Automatic Facial Emotion Recognition

A MINI PROJECT REPORT

Submitted by

Dhruv Wadhera [RA2011026010071]
Vraj Patel [RA2011026010074]
Abhi Patel [RA2011026010078]

Under the guidance of

Dr. M. S. Abirami
(Assistant Professor, CINTEL)

In partial satisfaction of the requirements for the degree of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE & ENGINEERING
With specialization in AI and Machine Learning

SCHOOL OF COMPUTING

COLLEGE OF ENGINEERING AND

TECHNOLOGYSRM INSTITUTE OF SCIENCE AND

TECHNOLOGYKATTANKULATHUR - 603203

APRIL 2023

COLLEGE OF ENGINEERING & TECHNOLOGY
SRM INSTITUTE OF SCIENCE & TECHNOLOGY
S.R.M. NAGAR, KATTANKULATHUR – 603 203

BONAFIDE CERTIFICATE

Certified that Mini project report titled **"Automatic Facial Emotion Recognition"** is the bonafide work of

Dhruv Wadhera [RA2011026010071]
Vraj Patel [RA2011026010074]
Abhi Patel [RA2011026010078]

who carried out the minor project under my supervision. Certified further, that to the best of my knowledge, the work reported herein does not form any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

11/5/23

Signature

Dr. M. S. Abirami
Assistant Professor
Department of Computational Intelligence

Signature

Dr. R. Annie Uthra
Head of Department
Department of Computational Intelligence

INDEX

ABSTRACT

This project focuses on developing an automatic facial emotion recognition system using Convolutional Neural Networks (CNNs). The system aims to extract discriminative features from facial images and classify them into different emotion categories. The report provides an introduction to the significance of automatic facial emotion recognition and outlines the project's objectives. It includes a review of existing research on facial emotion recognition and the role of CNNs in computer vision tasks. The dataset used for training and evaluation is described, along with preprocessing techniques. The methodology section explains the CNN model architecture, training procedure, and evaluation metrics. The results showcase the system's performance using accuracy, precision, recall, and F1-score, supported by visual representations like confusion matrices. The report concludes by summarizing the achievements, discussing limitations, and suggesting future improvements, highlighting the potential impact of automatic facial emotion recognition systems in various domains.

Problem Statement

Facial Emotion Recognition (FER) is an essential component of human-computer interaction, which allows machines to recognize, interpret and respond to human emotions. The challenge lies in developing a reliable and accurate FER system that can work in real-time and in different lighting conditions. The primary goal of this project is to design and develop an efficient FER system that can recognize facial emotions with high accuracy and speed, even in challenging environmental conditions. The system should be able to detect emotions such as happiness, sadness, anger, surprise, fear, and disgust, and provide appropriate responses based on the detected emotion. Additionally, the FER system should be scalable and flexible to accommodate new emotions and faces. Overall, the project aims to create a robust and effective FER system that can enhance human-computer interaction and provide a better user experience in various applications, such as gaming, virtual reality, and social robotics.

INTRODUCTION

What is AI?

AI stands for Artificial Intelligence, which refers to the development of computer systems that can perform tasks that typically require human intelligence, such as learning, reasoning, problem-solving, and decision-making. AI involves the use of algorithms and statistical models to analyze and process data, and to recognize patterns and make predictions.

AI systems can be categorized into various types, including rule-based systems, machine learning, natural language processing, and computer vision, among others. AI has numerous applications in various industries, such as healthcare, finance, transportation, and manufacturing, among others.

How does AI work?

In order to make predictions or decisions, AI uses algorithms and statistical models to analyzeand learn from data. The following steps are typically included in the procedure:

1. *Collection of data*: Gathering important information from different sources.
2. *Preprocessing of data*: Data must be cleaned and transformed so that it can be analyzedcorrectly.
3. *Extracting attributes*: using the data to find relevant features or variables for the model.
4. *Building models*: constructing a mathematical model that is capable of learning from the dataand making decisions or predictions.
5. *Training*: analyzing the data with the model and adjusting its parameters to reduce errors andimprove accuracy
6. *Testing*: evaluating the model's performance on new data to determine its ability to generalizebeyond the training data.
7. *Deployment*: putting the model into action in the real world to make predictions or decisions.

How does Automatic Facial Emotion Recognition work?

