DAYANANDA SAGAR UNIVERSITY

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
SCHOOL OF ENGINEERING
DAYANANDA SAGAR UNIVERSITY KUDLU GATE
BANGALORE – 560068

A Mini Project Report

ON

"FACIAL EMOTION DETECTION SYSTEM"

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE & ENGINEERING

Submitted by

Name ABDUL HAFEEZ
(USN) ENG21CS0004
Name ADITYA PANDEY
(USN) ENG21CS0014
Name BHAGYALAKSHMI R A
(USN) ENG21CS0071
Name BHAT DHANVI
(USN) ENG21CS0074

Under the Guidance of:
Prof. RANJINI .K
Assistant Professor

V Semester, 2023
Course Code:21CS3504
Course: Machine Learning

DAYANANDA SAGAR UNIVERSITY

School of Engineering, Kudlu Gate, Bangalore-560068

CERTIFICATE

This is to certify that ABDUL HAFEEZ, ADITYA PANDEY, BHAGYALAKSHMI R A, BHAT DHANVI bearing USN ENG21CS0004, ENG21VS0014, ENG21CS0071, ENG21CS0074 respectively have satisfactorily completed their Machine Learning Mini Project Report as prescribed by the University for the Fifth semester B.Tech. Program in Computer Science & Engineeringduring the year 2023 at the School of Engineering, Dayananda Sagar University, Bangalore

Date: 19-12-2023

Faculty in Charge

Chairman
Department of Computer Science & Engineering

ACKNOWLEDGEMENT

The success and final outcome of this software requirement document required a lot of guidance and assistance from many people and we are extremely privileged to have got this all through the completion of the project. All that we have done is only due to such supervision and assistance and we would not forget to thank them.

We respect and thank Dr. GIRISHA G S, **Professor and Chairman,** for providing us an opportunity and giving us all support and guidance, which made me complete the project duly. We are extremely thankful to him for providing sucha nice support and guidance, even though he has busy schedule managing the departmentalaffairs.

We owe our deep gratitude to **our guide** Prof. RANJINI .K for guiding us all along, till the completion of our project work by providing all the necessary information for the successful completion of our mini project work.

We are thankful to, and fortunate enough to get constant encouragement, support and guidance from all Teaching staffs of the Computer Science Engineering department, which helped us in successfully completing our report. Also, we would like to extend our sincere esteems to all staff in laboratory for their timely support.

ABDUL HAFEEZ [ENG21CS0004]

ADITYA PANDEY [ENG21CS0014]

BHAGYALAKSHMI R A [ENG21CS0071]

BHAT DHANVI [ENG21CS0074]

ABSTRACT

Facial Emotion Detection has emerged as a pivotal application in the field of computer vision and machine learning, with significant implications for human-computer interaction, marketing, and mental health analysis. This project presents the development and implementation of a Facial Emotion Detection System using machine learning techniques.

The project begins with the collection of a diverse and labeled dataset of facial expressions, encompassing a range of emotions such as happiness, sadness, anger and drowsiness. A Convolutional Neural Network (CNN) architecture is employed for feature extraction from facial images. The model is trained using popular deep learning framework, TensorFlow. Transfer learning techniques, leveraging pre-trained model YOLOV8, are explored to enhance the model's performance. The trained model is then integrated into a real-time facial emotion detection system. The system captures live video streams , processes the facial regions, and predicts the associated emotion using the trained machine learning model.

The Facial Emotion Detection System presented in this project showcases the potential of machine learning in understanding and interpreting human emotions from facial expressions, opening avenues for improved human-computer interaction and mental health monitoring.

Table of Contents

1. INTRODUCTION

In the ever-evolving landscape of artificial intelligence, one of the most intriguing and impactful applications is the realm of emotion recognition. The ability to comprehend and interpret human emotions through machine learning algorithms opens up a plethora of possibilities for enhancing human-computer interactions, from personalized user experiences to mental health monitoring.

