International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)
doi : https://doi.org/10.32628/IJSRST2183209

Emotion Recognition and Drowsiness Detection using Digital Image Processing and Python

Pramoda R[1], Arun P S[2], B S Athul[2], Bharath B[2], B Naveen Reddy[2]

[1] Assistant Professor, Department of Computer Science and Engineering Nagarjuna College of Engineering and technology Bengaluru, India

[2]B.E. Scholar, Department of Computer Science and Engineering Nagarjuna College of Engineering and technology Bengaluru, India

Article Info
Volume 8, Issue 3
Page Number : 1037-1043
Publication Issue
May-June-2021
Article History
Accepted : 20 June 2021
Published : 30 June 2021

ABSTRACT

We suggest a system that can detect a person's emotions as well as their level of sleepiness. The majority of our work is devoted to extracting information from the frontal face. The article goal is to create a product that is both reasonable and efficient in its operation. The system was created in Python using artificial intelligence and digital image processing technologies. Identifying eye blinking is essential in some situations, such as in the prevention of automobile accidents or the monitoring of safety vigilance.

Keywords: Digital Image Processing, Artificial Intelligence, OpenCV, Haar Cascade

I. INTRODUCTION

Artificial Intelligence and Digital Image Processing are two fields that are developing slowly but steadily in our nation right now. AI and DIP methods and applications are now being used in many industries, including many previously untapped ones. The project may also be used for marketing purposes since it provides us with information on how customers react to a product. It produces accurate findings and is simple to apply and understand in most standard systems, making it a popular choice. Furthermore, these functions are often implemented cost-effectively and efficiently in schools, universities, and other areas where monitoring is needed, but funds are limited, a significant element in their implementation. Using our suggested project, surveillance may be given, with the findings assisting in the maintenance of a regular health check and the understanding of a person's emotional state while at work. It may also be used to get feedback from employees after implementing changes at the workplace.

II. LITERATURE REVIEW

A. Emotion Recognition

Nowadays, Research work studies and Computer applications are making use of the Deep Learning technique. Even the Image Processing Field uses this technique within them. On the other hand, the precision of the whole method may be enhanced via computer-related programmes. In the proposed research, we want to develop as effective model which could extract the facial details effectively.

A traditional image could have an excessive amount of noise present, which may reduce the dataset training rate. Convolutional -Neural Network (CNN) was introduced as a replacement technique for recognition of expressions to overcome this problem. Before recognizing the face from the Primary image, it is necessary to reduce the range of the facial features. When this is done, Haar Cascade Features are frequently used to trace the many coordinates of the eyes and mouth, among other essential body components. Humans' tendency to communicate with one another via their expressions may even be shown by how they talk to one another. The many emotional expressions on the face, which may be seen in people, communicate specific information. The Detection of the face is essential in the early stages of the face recognition process. The usage of the AdaBoost cascade classifier is often used in the identification of a face on a genuine picture, and this method is detected the majority of the time.

In the last decade or so, facial expression recognition has gotten a lot of positive feedback, and the findings have been utilised in a number of situations. Due to changes in pictures caused by lighting, ageing, and other variables, achieving accuracy in a real time scenario is a major disadvantage. Emotions are produced without consciously attempting to do so, such as reflexing their eye expressions. Happiness, sorrow, surprise, anger, and stability are some of the common feelings that a person's figure develops in response to the various things that one may see oneself in (average). This may be a suggested method for utilising a deep Neural Network (CNN) to determine someone's feelings, and it could even analyze how rapidly the expression of a person changes.

B. Drowsiness Detection

Recognition of drowsiness is critical in some circumstances, such as driving a car or maintaining security vigilance to prevent an accident or disaster from occurring. For example, in a car, this function becomes critical since, anytime a driver begins to feel a little sleepy-eyed and drowsy, the pursuit of those feelings may immediately alert the user to stay awake and focused on driving.