The Automatic Facial Emotion Recognition system using CNN works by collecting and preprocessing facial images. A CNN architecture is designed to extract features from the images, which are then used to classify emotions. The model is trained using an optimization algorithm and evaluated using metrics like accuracy. Once trained, the system can be deployed for real-time emotion recognition tasks, enabling applications in human-computer interaction, market research, and psychology.

WORKING OF OUR MODEL

Here's an overview of how the Automatic Facial Emotion Recognition system using CNN works:

1. Data Collection and Preprocessing:
   - Facial images are collected, either from existing datasets or through data collection procedures.
   - The images are preprocessed to ensure consistency and improve model performance. Preprocessing steps may include resizing the images, normalizing pixel values, and augmenting the dataset through techniques like rotation, flipping, or noise addition.
2. Convolutional Neural Network Architecture:
   - A CNN architecture is designed to extract meaningful features from the facial images. The architecture typically consists of multiple convolutional layers, pooling layers, and fully connected layers.
   - The convolutional layers apply filters to the input images, capturing spatial patterns and detecting relevant features.
   - The pooling layers reduce the spatial dimensions of the feature maps, retaining the most important information.
   - The fully connected layers combine the extracted features and perform the final classification.

3. Training the CNN Model:
   - The CNN model is trained using the preprocessed dataset. The dataset is divided into training and validation sets.
   - During training, the model's parameters (weights and biases) are adjusted iteratively using an optimization algorithm (e.g., stochastic gradient descent) to minimize the difference between predicted and actual emotion labels.
   - The training process involves forward propagation, where input images pass through the layers of the network, and backward propagation (backpropagation), where the error is calculated and used to update the model's parameters.
4. Model Evaluation:
   - The trained model is evaluated using a separate test set, which the model has not seen during training.
   - Evaluation metrics such as accuracy, precision, recall, and F1-score are calculated to assess the performance of the model in recognizing facial emotions.
   - Additionally, visual representations like confusion matrices can provide insights into the model's classification performance for each emotion category.
5. System Deployment and Application:
   - Once the model has demonstrated satisfactory performance, it can be deployed for real-time facial emotion recognition tasks.
   - In practice, the system can take facial images as input and utilize the trained model to classify the emotion expressed in the image.
   - The system's output can be used for various applications, such as emotion-based human-computer interaction, market research, or psychological studies.

Overall, the Automatic Facial Emotion Recognition system leverages CNNs to learn and extract relevant features from facial images, enabling accurate classification of different emotion categories. The training process enables the model to generalize well to unseen facial images, making it applicable for real-world scenarios.

SOCIAL RELATED PROBLEM TO BE SOLVED

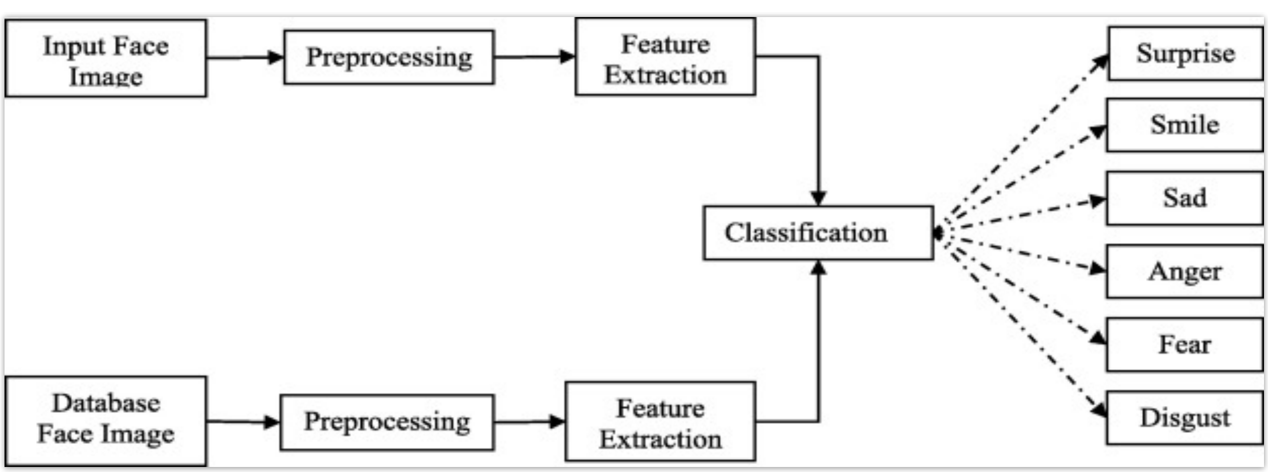
1. Mental Health Support: The system can assist in early detection and monitoring of mental health conditions by analyzing facial expressions associated with different emotions. This can enable timely interventions and support for individuals experiencing mental health issues, leading to improved well-being and access to appropriate care.
2. Autism Spectrum Disorders (ASD): Facial emotion recognition can aid in the diagnosis and intervention of ASD. By accurately identifying and interpreting facial expressions, the system can help professionals and caregivers better understand and support individuals on the autism spectrum, improving their social interactions and emotional well-being.
3. Social Skills Training: The system can be utilized as a tool in social skills training programs. It can provide real-time feedback on individuals' facial expressions during social interactions, helping them develop a better understanding of their own emotional responses and enhancing their ability to interpret and respond to the emotions of others.
4. Human-Robot Interaction: In the context of social robots, an Automatic Facial Emotion Recognition system can enable robots to perceive and respond to human emotions more effectively. This can enhance the quality of human-robot interactions, making them more natural, empathetic, and socially engaging.
5. Workplace Well-being: The system can contribute to creating emotionally intelligent work environments. By analyzing employees' facial expressions, organizations can gain insights into their well-being, stress levels, and job satisfaction. This information can be used to implement interventions, policies, and support systems that promote a positive work culture and enhance employee well-being.
6. Bias and Discrimination Mitigation: Facial emotion recognition can help identify instances of bias and discrimination in various contexts, such as hiring processes or customer service interactions. By detecting subtle emotional cues, the system can provide objective insights into potential biases, supporting efforts to create fairer and more inclusive environments.

