A picture containing text, clipart

Description automatically generated

Report file

Prepared by

|  |  |
| --- | --- |
| Name | Sap ID |
| Charu Gupta | 500090912 |
| Lakshay Agarwal | 500094127 |
| Naman Jain | 500090994 |

Submitted to:

Dr. Rohit Srivastava

**Table of Contents**

|  |  |  |
| --- | --- | --- |
| **Topic** | | **Page No** |
| 1 | Introduction |  |
|  | 1.1 Prologue |  |
|  | 1.2 Problem Statement |  |
|  | 1.3 Objectives |  |
|  | 1.4 System Requirements |  |
| 2 | System Analysis |  |
|  | 2.1 Motivation |  |
|  | 2.2 Proposed System |  |
| 3 | Design |  |
|  | 3.1 Flow Chart Diagrams |  |
| 4 | Implementation |  |
|  | 4.1 Methodology |  |
|  | 4.2 Algorithms |  |
|  | 4.3 Screenshots |  |
| 5 | Mathematical Model |  |
| 6 | References |  |

**Introduction**

* 1. Prologue

|  |  |  |  |
| --- | --- | --- | --- |
| **Paper** | **Algorithm** | **Results** | **Inference** |
| Facial emotion recognition using CNN[1] | CNN | In the FERC-2013 dataset, the average accuracy was 70.14%, surpassing the previous model's 67.02%. For the JAFFE dataset, the proposed model attained an impressive 98.65% accuracy, outperforming prior results. Emotion detection accuracy varied across different emotions, with happy emotions achieving the highest accuracy of 95% and anger the lowest at 56%. | The CNNs, demonstrate superior performance compared to traditional methods. The dataset  FERC-2013 achieved an accuracy of 70.14% and the JAFFE dataset achieved 98.65% accuracy. This indicates the effectiveness of CNNs in accurately recognizing emotions from facial expressions. Also, the paper tells the importance of deep learning techniques in enhancing emotion detection systems, in various fields such as human-computer interaction and emotional computing. |
| Facial emotion recognition  in real-time  and  static images [2] | SVM, K-Means | For static images from the Cohn-Kanade Database, SVM achieved an accuracy of approximately 93%. In real-time emotion detection, various classification algorithms like Linear SVM achieved the highest accuracy of 94.1%, followed by Polynomial SVM with 91.2%. K-Means Clustering achieved accuracy of 88.7% . Linear SVM outperformed other algorithms, including k-means clustering and Random Forest, indicating its effectiveness in classifying facial expressions. | The automatic recognition of facial emotions using computer vision and machine learning algorithms is feasible, with support vector machines (SVM) showing highest results among the other algorithms. Linear SVM achieved the highest accuracy of around 94.1% in real-time emotion detection, indicating its effectiveness in classifying emotions accurately. |
| Estimation of Emotion using CNN[3] | Adaboost and Viola-Jones for detecting face in the image and CNN for emotion detection | Precision = 1928/(1928+1651) = 53.86  Recall = 1928/(1928+1651) = 53.86  F1 Score = 2x(precisionxRecall) /(Precision+Recall) =53.86  Accuracy = (TN+TP)/Total = 59.19 | In this paper emotion detection is performed using Viola-Jones and Adaboost for face detection, coupled with a convolutional neural network implemented in Keras and TensorFlow, which predicts emotions like happy, sad, disgust, surprise, angry, and neutral, evaluated on a Kaggle dataset with metrics such as accuracy, precision, recall, and F-score.The CNN is trained using 28709 images and tested using 7178 images of Kaggle dataset. |
| Face Expression Recognition Using OpenCV And CNN  [4] | CNN | The paper describes the convolution process in mathematics as a means to blend two functions, serving as a buffer to focus only on essential facts for function mapping. It requires two elements: input data and convolution filters (kernels), resulting in a feature map. | CNN is used for categorizing emotions from facial expressions shows promising results over statistical approaches. This advancement has the potential to significantly enhance human-machine interaction by enabling machines to understand and respond to human emotions more effectively. |
| A real time face emotion classification and recognition using deep learning model[5] | VGG-16 | The performance measures are validated with the VGG-16 model designed with an accuracy of 88%. The results prove that the network architecture designed has better advancements than existing algorithms. | The paper presents deep learning algorithms for accurate identification and emotion classification. It outlines a three-phase process: face detection using haar cascade, feature analysis with convolutional neural networks, and emotion classification. OpenCV, datasets, and Python are utilized, with real-time efficacy demonstrated through experiments on students' emotions. The system's performance is evaluated based on accuracy in face detection and recognition. |
| Face Detection and Recognition Using OpenCV[6] | SVM and  Haar Cascade with CNN | SVM achieved 83% (head detect) and Haar Cascade got a maximum of 95% in the problem of mask detection on face. | The paper provides an overview of OpenCV, an image and video processing library in computer vision.. It discusses common applications like face detection, recognition, and object detection, along with its role in tasks such as recognizing facial expressions, gender, and age. |
| Using CNN and Open CV, Mood  Identification with Face Feature Learning[7] | CCN | The model achieved satisfactory accuracy in emotion recognition, with results indicating an accuracy rate of nearly 83%. | The Convolutional Neural Network with OpenCV effectively recognized emotions from facial images. Using Python libraries like NumPy, OpenCV, and OneHot Encoder enhanced image processing and feature extraction, improving emotion recognition accuracy. CNN's multi-layered architecture accurately captured facial features and subtle emotional nuances. The model demonstrated proficiency in identifying a wide range of emotions from facial expressions in real time, suggesting practical applications in human-computer interaction. |
| Emotion Detection with Facial Feature Recognition Using  CNN & OpenCV.[8] | CNN | The Emotion Detection Recognition (EDR) software was capable of providing the actual results of the emotion within a few seconds.  The accuracy of the emotion detector was close to 80%. | The EDR system is capable of recognizing and displaying a variety of human emotions with a certain level of accuracy. The paper mentions that the accuracy of the emotion detector is close to 80%, which suggests that the system is fairly reliable in detecting emotions. The paper discusses the importance of understanding human behavior and emotions for decision-making and interaction with technology. |
| A Survey on Human Face Emotion Recognition  using Machine Learning Models[9] | CNN, Multiple Pipelines, Transfer Learning | The Convolutional Neural Network method achieved an accuracy level of 83.4%, making it the most accurate among the tested methods. Following CNN, the Multiple Pipelines method achieved an accuracy of 79.04%, while Transfer Learning yielded an accuracy of 78.5% | The results indicate that among the methodologies evaluated, the traditional CNN approach demonstrated the highest accuracy in recognizing human emotions from facial expressions. This suggests that the CNN architecture, with its ability to extract high-level representations from images, remains a strong contender for HFER tasks. The comparatively lower accuracies observed with Multiple Pipelines and Transfer Learning methods imply that while these approaches may offer certain advantages, such as flexibility in model architecture or leveraging pre-trained models, they may not always outperform the more established CNN method in this specific task. |

