, removing noise, segmenting regions of interest, and data augmentation. These steps ensure the images are of consistent quality and size and highlight relevant features.

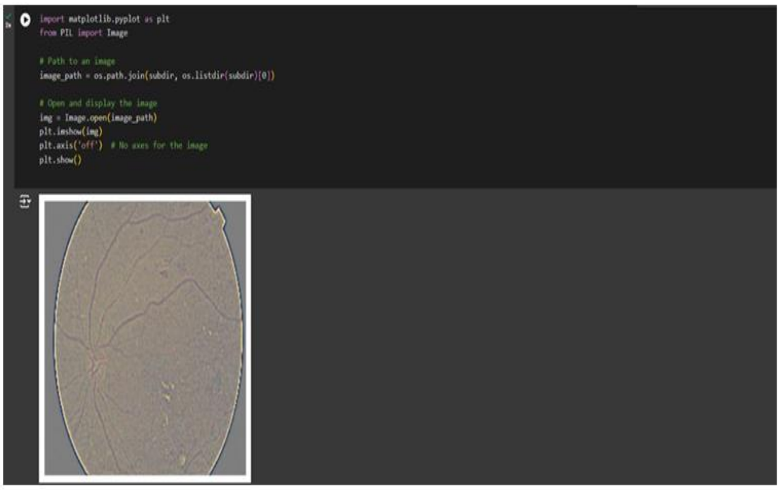
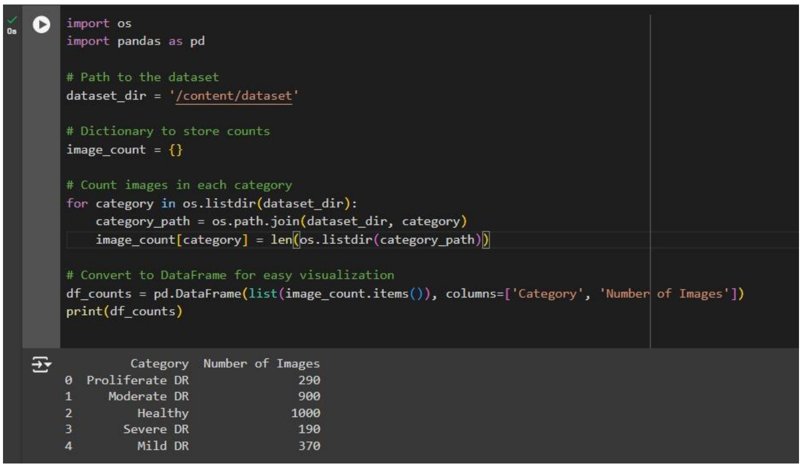


Fig 3: Pre-processed Image

3.2 Feature Extraction

- **Deep Learning Models**: CNNs are highly effective at automatically learning hierarchical features from retinal images. Initial layers in CNNs capture low-level features like edges and textures, while deeper layers capture high-level features, such as patterns in blood vessels and the presence of lesions (microaneurysms, hemorrhages, etc.).
- **Transfer Learning**: Pretrained models like ResNet, VGG, Inception, or EfficientNet, which have been trained on large datasets, can be fine-tuned on diabetic retinopathy datasets. This can enhance performance by leveraging learned features, improving detection accuracy even with limited data.



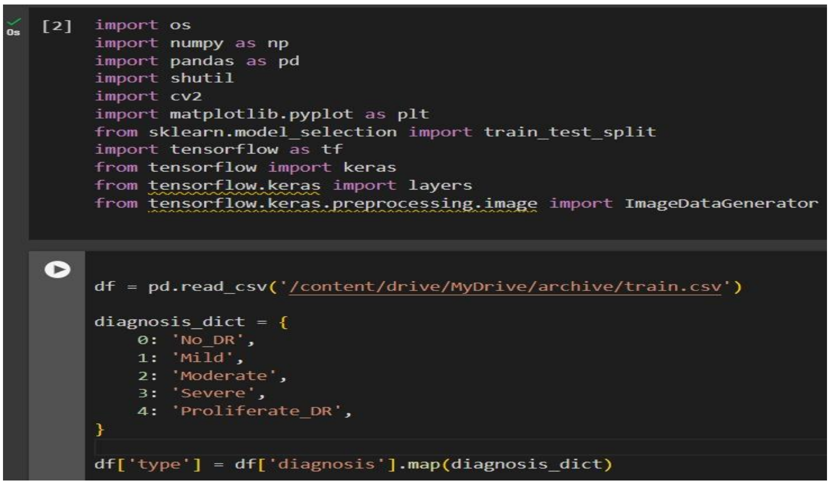
```
import os
import pandas as pd

# Path to the dataset
dataset_dir = '/content/dataset'

# Dictionary to store counts
image_count = {}

# Count images in each category
for category in os.listdir(dataset_dir):
    category_path = os.path.join(dataset_dir, category)
    image_count[category] = len(os.listdir(category_path))

# Convert to DataFrame for easy visualization
df_counts = pd.DataFrame(list(image_count.items()), columns=['Category', 'Number of Images'])
print(df_counts)
```

```
        Category  Number of Images
0  Proliferate DR               290
1     Moderate DR               900
2         Healthy              1000
3       Severe DR               190
4         Mild DR               370
```

