Emotion Classification using Convolutional Neural Network

Shamit Ranjan Barman
UG Student, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

Joy Dev Nath
UG Student, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

Sajib Kumar Mitra
UG Student, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

A Krishnamoorthy
Assistant Professor (Senior), School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

R Kannadasan
Assistant Professor (Senior), School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

N Prabakaran
Assistant Professor (Senior), School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

Abstract

In this paper, using two distinct media, video and text, the model will undertake emotion analysis in order to classify participants into various categories, which will result in the identification and categorization of different emotions that are being displayed by the participants. In this paper, Convolutional Neural Network, in assistance with OpenCV, Keras, Tenorflow and Numpy, has been used to train the emotion classifying model so that it could distinguish different emotions. The model has been trained and validated with over thirty-five thousand images to increase accuracy. Pre-processed images are supplied from the database to the CNN model and it undergoes all the layers to make sure the model can accurately identify the patterns of the different emotions and finally identify it from a emotion dictionary using prediction and matching the pattern matching with the highest percentage with the respective emotion which are: happy, angry, disgust, neutral, sad, fear and surprise.

Keywords: *Personality, video, text, analysis, convolutional neural network, opencv, tensorflow, keras, numpy, categorizing, emotion, medium, placements.*

1. Introduction

Facial expression is a type of nonverbal communication that conveys information about a person's feelings. In computer vision and artificial intelligence, detecting and understanding human emotion is a major difficulty. The major goal of this work is to provide a reliable method for detecting and identifying human emotions in real time, such as anger, sadness, happiness, surprise, fear, disgust, and neutral. This paper will use a CNN model to detect and classify the emotions being displayed in a customer's face in real-time while a survey is being conducted. OpenCV, Keras, TensorFlow and Numpy, has been used to train the emotion classifying model so that it could distinguish different emotions and classify it for the user. The model is trained using the dataset FER-2013 which includes approximately thirty-six thousand images of which 80% is used to train the model and 20% is used to validate and test the model. Post the training the model is ready to identify emotions from real–time video.

