Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Emotion Detection Using Visual Expression Analysis

Zaina MH, Dr. Cini Kurian

Department of Computer Science, Al-Ameen College, Edathala, Mahatma Gandhi University, India

Emotion recognition encompasses the process of discerning an individual's emotional condition. The spectrum of human emotions is characterized by both complexity and simplicity. Humans, as the pinnacle of the food chain, have the extraordinary capacity to express emotions through a blend of communication methods such as speech, writing, body language, gestures, and facial expressions. Different forms of human expression like text, physiology, audio, and video are used to discern emotional states. This can be challenging because they are sometimes evaluated based on human perception. Nearly every decision we make in life is influenced by our emotions. Predicting emotions accurately can greatly enhance emotional intelligence and promote personal development. Advances in machine and deep learning have enabled computers and machines to understand, recognize, and evaluate emotions. This concept falls under cognitive systems in data and machine learning. It involves mimicking the human cognitive process using training data, such as images and videos of humans, with the goal of identifying and classifying the emotions shown in the visual data.

1. Introduction

Visual emotion recognition systems use various techniques to study and classify facial expressions to determine the emotions being shown.. Visual expression recognition (VER) is intricate and demanding, yet it proves beneficial in diverse fields like healthcare, emotionally responsive robots, and human-computer interaction [1]. Although advancements in VER improve its effectiveness, achieving a high degree of accuracy continues to be a formidable task.. The six fundamental emotions experienced by humans are anger, happiness, sadness, disgust, fear, and surprise. In addition, the emotion known as contempt was included among the fundamental emotions [2]. The potential for visual emotion analysis is substantial, as understanding human emotions is crucial for the advancement of artificial intelligence. The prospects for visual emotion analysis are significant, given that comprehending human emotions is essential for AI development. The rapid progress of Convolutional neural networks (CNN), deep learning has become the preferred method for emotion analysis tasks[13]. The primary goal of study is to create a system that can recognize various human emotions by analyzing facial expressions, ultimately improving human-computer interaction (HCI)[14].

2. METHOD

Analyzing emotions entails capturing the image frame from a camera feed (IP, CCTV, USB camera), preprocessing the image (cropping, resizing, rotating, color correction), extracting crucial features using a model, and conducting emotion classification. For our analysis, we will focus on pre-recorded images and videos that depict emotions. A webcam has the ability to detect the user's expressions in order to understand their emotional state. This can also be valuable for assessing one's emotional resilience. The examination of human sentiments, also known as opinion mining or Emotion AI[12]. Neuroscience is the exploration of various states of the human brain[13]. The main goal of sentiment analysis is to assess the sentiment of the input and to analyze the polarity of the input, whether it is an image, video, facial expression, or any other form of data. To ascertain the predominant sentiment conveyed as positive, negative, or neutral. Visual expressions are indicators of one's emotional state such as joy, astonishment, sorrow, apprehension, revulsion, and ire collectively known as the seven universal facial expressions of emotion [3]. After conducting the initial analysis, programs often conduct a thorough analysis to identify emotions like enjoyment, happiness, disgust, anger, fear, and surprise. Two factors precede this analysis. The first step involves quantifying the input data for algorithms to analyze and interpret, while the second step focuses on utilizing psychological research to determine which expressions correspond to specific emotions.

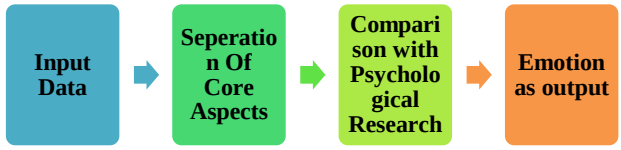

Fig.1 Demonstrates an analysis of sentiments performed programmatically.

The field of Cognitive Science focuses on exploring the intricacies of the human brain and unravelling the fundamental principles underlying intelligence. By leveraging human intelligence, computer systems are designed with the aspiration that they will mimic learning and cultivate intelligent behaviours akin to humans. Cognitive Science functions across three separate levels of analysis. In the computational theory, the analysis objectives are clearly defined at this stage and then input into the computer system. Representation and algorithms are commonly referred to as the training stage in the field of Machine Learning. The machine is provided with optimal input and output scenarios, and algorithms are implemented to facilitate the transformation process from input to output.