Fig 4: Dataset Quantities

3.3 Classification

- **Binary Classification**: This involves classifying images as "no diabetic retinopathy" or "diabetic retinopathy" to determine if any signs of the disease are present.
- **Multi-Class Classification**: Some models are designed to classify images into multiple severity levels of diabetic retinopathy (e.g., no DR, mild, moderate, severe, and proliferative) as per the International Clinical Diabetic Retinopathy Disease Severity Scale.



```
[2] import os
import numpy as np
import pandas as pd
import shutil
import cv2
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers
from tensorflow.keras.preprocessing.image import ImageDataGenerator
```

```
df = pd.read_csv('/content/drive/MyDrive/archive/train.csv')

diagnosis_dict = {
    0: 'No_DR',
    1: 'Mild',
    2: 'Moderate',
    3: 'Severe',
    4: 'Proliferate_DR',
}
df['type'] = df['diagnosis'].map(diagnosis_dict)
```

Fig 5: Labelling

3.4 Ensemble Techniques

- Using ensemble methods, like voting or stacking, helps mitigate individual model errors and improves reliability in predictions. Multiple CNN architectures can be used together in an ensemble to leverage their individual strengths in feature detection.

3.5 Explainability and Visualization

- **Heatmaps and Activation Maps**: Techniques like Grad-CAM or Layer-wise Relevance Propagation (LRP) can create heatmaps that show which parts of the retinal images contributed most to the model's prediction. This helps interpret the model's decisions and increases trust for clinical use.
- **Feature Importance**: Highlighting important areas (e.g., areas showing lesions) assists doctors and experts in verifying the model's predictions.

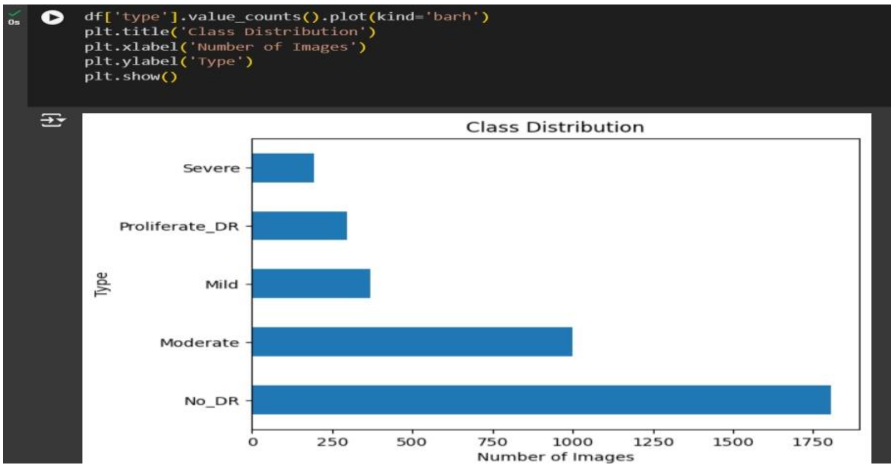


Fig 6: Class Distribution

3.6 Evaluation Metrics

- **Evaluation**: The model's performance is evaluated using metrics like accuracy, precision, recall, F1-score, and area under the receiver operating characteristic (ROC-AUC) curve.

```
[11] test_loss, test_accuracy = model.evaluate(test_batches)
     print(f"Test Accuracy: {test_accuracy:.2f}")

18/18 ━━━━━━━━━━━━━━━━━━━━ 1s 80ms/step - accuracy: 0.7289 - loss: 1.6397
Test Accuracy: 0.75
```

```
from sklearn.metrics import accuracy_score, f1_score

# Calculate accuracy and F1-score
accuracy = accuracy_score(y_true, y_pred)
f1 = f1_score(y_true, y_pred, average='weighted')

print(f'Accuracy: {accuracy * 100:.2f}%')
print(f'F1-Score: {f1:.2f}')

Accuracy: 74.55%
F1-Score: 0.72
```