1.2Problem Statement

The challenge in human-computer interaction lies in accurately analyzing user sentiment in real time. Facial expressions offer a valuable cue, but existing systems often struggle to balance real-time performance with accurate sentiment classification. This project aims to develop a robust system that utilizes OpenCV for efficient facial detection and feature extraction from live video streams and provide the sentiment of users.

1.3Objectives

1. Implement an efficient system using OpenCV to capture live video.
2. Utilize CNNs to analyze the captured facial images and classify them into distinct sentiment categories such as happiness, sadness, anger, etc.

1.4System Requirements

1. **Hardware Requirements:** A computer with sufficient processing power to handle real-time video processing, ideally equipped with a multi-core CPU or GPU for enhanced performance.Webcam or camera for capturing live video streams for facial analysis.
2. **Software Requirements:**Operating System: The system should be compatible with major OS such as Windows, macOS, or Linux
3. **Development Environment:**Python programming language for implementing the sentiment analysis system.OpenCV library for computer vision tasks, including facial detection and feature extraction.Deep learning frameworks such as TensorFlow or PyTorch for building and training the sentiment analysis model.
4. **System Performance:**The system should be optimized for real-time performance, and capable of processing live video streams with minimal latency.Adequate memory and storage resources to accommodate the application, datasets, and trained models.

**SYSTEM ANALYSIS**

2.1Motivation

The motivation behind the development of the real-time sentiment analysis system arises from the increasing demand for efficient human-computer interaction systems. Traditional methods cannot often accurately analyze user sentiment in real-time, which is crucial for applications such as virtual assistants, Driver Assistance Systems, and Smart Home Systems. By leveraging facial expressions as a cue for sentiment analysis, the proposed system aims to bridge this gap and provide users with immediate feedback based on their emotional expressions.

2.2Proposed System

The proposed system utilizes computer vision techniques, particularly facial detection and feature extraction, to capture live video streams and extract facial expressions in real time. These facial expressions are then analyzed using convolutional neural networks (CNNs) to classify them into distinct sentiment categories, such as happiness, sadness, anger, etc. The system also includes a user-friendly interface that displays real-time sentiment analysis results, providing users with immediate feedback on the emotional expressions detected.