The hardware implementation marks the culmination of the cognitive process. It involves implementing the algorithm in real-life scenarios and evaluating its performance compared to that of a human mind.

A. Methodology Process

- Face Detection in Images and Video Frames is covered in this section.

Initially, the camera's video is employed to identify and pinpoint the human face. The bounding box coordinate is utilized to pinpoint the precise location of a face in real-time. The task of face detection remains quite challenging as there is no assurance that every face within a particular input image will be successfully detected [4], particularly in uncontrolled settings

with difficult lighting, varying head angles, significant distances, or instances of obstruction.

- Recognition of visual expressions from images.

Visual expressions play a crucial role in communication for both humans and animals. Facial expressions provide a simple yet effective way to study human behaviors and psychological traits. It is commonly utilized in various medical treatments and therapies as well. In this part, we'll focus on analyzing facial expressions and portraits to interpret the emotions conveyed in the images. Later these stages will be followed again with a video-based input.

Fig.2 Stages of Image classification in emotion detection

B. Functioning of the Proposed System

The VER system collects input data from images and streaming videos to read and store the visual emotions dataset. The saved frames adjust the size of the captured image to create a suitable frame for streaming purposes. It then identifies the faces recorded in the video and proceeds to extract the facial area. Once the extraction process is completed, the classification is carried out by identifying one of the seven fundamental emotions displayed by the person on the webcam.

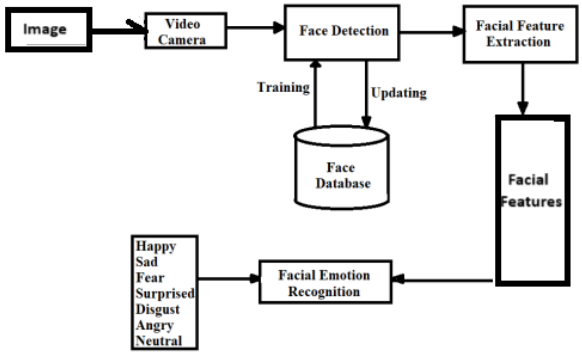

Fig.2 The system architecture employed for recognizing visual emotions.

The VER system is designed to detect the emotions on a person's face shown on the webcam, allowing the camera to then concentrate solely on capturing the facial image. Image processing techniques, such as Neural Networks, are utilized to identify the type of expression displayed by a person, along with determining the accuracy of that specific emotion category.

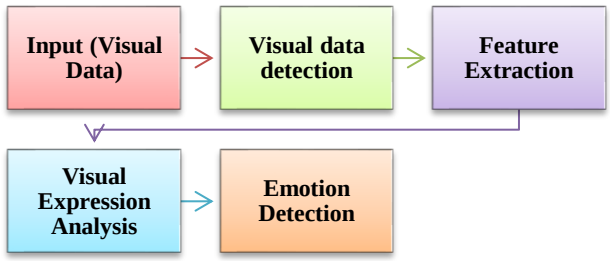

Fig.3 Block Diagram of proposed suggested System

C. System Implementation

- Performing image preprocessing.

Once the faces are recognized, the image data undergoes optimization prior to being input into the emotion classifier. This measure significantly enhances the precision of detection. Image preprocessing typically involves several sub steps to adjust the image for changes in lighting, remove noise, and apply techniques like image smoothing, rotation correction, resizing, and cropping for enhancement.

- Feature Extraction

Following pre-processing, the pertinent features are extracted from the pre-processed data that includes the identified faces. Recognition of visual expressions is likely due to the movement of facial muscles, causing changes in the appearance of facial features. Precisely extracting features poses a significant challenge that demands thorough examination for a capable facial expression recognition system. Various techniques are available for identifying a wide range of facial attributes. Consider, for instance, Action Units (AU), facial landmark movements, distances between facial features, gradient characteristics, facial texture, and other aspects. Typically, AI emotion recognition classifiers rely on Support Vector Machines (SVM) or Convolutional Neural Networks (CNN). Ultimately, the identified human face is categorized according to the facial expression employed for emotion classification with the help of extracted features.

