A Computer Vision based Real-Time Emotion Detection Technique using Eye and Lip Recognition

Dr. Sandeep Bhat[1], Prof. Bheema Shastry C.[2]
[1]Professor,Dept of CSE, SIT, Valachil,Mangaluru, Karnataka,India
[2]Professor,Dept of E&CE, SIT, Valachil, Mangaluru, Karnataka,India

Abstract

Face recognition technology recognises and classifies the emotions that a person is displaying through the examination of their facial expressions. Lip and eye detection, which analyses an individual's emotional state, is a popular technique in computer vision techniques for emotion recognition.

In this work, a computer-vision-based technique is used to identify the eye and lip; the recognised area is referred to as a region of interest (RoI). Depending on the user's emotional state, the ROI's overall pixel count varies. In this experiment, the relationship between lip movement and the eye is what determines one's capacity to identify emotions. This work achieved relatively good outcomes in diverse environmental and human circumstances without using any complex machine learning algorithms.

Keywords— Emotion, Face Detection, Image, RoI and Vision

1. Introduction

Over the past 20 years, automatic face expression recognition has attracted a lot of attention. This is due to the increasing necessity of behavioural biometric systems and human-machine interactions, in which emotion intensity and facial expression identification are key components [1]. The majority of the works that are currently in existence seldom ever simultaneously model data from various classes of facial activity, and they hardly ever encode the intensity of the observed facial emotion. Gabor filters, a histogram of oriented gradients (HOG), and local binary patterns (LBP) are the feature extraction techniques. The Nearest Neighbour Algorithm (kNN), Random Forest (RF), and Support Vector Machine (SVM) are the algorithms used for categorization. [1, 2].

It is possible to simulate human facial expressions as seen by the brain using video, electrical impulses, or pictures. Emotion detection by machines involves a variety of image processing algorithms for feature extraction because it is challenging for machines to discern emotions from images or videos and hard for people to do so. This research proposes a system with two key components: face detection and facial emotion identification. Through an experimental examination, the primary goal of this work is to recognise facial emotions [2].Through the processes of feature extraction, subset feature generation, and emotion classifiers, this research aims to construct a highly tuned real-time learning model of emotion recognition. The characteristic values are gathered when utilising the Haar Cascade method to identify the input figure. This work also aims to categorise human emotions such as fear, neutrality, anger, surprise, happiness, and sadness through the use of two classifiers—the Long Short Term Memory (LSTM) and Convolutional Neural Network (CNN)—as well as a precisely calibrated real-time emotion detection algorithm that makes use of virtual markers [3].

Emotion is the factor that influences what makes us human. Every aspect of our daily activities is affected, including our ability to focus, see, recall, make straightforward decisions, and communicate with others. People have mastered astounding skills for analysing and interpreting facial expressions. The Fusiform Face Region (FFA) was also developed as a distinct brain region because facial recognition is essential for human endurance. The key concept in this work is the use of a multi-facet Haar wavelet-based method, also known as the HAAR overflow face acknowledgment computation, which distinguishes various demeanours from significant face regions at various levels [4]. Human emotion and stressful psychological occurrences are closely related. According to a study in computational psychology, it is essential to comprehend how stress and emotions are related. Deep learning has primarily been used to recognise psychological stress, but studies have also found that it can also recognise emotional expressions in images. In this paper, a hybrid system that combines a regression classifier and the Haar Cascade Algorithm (HCA) is described. An HCA is taught to recognise a variety of facial expressions, such as anger, disdain, neutrality, fear, grief, delight, and surprise [5].

Here, one of the most well-known areas of research in human-computer interaction (HCI) is facial expression analysis. It has numerous uses in the study of human emotion, the modelling of behavioural and cognitive processes, and the creation of next-generation user interfaces. In this paper, a facial expression classification system is proposed. It uses the Haar classifier for face recognition, Local Binary Patterns (LBP) histograms of various block sizes of a face picture as feature vectors, and Principal Component Analysis (PCA) to categorise distinct facial expressions. Due to the low computational complexity of the method, expression classification can be done in real time.