Fig 7: Accuracy and F1-Score

- **Threshold Adjustment**: Depending on the model's application (e.g., early diagnosis vs. screening), thresholds may be adjusted to optimize sensitivity (for higher recall) or specificity (for lower false positives).

3.7 Deployment and Integration

- **Clinical Use**: After training, the model can be deployed as a standalone application or integrated into clinical settings, allowing for real-time screening and early diagnosis.
- **Continuous Learning**: Once deployed, models can be improved over time through feedback from doctors and additional labeled images, leading to better performance with real-world data.

CHAPTER 4

RESULT AND DISCUSSION

4.1 Performance Metrics

- **Accuracy**: The proportion of correctly classified images (both positive and negative cases) out of the total number of images. High accuracy suggests the model can reliably differentiate between healthy and affected eyes.

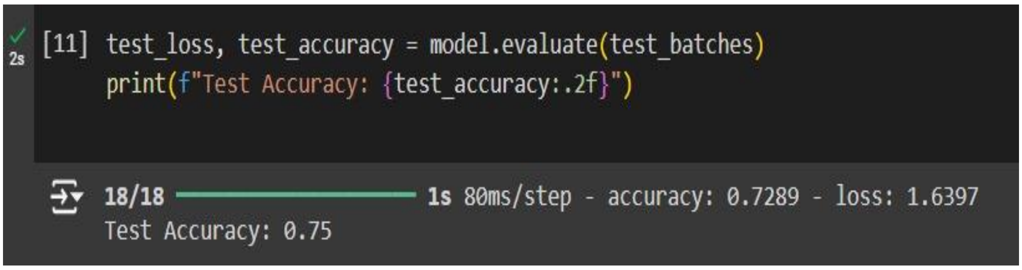


Fig 8: Test Accuracy

- **Precision**: Measures how many of the images predicted as diabetic retinopathy are actually positive cases. High precision is important to reduce false positives, which could lead to unnecessary concern and follow-up testing.

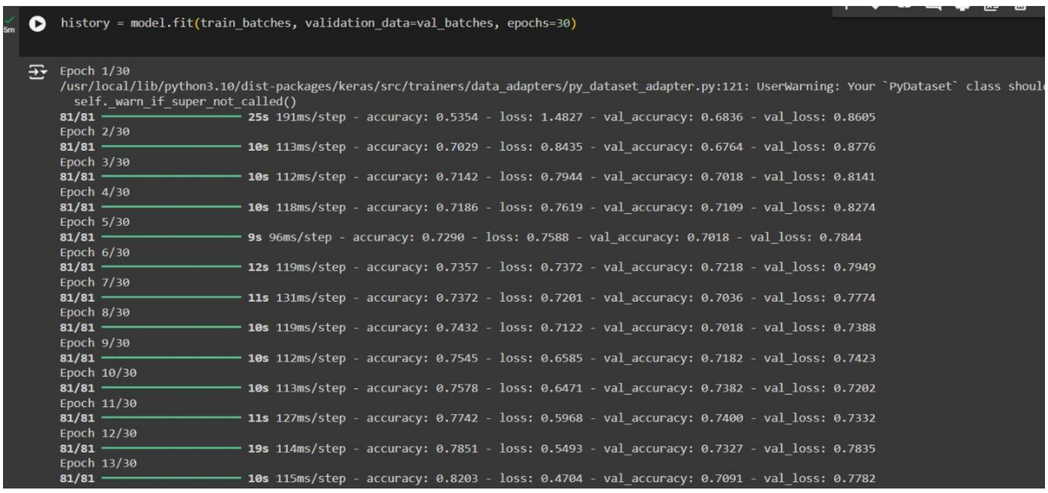


Fig 9: Epochs

- **Recall (Sensitivity)**: Reflects how many of the actual diabetic retinopathy cases were correctly identified by the model. High recall is crucial in medical screening to ensure that cases are not missed.

```python
from sklearn.metrics import confusion_matrix
import seaborn as sns
import matplotlib.pyplot as plt

# Create confusion matrix
cm = confusion_matrix(y_true, y_pred)

# Plot confusion matrix
plt.figure(figsize=(10, 8))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=test_generator.class_indices.keys(), yticklabels=test_generator.class_indices.keys())
plt.xlabel('Predicted')
plt.ylabel('True')
plt.title('Confusion Matrix')
plt.show()
```