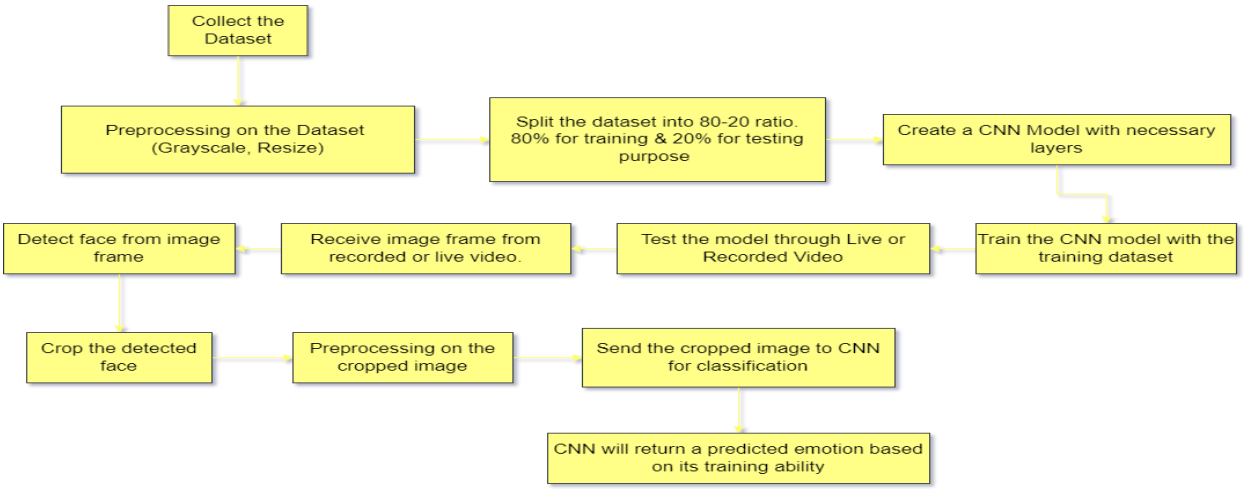
2. Related Work

[1]The paper "Real-time Convolutional Neural Networks for Emotion and Gender Classification." Uses the algorithm Convolutional Neural Network (CNN) with Xception pre-trained model. They have an accuracy of 66% on Emotion classification with 7 classes of Emotions (Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral). They have used the Dataset FER-2013. But their weak points include predicting "sad" instead of "fear" and predicting "angry" instead of "disgust". [2]The paper "A real-time face emotion classification and recognition using deep learning mode" uses the algorithm Convolutional Neural Network (CNN) with the VGG16 pre-trained model. They have an accuracy of 88% with 7classes of Emotions (Angry, Disgust, Fear, Happy, Sad, Surprise, Neutral). But their weak point was that they only used 2901 images from the whole dataset. [3]The paper "Human Emotion Detection and Classification Using Convolution Neural Network." Uses the algorithm Convolutional Neural Network (CNN) as classifier & GLCM and LPB based feature extraction. They only detected 6 classes of emotions: Happy, sad, surprise, angry, disgust, fear. Each class they have only 215 images. The model was tested from the image, not in real-time or recorded video. [4] The techniques Support Vector Machine (SVM) as a classifier and Histogram of Oriented Gradients (HOG) as a feature extractor are used in the study "Emotion Detection Through Facial Feature Recognition." They can classify five emotions with an accuracy of 81 percent: happy, sad, surprise, rage, and fear. However, their model was evaluated using a picture rather than a live or recorded video. [5] The algorithm Kaggle's Facial Expression Recognition Challenge and the Karolinska Directed Emotional Faces (KDEF) datasets were used in the publication "Emotion Recognition from Facial Expression using Deep Learning." This research shows how emotion may be used in a network teaching system. Emotion recognition is unaffected by wearing glasses on the face. However, the distance between the camera and the subject's face will have an effect on facial recognition. The impact of the human face on emotion detection is influenced by factors such as hearing, sitting positions, and light strength. [6] Deep-Learning Based FER Approaches are used in the study "A Brief Review of Facial Emotion Recognition Based on Visual Information." With an average of 72.65% to 63.2 percent, they were able to exceed traditional tactics. However, they utilised large-scale datasets and required a lot of computational power, which made it unsuitable for mobile platforms with limited resources. [7] The study "Automatic Emotion Recognition Using Facial Expression" was published in the journal Psychological Science. For feature selection and picture recognition, multi-modal deep learning is applied. We employed

segmentation, pre-processing, and feature extraction. This article presented a new architecture for future interactive television. A real-time emotion identification system underpins the suggested method. However, the rate of recognition varies depending on the type of face database employed. For real-time applications, there is a need to increase detection and timing performance. [8] The study "Facial Emotion Recognition: A Survey & Real-World User Experiences in Mixed Reality" was published in the journal "Facial Emotion Recognition: A Survey & Real-World User Experiences in Mixed Reality." Device for Mixed Reality In Augmented Reality, the Microsoft Holo-lens is utilized to observe emotion detection (AR). A wide variety of output is created, obviating the need for previous versions to be imitated. However, this necessitates extensive debugging and highly powerful sensors. As a result, it might not be appropriate for everyday usage. [9] In the paper "Fuzzy Emotion Recognition Using Semantic Facial Features and Knowledge-based Fuzzy" the proposed research is creating fluffy feeling recognition structure as speculation of blended feeling. Fluffy feeling recognition is the programmed feeling recognition from a facial expression that consents to the standards of the rapists in examining feelings by zeroing in on changes in facial segments just as obliging the part of uncertainty and FRS. Fuzzy terms use natural linguistic or language variables that are understood by humans in everyday life. Fuzzy linguistic variables have both quantitative and qualitative value; the quantitative value is a real value and the qualitative value is the linguistic value, e.g., low, medium, high. Subsequently, the recognition emulates the manner in which an analyst examination feeling from a facial expression and facial part changes. The aftereffect of the testing utilizing owns made Indonesian Mixed Emotion Dataset (IMED) shows that the proposed framework acquired a high exactness rate of 88.519%, accuracy. [10] The paper "Multimodal Emotion Detection Using Deep Learning" provides an overview of the emotional recognition of multimodal signals using deep learning, as well as a comparison of their applications based on current research. Multimodal affective computing systems are compared to unimodal solutions because they have a greater categorization accuracy. The findings suggest that a multimodal method based on biological signals may be used to identify emotional states with better precision and improvement. A new approach and architecture based on the kernel matrix and a deep neural network are suggested to improve the efficiency of multimodal fusion. On different given databases, this approach provides varying accuracy, ranging from roughly 63 percent when dealing with DEAP data set to less than 57 percent when dealing with DECAF data set.

