IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.125241

Design and implementation of Emotion Recognition System

Anjali Vijay Nikhate[1], Devyani Waman Channe[2], Sakshi Sanjay Sontakke[3], Isha Prabhakar Chaple[4], Prof . Manisha Jagdish More[5]

Final year Student, CSE Department, RCERT , Chandrapur, India[1-4]

CSE Department, RCERT , Chandrapur, India[5]

Abstract: Human emotion recognition has significant part in our day to day lives. Objective of the study is to develop and implement a system capable of analyzing, predicting, and classifying emotions in real-time using Convolutional Neural Network (CNN) algorithm, with the assistance of the OpenCV library. The proposed approach enables the classification of various emotions, including anger, disgusted, fear, happy, neutral, sad, and surprised, based on feature extraction. FER2013 dataset is utilized for performance evaluation, and pre-processing techniques such as facial landmark detection are employed during training and testing.

This dataset is utilized for training and testing purposes, as it is understood that while one-third of communication is conveyed verbally, the remaining two-thirds are conveyed through non-linguistic. means. Although there are existing emotion recognition systems, in real-life scenarios, consider the example of mental hospitals where this technology provides medical professionals with insights into patients' emotional states. By leveraging this technology, medical professionals can offer improved care and potentially enhance outcomes. Facial expression recognition remains a challenging problem in computer vision, as images of the same person in different expressions can change in brightness, background, and position.

Keywords: Emotion recognition, Convolutional Neural Network (CNN), OpenCV, Pre-processing.

I. INTRODUCTION

Human communique conveys essential information now not only about rationale but additionally approximately dreams and feelings as well. The most expressive way for humans to express their feelings is through facial expressions. Humans have little difficulty in finding and interpreting facial expressions in a scene. However, developing an automatic system to accomplish this challenge remains quite difficult. Facial expressions play a vital role in communication as they convey a wealth of information. According to Mehrabian, a researcher known for his work on communication, words alone account for just 7% the overall message conveyed, while voice intonation carries 38%, and facial expressions convey a significant 55% of the message. It's important to note that Mehrabian's model specifically refers to the relative importance of different communication channels in conveying emotions rather than overall communication. The field of emotion recognition has various applications across different industries.

In healthcare, it can be used to assess patient well-being and aid in diagnosing conditions such as autism or depression. A device that correctly performs those operations within the actual global could be a crucial step in reaching similar human interactions among people and machines.

Human communication conveys essential information now not only about rationale but additionally approximately dreams and feelings as well. Object detectors used for emotion detection . This study focus on application of deep learning algorithms for real-time facial emotion classification using a webcam. The implemented object detection models effectively identify facial expressions and classify them into different emotional categories. In unique, the importance of automatically spotting feelings from human speech and other conversation cues has grown with increasing function of spoken language and gesture interfaces in human-computer interactions and laptop mediated programs. The utilization of FER (Facial Expression Recognition) systems is increasingly prevalent across various industries, including the medical sector. One notable application is its implementation within mental health facilities, where it can assist in assessing patients' emotional states. Moreover, FER technology has found relevance in understanding tourist satisfaction during their visits to different locations. These examples demonstrate the limitless potential and broad range of applications for this innovative technology.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.125241

The CNN algorithm plays a crucial role in analysing images for emotion detection. It is responsible for both extracting relevant features from the images and classifying them based on facial characteristics, thereby aiding in the accurate detection of human emotions. Traditional FER methods gained significant popularity over time. These approaches involved initial face detection, followed by the extraction of facial features, and finally, the classification of these features using classifiers. However, the emergence of deep learning-based FER techniques, specifically Convolutional Neural Networks (CNNs), revolutionized the field. These advanced methods replaced the traditional approach, offering more accurate and efficient facial expression recognition capabilities.

II. LITERATURE REVIEW

The ability to learn is significantly impacted by a person's emotions. Children with high-functioning autism (HFA) frequently struggle to focus and pay attention in class because they are unable to control their emotions. Support vector machines (SVM) and the sliding window technique were used to build classifiers for identifying emotions. By training the DNN model, we may identify a collection of ideal characteristics for classifying face emotions from diverse facial viewpoints.