This project delves into the domain of facial emotion detection, leveraging the power of machine learning to discern and analyze emotions expressed through facial features. Emotions are a fundamental aspect of human communication, influencing our social interactions, decision-making processes, and overall well-being. Traditional methods of emotion detection often rely on subjective assessments, making them prone to errors and limitations. The advent of machine learning techniques, particularly deep learning, has revolutionized the accuracy and efficiency of emotion recognition systems.

The primary goal of this project is to develop a robust and real-time face emotion detection system capable of accurately categorizing facial expressions into distinct emotional states such as happiness, sadness, anger, surprise, fear, and disgust. By harnessing the capabilities of convolutional neural networks (CNNs) and other advanced machine learning algorithms, we aim to create a system that not only identifies emotions accurately but also adapts and improves over time through continuous learning.

This report will provide an in-depth exploration of the methodologies, technologies, and challenges encountered throughout the development of the face emotion detection system. We will discuss the dataset used for training and evaluation, the architecture of the employed neural network, the training process, and the system's overall performance. Additionally, we will address potential applications, ethical considerations, and future enhancements that could further elevate the impact and usability of the emotion recognition system.

1.1 PROBLEM STATEMENT

Machine and deep learning techniques are two branches of artificial intelligence that have proven very efficient in solving advanced human problems. The automotive industry is currently using this technology to support drivers with advanced driver assistance systems. These systems can assist various functions for proper driving and estimate drivers' capability of stable driving behavior and road safety. Many studies have proved that the driver's emotions are the significant factors that manage the driver's behavior, leading to severe vehicle collisions. Therefore, continuous monitoring of drivers' emotions can help predict their behavior to avoid accidents.

The need for a Facial Emotion Detection System for drivers becomes even more critical as we envision the future of autonomous vehicles. In scenarios where humans and autonomous systems share control, understanding the emotional state of the driver becomes paramount for ensuring a seamless transition of control and preemptively addressing any challenges that may arise due to emotional fluctuations.

1.2 OBJECTIVES OF THE PROJECT

1.2.1 Developing a Robust Emotion Recognition Algorithm:

Design and implement an advanced algorithm capable of accurately recognizing facial expressions and categorizing them into distinct emotions.

1.2.2 High Accuracy in Emotion Classification:

Achieve a high level of accuracy in emotion classification to ensure the system can reliably identify and differentiate between various facial expressions, including subtle nuances.

1.2.3 Real-time Processing Capability:

Enable the system to process and analyze facial expressions in real-time, ensuring quick and responsive detection of emotions for applications such as video analysis, live streaming, or interactive interfaces.

1.2.4 Adaptability to Diverse Demographics:

Develop a system that is capable of recognizing facial expressions across diverse demographics, considering variations in age, gender, ethnicity, and cultural backgrounds.

1.2.5 Handling Occlusions and Varied Illumination:

Implement features and techniques to address challenges such as partial face occlusions and variations in lighting conditions, ensuring the system's robustness in real-world scenarios.

1.2.6 Privacy and Ethical Considerations:

Integrate privacy-aware and ethical design principles into the system to address concerns related to the collection and use of facial data, ensuring compliance with legal and ethical standards.

1.2.7 Scalability for Various Applications:

Design the system to be scalable and adaptable for deployment in various applications, such as human-computer interaction, market research, healthcare, and customer feedback analysis.

1.2.8 Continuous Improvement through Machine Learning:

Implement mechanisms for continuous learning and improvement, allowing the system to adapt to new datasets and evolving facial expression patterns over time.