- Emotion Classification

Following pre-processing, the AI model extracts the pertinent features from the pre-processed data that includes the identified faces. Various techniques can be used to identify a variety of facial characteristics. For instance, Action Units (AU), facial landmark movements, facial landmark distances, gradient characteristics, facial texture, and various other factors. The classifiers utilized in AI emotion recognition are founded on Support Vector Machines (SVM)[6] or Convolutional Neural Networks (CNN) [5]. The identifiable human face is categorized through facial expressions, with pre-determined emotional classes being assigned.

The widely recognized Face Emotion Recognizer, commonly referred to as FER, is an open-source Python library utilized for analyzing the sentiments displayed in images and videos. The

Facial Expression Recognition system is developed with a version utilizing a Convolutional neural network that incorporates the given weights from the HDF5 data file when building the FER system's model. One way to override this is by utilizing the FER constructor during the model's invocation and initialization.

The MTCNN (multi cascade Convolutional network) is one of the parameters of the constructor. Identifying faces is a method utilized in recognizing individuals. The MTCNN model will be utilized for detecting faces else the function will default to use the OpenCV Haarcascade classifier. The function detect emotions is designed to categorize emotions with options including 'fear', 'neutral', 'happy', 'sad', 'anger', and 'disgust'. Each emotion is carefully assessed and then assigned a numerical value on a scale ranging from 0 to 6. The program initiates by receiving the image or video requiring analysis. To initialize the FER constructor, you only need to supply a face detection classifier. Next, we invoke the detect emotions function of this constructor by providing it with the input object (image or video). The outcome obtained presents a range of emotions, each accompanied by a specific value. In the end, the top emotion function is able to isolate the most prominent emotion detected in the object and provide it as output.

- Identifying emotions displayed on faces in videos.

In this section, just like how we process images to extract sentiment, we will be focusing on working with videos. In essence, a video is formed by seamlessly stringing together a series of consecutive image frames to create motion. Basically, the operation of algorithms for videos and images is fundamentally identical. An extra step involved in video processing is to divide the video into its individual frames and then apply image processing algorithms to each frame. Though the fundamental algorithm remains consistent between images and videos, there are several crucial adjustments tailored specifically for videos. The Video analyze function smoothly handles the extraction of individual image frames from a video for independent analysis. Each frame processed by this function is saved as an individual image in the main directory folder while the code is in operation. Additionally, this function generates a duplicate of the initial video by framing the face with a box and displaying real-time emotions captured within the video. The Pandas Data Frame is generated from the analyzed values, followed by plotting this data frame using matplotlib, displaying each emotion is depicted over time. We can enhance our examination of this data frame by extracting individual emotion values identified by the model, and determining the prevalent sentiment throughout the entire video. In this manner, we have the capacity to analyze videos by isolating specific image frames and examining them. The diagram below illustrates the process, demonstrating the incorporation of an extra step for video processing. We'll encounter this implementation in the section that follows. The fig.4 demonstrates the ability to identify emotions displayed on faces in images and Videos.

