e-ISJN: A4372-3114
p-ISJN: A4372-3115

ISSN: 2321-7782 (Online)
ISSN: 2347-1778 (Print)
Impact Factor: 7.529

Volume 10, Issue 9, September 2022

International Journal of Advance Research in Computer Science and Management Studies

Research Article / Survey Paper / Case Study

Available online at: www.ijarcsms.com

# *Emotion Recognition Using Facial Speech and Biomodel Data Algorithms*

**Dr. D J Samatha Naidu**[1]
Professor and Principal,
Annamacharya PG College of Computer Studies,
Rajampet, India.

**B. Manjusha**[2]
Student,
Annamacharya PG College of Computer Studies,
Rajampet, India.

***Abstract:*** ***The interaction between human beings and computers will be more natural if computers are able to perceive and respond to human non-verbal communication such as emotions. Although several approaches have been proposed to recognize human emotions based on facial expressions or speech, relatively limited work has been done to fuse these two, and other, modalities to improve the accuracy and robustness of the emotion recognition system. This paper analyzes the strengths and the limitations of systems based only on facial expressions or acoustic information. It also discusses two approaches used to fuse these two modalities: decision level and feature level integration. Using a database recorded from an actress, four emotions were classified: sadness, anger, happiness, and neutral state. By the use of markers on her face, detailed facial motions were captured with motion capture, in conjunction with simultaneous speech recordings. The results reveal that the system based on facial expression gave better performance than the system based on just acoustic information for the emotions considered. Results also show the complementarily of the two modalities and that when these two modalities are fused, the performance and the robustness of the emotion recognition system improve measurably.***

***Keywords: Emotion recognition, speech, vision, PCA, SVC, decision level fusion, feature level fusion, affective states, human-computer interaction (HCI).***

## I. INTRODUCTION

Inter-personal human communication includes not only spoken language but also non-verbal cues such as hand gestures, facial expressions and tone of the voice, which are used to expressfeeling and give feedback. However, the new trends in human-computer interfaces, which have evolved from conventional mouse and keyboard to automatic speech recognition systems and special interfaces designed for handicapped people, do not take complete advantage of these valuable communicative abilities, resulting often in a less than natural interaction. If computers could recognize these emotional inputs, they could give specific and appropriate help to users in ways that are more in tune withthe user's needs and preferences. Facial emotions are important aspects in human communication that help us to understand the intentions of others. Facial emotion detection became a well attempted research topic now days due to its prospective accomplishments in many Domains such as Medical engineering, Vehicles, Robotics and Forensic applications etc. Emotion Recognition will help to understand the inner feelings for people by using their facial expression.

### 1.1 MOTIVATION

This task is done by detecting the facial actions per every unit of face measurements as a sub part of facial action coding system. This project offers light on utilization of CNN from a live video stream as an input. Using various machine learning libraries like tensor flow and many more. With this development it's advantageous to various domains such as medical engineering, technology, marketing etc.

## 1.2 OBJECTIVE

Detecting emotion using speech data and facial expression images and to implement this project we have trained Convolution Neural network(CNN) algorithm with RAVDESS Audio Dataset for speech emotion recognition and for face expression we have used Emotion Facial Expression images dataset.

## 1.3 EXISTING WORK

In the existing system affective computing is the "computing that relates to, arises from, or influence emotions", or in the other words, any form of computing that has something to do with emotions. The creation of automatic classifier involves collecting information, extracting the features which are important and finally training the data, so it classify and recognize some patterns.

### LIMITATIONS

Creation of model in real life is difficult.

Voice recognition software won't always put your words on the screen completely accurately.

Programs cannot understand the context of language the way that humans can, leading to errors that are often due to misinterpretation.

## 1.4 PROPOSED WORK

A neural network is a type of machine learning algorithm allows the computer to learn by incorporating new data.

Convolution Neural networks is very useful in image recognition in order to analyze visual imagery and are frequently used in classifying the images.