Computer vision technologies are currently undergoing active development, and one of the key challenges they aim to tackle is recognition. Developers now have access to a wide range of libraries that can assist them in addressing various computer vision-related issues. To ascertain the functionality of these libraries, it is crucial to evaluate their efficacy. The easiest method to learn everything there is to know about a certain system, library, or API is to read the documentation carefully. As a result, the purpose of this article is to perform research that focuses on the theoretical underpinnings of developing a trustworthy face recognition system. It will go into detail on how technologies and tools are used for implementation, analyse their benefits and drawbacks, and do performance assessments to address any related issues. The previous ten years have seen the development of a number of expression recognition techniques, and recent years have seen significant advancements in this field of study. The network training process took roughly 8 days. The quantity of generated maps and the size of the kernels serve as the primary parameters for convolutional layers.

Camera which uses real-time pictures have been widely employed as effective tools for both face detection and human identification systems. These cameras utilize face detection algorithms to capture images of faces present in the captured items. The resulting image obtained from the camera lens comprises both the face and the background. In the face identification process, specific facial characteristics or traits are identified and sought in the camera images. The system then determines whether these patterns correspond to faces or not. By analysing the position of the Region of Interest (ROI) and also the coordinates of the face image, the system identifies the detected face image within the backup image generated by each camera lens..

The face is one part of the face with a distinctive shape from the contour of the human body, beginning with arched eyebrows, squinting, producing curves on the lips, the direction of the glance with the eyes, etc. The Haar Cascade Classifier technique is used in this study's face detection process. The mathematical function (Haar Wavelet) with a box shape is referred to as Haar. The RGB value of each pixel was the sole factor considered at first when processing images, but this approach proved to be inefficient. The Haar-Like characteristic was created via an image processing method created by Viola and Jones.

The process known as Haar-like feature extraction operates on images by analyzing them in square regions, with each box containing multiple pixels. These regions produce various values after processing, which indicate the presence of light and dark areas. These values form the foundation for image processing tasks. The software employed in this study automatically detects the face in the provided image, Zortion, and adjusts its position based on facial data within each image. The categorization of human facial expressions encompasses three key steps: facial expression categorization, feature extraction, and face detection. The researchers created a system that can categorise face expressions on a broad scale, focusing on seven essential human expressions:

1. **Happy**: The recognition of a positive emotional state characterized by the activation of cheek muscles, with an upward movement seen at the side and edge of the lips, conveying happiness and joy.

2. **Angry**: The identification of an intense emotional state associated with the activation of facial muscles, typically involving a downward or tense movement of the cheek muscles, with lips pressed together or turned downward at the edges, expressing anger and displeasure.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.125241

3. **Sad**: The detection of a profound emotional state characterized by the activation of cheek muscles, often accompanied by a downward movement at the side and edge of the lips, indicating sadness, sorrow, or a sense of unhappiness.

4. **Fear**: The discernment of an emotional state triggered by fear or anxiety, where the cheek muscles may appear tense or contracted, with the lips slightly parted or pulled back at the corners, reflecting a state of fear and concern.

5. **Disgust**: The recognition of an emotional state characterized by the activation of cheek muscles, potentially resulting in a wrinkling or raising of the upper lip, coupled with a downward or curled movement at the edge of the lips, expressing disgust or revulsion.

6. **Surprise**: The identification of an emotional reaction to unexpected or startling stimuli, often accompanied by the activation of cheek muscles, with the mouth opening slightly and the lips parting at the sides or edges, displaying surprise or astonishment.

7. **Neutral**: The categorization of a facial expression and overall emotional state that shows minimal activation of cheek muscles, with lips in a relaxed or neutral position, indicating a lack of any specific or dominant emotional expression.

In this section, we will discuss the works connected to the suggested system, which encompass various techniques that aim to address different issues and advance the field of emotion recognition. While this research primarily focuses on the CNN model, it also incorporates several feature extraction and preprocessing approaches to compare how their inclusion affects the accuracy of the model. Over the years, researchers have extensively studied facial expressions, continually seeking improvements in each project undertaken. This ongoing pursuit has created numerous opportunities within the field. The main goal of their research is to enhance the accuracy of the FER2013 dataset. To achieve this, they have employed a convolutional neural network as the framework for categorizing the basic seven emotions in their proposed model.