**DESIGN**

3.1Flowchart

YES

KNOWN

FACIAL EXPRESSION DETECTION

DATABASE

UNKNOWN

NO

START

IMAGE ACQUISITION USING CAMERA

FACE DETECTION USING

HAAR -CASCADE CLASSIFIER

END

EXTRACT FEATURES

IMAGE RECOGNITION USING CNN

IMAGE MATCH?

**References**

**IMPLEMENTATION**

4.1Methodology

1. **Data Collection and Preprocessing:**
   1. Dataset Acquisition: Collect a dataset of facial expression images with emotion labels.
   2. Data Augmentation: Enhance dataset diversity through techniques like flipping.
2. **Model Development:**
   1. CNN Architecture: Design a CNN model for emotion classification.
   2. Training: Train the CNN model using augmented data, optimizing hyperparameters.
3. **Real-time Emotion Detection:**
   1. OpenCV Integration: Integrate OpenCV for real-time facial detection.
   2. Model Inference: Utilize the trained CNN model for real-time emotion prediction.
4. **Evaluation and Validation:**
   1. Model Evaluation: Assess model performance on a test set with metrics and visualizations.
   2. Real-time Performance: Measure system latency and optimize for real-time inference.

4.2Algorithms

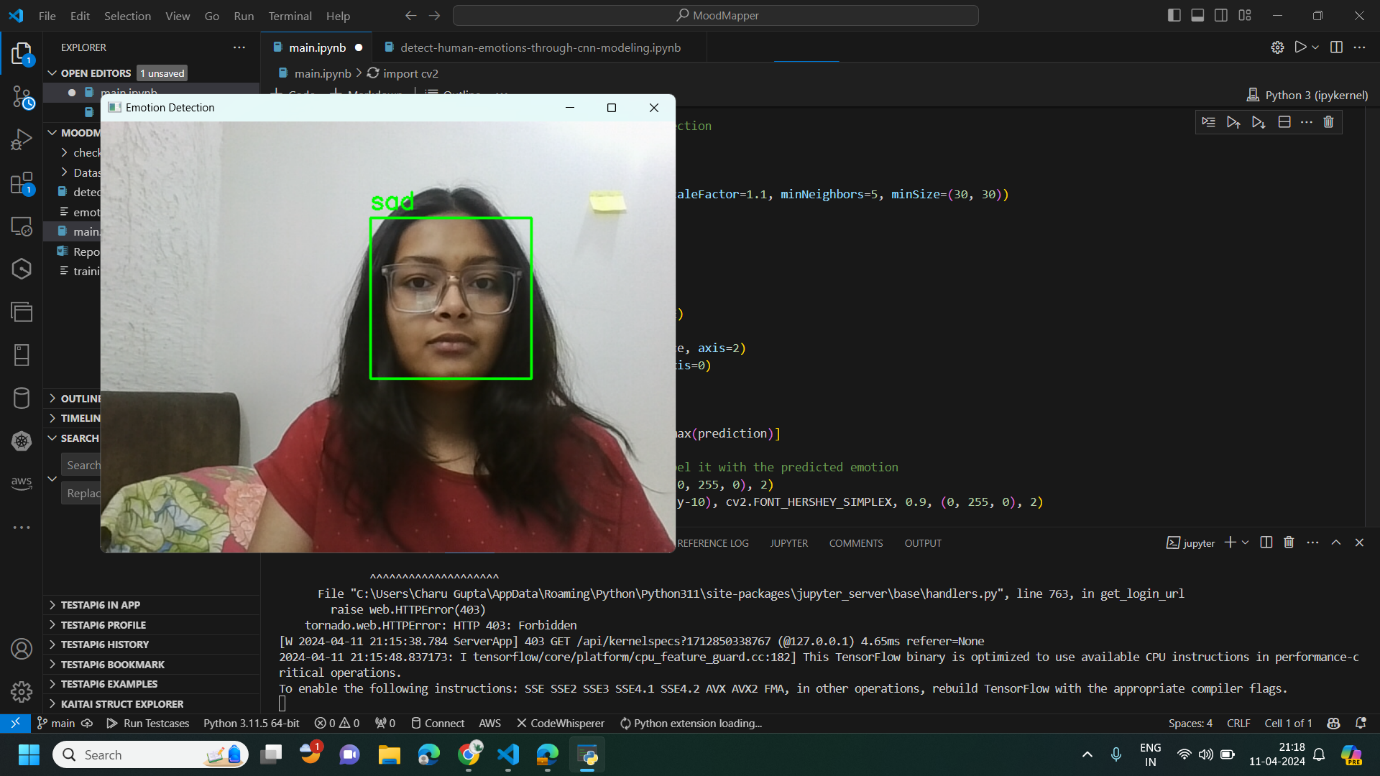