### ADVANTAGES

Automatically detects the important features without any human supervision.

Gives good accuracy.

Computationally Efficient.

## II. RELATED WORK

Several approaches to recognize emotions from speech have been reported. A comprehensive review of these approaches can be found.

Creation of model in real life is difficult.

Voice recognition software won't always put your words on the screen completely accurately.

Programs cannot understand the context of language the way that humans can, leading to errors that are often due to misinterpretation.

 

### 2.1 METHODALOGY FOR WORK FLOW

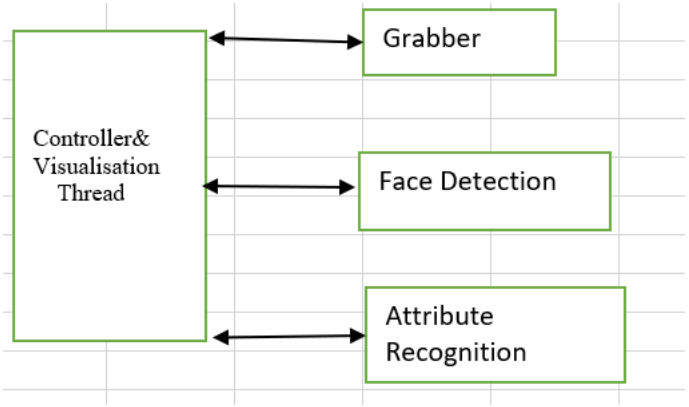


Fig 1: Workflow Methodology

## III. MODULES

we are applying PCA feature extraction algorithm to extract important features from dataset and then train CNN and RNN algorithm and then compare accuracy of both algorithms .we are using above dataset and deep learning CNN and RNN algorithms to build emotion prediction model and then calculate and compare accuracy between CNN and RNN. we are using combination of YOLO face detection and CNN to build emotion prediction model and this title gave better prediction result and accuracy compare to title1 and title2. we are using combination OPENCV and CNN to build deep learning to predict human faces emotion. In this title we are reading dataset images using OPENCV and then training with CNN using SOFTMAX function and then calculate loss/error rate. This module give better prediction accuracy and less error rate with good emotion detection result compare to above 3 algorithms.

## IV. IMPLEMENTATION SCREEN

To run project double click on 'run.bat' file to get below screen.

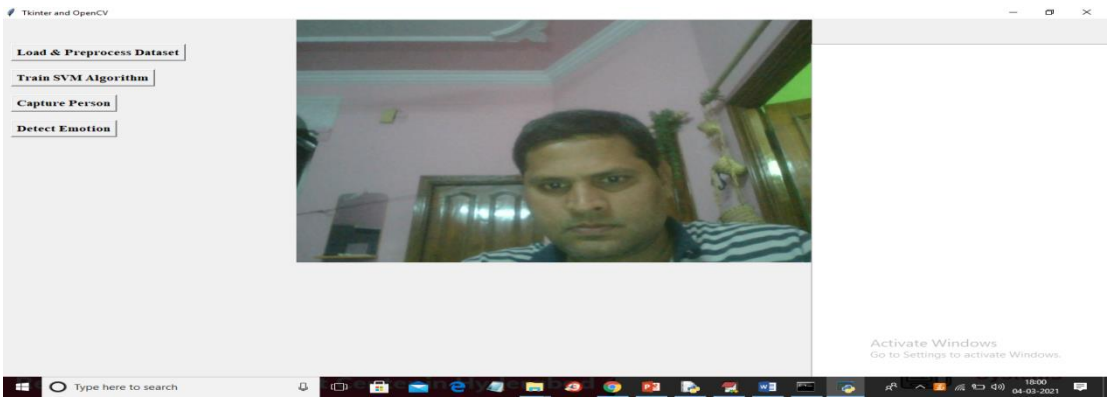


Fig 2:run project

In above screen web cam started and now clicks on 'Load & Preprocess Dataset' button to read images and process them. This process may take 2 to 3 minutes of time.