Since every individual has a unique range and intensity of facial expressions, a flexible method is advised for facial expression analysis [6]. Here, convolutional neural networks (CNNs)' deep semantic features are combined with modified superficial classical characteristics to produce the final recognition outcome [7]. The variations in each face feature are used to differentiate between the various facial moods. Machine learning techniques are used to distinguish and classify numerous types of facial emotions by training on various sets of images. We can assist with a variety of identification obstacles, psychological studies, and practical problems by putting these algorithms into practice. The proposed approach is implemented using open-source computer vision software (OpenCV) and Python-based machine learning [8].

But emotions play a significant role in many occupations, including teaching. Despite the fact that people react differently depending on their culture, environment, language, and other circumstances, it is widely acknowledged that all members of the human race experience emotions. Emotions play a vital role in education, just as the environment or the language used to communicate information can. It is essential for cognitive functions, including learning and assimilating new information. The attitudes of the students during lectures have a big impact, whether in real classrooms or online learning environments [9]. Emotional detection is a crucial field since emotions differ considerably from person to person and are challenging to recognise with frequent lighting and illumination changes. The majority of this field's study focuses on identifying emotions in the region of the face. The dataset contains video frames that show mouth expressions in various moods, including happy, normal, and shocked. The mouth regions are recognised using the Haar-Based Cascade classifier at a frame rate of 20 fps. The three emotions of joyful, normal, and shocked are identified by extracting the HOG features as the next step [10]. Next, automated facial emotion identification is a challenging task during human-computer interaction. Here, hybrid CNN techniques to recognise human emotions and sub-classify them according to their traits are used. The FER13 dataset is used in this study to recognise emotions in order to get the best results in terms of accuracy and loss. The approach improves accuracy by 88.10% on average. The approach has been successful in classifying emotions into seven basic categories [11]. Thus, it has been proven that the suggested method by CNN is effective in recognising facial expressions [12].

Although mobile applications that respond to user emotions are growing in popularity as a result of their cutting-edge functionality and positive user feedback, to create such an application, a highly accurate, real-time emotion identification system is necessary. Because a mobile device's processing power is constrained, the algorithm utilised in the emotion recognition system should require as few calculations as is practical. The suggested method uses the camera built into a smart phone to record face-to-face footage. After obtaining a few sample frames from the movie, a face detection module is utilised to extract the face regions in the collected representative frames [13]. For the system's video component, a brand-new emotion recognition technique that depends on image set matching has been developed. The recommended solution eliminates the need to identify and track certain face landmarks throughout the provided video sequence, strengthening the video processing chain and removing a common source of error in video-based emotion recognition systems. Contrarily, the system's audio component makes use of utterance-specific Gaussian Mixture Models (GMMs), which were altered from a Universal Background Model [14]. In computer vision and artificial intelligence, identifying and detecting human emotion is a significant difficulty. Human speech is significantly influenced by emotions. Most of the time, emotions are used in communication. There are many universal emotions, such as fury, grief, love, surprise, fear, disdain, and indifference. In the facial recognition procedure, the Haar Cascade characteristics are extracted from a face using the Viola Jones algorithm, and the emotion is confirmed and distinguished using a deep neural network. This has applications in a number of fields, including robotics, security, and surveillance [15].

The process of emotion detection using face recognition typically involves the following steps: Face detection is the first step to detecting and locating faces within an image or video frame. This can be done using face detection algorithms or libraries like OpenCV, Dlib, or face recognition. Once the face is detected, the next step is to identify the facial landmarks or key points on the face, such as the position of the eyes, nose, mouth, and eyebrows. These landmarks provide crucial information for analysing facial expressions accurately and are called facial landmark detection. Libraries like Dlib or facial landmark models can be used for this purpose. After obtaining the facial landmarks, various features can be extracted from the face, such as the distances between specific landmarks, the angles of facial components, or the intensity of certain facial regions. These features are used to represent facial expression patterns. This step is called feature extraction.

It is important to note that while emotion detection using face recognition can provide valuable insights, it is not a perfect science. The accuracy of the emotion classification can vary depending on factors such as lighting conditions, pose variations, occlusions, and individual differences in facial expressions. Continuous research and advancements in computer vision and machine learning techniques aim to improve the accuracy and robustness of emotion detection systems.

2. The block diagram and its description

The block diagram of the emotion detection system is depicted in figure 1.