In this case, the algorithmic rule utilized to identify attention blinking is being detected via a camera placed in the victimization live video feeds. The landmarks are recognized precisely enough so that the attention gap may be measured with accuracy. In addition to the problems already stated, facial recognition and identification via live video victimization cameras will provide a slew of additional challenges. The position of the face is likely to be the most significant factor. The face should be aligned properly to extract details of face more accurately and, as a consequence, choose which aspects to record and to work with.

III. PROPOSED SYSTEM

A. Emotion Recognition

The system will recognize a face through which it will be able to deduce the emotion. Face functions refer to how people communicate with one another via their faces rather than simply speaking or writing. Facial expressions are formed by facial muscle movements. Then we assign the different emotions that they portray various names. Following that, we load models, get ability to incorporate shapes for inference, and finally, we begin lists for calculating modes, which is the last phase.

We convert the bgr image to a grayscale image to an RGB image as soon as the camera is turned on. After that, we scale the grey image. We put the frame through emotion classifier as soon as it is converted to grayscale, which is designed to match the mood with previously recorded emotions and then classify emotion based on probability. Then we label the classified image. It's possible that the emotion won't be recognised if the sensation can't be noticed. If the given emotions are present in the database, the probability of a particular emotion being present is computed. After analyzing the possibilities and selecting a particular sensation to experience, the emotions with the highest likelihood is chosen to be given.

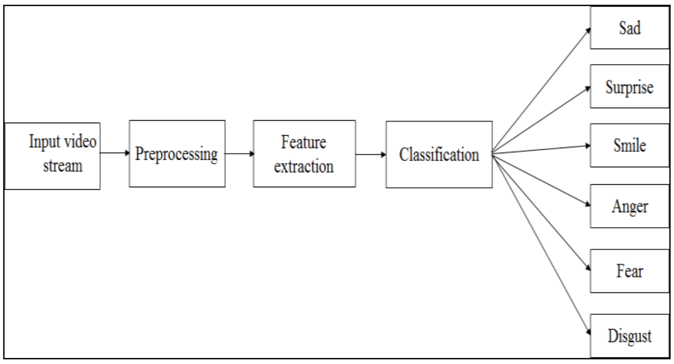

Figure 1: Emotion Recognition Architecture

B. Drowsiness Detection

Drowsiness detection could be achieved with the help of Computer Vision Technologies. With the help of this technology we can analyse and extract different features from images and videos. We analyse the eyes of a person continuously through a live camera from which the eye aspect ratio is calculated.

Video processing entails a variety of procedures and methods that modern computers use for a broader range of video accessibility, such as video tracking, picture stabilization, object recognition, and so on. Detection and tracking algorithms detect and track items or people and identify forms and figures in their environment. It may also be used to identify any flaws present in the supplied picture or video. Morphological transformation is a phrase that refers to how pictures are shaped and how some of the fundamental processes are carried out. In the pictures, these procedures are carried out. A structural element that determines the type of the operation to be carried out. Erosion and Dilation are two fundamental morphological operators. Image Retouching, We may utilize image enhancement procedures such as noise reduction, grey level transformations, and colour conversions to improve the effects of a picture or make it appear better than it did before. The face landmarks are located using the shape predictor algorithm. The left and right eye indexes are discovered and utilized to extract the eye region from the frame. Pre-processing is done to all of the frames initially, which includes shrinking and grayscale conversion. The aspect ratio of both eyes is calculated. The final eye aspect ratio is the combination of both of these numbers.

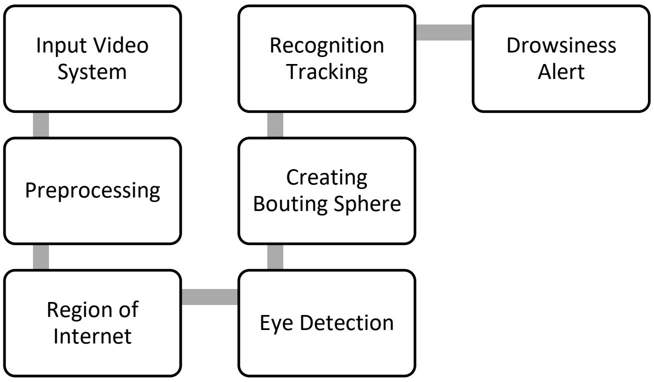

Figure 2 : Drowsiness Detection Architecture

IV. IMPLEMENTATION

A. Emotion Recognition

1. Input Live Video Stream

A video Stream is given as input, from which input faces are taking, and emotion is classified.