The study demonstrates that convolutional neural networks can significantly improve the precision of biometric applications. However, it should be noted that the recognition rate for each emotion class may vary, as achieving an entirely a similar or almost identical rate of identification for each class proved challenging. To address this, a face expression dataset called RAF-DB has been introduced. This dataset comprises photos of individuals across different age groups and in diverse dynamic environments. The researchers have utilized a maintaining a deep locality CNN classification technique the seven basic emotions within this dataset.

Figure 1 : 7 Different Emotions

III. METHODOLOGY

This paper introduces a new approach to enhance real-time emotion recognition. The proposed method incorporates addinal feature extraction techniques to improve training accuracy in conjunction with a Convolutional Neural Network (CNN). CNNs are widely used for analyzing images in emotion recognition tasks and consist of various layers, including convolutional, pooling, fully connected, and normalization layers. The proposed method employs a CNN's two-level structure with four filters for pattern detection. The initial part of the CNN focuses on background removal by identifying shapes, edges, textures, and facial objects. Specifically, edge detectors, circle detectors, and corner detectors are utilized at the beginning of the convolutional layer. To start with Face is detected, then we use the CNN filter which will predict the facial features which consist of ear, lips, nose, eyes, and cheeks, and from this minimum of three features should be there on the face.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.125241

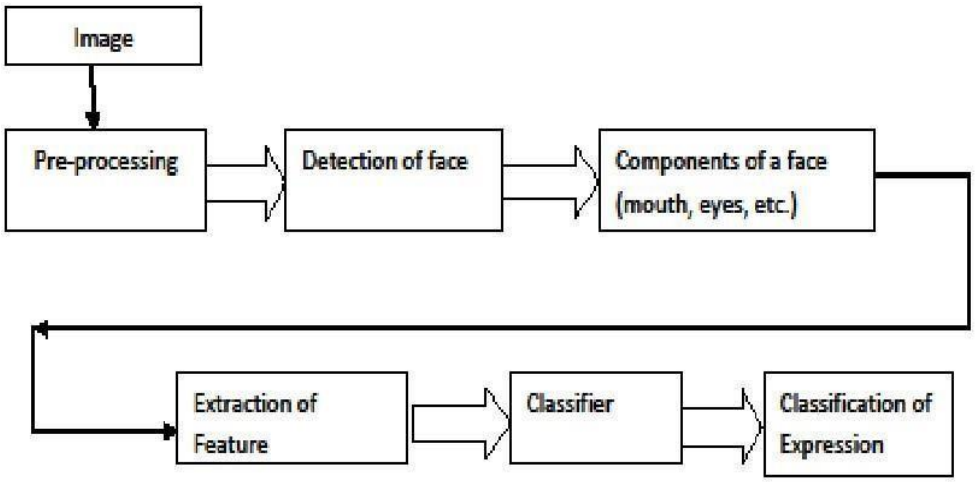


Figure 2 : Basic steps involved shows in architecture of the proposed system.

This research presents a system consisting of four modules. The initial module utilizes a webcam to capture real-time video and applies Local Binary Patterns (LBP) for face detection. The subsequent module focuses on image preprocessing, which involves cropping, resizing, and intensity normalization techniques.

Feature extraction methods are then applied to extract relevant features from the processed image, and face detection is performed based on these selected features. It is worth noting that pre-processing plays a significant role in this system. The following section provides a detailed explanation of the pre-processing techniques employed in the suggested setup. The data preprocessing stage encompasses the following steps:

A. Face Recognition
B. Color space conversion
C. Data Scaling
D. Data Augmentation

A. **Face Recognition** : In order to identify the location of faces within an image, a face detection process called Face Registration is employed. OpenCV Cascade classifier is commonly used for this purpose, enabling the detection of faces. Once a face is detected, the corresponding portion is cropped out, eliminating background complexity and facilitating more efficient model training.

B. **Color space conversion** : Initially, images are composed of three channels - red, green, and blue - with a size of 48x48 pixels. To reduce pixel value complexity, the dataset images are converted into grayscale, which results in a single-channel representation.

C. **Data Scaling** : Normalization is applied to the dataset images to ensure consistency and facilitate effective training. This process modifies the range of pixel intensity values within a certain limit. By stretching the contrast or histogram of the images, deep networks can better analyze them, leading to improved performance.