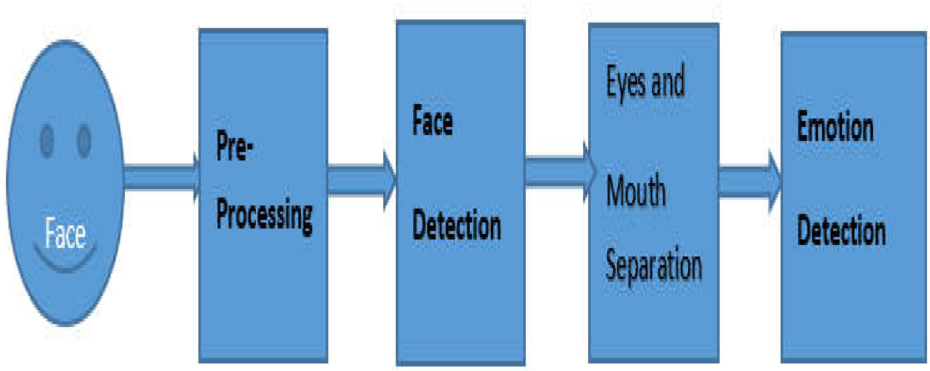


Figure 1: The emotion detection system.

The block diagram illustrates that the first step in an emotion detection system is to capture the image of a person through a video camera. Next, the captured frame undergoes an appropriate filtering operation to remove the noise content in it. After this phase, the face of the person is detected using a face detection algorithm. Once the face is detected, the next step is to separate the eyes and mouth using a proper algorithm.

The emotion is identified by the Pixel Count Based Algorithm, which counts the number of pixels (white or black) in the region of interest, that is, the mouth and eyes in the detected face. The number of pixels in the region of interest increases when a person expresses an emotional situation like anger or surprise. When a person is sad (eye only), the number of pixels in the region of interest is less. This process will be carried out for real-time face detection and the identification of the status of the emotion.

3. The working principle

The working principle of the Emotion Detection System is illustrated in Figure 2. At first, a camera should turn on and begin to capture the image of a person. Once the capture of the image of a person is done, the next step is to detect the face of the person. After detecting the face, the cropping of the eyes and mouth in the face is essential.

The Pixel Count Based (PCB) algorithm is utilised here to find the number of white pixels in an eye and the mouth region. During happy and surprised situations, the eye opens more compared to sad and other emotional statuses. Hence, this notion is implemented in this algorithm, and the necessary white pixels are computed. Here, in the region of interest (RoI), if the pixel count P exceeds the threshold value of 100, this emotion detection system detects that the resultant emotion is either ‘happy’ or 'surprised'. And the concept is that the mouth is closed and the eyes are partially opened when a person is ‘sad’ or under tension. In this way, the emotional status can be identified by the emotional detection system.