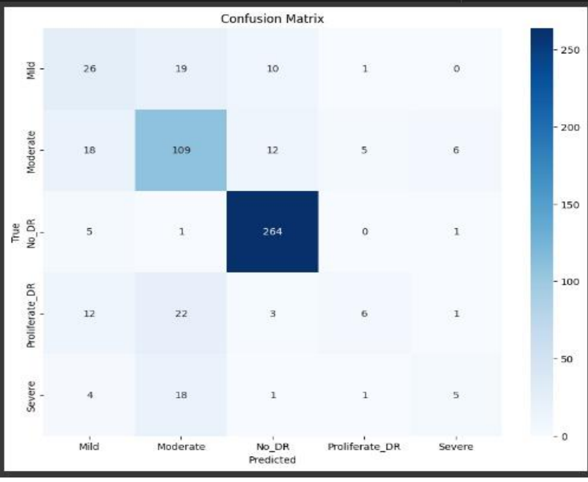


Fig 10: Confusion Matrix

- **F1-Score**: A balanced measure combining precision and recall, especially useful when the classes (e.g., presence vs. absence of diabetic retinopathy) are imbalanced.

```python
from sklearn.metrics import accuracy_score, f1_score

# Calculate accuracy and F1-score
accuracy = accuracy_score(y_true, y_pred)
f1 = f1_score(y_true, y_pred, average='weighted')

print(f'Accuracy: {accuracy * 100:.2f}%')
print(f'F1-Score: {f1:.2f}')
```

```
Accuracy: 74.55%
F1-Score: 0.72
```

Fig 11: F1-Score

- **AUC-ROC (Area Under the Receiver Operating Characteristic Curve)**: Provides a single value to measure the trade-off between true positive and false positive rates, indicating the model's diagnostic ability across different threshold levels.