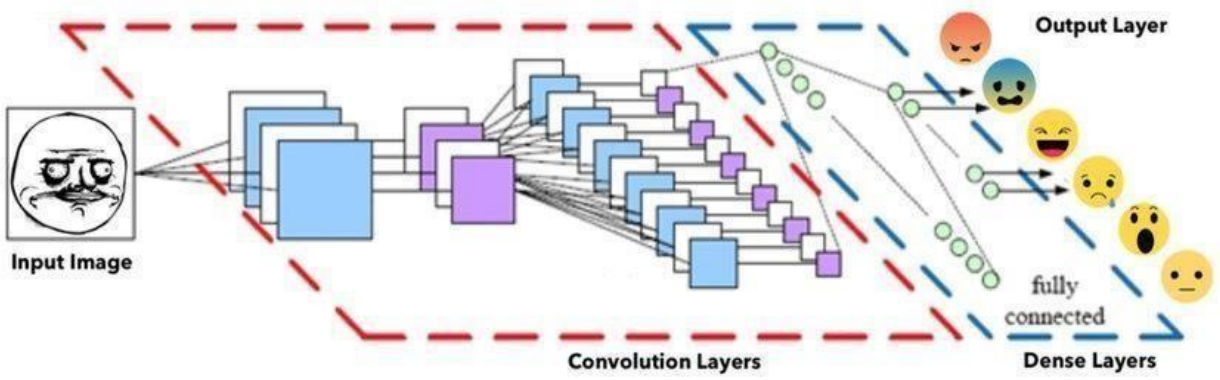
D. **Data Augmentation** : Convolutional neural networks often benefit from larger datasets, as they can extract more features and match them with unlabeled data. In situations where collecting sufficient data is challenging, image augmentation can be employed. This technique generates additional images by applying various operations, such as random rotation, shifts, shear, flips, among others, to the existing dataset.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.125241



IV. CONCLUSION

Finding the appropriate expression is just as crucial for effective communication as comprehending the message itself since facial emotions play a significant role. Distinguishing between various types of facial emotions is essential for a range of applications, including robot vision, video surveillance, digital cameras, security, and human-computer interfaces. The objectives of this project are to achieve face recognition and emotion recognition using computer vision techniques. Additionally, advanced feature extraction and classification in face expression recognition will be enhanced. The OpenCV library has been found to be more effective and perform better for face detection and recognition, as indicated by the study mentioned in this article. Furthermore, it suggests that developing recognition applications for the Internet of Things (IoT) platform with OpenCV is preferable.

In this study, we present a system that combines established techniques like Convolutional Networks with specific image pre-processing procedures to recognize facial expressions. The proposed method utilizes a real-time camera to detect individuals' emotions. We analyzed seven different facial expressions from two datasets, namely JAFFE and FER2013, using a Haar-like detector. The system detects faces in input images, categorizing them into one of the following seven facial expressions: surprise, fear, disgust, sadness, anger, happiness, and neutrality. The Convolutional Neural Network approach with data augmentation has demonstrated higher efficiency compared to other machine learning methods in image processing. The proposed model achieved superior validation accuracy compared to existing models, with an accuracy rate of 80% on the FER2013 database.

ACKNOWLEDGMENT

We would like to express our deepest appreciation and respect to **Dr. Nitin Janwe**, the Head of Department of the Computer Science and Engineering department at RCERT in Chandrapur, Maharashtra. Additionally, we extend our sincere gratitude to **Prof. Manisha More**, our dedicated project guide.

We take this opportunity to publicly acknowledge and express our gratitude to all those who contributed to the design and implementation of the Emotion Recognition System. It is through the collaborative efforts and unwavering support of numerous individuals and organizations that this project has reached its successful completion.

REFERENCES

[1] M. Shidujaman, S. Zhang, R. Elder, and H. Mi, ""roboquin": A mannequin robot with natural humanoid movements," in *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2018, pp. 1051–1056.