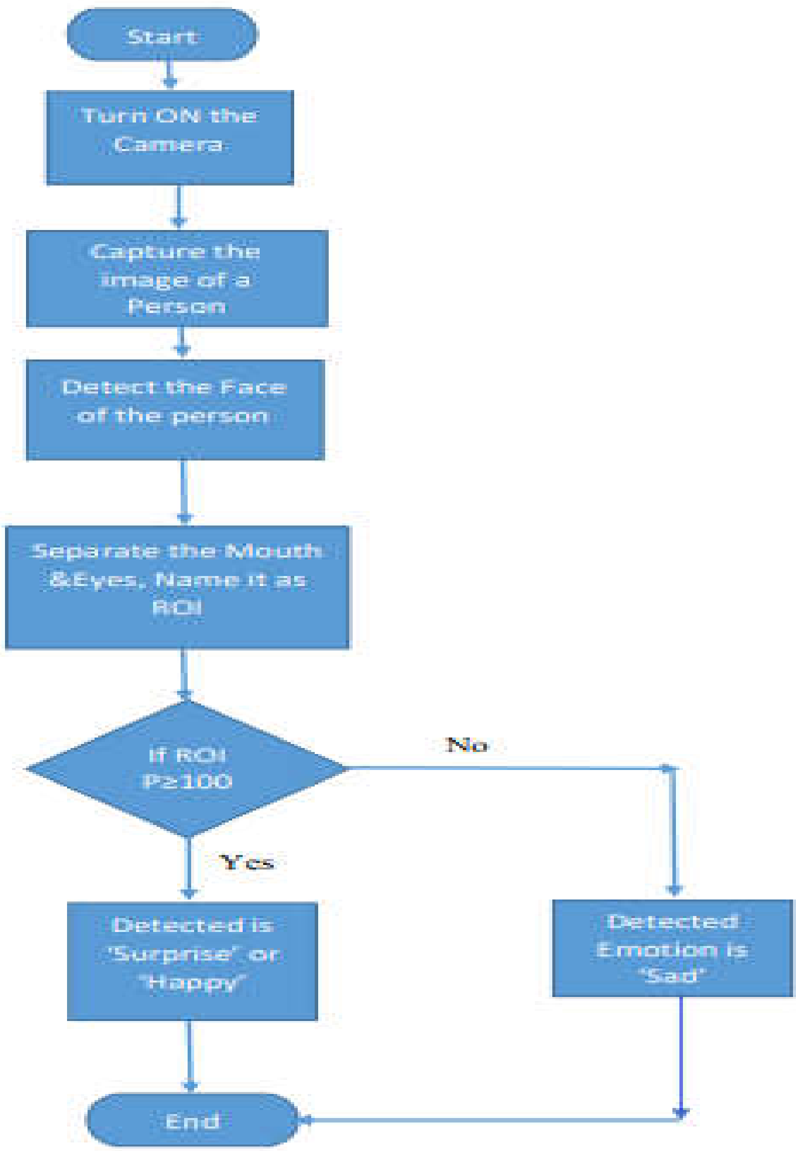


Figure 2: The working principle of the emotion detection system.

4. Results and discussions

The emotion is identified by the Pixel Count Based (PCB) Algorithm, which counts the number of pixels (white or black) in the region of interest is, the mouth and eyes in the detected face. The emotion is sorrow, as depicted in figure 3.

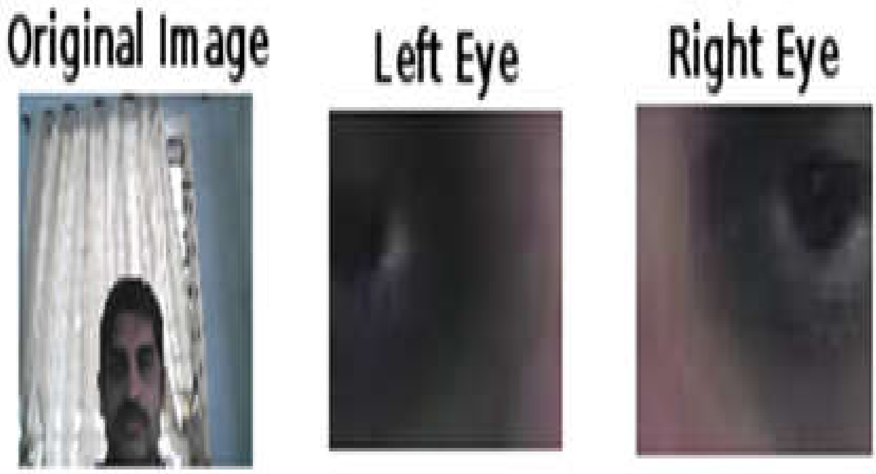


Figure 3: The Emotion Detection Process.

The number of pixels in the region of interest increases when a person expresses an emotional situation like anger or surprise. When a person is sad (eye only), as illustrated in Figure 4.The number of pixels in the region of interest is less than the threshold value of 50; hence, the detected emotion is sorrow, as shown in figure 5. This process will be carried out for real-time face detection and the identification of the status of the emotion.