 

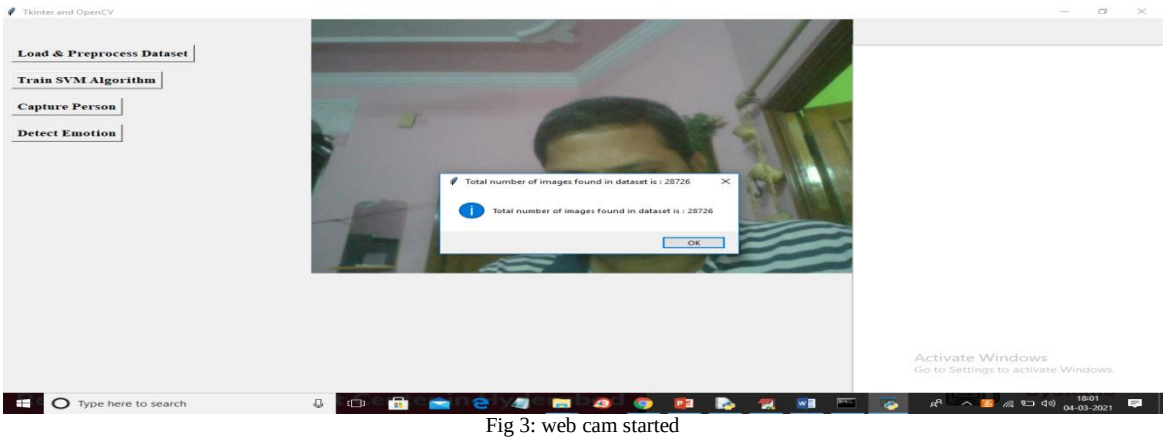

Fig 3: web cam started

In above screen we can see application process 28726 images and now click on 'Train SVM Algorithm' button to train SVM with all those images.

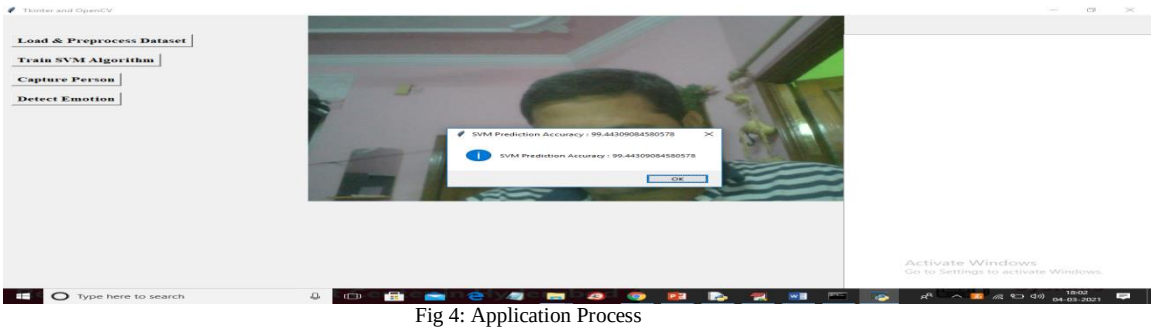

Fig 4: Application Process

In above screen SVM model trained with prediction accuracy as 99%. Now make some expression and then click on 'Capture Image' button to get below screen.

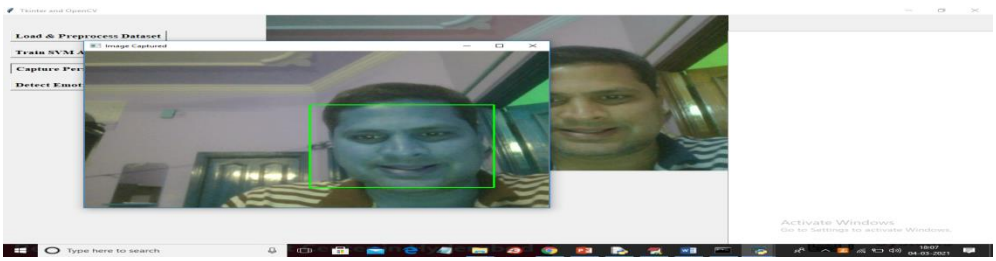

Fig 5: SVM model trained with prediction

Below is the prediction result.