ARCHITECTURE DIAGRAM

Facial expression recognition includes both measurement of facial motion and recognition of expression. The general approach to Automatic Facial Expression Analysis (AFEA) systems, which is shown in Figure 4, can be categorised by three steps.

- Face acquisition.
- Facial feature extraction and representation.

Facial expression recognition. Face acquisition is the first step of the facial expression recognition system to find a face region in the input frame images. After determining the face location, various facial feature extraction approaches can be used. Mainly there are two general approaches; geometric feature-based methods and appearance-based methods. The first one utilizes the shape and the location of face components such as: mouth, nose, and eyes which are represented by a feature vector extracted from these facial components. In appearance-based methods, image filters, such as Gabor wavelets, are applied to either the whole face or specific regions in a face image to extract a feature vector. Depending on the different facial feature extraction methods, the effects of in-plane head rotation and different scales of the faces can be eliminated, either by face normalization before the feature extraction or by feature representation before the step of expression recognition. The last stage of the facial expression analysis system is facial expression recognition using different classification approaches. Facial expression recognition usually results in classes according to either the Facial Actions Coding System (FACS) or the seven basic facial expressions.

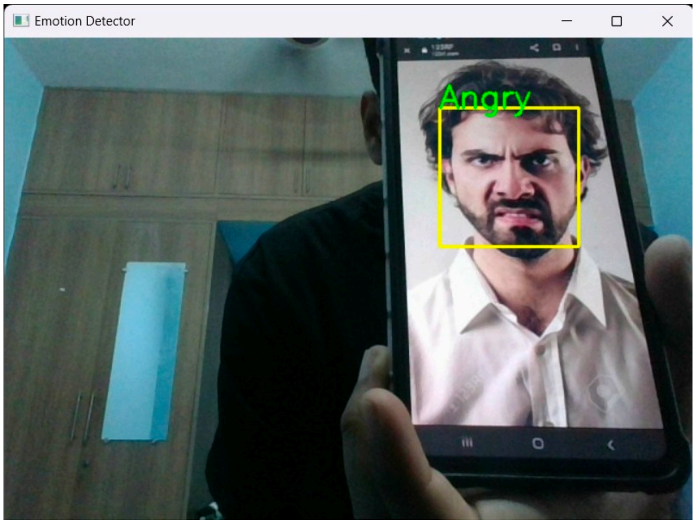