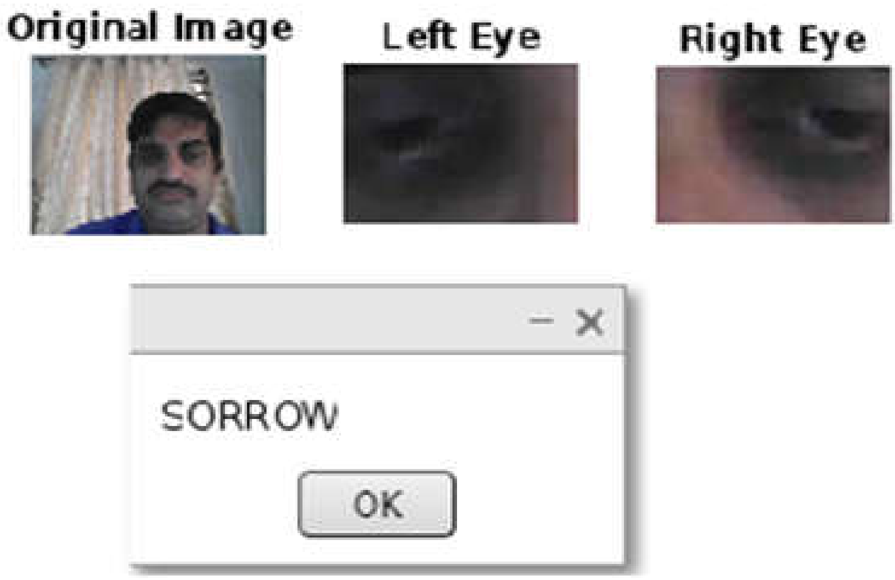


Figure 4: The Detected Emotion is Sorrow.

The real-time PCB algorithm's result is depicted in figure 5.

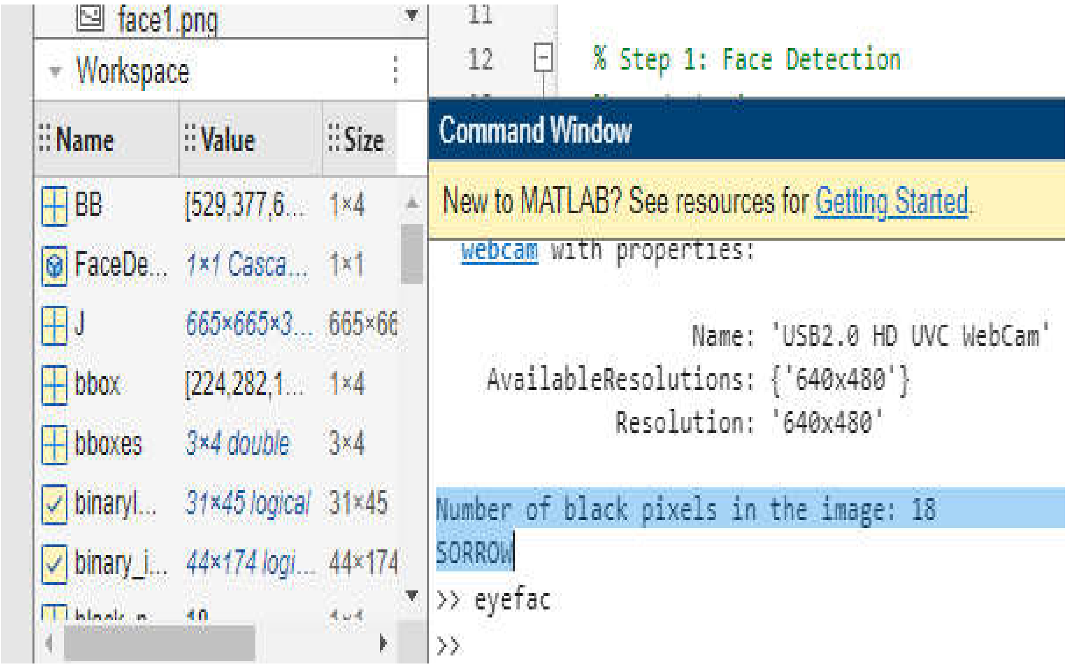


Figure 5: The PCB Algorithm Result.

The number of pixel count in the region of interest is more than the threshold value of 52(considering the average results); hence the detected emotion is Happy as shown in figure 6.