3. Proposed Architecture

Figure.1. Proposed architecture diagram



The dataset is divided into seven categories: anger, sorrow, happiness, surprise, fear, disgust, and neutral. The dataset is further divided into two parts where 80% of the images are used to train the CNN model and 20% are used to test the model. The images are preprocessed and made into grayscale so the patterns in the facial expressions will have more clarity. The model now undergoes training by going through the dataset. In the CNN Model, a filter runs throughout the image and extracts pixels according to the model. Then these pixels undergo RELU and pooling. This is a layer of CNN. The more layers there are, the more accurate the classification of the emotions will be. The model is first tested with still images and after that, it is tested with live videos where frames are taken from the video, and front here the samples are extracted which trains the model. After the training is completed, the rest 20% of the dataset is used to test the CNN model. After this testing is done, live video testing is done to check the real-time application of the model. The process is similar, frames are extracted from the video and the faces are highlighted. The image undergoes preprocessing where it is made into grayscale and undergoes the layers of the CNN model. The model then gives us the emotion that the person on the video is experiencing

4. Further Work

The further scope of this paper is to detect facial emotions in a different angle when the user is facing the camera. Currently facial emotions most accurately detected when the user is facing directly into the camera. More work needs to be done to increase the accuracy of emotion detection in some aspects like differentiating “Anger” from “Disgust”. Also, another major future scope of this paper is to apply it into a commercial/real-world situation to collect facial emotion data & measure its accuracy and viability. Facial recognition technology has a promising future. Forecasters predict that this technology will expand at a rapid pace and create significant money in the next years. The primary categories that will be heavily affected are security and surveillance. Private companies, public structures, and schools are among the other domains that are now welcome it with open arms. It is expected that businesses and financial institutions would embrace it in the future years to prevent fraud in debit/credit card transactions and payments, particularly those made online. This method

would close the gaps in the widely used yet insecure password system. Robots using face recognition technologies may potentially make an appearance in the future. They can assist in the completion of activities that are inconvenient or difficult for humans to do. Embed systems such as Arduino, Raspberry Pi, and others may be readily included and profit from this system not only as an attendance system but also as a security system, and it can be created in a variety of different sectors since AI has become a need for all businesses.

5. Result and Discussion

After running the model, it is detecting different expressions of the user and its specified name. A 89% accuracy of 7 distinct of emotions has been successfully achieved.

Here are some screenshots with the emotion classifying model output.

In this image, the model is detecting that the person is happy after going through the prediction and recognizing the patterns and comparing the percentage match with the other emotions.