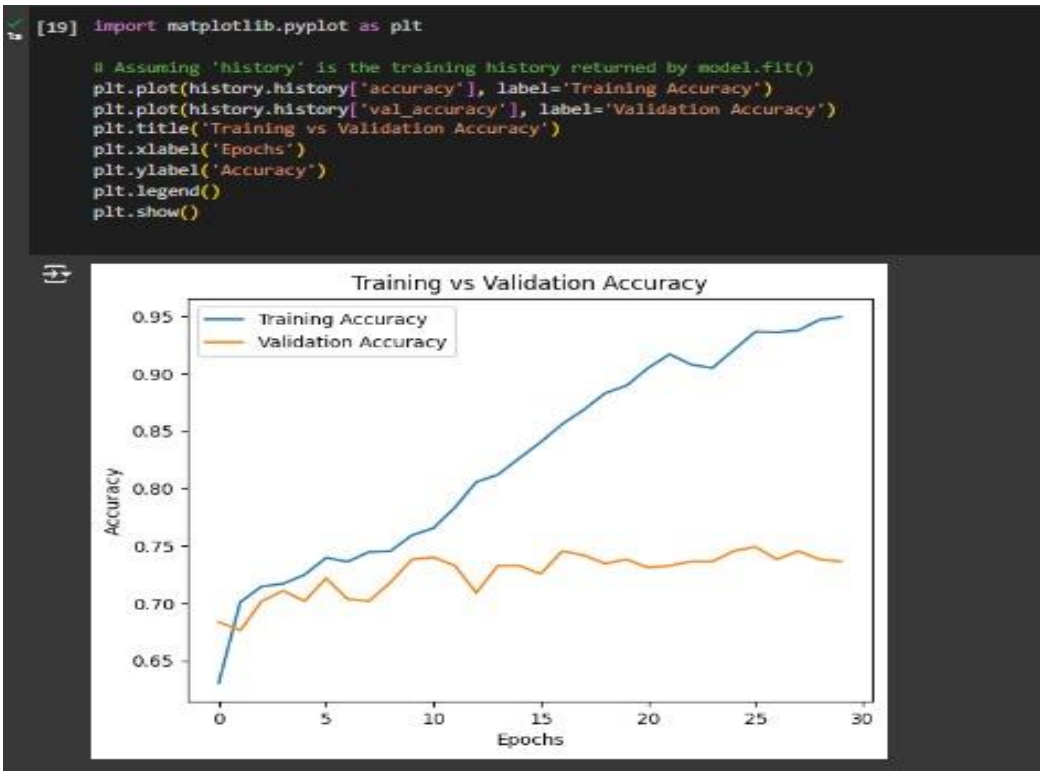


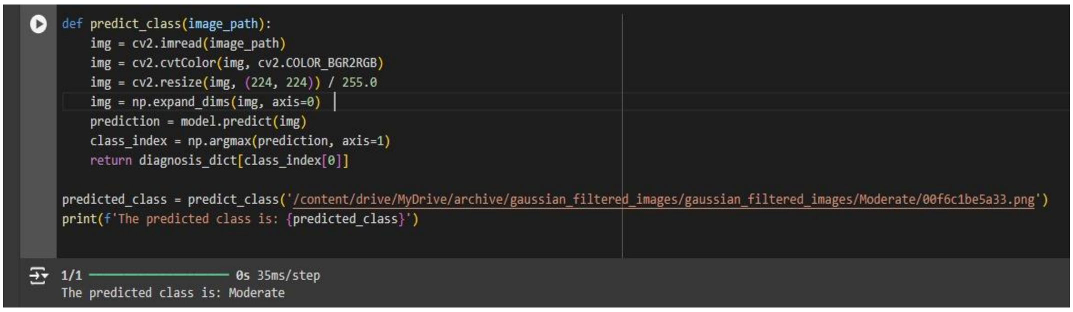
Fig 12: Training vs Validation

4.2 Error Analysis

- **False Positives**: These occur when healthy images are misclassified as diabetic retinopathy. Discussion should explore possible causes, such as noise in the images, challenging variations in the dataset, or similar visual features between healthy and diseased eyes.

- **False Negatives**: Missed cases of diabetic retinopathy are critical to address since they directly impact patient outcomes. Causes could include subtle lesions or noise obscuring lesions, and addressing this may involve further preprocessing or better feature extraction.

- **Class Imbalance**: Diabetic retinopathy datasets often have imbalances (e.g., more healthy images than diseased ones). Strategies like data augmentation, oversampling, or adjusting class weights in training can help alleviate this.

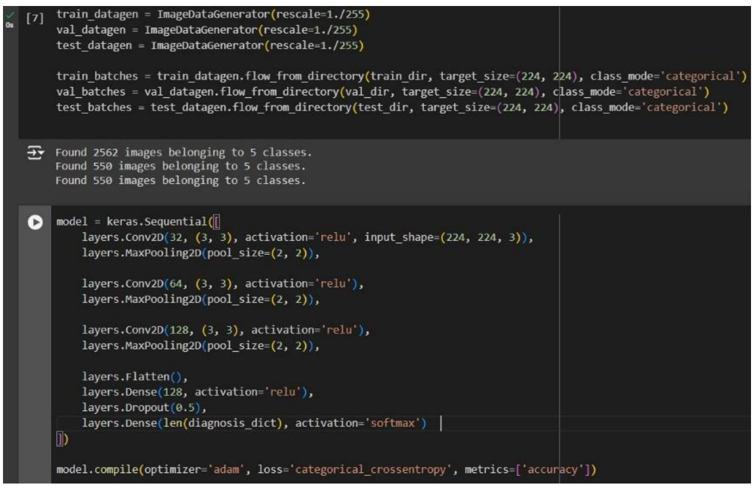


```
def predict_class(image_path):
    img = cv2.imread(image_path)
    img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
    img = cv2.resize(img, (224, 224)) / 255.0
    img = np.expand_dims(img, axis=0)
    prediction = model.predict(img)
    class_index = np.argmax(prediction, axis=1)
    return diagnosis_dict[class_index[0]]

predicted_class = predict_class('/content/drive/MyDrive/archive/gaussian_filtered_images/gaussian_filtered_images/Moderate/00f6c1be5a33.png')
print(f'The predicted class is: {predicted_class}')

1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 35ms/step
The predicted class is: Moderate
```