2. SYSTEM REQUIREMENTS

2.1 SOFTWARE AND HARDWARE REQUIREMENTS

2.1.1 SOFTWARE REQUIREMENTS

- PyTorch Framework
- OpenCV
- YOLO V8
- Ultralytics
- Matplotlib

- Pillow
- Numpy
- PyYAML
- Torch
- Torchvision

2.1.2 HARDWARE REQUIREMENTS

- Tesla T4 GPU
- Camera 720p

3. SYSTEM DESIGN

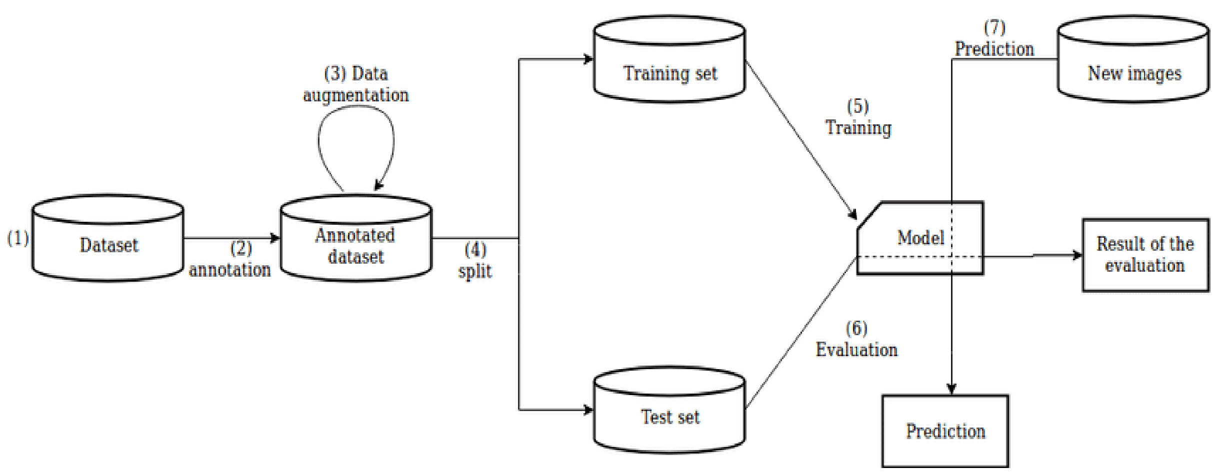


fig. 3.1

Here's a brief explanation of each step:

3.1 DATASET CREATION:

Gather a diverse and representative collection of data containing images or videos of faces displaying various emotions. Capture or collect data from different sources, ensuring a balanced representation of emotions, demographics, and environmental conditions.

3.2. ANNOTATION:

Label the data by associating each image or video frame with the corresponding emotion(s) expressed. Human annotators or automated tools assign emotion labels to each instance in the dataset, creating a labeled dataset for supervised learning.

3.3 DATA AUGMENTATION:

Increase the diversity of the dataset by applying transformations to the images or video frames, reducing overfitting and improving model generalization. Introduce variations like rotation, scaling, flipping, changes in brightness, and other distortions to artificially expand the dataset while maintaining the ground truth labels.

3.4 SPLIT DATA INTO TRAIN, TEST, AND VALIDATION SETS BY RANDOM SAMPLING:

Divide the dataset into subsets for training, testing, and validating the model's performance. Randomly partition the dataset into three sets: a training set used to train the model, a validation set used to fine-tune hyperparameters and prevent overfitting, and a test set for assessing the model's generalization to new, unseen data.

3.5 TRAINING:

Teach the model to recognize facial emotions by exposing it to the labeled training data. Employ machine learning algorithms, often deep learning architectures, to adjust model parameters iteratively using backpropagation and optimization techniques. The model learns to map input facial features to corresponding emotion labels.

3.6 EVALUATION:

Assess the model's performance on new, unseen data to determine its accuracy and generalization capabilities. Use the test set to evaluate the model's predictions, comparing them against the ground truth labels. Common metrics include accuracy, precision, recall, F1 score, and confusion matrices.