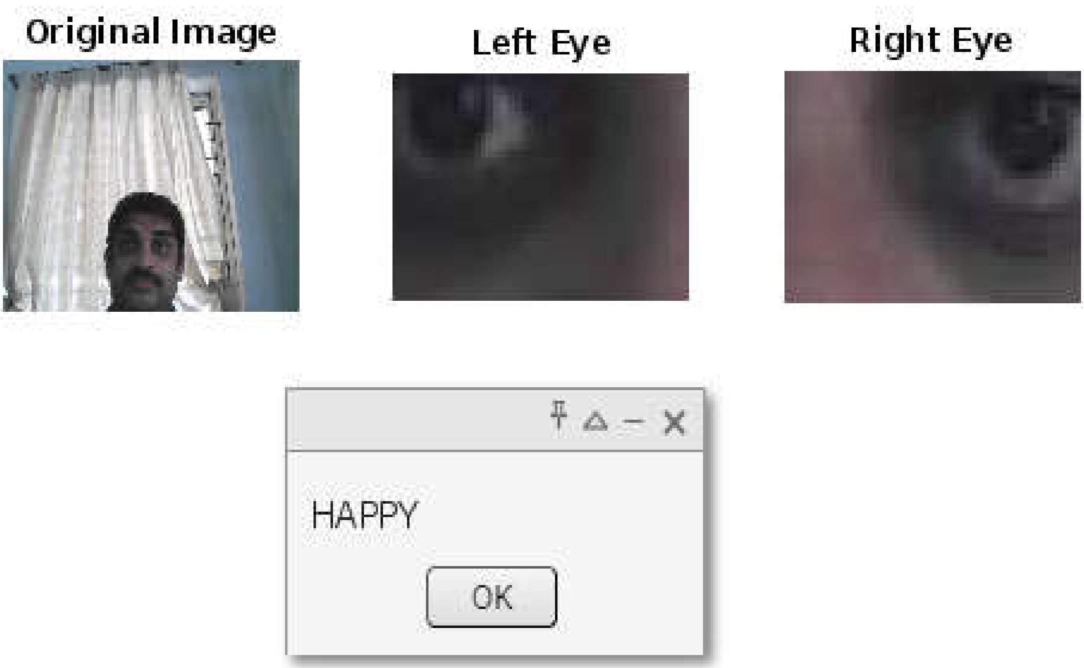


Figure 6: The Detected Emotion is Happy.

The PCB algorithm is applied to detect the emotion from the mouth is also illustrated in figures 7 and 8 for Happy and Sad respectively.

Figure 7: The Detected Emotion is Happy.

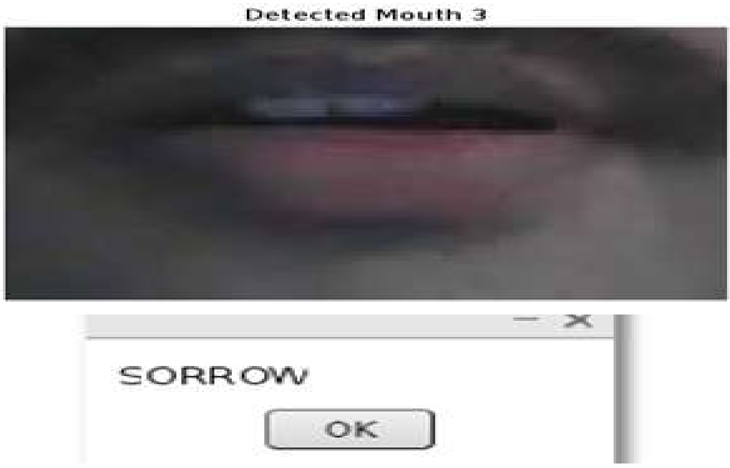
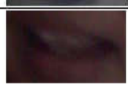


Figure 8: The Detected Emotion is Sorrow.

The emotion detection of different persons under various circumstances is depicted in table 1.

Table 1: Emotion detection of different persons under various circumstances

S.No.	Data Set	Eye	Lip	Concussion	Success Rate (%)	Remarks
1	17			Sorrow	13/17=76.4	24% of results were Happy
2	18			Sorrow	17/18=94.4	Justified as **Crying** not only Sorrow
3	21			Neutral	18/21=85.7	Between sorrow and happy
4	17			Happy	12/17=70.5	30% of results were sorrow.

The results indicate a varied distribution of emotions across the dataset. However, it's important to note that this emotion detection technique may be influenced by factors such as lighting conditions, image quality, and the choice of algorithm.

5. Conclusion and future scope

This research has been successfully applied to eye and lip detection techniques to infer emotional states in images and video frames. Eventually, this work gave the emotion detection of a happy, sorrow (sad) and neutral distribution of 70.5%, 85.4% and 85.7% respectively. These discoveries could be used in a variety of disciplines, such as human-computer interaction and sentiment analysis. Future research could focus on improving accuracy and robustness in real-world scenarios.