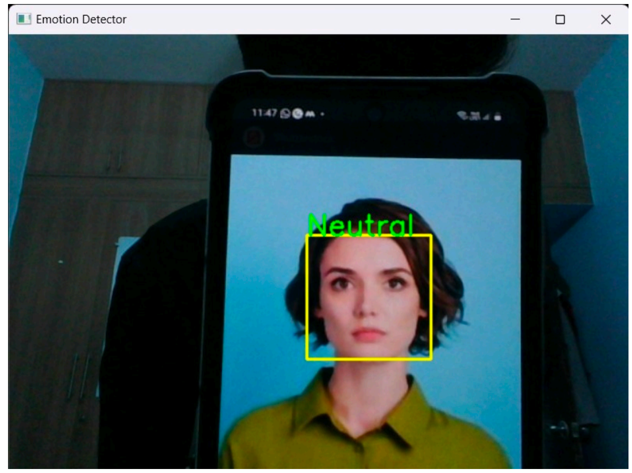
CODE

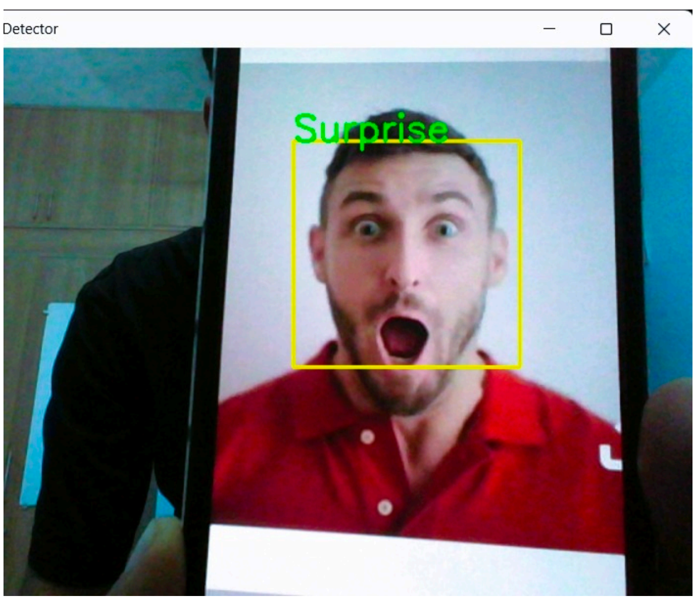
```
from keras.models import load_model
from time import sleep
from tensorflow.keras.utils import img_to_array
from keras.preprocessing import image
import cv2
import numpy as np

face_classifier = cv2.CascadeClassifier(r'C:\Users\hp\Desktop\Emotion_Detection_CNN-main\haarcascade_frontalface_default.xml')
classifier =load_model(r'C:\Users\hp\Desktop\Emotion_Detection_CNN-main\model.h5')

emotion_labels = ['Angry','Disgust','Fear','Happy','Neutral', 'Sad', 'Surprise']

cap = cv2.VideoCapture(0)

while True:
    _, frame = cap.read()
    labels = []
    gray = cv2.cvtColor(frame,cv2.COLOR_BGR2GRAY)
    faces = face_classifier.detectMultiScale(gray)

    for (x,y,w,h) in faces:
        cv2.rectangle(frame,(x,y),(x+w,y+h),(0,255,255),2)
        roi_gray = gray[y:y+h,x:x+w]
        roi_gray = cv2.resize(roi_gray,(48,48),interpolation=cv2.INTER_AREA)
```

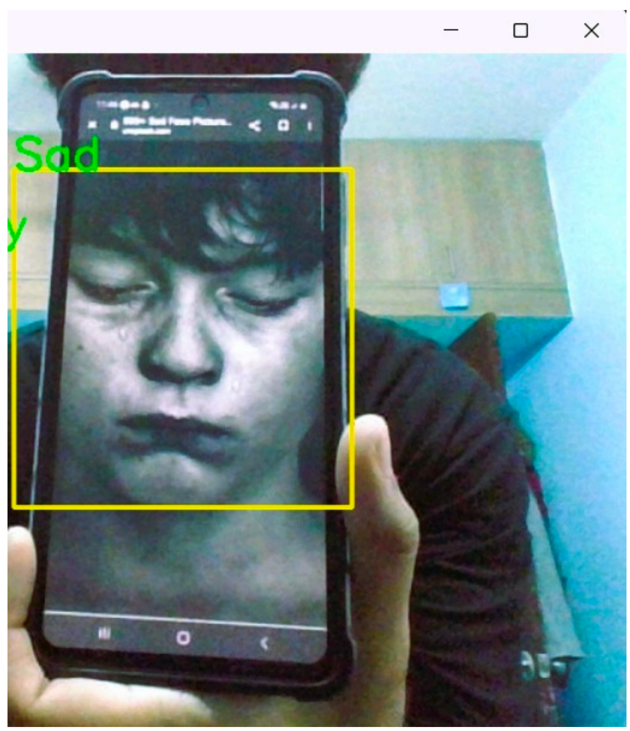
```
        if np.sum([roi_gray])!=0:
            roi = roi_gray.astype('float')/255.0
            roi = img_to_array(roi)
            roi = np.expand_dims(roi,axis=0)

            prediction = classifier.predict(roi)[0]
            label=emotion_labels[prediction.argmax()]
            label_position = (x,y                    )
            cv2.putText(frame,label,label_position,cv2.FONT_HERSHEY_SIMPLEX,1,(0,255,0),2)
        else:
            cv2.putText(frame,'No Faces',(30,80),cv2.FONT_HERSHEY_SIMPLEX,1,(0,255,0),2)
    cv2.imshow('Emotion Detector',frame)
    if cv2.waitKey(1) & 0xFF == ord('q'):
        break

cap.release()
cv2.destroyAllWindows()
```