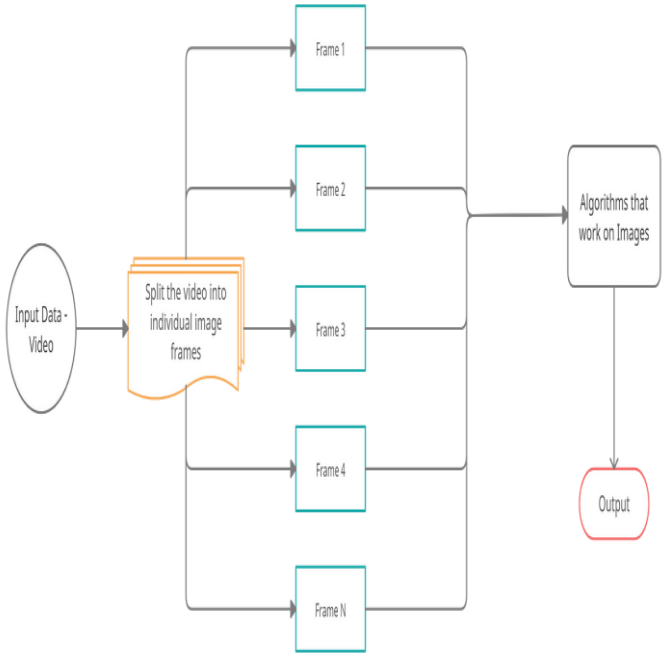

Fig.4 The Image Processing Algorithm and extending it to work for Videos

3. Result and Discussions

The code elegantly processes images individually, extracting diverse emotions along with their corresponding intensity levels as the output. We can utilize the top emotion function to identify the prevailing sentiment portrayed in the image.

TABLE1. EMOTIONAL INTENSITIES DERIVED FROM THE OUTCOME

Label	Human Emotion	Emotion Value
0	Angry	64
1	Disgust	58
2	Fear	60
3	Happy	84
4	Sad	53
5	Surprise	64
6	Neutral	74

We have just witnessed the fascinating process of examining images to identify the emotions and facial expressions of people captured within them. Next, we will proceed with a similar analysis, this time utilizing videos.

We now wrap up our examination of both images and videos for Emotion Recognition. We

successfully engaged with human faces to comprehend the emotions conveyed through facial expressions.

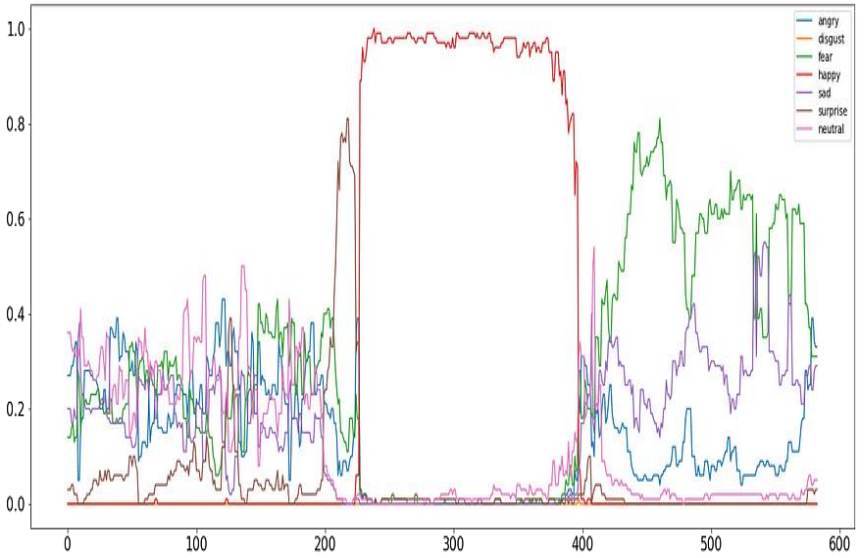

Fig.5 The sequence of Emotions throughout the video's length is plotted on a graph

The plotted line graph illustrates how the intensity of different emotions changes over time. each emotion is represented by a line with a specific color. The table below shows the emotion and the color plotted for the emotion.

TABLE 2: THE EMOTIONS PLOTTED ALONG WITH THEIR CORRESPONDING COLORS

EMOTION DETECTED	COLOR
ANGRY	BLUE
DISGUST	GREEN
HAPPY	RED
SAD	PURPLE
SURPRISE	PINK
NEUTRAL	BLACK

The graph effectively illustrates the temporal elaboration of colorful feelings, with happiness arising as the most current and violent emotion. in discrepancy, fear and surprise show notable peaks, pressing moments of significant emotional impact. The harmonious neutral state provides a stable reference for interpreting other feelings oscillations. this comprehensive view can be abused for operations in psychology, stoner experience exploration, and sentiment analysis, offering precious perceptivity into mortal emotional dynamics.