[2] Jung H, Lee S, Yim J, Park S, Kim J (2015) Joint fine-tuning in deep neural networks for facial expression recognition. In: Proceedings of the IEEE international conference on computer vision. pp 2983–2991

[3] Zhang K, Huang Y, Du Y, Wang L (2017) Facial expression recognition based on deep evolutional spatial-temporal networks. IEEE Trans Image Process 26(9):4193

[4] Wu YL, Tsai HY, Huang YC, Chen BH (2018) Accurate emotion recognition for driving risk prevention in driver monitoring system. In: 2018 IEEE 7th global conference on consumer electronics (GCCE). IEEE, pp 796–797

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.125241

[5] Y.-H. Byeon and K.-C. Kwak*, "Facial expression recognition using 3d convolutional neural network," in *International Journal of Advanced Computer Science and Applications*, vol. 5, no. 12, 2014.
[6] M. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan, "Recognizing facial expression: machine learning and application to spontaneous behavior," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005*, vol. 2, 2005, pp. 568–573 vol. 2.
[7] Z. Zhang, M. Lyons, M. Schuster, and S. Akamatsu, "Comparison between geometry-based and gabor-waveletsbased facial expression recognition using multi-layer perceptron," in *Third IEEE International Conference on Automatic Face and Gesture Recognition, 1998. Proceedings*, 1998, pp. 454–459.
[8] P. Yang, Q. Liu, and D. Metaxas, "Boosting coded dynamic features for facial action units and facial expression recognition," in *IEEE Conference on Computer Vision and Pattern Recognition, 2007. CVPR '07*, 2007, pp. 1–6.
[9] B. Fasel, "Robust face analysis using convolutional neural networks," in *16th International Conference on Pattern Recognition, 2002. Proceedings*, vol. 2, 2002, pp. 40–43 vol.2.
[10] P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion," Journal of Personality and Social Psychology, vol. 17, no 2, p. 124-129, 1971.
[11] A. Savva, V. Stylianou, K. Kyriacou, and F. Domenach, "Recognizing student facial expressions: A web application," in 2018 IEEE Global Engineering Education Conference (EDUCON), Tenerife, 2018, p. 1459-1462.
[12] Krithika L.B and Lakshmi Priya GG, "Student Emotion Recognition System (SERS) for e-learning Improvement Based on Learner Concentration Metric," Procedia Computer Science, vol. 85, p. 767-776, 2016.
[13] U. Ayvaz, H. Gürüler, and M. O. Devrim, "USE OF FACIAL EMOTION RECOGNITION IN E-LEARNING SYSTEMS," Information Technologies and Learning Tools, vol. 60, no 4, p. 95, sept. 2017.
[14] D. Yang, A. Alsadoon, P. W. C. Prasad, A. K. Singh, and A. Elchouemi, "An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environment," Procedia Computer Science, vol. 125, p. 2-10, 2018.
[15] A. Fathallah, L. Abdi, and A. Douik, "Facial Expression Recognition via Deep Learning," in 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA), Hammamet, 2017, p. 745-750.
[16] OpenCV: OpenCV Tutorials [Electronic resource] – Access mode: https://docs.opencv.org/master/d9/df8/tutorial_root.html
[17] S. Is¸ık-Ulusoy, S. Gulseren¨ , N. Ozkan¨ , and C. Bilen, "Facial emotion recognition deficits in patients with bipolar disorder and their healthy parents," *General Hospital Psychiatry*, vol. 65, 04 2020.
[18] S. Lenzoni, V. Bozzoni, F. Burgio, Beatrice, A. Wennberg, A. Botta, E. Pegoraro, and C. Semenza, "Recognition of emotions conveyed by facial expression and body postures in myotonic dystrophy (dm)," *Cortex*, vol. 127, pp. 58–66, Jun. 2020.
[19] S. Passardi, P. Peyk, M. Rufer, T. S. H. Wingenbach, and M. C. Pfaltz, "Facial mimicry, facial emotion recognition and alexithymia in posttraumatic stress disorder." *Behaviour research and therapy*, vol. 122, p. 103436, 2019.
[20] S. Jaiswal and G. Nandi, "Robust real-time emotion detection system using cnn architecture," *Neural Computing and Applications*, 10 2019. [7] A. Mollahosseini, D. Chan, and M. Mahoor, "Going deeper in facial expression recognition using deep neural networks," 11 2015.
[21] C. Pramerdorfer and M. Kampel, "Facial expression recognition using convolutional neural networks: State of the art," 12 2016.