References

[1]. Dhwani Mehta, Mohammad Faridul Haque Siddiqui and Ahmad Y. Javaid, "Recognition of Emotion Intensities Using Machine Learning Algorithms: A Comparative Study", Sensors; (http://www.mdpi.com/journal/sensors),2019,pp.1-24, 19, 1897, 2019.

[2]. Prof.Sarika Pawar,Harshada Pisal,Mangesh Shelk and Nitin Nimbalka,"Emotion Detection System Using Machine Learning", IJARIIE-ISSN(O)-2395-4396,Vol-8 Issue-2, 2022.

[3]. Vijay Kumar Sharma, Swati Sharma, Umang Arora and Anant Gupta,"Finely-Grained Real-Time Facial Emotion Recognition towards Neural Network", Biosc.Biotech.Res.Comm. Special Issue, Vol. 14, No. 05, pp-176-181, 2021.

[4].Bhavya Alankar,Mohammad Sharay Ammar & Harleen Kaur, "Facial Emotion Detection Using Deep Learning and Haar Cascade Face Identification Algorithm",art of the Lecture Notes in Networks and Systems book series (LNNS),volume 202, May 2021.

[5]. M. Harshinni; V. Sindhu; S. Jegan; J D Joel; A R Kannathal, "A Deep Learning Approach for Human Stress Detection using Haar-Cascade Algorithm"Publisher: IEEE 7th International Conference on Trends in Electronics and Informatics (ICOEI), 2023.

[6]. S L Happy, Anjith George, and Aurobinda Routray "A Real Time Facial Expression Classification System Using Local Binary Patterns", 4th International Conference on Intelligent Human Computer Interaction (IHCI), 2012.

[7]. Hao Meng , Fei Yuan ,Yue Wu , and Tianhao Yan, "Facial Expression Recognition Algorithm Based on Fusion of Transformed Multilevel Features and Improved Weighted Voting SVM" , Hindawi Mathematical Problems in Engineering, A journal, Volume 2021,pp.107-124, 2021.

[8]. Rajesh K M,Naveenkumar M, "A Robust Method for Face Recognition and Face Emotion Detection System using Support Vector Machines", IEEE International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT),pp.1-7, 2016.

[9]. Dr.S.Suriya and Babuvignesh C, "Deep Learning Based Hybrid Approach for Facial Emotion Detection", ICCAP 2021, December 07-08, Chennai, India, pp.203-214, 2021.

[10]. J. Sujanaa and S. Palanivel, "Hog-Based Emotion Recognition Using One-Dimensional Convolutional Neural Network", ICTACT Journal on Image and Video Processing, November 2020, Volume 11, Issue 02, ISSN: 0976-9102 (online), pp.2310-2315, 2020.

[11]. Nehmat Sandhu , Aksh Malhotra, Mandeep Kaur Bedi, "Human Emotions Detection Using Hybrid CNN Approach",International Journal of Computer Science and Mobile Computing, Vol.9, Issue.10, pp. 1-9, October- 2020.

[12]. P A Riyantoko, Sugiarto and K M Hindrayani, "Facial Emotion Detection Using Haar-Cascade Classifier and Convolutional Neural Networks", Second International Conference on Science & Technology (ICoST) Journal of Physics: Conference Series 1844 (2021) 012004, IOP Publishing, pp.1-5, 2021.

[13]. M. Shamim Hossain and Ghulam Muhammad, "An Emotion Recognition System for Mobile Applications", Special Section on Emotion-Aware Mobile Computing, IEEE Access, March 2017, pp.2281-2287, 2017.

[14]. Simon Dobrišek1 , Rok Gajšek1 , France Mihelič1 , Nikola Pavešić1 and Vitomir Štruc "Towards Efficient Multi-Modal Emotion Recognition", Int. J. Adv. Robotic Sy,(http://www.intechopen.com) 2013, Vol. 10, Issue 53 ,2013.

[15]. J. Jayapradha Soumya Sharma and Yash Dugar, "Detection and Recognition of Human Emotion using Neural Network",International Journal of Applied Engineering Research, ISSN 0973-4562 Volume 13, Number 8, pp. 6472-6477, 2018.