Figure 2: Model detecting Happy emotion

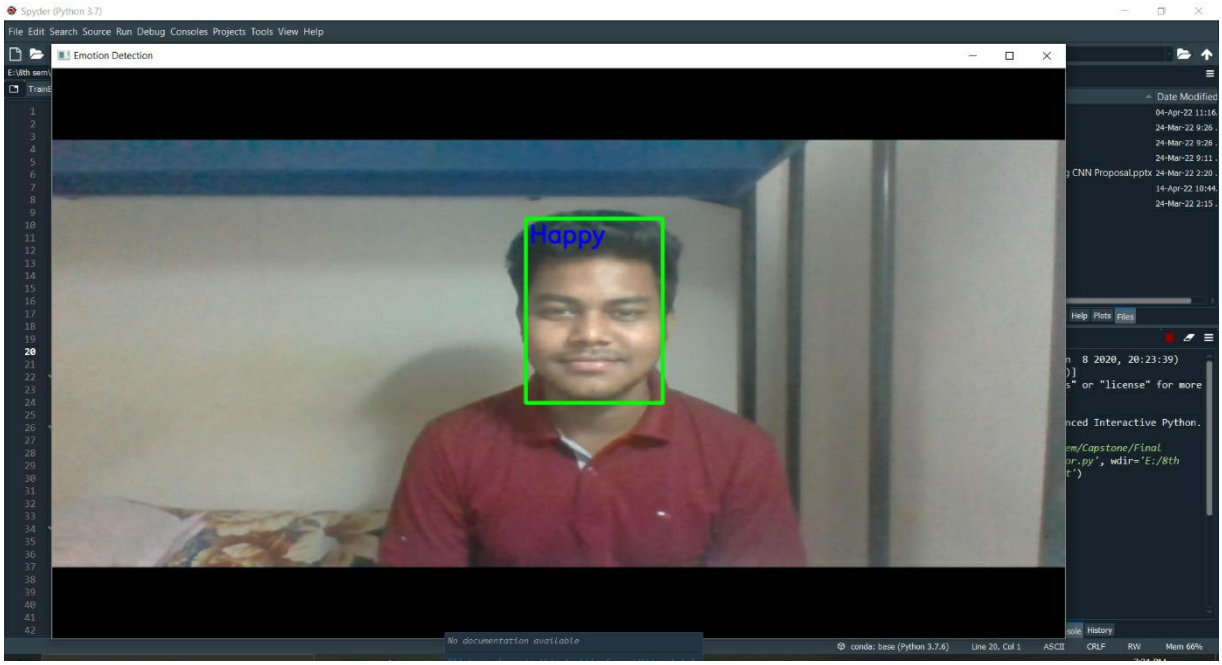


Figure 3: The graph shows that the pattern matches that of happy the most.

In this one, the model is detecting a neutral emotion.

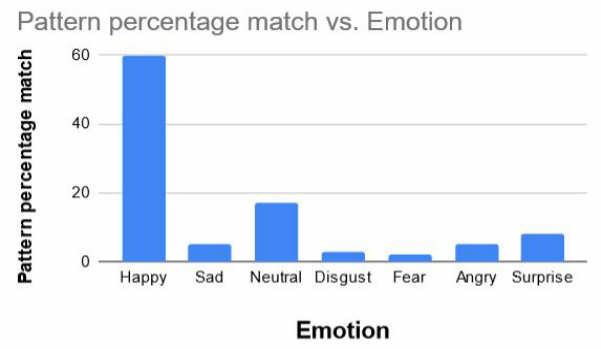


Figure 4: Model detecting Neutral emotion

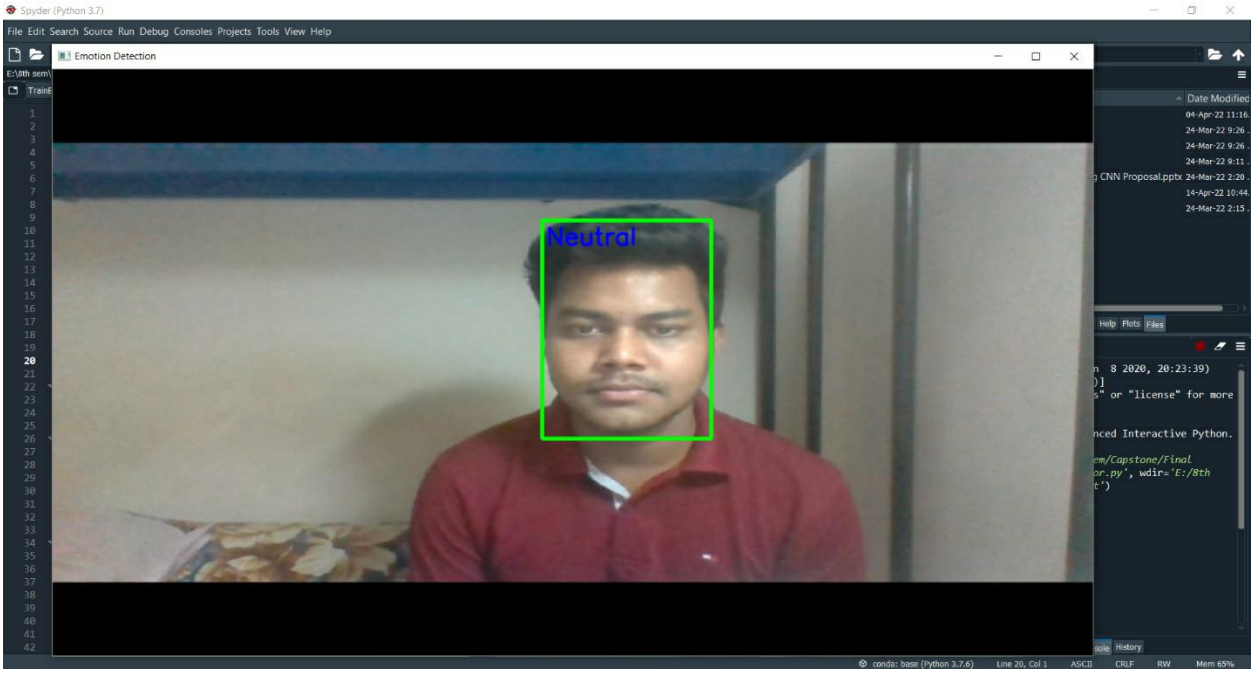
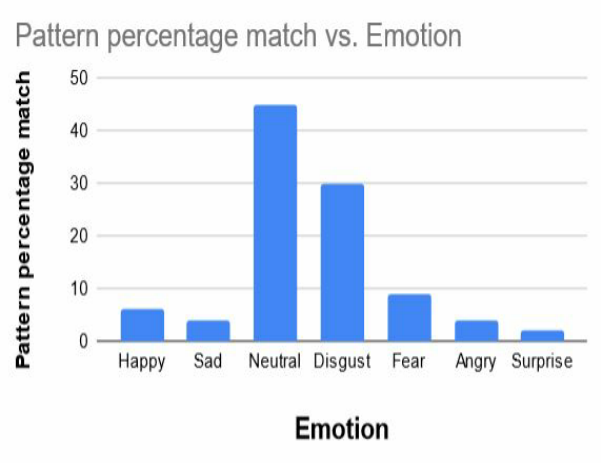