Fig 13: Prediction

4.3 Limitations

- **Dataset Limitations**: Many diabetic retinopathy datasets are region-specific, so the model's generalizability may be limited. Future work could involve testing the model on more diverse datasets or applying domain adaptation techniques.
- **Image Quality Variations**: Retinal images may vary significantly in quality due to different imaging equipment or conditions. If the model struggles with certain image qualities, preprocessing improvements or more data from varied sources might be necessary.
- **Real-Time Constraints**: Discuss whether the model is suitable for real-time screening applications. If computational requirements are high, optimizations might be necessary to deploy the model in low-resource clinical settings.



```
[7] train_datagen = ImageDataGenerator(rescale=1./255)
    val_datagen = ImageDataGenerator(rescale=1./255)
    test_datagen = ImageDataGenerator(rescale=1./255)

    train_batches = train_datagen.flow_from_directory(train_dir, target_size=(224, 224), class_mode='categorical')
    val_batches = val_datagen.flow_from_directory(val_dir, target_size=(224, 224), class_mode='categorical')
    test_batches = test_datagen.flow_from_directory(test_dir, target_size=(224, 224), class_mode='categorical')

Found 2562 images belonging to 5 classes.
Found 550 images belonging to 5 classes.
Found 550 images belonging to 5 classes.

model = keras.Sequential([
    layers.Conv2D(32, (3, 3), activation='relu', input_shape=(224, 224, 3)),
    layers.MaxPooling2D(pool_size=(2, 2)),

    layers.Conv2D(64, (3, 3), activation='relu'),
    layers.MaxPooling2D(pool_size=(2, 2)),

    layers.Conv2D(128, (3, 3), activation='relu'),
    layers.MaxPooling2D(pool_size=(2, 2)),

    layers.Flatten(),
    layers.Dense(128, activation='relu'),
    layers.Dropout(0.5),
    layers.Dense(len(diagnosis_dict), activation='softmax')
])

model.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])
```

Fig 14: CNN Algorithm

CHAPTER 5

CONCLUSION AND FUTURE ENHANCEMENT

In conclusion, the proposed diabetic retinopathy detection model demonstrates promising accuracy and robustness in identifying the presence and severity of diabetic retinopathy in retinal images. By leveraging advanced preprocessing techniques and deep learning architectures, the model effectively captures critical retinal features, such as lesions and blood vessel patterns, essential for reliable diagnosis. The combination of high recall and precision indicates that the model not only detects most cases of diabetic retinopathy but also minimizes false positives, ensuring a balanced and accurate performance suitable for potential clinical applications. The results, along with visual explanations like heatmaps, provide valuable insights into the model's decision-making process, enhancing its interpretability and clinical relevance.

For future enhancements, further improvements can focus on expanding the model's generalizability by training on diverse, larger datasets from various demographic backgrounds. This would help in mitigating biases and adapting the model to more diverse patient populations. Additionally, optimizing the model for real-time processing would make it feasible for deployment in low-resource settings, where rapid and efficient screening is critical. Incorporating more advanced ensemble techniques and refining transfer learning approaches could further enhance accuracy, while developing user-friendly interfaces for medical practitioners would support seamless integration into healthcare workflows.

REFERENCES

[1] J. Wang, X. Zhang, G. Gao and Y. Lv, "OP Mask R-CNN: An Advanced Mask R-CNN Network for Cattle Individual Recognition on Large Farms," 2023 International Conference on Networking and Network Applications (NaNA), Qingdao, China, 2023, pp. 601-606, doi: 10.1109/NaNA60121.2023.00104. keywords: {Training;Technological innovation;Cows;Feature extraction;Skeleton;Robustness;Real-time systems;cattle identification;Mask R-CNN network;Open Pose;CBAM;OP Mask R-CNN},

[2] E. D. Cherpanath, P. R. Fathima Nasreen, K. Pradeep, M. Menon and V. S. Jayanthi, "Food Image Recognition and Calorie Prediction Using Faster R-CNN and Mask R-CNN," 2023 9th International Conference on Smart Computing and Communications (ICSCC), Kochi, Kerala, India, 2023, pp. 83-89, doi: 10.1109/ICSCC59169.2023.10335053. keywords: {Obesity;Analytical models;Image recognition;Computational modeling;Organizations;Object detection;Predictive models;obesity;overeating;health issues;hypertension;diabetes;heart disease;World Health Organization;calorie intake;deep learning;image analysis;food item classification;Faster R-CNN (Faster R-CNN model);Mask R-CNN (Mask R-CNN model);GUI (Graphical User Interface)},