TABLE 3: TIME-INTENSITIES RELATIONSHIP

Time Interval	Angry	Disgust	Fear	Happy	Sad	Surprise	Neutral
1-5 seconds	60%	55%	50%	80%	45%	60%	70%
5-10 seconds	65%	60%	55%	85%	50%	65%	75%
10-15 seconds	70%	60%	60%	90%	55%	70%	80%
15-20 seconds	65%	60%	65%	85%	60%	65%	75%
20-25 seconds	60%	55%	70%	80%	55%	60%	70%

Table 3 provides a detailed breakdown of each emotion's intensity over different time intervals. The intensities represent the likelihood or intensity of each emotion being detected at a given time. This structured representation can be helpful in visualizing how emotions evolve over time and understanding their fluctuations.

4. CONCLUSION

Sentiment analysis and image detection have numerous real-world applications in today's society. Object detection algorithms can be observed in various settings such as public parking lots and traffic monitoring systems. Capturing images of individuals operating vehicles for record-keeping purposes. Sentiment analysis is also utilized in therapy when in-person sessions between the therapist and patient are unfeasible. Research into human cognition has contributed to the advancement of medication. In the realm of technology, there are virtual assistants, profile evaluation assistants, and automation bots that have been designed to imitate human actions in order to enhance precision and reduce mistakes [8][9]. Consequently, it plays a significant role in the contemporary world driven by advancements in artificial intelligence. For a deeper and intricate method in computer vision utilizing cloud-based solutions like Azure Cognitive Services [10] or incorporating deep learning mechanisms. They could prove invaluable for tackling complex situations. We harnessed artificial intelligence to explore cognitive science and analyze human faces, commonly known as computer vision. We successfully captured emotions from photographs and videos featuring human faces.

References

[1] Zhao, X., Wu, Y., & Yu, W. (2021). Applications of emotion recognition based on deep learning in health care: A review. Journal of Medical Systems, 45(5), 1-10. https://doi.org/10.1007/s10916-021-01727-8

[2] Ekman, P. (1999). Basic emotions. Handbook of Cognition and Emotion, 45–60. https://doi.org/10.1002/0470013494.ch3

[3] Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. Journal of Personality and Social Psychology, 17(2), 124-129. https://doi.org/10.1037/h0030377

[4] Ko, B. C. (2018). A brief review of facial emotion recognition based on visual information. Sensors, 18(2), 401.

[5] Li, S., & Deng, W. (2020). Deep facial expression recognition: A survey. IEEE Transactions on Affective Computing, 13(1), 119-134. https://doi.org/10.1109/TAFFC.2020.2981446

[6] H.-H. Tsai and Y.-C. Chang, "Facial expression recognition using a combination of multiple facial features and support vector machine," Soft Computing, vol. 22, no. 13, pp. 4389–4405, 2018.

[7] E. Cambria, Affective Computing and Sentiment Analysis, IEEE Intelligent Systems.31(2016)102-107.
[8] He, Y., Hong, L., & Zhang, Y. (2020). Real-time emotion detection for human-computer interaction using deep learning. Neurocomputing, 395, 208-218. https://doi.org/10.1016/j.neucom.2020.02.037
[9] Chen, J., Le, Q., Tran, D., & Pinkus, A. (2019). AI in the real world: Applications of sentiment analysis and emotion detection. IEEE Transactions on Affective Computing, 10(4), 545-562. https://doi.org/10.1109/TAFFC.2018.2873158
[10] Microsoft. (n.d.). Azure Cognitive Services: Emotion API. Microsoft Azure. https://azure.microsoft.com/en-us/services/cognitive-services/emotion
[11] R. Chandrasekaran, Thamizhvani T. R., Kishore Kanna R.. "chapter 9 Emotional Emotional Behaviour Study Using AI Tools
[12] Central Asian Journal of Medical and Natural Science. (2024). Emotion AI in business intelligence: Understanding customer sentiments. https://cajmtcs.centralasianstudies.org
[13] Neuroscience News. (2024). AI tunes into emotions: The rise of affective computing. https://neurosciencenews.com
[14] IEEE Xplore. (2024). Emotion Recognition by Facial Expressions Using Deep Learning. https://ieeexplore.ieee.org/document/10469626
[15] MDPI Sensors. (2024). Systematic Review of Emotion Detection with Computer Vision and Deep Learning. https://www.mdpi.com/s24113484