Figure 5: The graph shows that the pattern matches that of neutral the most



In this one, the model is detecting a neutral emotion.

In this one, the model is detecting a neutral emotion.

Figure 6: Model detecting Surprise emotion

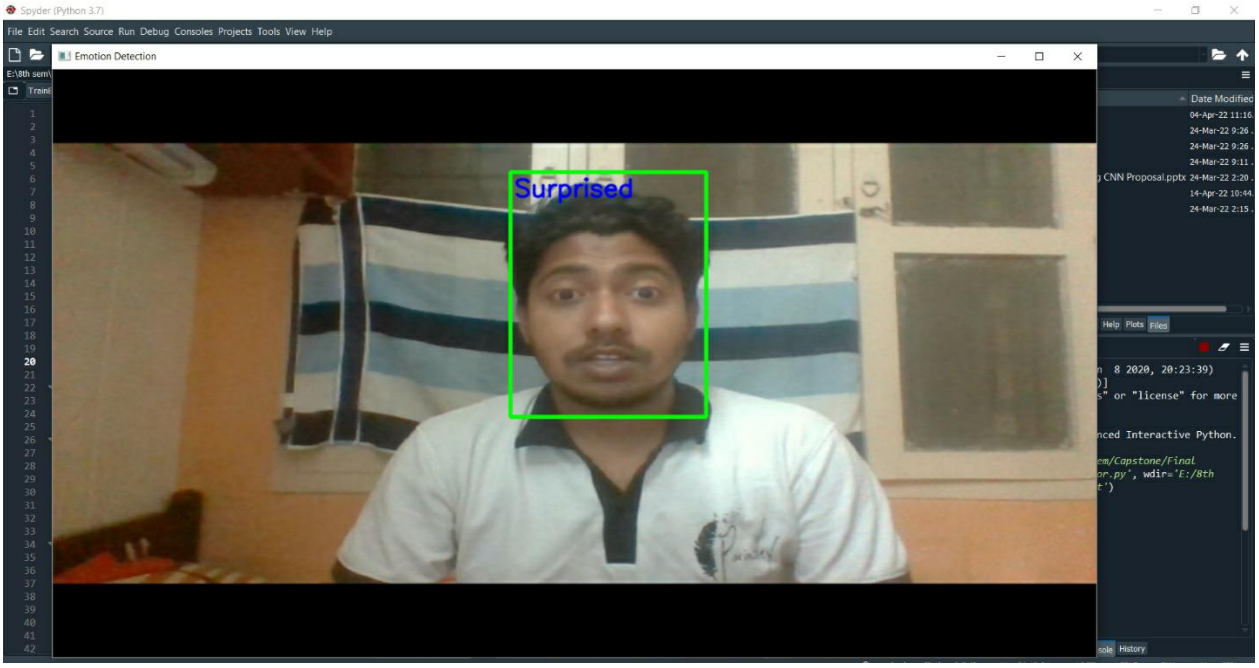
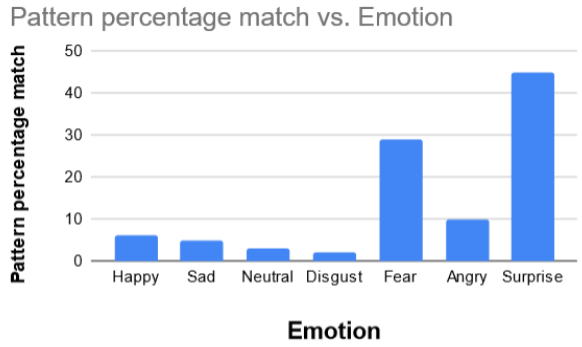