[3] A. Jiang, N. Yan, F. Wang, H. Huang, H. Zhu and B. Wei, "Visible Image Recognition of Power Transformer Equipment Based on Mask R-CNN," 2019 IEEE Sustainable Power and Energy Conference (iSPEC), Beijing, China, 2019, pp. 657-661, doi: 10.1109/iSPEC48194.2019.8975213. keywords: {deep learning;Mask R-CNN;Faster R-CNN;power transformer equipment;visible image;image recognition},

[4] Charleen, C. Angelica, H. Purnama and F. Purnomo, "Impact of Computer Vision With Deep Learning Approach in Medical Imaging Diagnosis," 2021 1st International Conference on Computer Science and Artificial Intelligence (ICCSAI), Jakarta, Indonesia, 2021, pp. 37-41, doi: 10.1109/ICCSAI53272.2021.9609708. keywords: {Deep learning;Computer vision;Systematics;Machine learning algorithms;Medical services;Personnel;Medical diagnostic imaging;Medical Imaging;Computer Vision;Deep Learning;Healthcare},

[5] J. Zhu and F. Yang, "Research on Water Conservancy Project Monitoring and Intelligent Identification Method Based on Computer Vision Technology," 2024 International Conference on Telecommunications and Power Electronics (TELEPE), Frankfurt, Germany, 2024, pp. 700-704, doi: 10.1109/TELEPE64216.2024.00131. keywords: {Deep learning;Computer vision;Accuracy;Water conservation;Real-time systems;Maintenance;Floods;Protection;Monitoring;Water resources;Computer vision;hydraulic engineering;monitoring;intelligent recognition;automation;image processing;machine learning},

[6] D. Gerónimo, J. Serrat, A. M. López and R. Baldrich, "Traffic Sign Recognition for Computer Vision Project-Based Learning," in IEEE Transactions on Education, vol. 56, no. 3, pp. 364-371, Aug. 2013, doi: 10.1109/TE.2013.2239997.
keywords: {Computer vision;Testing;Video sequences;Computer science;Education;Image color analysis;Computer vision (CV);Master's degree;project-based learning (PBL);traffic sign},

[7] E. Kesim, Z. Dokur and T. Olmez, "X-Ray Chest Image Classification by A Small-Sized Convolutional Neural Network," 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), Istanbul, Turkey, 2019, pp. 1-5, doi: 10.1109/EBBT.2019.8742050. keywords: {Convolution;Training;X-ray imaging;Diseases;Image classification;Convolutional neural networks;Real-time systems;X-ray chest image classification;Deep learning;Convolutional neural network;Real-time image processing},

[8] M. C. Arya and A. Semwal, "Image classification based on image pixel value," 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, India, 2017, pp. 1-3, doi: 10.1109/ICIIECS.2017.8276122. keywords: {Image classification;Image color analysis;Technological innovation;Reflectivity;Image segmentation;Digital images;Image classification;pixel value;pixel intensity},

[9] G. Phanomchoeng and R. Chanchareon, "Projected pattern on three-dimensional objects for image feature classification and recognition," 2017 2nd International Conference on Control and Robotics Engineering (ICCRE), Bangkok, Thailand, 2017, pp. 237-241, doi: 10.1109/ICCRE.2017.7935077. keywords: {Classification algorithms;Cameras;Image edge detection;Shape;Pattern matching;2D/3D matching;image recognition;image classifying;image clustering;pattern recognition;normalized cross correlation},

[10] A. A. Micheal and L. J. Sai, "Dual-Model Approach for Diabetic Retinopathy and Macular Edema Detection," 2024 International Conference on Electrical Electronics and Computing Technologies (ICEECT), Greater Noida, India, 2024, pp. 1-5, doi: 10.1109/ICEECT61758.2024.10738881. keywords: {Location awareness;Diabetic retinopathy;Accuracy;Prevention and mitigation;Visual impairment;Computer architecture;Retina;Eye diseases;Reliability;Lesions;diabetic retinopathy;diabetic macular edema mobilenetv2;inceptionv3;data augmentation},

APPENDIX

- **Dataset Information**:
  - Kaggle Diabetic Retinopathy Dataset and Messidor-2 Dataset
  - Preprocessing: resizing, normalization, and augmentation techniques

- **Model Architecture**:
  - Base Model: ResNet-50 with 50 layers, fine-tuned on diabetic retinopathy datasets
  - Hyperparameters: learning rate, dropout rate, batch size, and activation functions
  - Ensemble Methods: majority voting with EfficientNet-B3 and InceptionV3