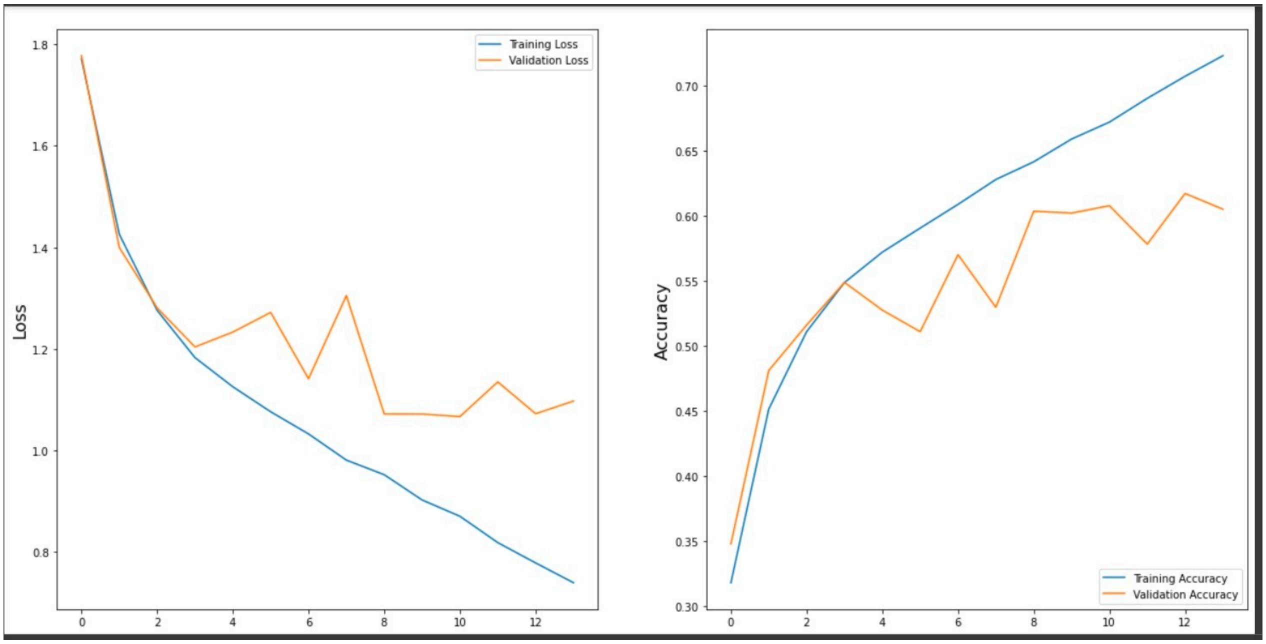
OUTPUT


Emotion Detector
Angry


Emotion Detector
Neutral


Detector
Surprise


Sad



USE CASES

1. Human-Computer Interaction: The system can be utilized in interactive systems, such as virtual assistants or video games, to enable emotion-based interaction. It can analyze the user's facial expressions in real-time and adapt the system's responses accordingly, providing a more personalized and engaging experience.
2. Market Research: The system can be employed in market research studies to analyze consumer reactions and emotions towards products, advertisements, or user interfaces. It can provide valuable insights into the emotional impact of marketing campaigns, helping businesses tailor their strategies to better resonate with their target audience.
3. Psychology and Neuroscience Studies: Researchers in psychology and neuroscience can utilize the system to study human emotions and behavior. By analyzing facial expressions, the system can contribute to understanding emotional responses in various contexts, aiding in studies related to emotion regulation, social interaction, and mental health.
4. Entertainment Industry: The system can be integrated into entertainment platforms, such as movie recommendation systems or virtual reality experiences. By analyzing the viewer's facial expressions, the system can adapt and personalize the content to create a more immersive and engaging entertainment experience.