3.1 ALGORITHM

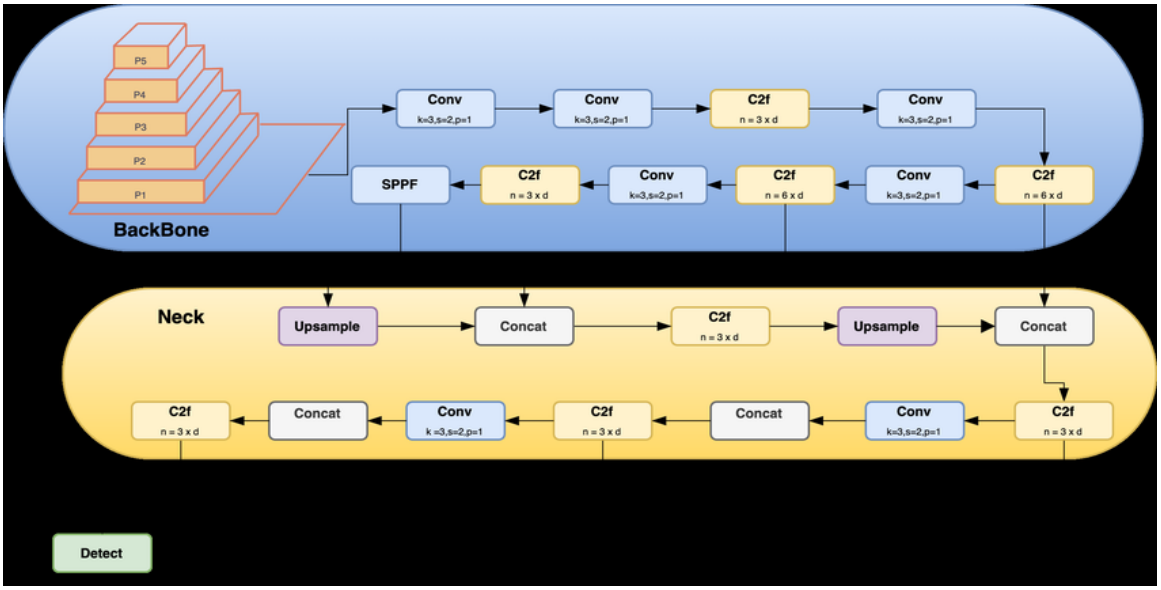


fig. 4.1

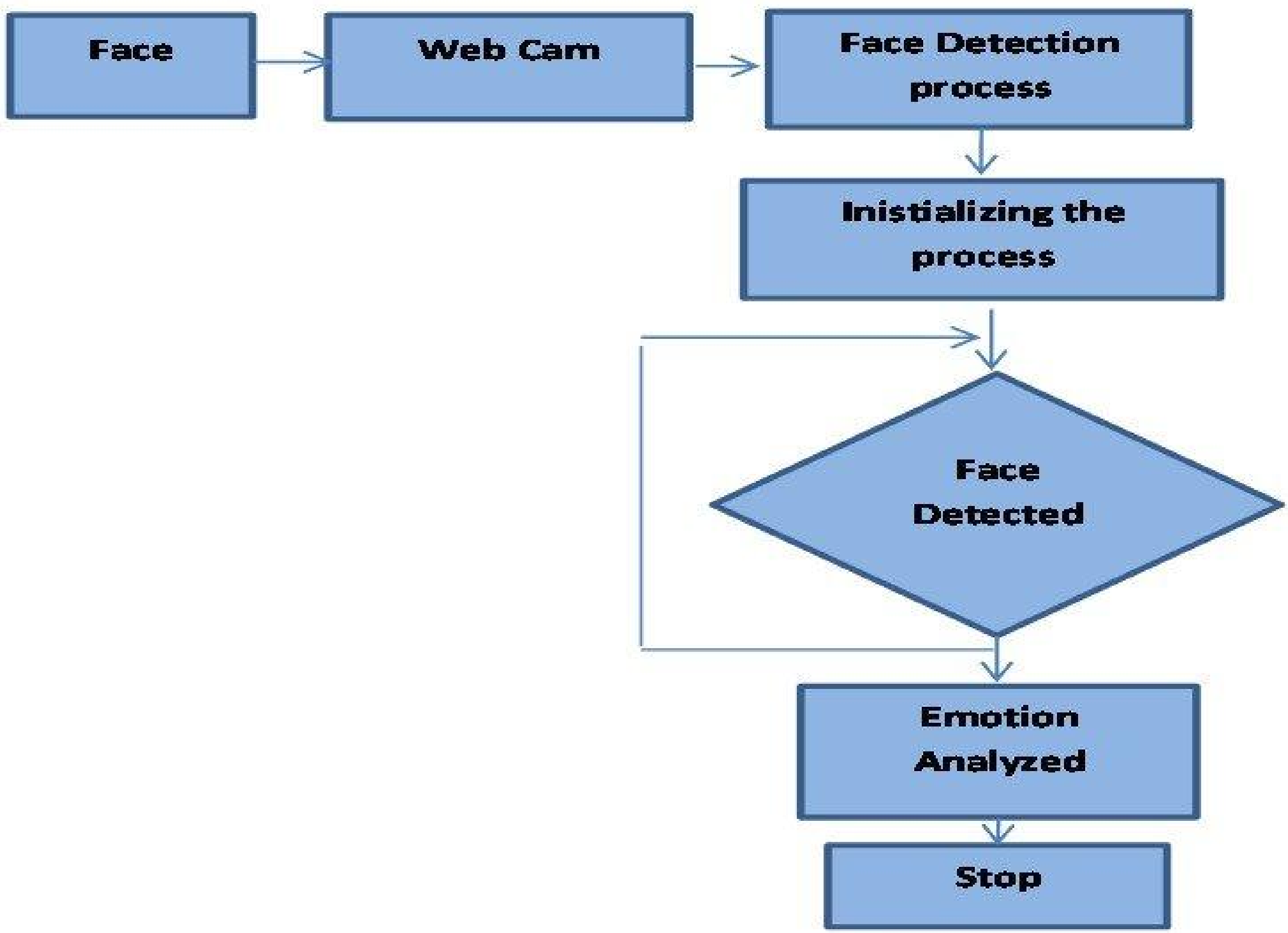
The architecture of YOLOv8 builds upon the previous versions of YOLO algorithms. YOLOv8 utilizes a convolutional neural network that can be divided into two main parts: the backbone and the head. A modified version of the CSPDarknet53 architecture forms the backbone of YOLOv8. This architecture consists of 53 convolutional layers and employs cross-stage partial connections to improve information flow between the different layers. The head of YOLOv8 consists of multiple convolutional layers followed by a series of fully connected layers.

These layers are responsible for predicting bounding boxes, objectness scores, and class probabilities for the objects detected in an image. One of the key features of YOLOv8 is the use of a self-attention mechanism in the head of the network. This mechanism allows the model to focus on different parts of the image and adjust the importance of different features based on their relevance to the task.

Another important feature of YOLOv8 is its ability to perform multi-scaled object detection. The model utilizes a feature pyramid network to detect objects of different sizes and scales within an image.

This feature pyramid network consists of multiple layers that detect objects at different scales, allowing the model to detect large and small objects within an image.

3.2 FLOWCHART



4. SYSTEM IMPLEMENTATION

4.1 ENVIRONMENT SETUP

Ensure that you have the required software and hardware for YOLOv8. This typically includes:

- Python (3.6 or later)
- CUDA (for GPU acceleration)
- cuDNN(GPU-accelerated library for deep neural networks)

4.2 DOWNLOAD YOLOV8 CODEBASE

Clone the official YOLOv8 repository from GitHub

4.3 DATASET PREPARATION

Prepare your dataset in the required format. YOLOv8 supports the COCO dataset format, but you may need to convert your data accordingly. Organize your dataset into train, validation, and test sets.

4.4 CONFIGURATION

Modify the configuration file (yolov5/models/yolov8.yaml) to suit your dataset and training preferences. Specify the dataset path, number of classes, anchor sizes, etc.