2.Face Detection

In this phase, the input data is compared to the pictures in the dataset. The system then checks to see whether the supplied data includes a face. Following this, the result is submitted to be pre-processed to retrieve the countenance from the facial picture.

3.Feature Extraction

This is a crucial step since it extracts the features using the feature extraction algorithm applied. Information is compressed, irrelevant characteristics are reduced, and data noise is removed using these processes. The face region is then transformed into a vector with a specified dimension. The trained model is then submitted to the test using an input picture, and the pre-processing and other stages are repeated.

4.Classification

This is the last step of the process, during which the emotions are categorized. The pictures may be categorized in a variety of ways. A neural network could be a robust classification method. It may be used with both linear or nonlinear data. Because it's a self-learning model with multiple hidden layers, it works even for pictures that aren't in the dataset.

However, a conclusion may summarise the paper's primary arguments. It should not be a repeat of the abstract. A conclusion may discuss the work's significance or propose applications and expansions. The use of numerous images or tables, in conclusion, is highly discouraged; they should be cited in the body of the article.

B. Drowsiness Detection

1. Video Streaming

First, position the camera in the car to clearly show the driver's face and then use a facial mask snag to block the mouth and eyes. Get the driver's head's condition as well. We are obtaining video from a web camera placed on the dash of a car to get pictures of a driver for video acquisition and a sound planned and executed structure for the dormancy zone.

2. Face Detection

The suggested structure starts by obtaining individual video traces. OpenCV is a powerful tool for preparing live files. For each bundle, the structure views the facial in the edge image. Viola-Jones object pioneer, an AI method for visual article disclosure, is used in this structure. The Haar estimate for facing a region is used to achieve this. Haar course is a significant liberal portion predicated on the assumption that you can see the facial image well enough. Haar figure designed to eject the nonface competitors by using jobs that are indisputable of stages. Furthermore, each step combines various Haar highlights, and each component is therefore referred to as a Haar merge classifier.

3. Aspect Ratio of Eyes

As a result, understanding the tactics of the particular face components suggests eliminating facial features that are difficult to communicate in other ways. The degree of the statement indicates that we are most excited about two frameworks of facial features — the eyes and the mouth. 8 At each eye are six (x, y)-animates, with the first one starting in the left corner of the eye (commensurately as if you were staring at someone) and the remainder of the area being worked on over a short period clockwise all-around rest of the region. For each video plot, the eye accomplishments are seen. The degree of eye point of view (EAR) is determined by comparing the height and breadth of the eye. Figure 1 shows the 2D accomplishment zones, and the numbers 1, 2, 3, 4, 5, and 6 represent the levels of achievement. In most cases, while one eye is open, the SEAR is solid; nevertheless, when one eye is closed, the SEAR is inclining toward zero.

It is commonly individual, and the head presents wanton. It is common among individuals that the perspective dimension of the open eye has a somewhat multidimensional character, and dimension varies based on the movement of eyes. Given that the two eyes simultaneously perform eye squinting, the ear of the two eyes is brought together at a point in between the vertical eye spots of intrigue and the level eye accomplishments. The denominator depicts the area among level eye accomplishments, with the denominator is appropriately weighted because only a single heap of stage thinks about any rate two approaches. The degree of eye perspective is self-evident, by which time it rapidly decreases to near zero, by which time it develops once again, giving the appearance of a single squint occurring.