Fig 6:prediction result

Below is next emotion.


Fig 7:Next emotion

Below is the prediction result.

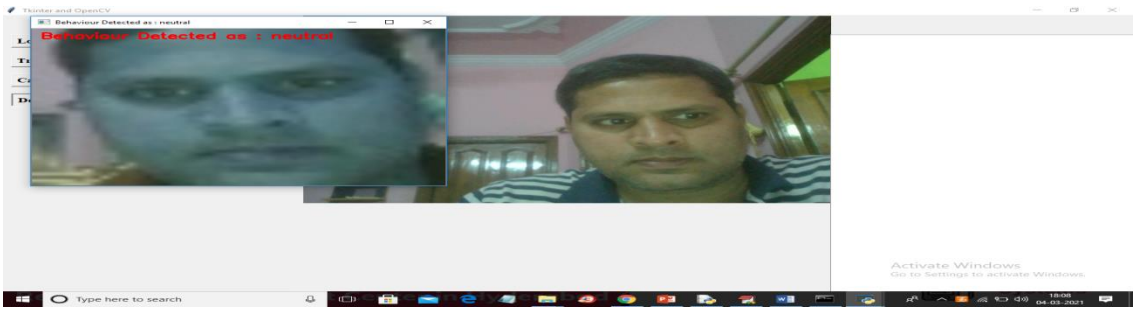


Fig 8:Prediction Result

## V. CONCLUSION

This research analyzed the strengths and weaknesses of facial expression classifiers and acoustic emotion classifiers. In these unimodal systems, some pairs of emotions are usually misclassified. However, the results presented in this paper show that most of these confusions could be resolved by the use of another modality. Therefore, the performance of the bimodal emotion classifier was higher than each of the unimodal systems.

Two fusion approaches were compared: feature-level and decision-level fusion. The overall performance of both approaches was similar. However, the recognition rate for specific emotions presented significant discrepancies. In the feature-level bimodal classifier, anger and neutral state were accurately recognized compared to the facial expression classifier, which was the best unimodal system. In the decision-level bimodal classifier, happiness and sadness were classified with high accuracy. Therefore, the best fusion technique will depend on the application.

The results presented in this research show that it is feasible to recognize human affective states with high accuracy by the use of audio and visual modalities. Therefore, the next generation of human-computer interfaces might be able to perceive humans feedback, and respond appropriately and opportunely to changes. Recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies.

## References

1. K. Graves, Ceh: Official certified ethical hacker review guide: Exam 312-50. John Wiley & Sons, 2007.
2. R. Christopher, “Port scanning techniques and the defense against them,” SANS Institute, 2001.
3. M. Baykara, R. Das¸, and I. Karado ˘gan, “Bilgi g ¨uvenli ˘gi sistemlerinde kullanilan arac¸larin incelenmesi,” in 1st International Symposium on Digital Forensics and Security (ISDFS13), 2013, pp. 231–239.
4. RashmiT V. “Predicting the System Failures Using Machine Learning Algorithms”.International Journal of Advanced Scientific Innovation, vol. 1, no. 1, Dec. 2020, doi:10.5281/zenodo.4641686.
5. S. Robertson, E. V. Siegel, M. Miller, and S. J. Stolfo, “Surveillance detection in high bandwidth environments,” in DARPA Information Survivability Conference and Exposition, 2003. Proceedings, vol. 1. IEEE, 2003, pp. 130–138.

 