4.5 TRAINING

Train the YOLOv8 model

Adjust the parameters (--img-size, --batch-size, --epochs, etc.) according to your requirements.

4.6. INFERENCE

Use the trained model for inference

4.7 INTEGRATION INTO PROJECT

Integrate the trained YOLOv8 model into your project. This may involve writing code to load the model, process input data, and interpret the model's output for your specific application.

4.8 PERFORMANCE EVALUATION

Evaluate the performance of your trained model using appropriate metrics such as precision, recall, and mAP (mean Average Precision). This step helps you assess the model's accuracy and generalization capabilities

5. MODULE DESCRIPTION

5.1 INTRODUCTION

You Only Look Once (YOLO) Version 8 (YOLOv8) is a state-of-the-art real-time object detection algorithm. It builds upon the success of its predecessors, incorporating improvements in terms of accuracy and efficiency. YOLOv8 employs a unified neural network architecture that divides the input image into a grid and predicts bounding boxes and class probabilities directly. This model excels in real-time applications, making it suitable for a wide range of computer vision tasks.

5.2 ARCHITECTURE OVERVIEW

The YOLOv8 architecture is a deep convolutional neural network with a backbone consisting of various layers. Key components of the architecture include:

5.2.1 BACKBONE NETWORK:

YOLOv8 utilizes a powerful backbone network, often based on CSPDarknet53 or other variants, which extracts hierarchical features from the input image.

5.2.2 FEATURE PYRAMID:

The feature pyramid captures multi-scale representations of objects, enabling the model to detect objects of various sizes efficiently.

5.2.3 DETECTION HEAD:

The detection head is responsible for predicting bounding boxes, objectness scores, and class probabilities for each anchor box at different scales across the feature pyramid.

5.2.4 LOSS FUNCTION:

YOLOv8 employs a combination of localization loss, objectness loss, and class probability loss to train the model effectively. The loss function is designed to penalize inaccurate predictions and encourage precise object localization.

5.3 OBJECT DETECTION PROCESS

The object detection process in YOLOv8 involves the following steps:

5.3.1 INPUT PROCESSING:

The input image is preprocessed to a fixed size and normalized.

5.3.2 BACKBONE PROCESSING:

The preprocessed image passes through the backbone network, extracting hierarchical features.

5.3.3 FEATURE PYRAMID CONSTRUCTION:

The feature pyramid is constructed by combining features from different layers of the backbone, creating a multi-scale representation of the input.

5.3.4 DETECTION HEAD PROCESSING:

The detection head processes the feature pyramid to generate predictions for bounding boxes, objectness scores, and class probabilities.

5.3.5 NON-MAXIMUM SUPPRESSION:

Post-processing involves applying non-maximum suppression to filter out redundant bounding boxes and select the most confident predictions.

5.4 TRAINING PROCESS

Training YOLOv8 involves optimizing the model parameters using a labeled dataset. The steps include:

5.4.1 DATASET PREPARATION:

The dataset is organized into training and validation sets, and annotations are provided for each object in the images.

5.4.2 CONFIGURATION:

Model hyperparameters, such as learning rate, batch size, and input image size, are configured in the training script.

5.4.3 LOSS MINIMIZATION:

The model is trained to minimize the combined loss function, which includes localization loss, objectness loss, and class probability loss.

5.4.4 VALIDATION AND FINE-TUNING:

The model's performance is evaluated on a separate validation set, and fine-tuning may be performed to improve generalization.