4. Head Pose Estimation

The accompanying step is used to determine whether or not the driver's head is present. It has been determined that the driver's facial image from head sighting confirmation while following has indeed been matched for head tilting. Following the head's movement or the development of highlights on the face is a structure used to think about the head at the current time. The use of optical stream estimation provides an observable framework for obtaining the movement. The optical stream technique is being used to track the movements of a firm's head on video. Following that, they used system gathered improvement regularisation using an ellipsoid model as a foundation for their subsequent approach. The most important consideration is identifying the aggressive movement of a head model that best accounts for the optical stream. However, after the optical progression of each point has been determined, they use an edge drop structure to determine the optimal advancement of the head. While using groups with a modest bundling rate and raucous images, the tracker in their test is completely suffocating over an unlimited amount of time. The suggested optical stream-based structure was developed to demand the kinds of progress required by a deformable model. They combined optical stream data with edge data to simulate the gliding motion achieved by the optical stream. Alert Unit

The appearance of the prepared unit occurs whenever the driver is in an idle state. The layout of the depiction yield is either a 1 or a 0, and the prepared unit uses this number to determine its position. The framework design for the proposed building is shown in figure1 to the right of this paragraph. The Video Capture Unit, Face Detection Unit, Head Present Estimation Structure, Facial Extraction, shut-eye territory, yawning statement, laziness disclosure, and Alert Unit are included. The OpenCV, Dlib, and SciPy packages are the focal point of the social activities here.

V. RESULTS

The findings of Emotion Recognition are shown in the illustration below. We propose a face recognition method based on Convolutional Neural Networks (CNN) assisted by machine learning. For example, we take a picture as input and then use CNN to predict which facial expression label should be shown, which should be one of the following labels: astonishment, happy, fear, sorrow, surprise, or neutral.

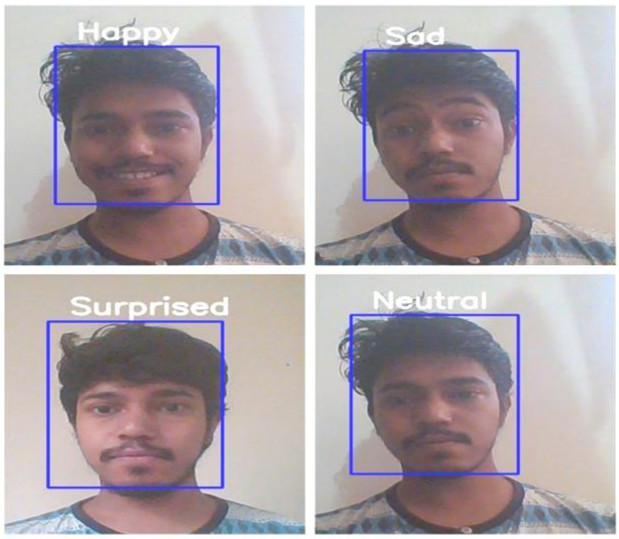

Figure 3: Emotions

The system monetarizes the eye aspect ratio and alerts the user if the person is drowsy. The aspect ratio is calculated individually for the left and right eye. Both eye aspect ratio is average to determine final eye aspect ratio. The proposed system also sends the person's location via SMS if the person is drowsy above the threshold time.

