Image Processing Based Driving Assistant System

GUNJAN GALA[1], GANESH CHAVAN[2], NIBHA DESAI[3]

[1, 2, 3] Shah and Anchor Kutchhi Engineering College, Mahavir Education Trust Chowk, Waman Tukaram Patil Marg, Next to Dukes Company, Chembur, Mumbai, Maharashtra, India.

Abstract- Rapid industrialization and the consequent urbanization has brought about an unprecedented revolution in the growth of motor vehicles all over the world and India is no exception. Such growing urbanization, combined with rising number of vehicle ownership, has led in recent years to an increased traffic related stress and overnight driving exhaustion which lead to accidents in long term and short term period of time. To avoid such ambiguity we are developing a system which will detect autonomously and continuously track drivers' faces and determine its emotion. With the help of which we can pinpoint if the driver falls to sleep. If the driver is falling asleep an alarming sound will play to wake up the driver and we can make speed adjustments in the car. Based on the emotion we can play music to make the driver journey more joyful.

Indexed Terms- Arduino, Image processing, Machine Learning, YOLO

I. INTRODUCTION

In the beginning of the 21st century advanced features such as collision warning and avoidance systems were introduced into motor vehicles. However, there are many issues that need to be addressed before the driving assistance system can be widely introduced in the future vehicles. The theoretical and experimental research on control of such issues is in a developing stage. The main challenge in the driver assistance system is the sensory issues. Today's technology has addressed many of the sensory issues with many still to be solved. The impact of automation on the driver necessitates an understanding of human factors in relation with the automated driving controls or assists. Research on the human factor is very important and demands a lot more work. Legal and institutional aspects of automated vehicles are very important concerns.

In the late 1980s and beginning of 1990s, state and private funded programs started more focused research in the United States, Europe and Japan, to bring the idea of automated vehicles closer to reality. The main initiative was to improve safety with automation. The very well organized and futuristic thorough research in this era, along with the rapid advancements in electronics and sensor technology, contributed to a more vivid understanding of the difficulties and potentials of such systems. Although the research in this period was focused more on advanced highways, it later switched to intelligent vehicle initiative (IVI). While a lot has been said about improved safety and higher comfort level with automation in different papers, sometimes inconsistencies exist between different points of views on these matters.

II. LITERATURE REVIEW

In dynamic scenes, tracking is used to follow a face through the sequence. In order to incorporate the face changes over time, in terms of changes in scale, position and to localize the search for the face, it is essential to exploit the temporal correspondence between frames. Tracking exploits the temporal content of image sequences. Face tracking can be divided into two categories.

1. Head tracking.
2. Facial feature tracking.

Feature tracking methods track contours and points or follow eyes and mouth, and require independent trackers for each feature. Head tracking methods use the information from the entire head and can be region-based color-based or shape-based. Color-based approaches are not robust to lighting changes and approaches that use information from the entire head are, in general, unable to handle occlusion. Tracking involves prediction and update for which filters like Kalman filter and Condensation filter have been used.

Tracking approaches can also be model-based, for example, using statistical models, or exemplar-based. A combination of feature and head tracking methods, together with filtering, have tried to eliminate the problems of the individual approaches. [1]

III. APPROACHES

The task of emotion recognition often involves the analysis of human expressions in multimodal forms such as texts, audio, or video. Different emotion types are detected through the integration of information from facial expressions, body movement and gestures, and speech. The technology is said to contribute in the emergence of the so-called emotional or emotive Internet.

The existing approaches in emotion recognition to classify certain emotion types can be generally classified into three main categories: knowledge-based techniques, statistical methods, and hybrid approaches.

A. Knowledge-based techniques

Knowledge-based techniques, utilize domain knowledge and the semantic and syntactic characteristics of language in order to detect certain emotion types.[2] In this approach, it is common to use knowledge-based resources during the emotion classification process such as WordNet, SenticNet, ConceptNet, and EmotiNet, to name a few. One of the advantages of this approach is the accessibility and economy brought about by the large availability of such knowledge-based resources. A limitation of this technique on the other hand, is its inability to handle concept nuances and complex linguistic rules.

Knowledge-based techniques can be mainly classified into two categories: dictionary-based and corpus-based approaches Dictionary-based approaches find opinion or emotion seed words in a dictionary and search for their synonyms and antonyms to expand the initial list of opinions or emotions. Corpus-based approaches on the other hand, start with a seed list of opinion or emotion words, and expand the database by finding other words with context-specific characteristics in a large corpus. While corpus-based approaches take into account context, their performance still vary in different domains since a word in one domain can have a different orientation in another domain.

B. Statistical methods

Statistical methods commonly involve the use of different supervised machine learning algorithms in which a large set of annotated data is fed into the algorithms for the system to learn and predict the appropriate emotion types. This approach normally involves two sets of data: the training set and the testing set, where the former is used to learn the attributes of the data, while the latter is used to validate the performance of the machine learning algorithm. Machine learning algorithms generally provide more reasonable classification accuracy compared to other approaches, but one of the challenges in achieving good results in the classification process, is the need to have a sufficiently large training set.[5]

Some of the most commonly used machine learning algorithms include Support Vector Machines Naive Bayes, and Maximum Entropy. Deep learning, which is under the unsupervised family of machine learning, is also widely employed in emotion recognition Well-known deep learning algorithms include different architectures of Artificial Neural Network (ANN) such as Convolutional Neural Network (CNN), Long Short-term Memory (LSTM), and Extreme Learning Machine (ELM). The popularity of deep learning approaches in the domain of emotion recognition may be mainly attributed to its success in related applications such as in computer vision, speech recognition, and Natural Language Processing (NLP).[4]

C. Hybrid approaches

Hybrid approaches in emotion recognition are essentially a combination of knowledge-based techniques and statistical methods, which exploit complementary characteristics from both techniques. Some of the works that have applied an ensemble of knowledge-driven linguistic elements and statistical methods include sentic computing and iFeel, both of which have adopted the concept-level knowledge-based resource SenticNet. The role of such knowledge-based resources in the implementation of hybrid approaches is highly important in the emotion classification process. Since hybrid techniques gain from the benefits offered by both knowledge-based and statistical approaches, they tend to have better classification performance as opposed to employing knowledge based or statistical methods independently. A downside of using hybrid techniques however, is the computational complexity during the classification process.