1. **Facial Detection Algorithms:**

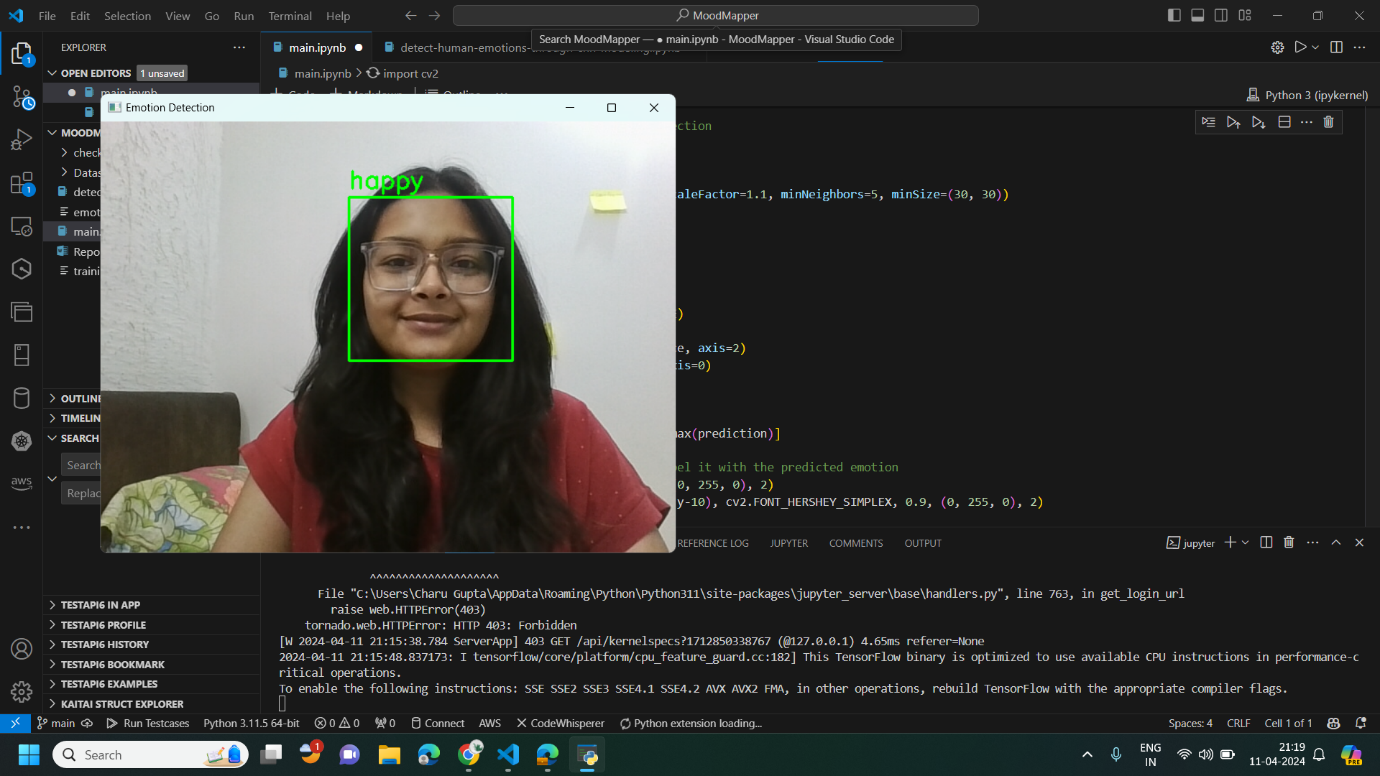
**Haar Cascades**: Haar cascades are a popular method for detecting objects in images, particularly faces. The algorithm involves training a cascade classifier with positive and negative image samples. Haar features are used to describe the characteristics of objects, and the classifier employs a series of stages to progressively narrow down the search for the target object. Once trained, the cascade classifier can efficiently detect faces in real-time by analyzing image regions at multiple scales. Haar cascades offer a computationally efficient solution for facial detection, making them suitable for real-time applications like sentiment analysis.

1. **Sentiment Analysis Algorithm:**

**Convolutional Neural Networks (CNNs):** In this project, CNN architectures are designed to process facial images and classify them into distinct sentiment categories such as happiness, sadness, anger, etc. The architecture typically consists of convolutional layers followed by max-pooling, batch normalization, and dense layers, with the final layer using softmax activation to output probability scores for each sentiment class.