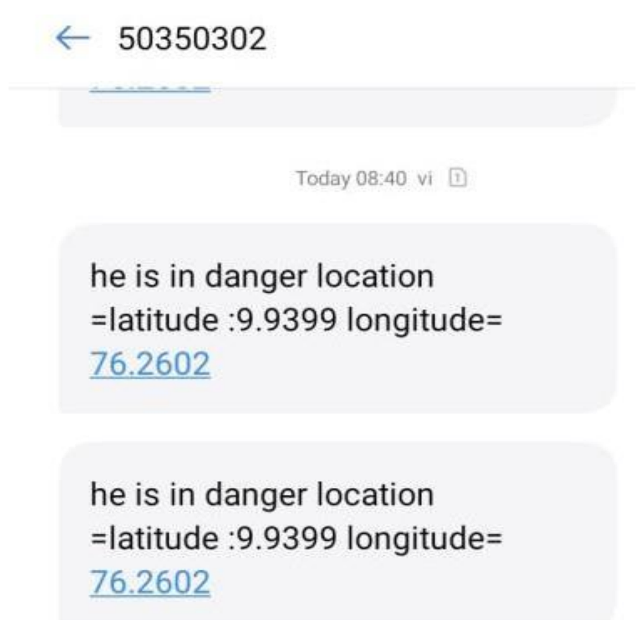

Figure 4: Drowsiness Detection

The figure below shows the message received by registered mobile when the person is drowsy.

Figure 5: Drowsiness Alert

VI. CONCLUSION

The automation of the many components in our environment has been completed. To unravel the problems of emotion recognition assisted face recognition in virtual learning settings, a suggested model is being examined. The efficiency and accuracy are both being addressed simultaneously, as is the case with traditional face recognition. Using Haar Cascades to detect the eyes and mouth and a neural network approach to recognize all kinds of emotion is being explored.

The use of emotion detection in virtual environments is a subject that has received a great deal of attention recently. Aside from changing uncertainty factors, the face patterns of instructors and students are becoming more complicated, making emotion recognition inside the online education network application mode a challenging subject to master in the current environment. Due to its ability to distinguish between regular eyes blinking and sleepiness, this system may be utilized in vigilant monitoring, sleepy driving detection, and other similar ones in the future.

VII. REFERENCES

[1]. SeongJoon Baek and Koeng-Mo Sung, "Fast K-nearest-neighbour search algorithm for nonparametric classification.", ELECTRONICS LETTERS 12th October 2000 Vol. 36, No 21 [3] Wen-Jyi Hwang and Kuo-Wei Wen. "Fast kNN classification algorithm based on partial distance search.", ELECTRONICS LETTER 15th October 1998 Vol. 34, No. 21

[2]. Bilel Ameur, Sabeur Masmoudi, Amira Guidara Derbel, Ahmed Ben Hamida. "Fusing Gabor and LBP Feature Sets for KNN and SRC- based Face Recognition." 2nd International Conference on Advanced Technologies for Signal and Image Processing

[3]. Vedat TÜMEN, Ömer Faruk SÖYLEMEZ, Burhan ERGEN. "Facial Emotion Recognition on a Dataset Using Convolutional Neural Network."

[4]. Lin-Lin Xu, Shu-Mei Zhang, Fu-Xing Wang. "CNN Expression Recognition Based on Feature Graph."

[5]. Xiao Sun, Man Lv, Changquin quang. "Improved Facial Expression Recognition Method Based on ROI Deep Convolutional Neutral Network."

[6]. Rajesh Kumar G A, Ravi Kant Kumar, Goutam Sanyal. "Facial Emotion Analysis using Deep Convolution Neural Network." International

Conference on Signal Processing and Communication (ICSPC'17) – 28th & 29th July 2017.

[7]. Tereza Soukupov´a and Jan Cech. "Real-Time Eye Blink Detection using Facial Landmarks." 21st Computer Vision Winter Workshop Luka Cehovin, Rok Mandeljc, Vitomir Struc (eds.) Rimske Toplice, Slovenia, February 3–5, 2016

[8]. Hongwei Ng, "Deep Learning for Emotion Recognition on Small Datasets using Transfer Learning",2015

Cite this article as :

Pramoda R, Arun P S, B S Athul, Bharath B, B Naveen Reddy" Emotion Recognition and Drowsiness Detection using Digital Image Processing and Python", International Journal of Scientific Research in Science and Technology(IJSRST), Print ISSN : 2395-6011, Online ISSN : 2395-602X, Volume 8, Issue 3, pp.1025-1031, May-June-2021. Available at doi : https://doi.org/10.32628/IJSRST2183209
Journal URL : https://ijsrst.com/IJSRST2183209