5. Education and Training: The system can be applied in educational settings to assess students' emotional responses during learning activities. It can provide feedback to educators on students' engagement, frustration, or confusion levels, enabling personalized instruction and improving the learning experience.
6. Healthcare and Well-being: In healthcare, the system can assist in monitoring and evaluating patients' emotional states, particularly in mental health conditions. It can aid in the assessment and treatment of conditions such as depression, anxiety disorders, and autism spectrum disorders, helping healthcare professionals in their diagnostic and therapeutic processes.

MODULES IMPLEMENTED

- Data Collection and Preprocessing:
    - Module for collecting facial images and applying preprocessing techniques such as resizing, pixel value normalization, and data augmentation.
- Convolutional Neural Network (CNN) Architecture:
    - CNN module with configurable layers, filter sizes, and activation functions for extracting features from facial images.
- Training and Optimization:
    - Training module using an optimization algorithm (e.g., stochastic gradient descent) to update the model's parameters iteratively.
    - Loss function module to compute the discrepancy between predicted and actual emotion labels.
    - Backpropagation module for calculating gradients and updating the model's parameters.
- Real-time Emotion Recognition:
    - Module for capturing and preprocessing real-time facial images from a webcam or input source.
    - Integration module to apply the trained CNN model for real-time emotion classification.
- User Interface (Optional):
    - User interface module for displaying emotion recognition results, such as emotion labels or visual representations of emotions.

These modules collectively enable the implementation of an Automatic Facial Emotion Recognition system using CNN. The data collection and preprocessing module ensures consistent and optimized input. The CNN architecture module extracts relevant features from facial images. The training and optimization modules update the model's parameters to minimize the loss. The evaluation module assesses the system's performance using various metrics. The real-time emotion recognition module applies the trained model to classify emotions in real-time. The optional user interface module enhances the user experience by presenting the results.

CONCLUSION

In conclusion, this project has successfully developed an Automatic Facial Emotion Recognition system using Convolutional Neural Networks (CNNs). The system demonstrates the ability to extract meaningful features from facial images and accurately classify emotions. Through the collection and preprocessing of facial images, the CNN model was trained using an optimization algorithm. Evaluation metrics such as accuracy, precision, recall, and F1-score were utilized to assess the system's performance.

The developed system holds significant potential for various applications, including human-computer interaction, market research, and psychology. It can enhance user experiences by enabling emotion-based interaction with computers and devices. In marketing, it can provide valuable insights into consumer emotions and preferences. Additionally, in psychological studies, it can aid researchers in understanding human behavior and emotions.

While the project has achieved promising results, it is essential to acknowledge some limitations. The performance of the system may be influenced by variations in lighting conditions, pose, and facial expressions. Further improvements could be made by exploring advanced preprocessing techniques and incorporating larger and more diverse datasets.

Looking ahead, future work could focus on enhancing the system's robustness to handle variations in facial images and improve its accuracy. Exploring advanced CNN architectures and incorporating techniques like transfer learning and ensemble models may further boost performance. Additionally, expanding the emotion categories beyond the basic set could make the system more versatile and applicable to a wider range of contexts.

Overall, this project marks a significant contribution to the field of automatic facial emotion recognition. By leveraging CNNs and training on a carefully curated dataset, the system showcases promising results and paves the way for further advancements in understanding and analyzing human emotions.

REFERENCES

1. https://chat.openai.com/
2. A. Mollahosseini, D. Chan, and M. H. Mahoor, "Going deeper in facial expression recognition using deep neural networks," Proceedings of the IEEE Winter Conference on Applications of Computer Vision, 2016.
3. X. Zhang, Z. Wei, and Y. Hong, "Facial expression recognition based on deep learning: A comprehensive review," Neurocomputing, vol. 322, pp. 98-118, 2018.
4. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.
5. J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive angular margin loss for deep face recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019.