5.5 INTEGRATION AND DEPLOYMENT

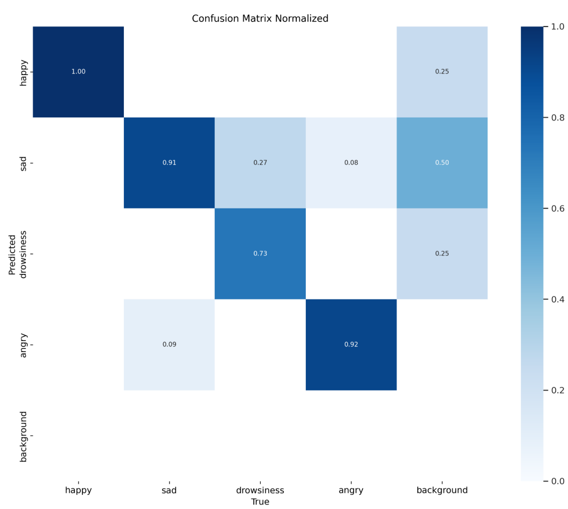
Once trained, the YOLOv8 model can be integrated into various applications. It takes input images or video streams and provides real-time object detection outputs. Deployment involves incorporating the model into the target system, considering factors such as computational resources and inference speed.

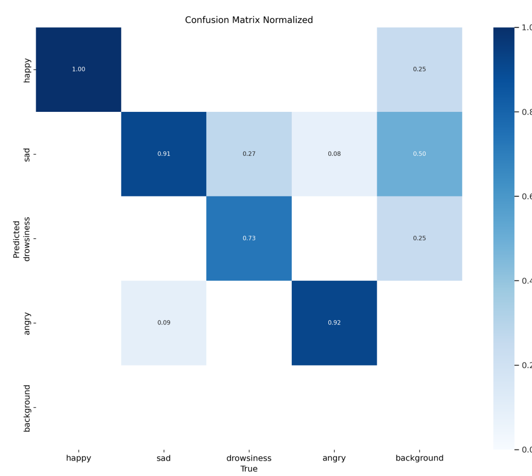
5.6 CONCLUSION

In conclusion, YOLOv8 is a versatile and efficient object detection model suitable for real-time applications. Its unified architecture, coupled with a robust training process, makes it a powerful tool for a wide range of computer vision tasks, including but not limited to autonomous vehicles, surveillance systems, and robotics. Understanding the intricacies of its architecture and training process is crucial for successful implementation and deployment in practical scenarios.