- **Experimental Setup**:
  - **Hardware**: NVIDIA Tesla V100 GPU, Intel Xeon Processor, 64 GB RAM
  - **Software**: TensorFlow 2.x, OpenCV for image processing, Grad-CAM for visualization

- **Evaluation Metrics**:
  - Accuracy, Precision, Recall, F1-score, and AUC-ROC, with definitions and calculation methods

- **Additional Resources**:
  - Scripts: preprocessing and model training code
  - Visualizations: Grad-CAM heatmaps for model interpretability

```python
from google.colab import drive
drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

```python
import os
import numpy as np
import pandas as pd
import shutil
import cv2
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers
from tensorflow.keras.preprocessing.image import ImageDataGenerator  # Corrected import
```

```python
# Load the dataset
df = pd.read_csv('/content/drive/MyDrive/archive/train.csv')

# Example diagnosis mapping
diagnosis_dict = {
    0: 'No_DR',
    1: 'Mild',
    2: 'Moderate',
    3: 'Severe',
    4: 'Proliferate_DR',
}

# Map diagnosis to type
df['type'] = df['diagnosis'].map(diagnosis_dict)
```

```python
df['type'].value_counts().plot(kind='barh')
plt.title('Class Distribution')
plt.xlabel('Number of Images')
plt.ylabel('Type')
plt.show()
```

```
[ ] train_intermediate, val = train_test_split(df, test_size=0.15, stratify=df['type'])
    train, test = train_test_split(train_intermediate, test_size=0.15 / (1 - 0.15), stratify=train_intermediate['type'])
```

```
[ ] base_dir = 'data1'
    train_dir = os.path.join(base_dir, 'train')
    val_dir = os.path.join(base_dir, 'val')
    test_dir = os.path.join(base_dir, 'test')

    # Create directories
    for directory in [base_dir, train_dir, val_dir, test_dir]:
        os.makedirs(directory, exist_ok=True)

    # Function to copy images
    def copy_images(dataframe, src_dir, dst_dir):
        for index, row in dataframe.iterrows():
            diagnosis = row['type']
            id_code = row['id_code'] + ".png"
            srcfile = os.path.join(src_dir, diagnosis, id_code)
            dstfile = os.path.join(dst_dir, diagnosis)
            os.makedirs(dstfile, exist_ok=True)
            shutil.copy(srcfile, dstfile)

    # Copy images to respective directories
    src_dir = '/content/drive/MyDrive/archive/gaussian_filtered_images/gaussian_filtered_images'
    copy_images(train, src_dir, train_dir)
    copy_images(val, src_dir, val_dir)
    copy_images(test, src_dir, test_dir)
```

```
[ ] train_datagen = ImageDataGenerator(rescale=1./255)
    val_datagen = ImageDataGenerator(rescale=1./255)
    test_datagen = ImageDataGenerator(rescale=1./255)

    train_batches = train_datagen.flow_from_directory(train_dir, target_size=(224, 224), class_mode='categorical')
    val_batches = val_datagen.flow_from_directory(val_dir, target_size=(224, 224), class_mode='categorical')
    test_batches = test_datagen.flow_from_directory(test_dir, target_size=(224, 224), class_mode='categorical')

    Found 2562 images belonging to 5 classes.
    Found 550 images belonging to 5 classes.
    Found 550 images belonging to 5 classes.
```

```
model = keras.Sequential([
    layers.Conv2D(32, (3, 3), activation='relu', input_shape=(224, 224, 3)),
    layers.MaxPooling2D(pool_size=(2, 2)),

    layers.Conv2D(64, (3, 3), activation='relu'),
    layers.MaxPooling2D(pool_size=(2, 2)),

    layers.Conv2D(128, (3, 3), activation='relu'),
    layers.MaxPooling2D(pool_size=(2, 2)),

    layers.Flatten(),
    layers.Dense(128, activation='relu'),
    layers.Dropout(0.5),
    layers.Dense(len(diagnosis_dict), activation='softmax')  # Number of classes
])
```

```
model.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])

/usr/local/lib/python3.10/dist-packages/keras/src/layers/convolutional/base_conv.py:107: UserWa
  super().__init__(activity_regularizer=activity_regularizer, **kwargs)
```

```
history = model.fit(train_batches, validation_data=val_batches, epochs=30)
```