Figure 7: The graph shows that the pattern matches that of surprise the most.



6. Conclusion

In this paper, the user's emotions will be recognized using facial expression. Face detection and interpretation in real time, including happy, sad, angry, terrified, surprised, disgusted, and neutral expressions. During a chaotic scenario, an automatic facial expression Recognition system will recognize and locate faces, extract facial features, and classify facial expressions. It proposes a methodology to handle the difficulties of emotion detection based on face recognition in virtual learning environments, while also considering efficiency and accuracy. This system has the capacity to monitor people's emotions, to distinguish between them and label them properly, and to utilize that knowledge to direct a person's thoughts and conduct.

References

[1] Arriaga, Octavio, Matias Valdenegro-Toro, and Paul Plöger. "Real-time convolutional neural networks for emotion and gender classification." arXiv preprint arXiv:1710.07557 (2017).

[2] Hussain, Shaik Asif, and Ahlam Salim Abdallah Al Balushi. "A real time face emotion classification and recognition using deep learning model." In Journal of Physics: Conference Series, vol. 1432, no.

1, p. 012087. IOP Publishing, 2020.

[3] Komala, K., D. Jayadevappa, and G. Shivaprakash. "Human Emotion Detection And Classification Using Convolution Neural Network." European Journal of Molecular & Clinical Medicine 7, no. 06 (2020).

[4] Madinah, Saudia Arabia. "Emotion detection through facial feature recognition." International Journal of Multimedia and Ubiquitous Engineering 12, no. 11 (2017): 21-30.

[5] Mehta, D., Siddiqui, M. F. H., & Javaid, A. Y. (2018). Facial emotion recognition: A survey and real-world user experiences in mixed reality. Sensors, 18(2), 416.

[6] Tian, Y., Kanade, T., & Cohn, J. F. (2011). Facial expression recognition. In Handbook of face recognition (pp. 487-519). Springer, London.

[7] Lakshitha V., Dalal R., Krishnamoorthy A, A novel approach to smart hydroponics system using IoT (2020), Journal of Advanced Research in Dynamical and Control Systems, 12 (5 Special Issue), pp. 890 - 898,DOI: 10.5373/JARDCS/V12SP5/20201832

[8] Krishnamoorthy A., Vijayarajan V, A novel and energy efficiency approach for enhancing the lifespan of cluster using mobile sink (2019) International Journal of Advanced Science and Technology, 28 (13), pp. 290 – 296.

[9] Raghuvanshi, A., & Choksi, V. (2016). Facial expression recognition with convolutional neural networks. CS231n Course Projects, 362.

[10] Mehta, Dhwani & Siddiqui, Mohammad Faridul Haque & Javaid, Ahmad. (2018). Facial Emotion Recognition: A Survey and Real-World User Experiences in Mixed Reality. Sensors (Basel, Switzerland). 18. 10.3390/s18020416.

[11] Liliana, D. Y., Basaruddin, T., &Widyanto, M. R. (2019). Fuzzy Emotion Recognition Using Semantic Facial Features and Knowledge-based Fuzzy. International Journal of Engineering and Technology, 11, 177-186.

[12] Abdullah, S. M. S. A., Ameen, S. Y. A., Sadeeq, M. A., &Zeebaree, S. (2021). Multimodal emotion recognition using deep learning. Journal of Applied Science and Technology Trends, 2(02), 52-58.