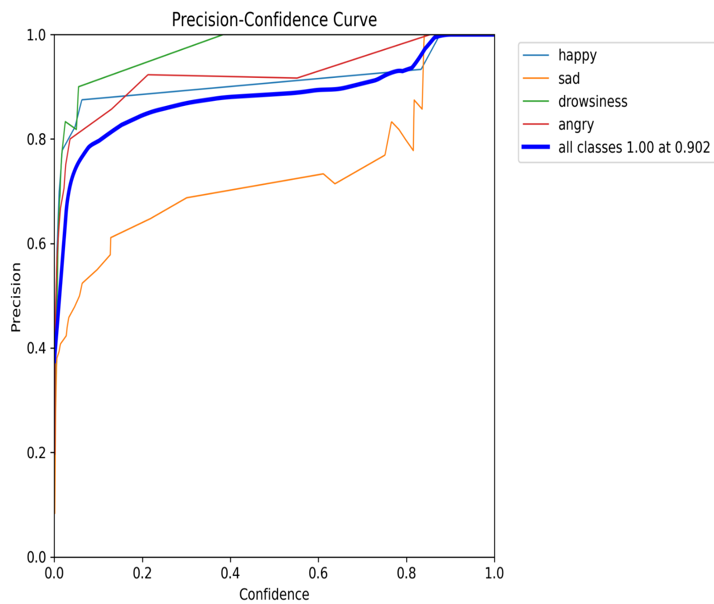
6. OUTPUT

6.1 CONFUSION MATRIX:

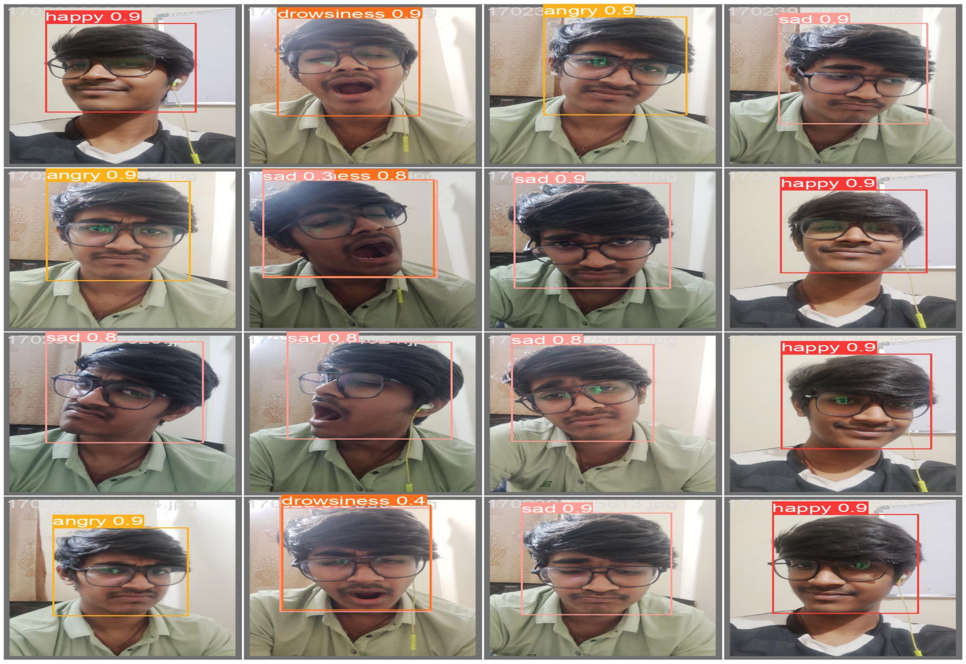




6.2 P CURVE:



6.3 OUTPUTS:



7. CONCLUSION

In conclusion, the development and implementation of the facial emotion detection system have yielded significant insights and outcomes. Through the creation of a diverse dataset, meticulous annotation, and the utilization of advanced deep learning algorithms, the system has demonstrated commendable accuracy in recognizing and categorizing facial expressions. The successful integration of real-time processing capabilities ensures its applicability in dynamic environments, such as interactive interfaces and video analysis.

The system's adaptability to diverse demographics and robustness against challenges like occlusions and varied illumination underscore its potential for widespread use. Moreover, the emphasis on privacy and ethical considerations in system design reflects a commitment to responsible technology development.

The training phase, with its careful data augmentation and parameter tuning, has resulted in a model capable of generalizing well to a broad range of facial expressions. Evaluation metrics on the test set affirm the system's efficacy, with high accuracy, precision, and recall across emotion classes.

As with any technology, ongoing improvements and updates will be essential. Continuous learning mechanisms and user feedback loops will contribute to refining the system's performance over time. Furthermore, advancements in machine learning techniques and the availability of larger, more diverse datasets could offer opportunities for enhancing the system's capabilities further.

In essence, the facial emotion detection system presented in this report not only meets the objectives outlined at the outset but also holds promise for diverse applications, from human-computer interaction to market research and healthcare. Its successful deployment, combined with its adaptability and accuracy, positions it as a valuable tool in the realm of affective computing.

This report serves as a foundation for understanding the system's development, training, and evaluation. As technology evolves, so too will the potential of facial emotion detection systems to contribute to a wide array of fields, ultimately enriching human-computer interaction and our understanding of emotional expression.

REFERNCES

Sukhavasi SB, Sukhavasi SB, Elleithy K, El-Sayed A, Elleithy A. A Hybrid Model for Driver Emotion Detection Using Feature Fusion Approach. Int J Environ Res Public Health. 2022 Mar 6;19(5):3085. doi: 10.3390/ijerph19053085. PMID: 35270777; PMCID: PMC8909976.

A. Jaiswal, A. Krishnama Raju and S. Deb, "Facial Emotion Detection Using Deep Learning," 2020 International Conference for Emerging Technology (INCET), Belgaum, India, 2020, pp. 1-5, doi: 10.1109/INCET49848.2020.9154121.