4.3Screenshots





**Mathematical Model**

The model will describe the process from capturing a video frame to predicting the emotion of a detected face. During which the components involved are Haar cascades, preprocessing, and emotion classification using a Convolutional Neural Network (CNN).

### Workflow Overview

* **Capture Video Frame**: Capture a frame from the video stream.
* **Face Detection**: Use Haar cascades to detect faces in the frame.
* **Preprocessing**: Resize and normalize the detected face to match the input size expected by the CNN.
* **Emotion Classification**: Pass the pre-processed face through the CNN to predict the emotion.

### Mathematical Model

#### Step 1: Capture Video Frame

Let's denote the captured video frame as (V), where (V) is a 3D tensor representing the pixel intensities of the frame.

#### Step 2: Face Detection

For each detected face in the frame, we apply the Haar cascade classifier. The output of the Haar cascade classifier for a detected face is a 2D activation map (A). The activation map is computed by convolving the input image (V) with a set of Haar-like features (), where (i) indexes each feature. The mathematical representation of applying a Haar-like feature () to the image (V) is:

where:

* (Ai) is the activation map for feature (i),
* (() (x, y)) are the weights of the feature (i),
* are the pixel values of the image within the region defined by the feature (i),
* (w) and (h) are the width and height of the feature (i).

#### Step 3: Preprocessing

For each detected face, we resize it to match the input size expected by the CNN ((IMAGE\_SIZE)) and normalize the pixel values. The resized and normalized face is denoted as (X).

#### Step 4: Emotion Classification

The pre-processed face (X) is passed through the CNN to predict the emotion. The CNN consists of several convolutional layers, pooling layers, and fully connected layers. The output of the CNN is a probability distribution over the classes (emotions), denoted as (P).

The mathematical representation of the CNN can be summarized as follows:

* **Convolutional Layer**:

where (A) is the activation map, (F) are the filter coefficients, (X) is the input volume, and (f) is the filter size.

* **Pooling Layer**:

where (y) is the output element, and (x) represents the input elements within the pooling window.

* **Fully Connected Layer**:

**)**

where () is the output of neuron (i), are the weights connecting neuron (j) in the input layer to neuron (i) in the output layer, () are the inputs to the neuron, () is the bias term for neuron (i), and () is the activation function (e.g., ReLU).

The final output of the CNN, (P), is a probability distribution over the classes (emotions), which can be represented as:

***P = softmax (WX + b)***

where:

* (P) is the probability distribution over the classes,
* (W) are the weights of the fully connected layer,
* (X) is the input to the fully connected layer (the pre-processed face),
* (b) is the bias term for the fully connected layer,
* Softmax function is which converts the output of the fully connected layer into a probability distribution.

**REFERENCES**

[1] Mehendale, N. Facial emotion recognition using convolutional neural networks (FERC). SN Appl. Sci. 2, 446 (2020). https://doi.org/10.1007/s42452-020-2234-1

[2]S. Gupta, "Facial emotion recognition in real-time and static images," 2018 2nd International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, 2018, pp. 553-560, doi: 10.1109/ICISC.2018.8398861

[3]K. M. Goud and S. J. Hussain, "Estimation of Emotion using CNN," 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2021, pp. 1561-1565, doi: 10.1109/ICESC51422.2021.9532763.

[4]Mahanthi, Lakshmi & Sesetti, Anuradha & Manjeti, Vijaya & Publication Hub, Tjcse. (2023). FACE EXPRESSION RECOGNITION USING OPEN CV AND CONVOLUTIONAL NEURAL NETWORK. 10.13140/RG.2.2.27984.69128.

[5]Shaik Asif Hussain and Ahlam Salim Abdallah Al Balushi 2020 J. Phys.: Conf. Ser. 1432 012087

[6]TH. Hasan, R., & Bibo Sallow, A. . (2021). Face Detection and Recognition Using OpenCV. Journal of Soft Computing and Data Mining, 2(2), 86-97.

[7]H. L. Sharma and M. Sharma, "Using CNN and Open CV, Mood Identification with Face Feature Learning," 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART), Moradabad, India, 2022, pp. 1109-1112, doi: 10.1109/SMART55829.2022.10047674.

[8]S. Giri et al., "Emotion Detection with Facial Feature Recognition Using CNN & OpenCV," 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2022, pp. 230-232, doi: 10.1109/ICACITE53722.2022.9823786.

[9]S. Yuvaraj, J. V. Franklin, V. S. Prakash, A. Anandaraj and R. Subha, "A Survey on Human Face Emotion Recognition using Machine Learning Models," 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2023, pp. 1993-1996, doi: 10.1109/ICACCS57279.2023.10112955.