IV. IMPLEMENTATION

As shown in the figure 1 there are four main modules in our project. The first one is camera module from which we take input. The input is in the form of live video feed from the camera. The second module is face tracking module where using YOLO training module we continually track face and face emotions.

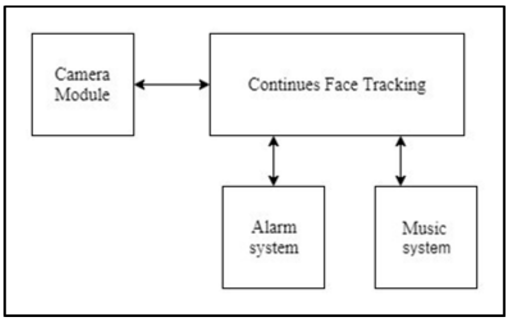

Figure 1. Block Diagram

In coincide with face tracking module there are music system alarm system connected to it which work based on the values detection by modules. If the eye where closed for more than 5-10 seconds the alarm system get activated and based on the detected emotions music system tries to maintain optimal emotion level for driving.

V. RESULT

Based on the different selection of approaches to detect track faces we are able to successfully detect and track faces in a video stream using a mobile application result of which are shown in figure 2. Using this application driver moment can be continuously monitor and if sleep is detected an alarm get triggered vai hardware module connected via bluetooth.

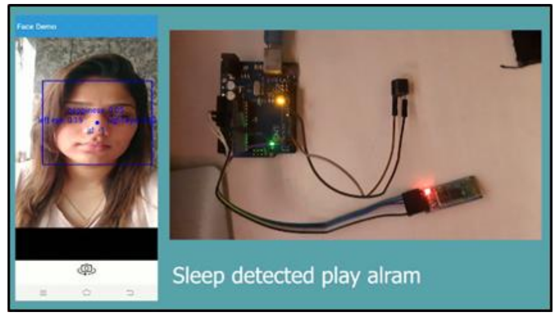

Figure 2. Result Screen

The hardware module consists of an arduino uno, connected with HC-05 bluetooth module which is connected to the mobile application and continuously fetches the data stream coming from the application.

Face 1 of 3		
Bounding polygon	(884.880004882812, 149.546676635742), (1030.77197265625, 149.546676635742), (1030.77197265625, 329.660278320312), (884.880004882812, 329.660278320312)	
Angles of rotation	Y: -14.054030418395996, Z: -55.007488250732422	
Tracking ID	2	
Facial landmarks	Left eye	(945.869323730469, 211.867126464844)
	Right eye	(971.579467773438, 247.257247924805)
	Bottom of mouth	(907.756591796875, 259.714477539062)
	... etc.	
Feature probabilities	Smiling	0.88979166746139526
	Left eye open	0.98635888937860727
	Right eye open	0.99258323386311531

Table 1. Facial feature contours

Facial feature contours probability can be illustrated in the table below which show probalisty for one face out of three face detected in the test stream.

CONCLUSION

The field of image processing technology is growing very fast, there are lots of recent trends in object and face detection systems. In our project we focus on using image processing to be used as a digital assistant for drivers. This system with slight modification can be utilized in different industries for various applications. Also with help OpenAI projects we are able to train our model to detect emotions in faces. We successfully created a web based platform on which we can track multiple human face with their face trailing lines with this we have created an mobile application which will provide a safer journey experience for drivers.

REFERENCES

[1] Communication and Aerospace Technology (ICECA) Face detection and tracking: Using OpenCV, 2017 International conference of Electronics.

[2] Face Detection and Tracking in a Video by Propagating Detection Probabilities. IEEE Transactions On Pattern Analysis And Machine Intelligence, VOL. 25, NO. 10, OCTOBER 2003.

[3] Georgios Georgakis, Md Alimoor Reza, Arsalan Mousavian, Phi-Hung Le, "Multiview RGB-D Dataset for Object Instance Detection", 3D Vision (3DV) 2016 Fourth International Conference on, pp. 426-434, 2016.

[4] Tian Wang, Yang Chen, Mengyi Zhang, Jie Chen, Hichem Snoussi, "Internal Transfer Learning for Improving Performance in Human Action Recognition for Small Datasets", Access IEEE, vol. 5, pp. 17627-17633, 2017.

[5] *Ijaiem, Traffic Surveying And Analysis*, International Journal of Application or Innovation in Engineering & Management.

[6] Smita Desai et al, International Journal of Computer Science and Mobile Computing, Vol.6 Issue.9, September- 2017, pg. 46-50, © 2017, IJCSMC All Rights Reserved 47.

ACKNOWLEDGMENT

We are grateful to the management of Shah & Anchor Kutchhi Engineering College for providing us the facility for the completion of our task. Firstly we extend our gratitude to Dr. Bhavesh Patel, Principle of Shah and Anchor Kutchhi Engineering College for his continuous support. We express our heartfelt thanks to our guide Mrs. Nibha desai for her valuable guidance and advice related to this work.

Our note of thanks goes to our most cherished Head of Department Dr. Shubha Subramaniam for her undiminished trust and his support throughout our tenure and for giving us the opportunity to work on our project which made us capable of handling assignments on our own and become relevant.