



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** X **Month of publication:** October 2023

DOI: <https://doi.org/10.22214/ijraset.2023.56171>

www.ijraset.com

Call: ☎ 08813907089

E-mail ID: ijraset@gmail.com

A Real-Time Emotion Detection System from Facial Expression Using Convolutional Neural Network

Aparna P S¹, Devika K B², Gopika P S³, Salu K⁴, Panchami V U⁵

Dept. of Computer Science and Engineering, Government Engineering College Palakkad Sreekrishnapuram

APJ Abdul Kalam Technological University, Kerala, India

Abstract: *This paper presents an abstract-real-time facial emotion detection system utilizing convolutional neural networks (CNNs) for robust face recognition, addressing the demand for rapid and accurate emotion classification in dynamic contexts. The proposed CNN-based approach ensures real-time processing while maintaining high accuracy, detailing the architecture of a specialized CNN model tailored for real-time emotion detection. The study optimizes network layers and parameters for efficient facial feature extraction and analysis, substantiated through extensive experiments on diverse datasets, showcasing the system's instant emotion detection proficiency. Rigorous quantitative analysis demonstrates its superior performance compared to existing methods, underscoring its effectiveness. The system's potential to precisely detect real-time emotions holds applications in interactive interfaces, spanning from immersive virtual environments to highly responsive human-computer interactions.*

Keywords: *Convolutional Neural Network, Face Recognition, Deep Learning*

I. INTRODUCTION

Facial expression recognition technology assumes a pivotal role in the realm of emotion computing research. Within both human-computer interaction and emotion computing research, the emergence of real-time facial emotion recognition constitutes a dynamic and indispensable avenue for exploration. This technology's relevance extends across diverse sectors—such as healthcare, education, marketing, and entertainment—proffering transformative potential by promptly perceiving and responding to human emotions. Central to unraveling human emotions is the precise identification of facial expressions.

Traditionally, recognition methods have revolved around feature extraction, a pivotal component in classifier-driven emotion identification. However, historically, this process has necessitated specialized expertise, consuming substantial time and effort and yielding only moderate levels of recognition accuracy. In contemporary times, the advent of deep learning, a subset of machine learning, has engendered revolutionary shifts across various applications, including facial emotion recognition. A particularly noteworthy approach within deep learning is the Convolutional Neural Network (CNN), which has garnered widespread attention due to its direct processing and analysis of images, bypassing intricate preprocessing stages. By fabricating neural network architectures that mirror the hierarchical and localized characteristics of human brain processing, CNNs have showcased remarkable proficiencies across a spectrum of tasks, ranging from facial recognition to target tracking.

This paper undertakes a comprehensive exploration at the juncture of real-time facial emotion recognition and CNN algorithms for facial recognition. We delve extensively into the intricacies of this pioneering approach, accentuating its potential to augment precision, efficiency, and responsiveness. By infusing CNNs, our objective is to bridge the divide between real-time applications and the intricate subtleties of human facial expressions, thereby laying the groundwork for interactions between humans and computers that exude greater immersion, empathy, and contextual awareness. Through rigorous experimentation and meticulous evaluation, our aim is to showcase the efficacy of our approach within real-time scenarios and underscore its significance within the broader panorama of emotion-aware computing.

II. LITERATURE REVIEW

A. A Lightweight Convolutional Neural Network for Real-Time Facial Expression Detection [1]

The researchers propose and create a basic convolutional neural network to detect facial expressions. They design a visual system that can be integrated into low-powered devices to classify facial expressions, reducing parameter complexity. While the model shows certain outcomes, real-world facial expression data might have noise due to factors like extreme lighting, hazy images, obscured faces, and other challenges affecting accurate recognition. Additional efforts are needed to address these issues comprehensively.

B. Facial Expression Recognition - A Real Time Approach [2]

In light of recent advancements in computing technology, which unmistakably prioritize enhancing user experiences, the realm of automatic facial expression detection emerges as a crucial avenue of research. Despite the challenges posed by intrinsic variations in posture and illumination conditions in 2D photography, it remains a predominant approach in this field. In response to these challenges, the utilization of 3D and 4D (dynamic 3D) recordings in expression analysis studies is on the rise. This study delves into the latest developments in the acquisition and tracking of 3D facial data. It also presents the current usage of 3D/4D face databases for analyzing facial expressions, along with the facial expression recognition systems that extensively exploit either 3D or 4D data. However, various unresolved issues must be addressed to pave the way for the future application of 3D facial expression detection systems.

C. The First Facial Expression Recognition and Analysis Challenge [3]

For more than twenty years, computer scientists have been actively researching automatic facial expression recognition and analysis, particularly focusing on identifying FACS Action Units (AUs) and detecting discrete emotions. Several widely used facial expression databases now exist, indicating substantial progress in standardization and comparability. However, the challenge lies in comparing different systems due to the lack of a standardized evaluation process and insufficient information to replicate reported individual outcomes, impeding the field's advancement. To address this, regular challenges for facial expression recognition and analysis could offer a solution. The inaugural challenge took place at the IEEE conference on Face and Gesture Recognition in 2011, in Santa Barbara, California. This challenge comprised two sub-challenges centered on AU detection and discrete emotion detection, providing an evaluation protocol, utilized data, and baseline method results for these sub-challenges.

d. Static and Dynamic 3D facial expression recognition: A comprehensive Survey [4]

Conditional Random Fields (CRFs) serve for sequence segmentation and frame labelling Latent-Dynamic Conditional Random Fields (LDCRFs) improve CRFs by integrating hidden variables for modelling motion patterns and label dynamics. Our approach utilizes LDCRFs to enhance automatic facial expression recognition from video, outperforming CRFs. Using Principal Component Analysis (PCA), we examine expression separability in projected spaces. Comparative analysis highlights the role of temporal shape variations in expression classification, especially for emotions like anger and sadness. Empirical evidence underscores the importance of considering shape alongside appearance for accurate facial expression recognition.

D. Facial Expression Recognition with Temporal Modelling of Shapes [5]

Automatic facial expression recognition poses significant challenges encompassing face localization, feature extraction, and modelling. This project introduces a novel approach to facial expression recognition. The process commences with face identification, achieved by extracting head contour points through motion information. Subsequently, these extracted contour points are utilized to create a rectangular bounding box around the facial region. Notably, among various facial features, eyes play a pivotal role in determining facial dimensions. Hence, the method proceeds by eye localization and extracting visual features based on their positions. These extracted visual features undergo modeling using a support vector machine (SVM) for facial expression recognition. By leveraging the SVM, an optimal hyperplane is determined, enabling the discrimination of distinct facial expressions with a remarkable accuracy level of 98.5%.

III. PROPOSED METHOD

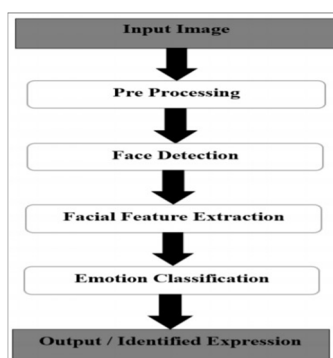


Figure 1: Methodology of the system

Pre-processing involves fundamental image operations, working with intensity images as both input and output. Common pre-processing steps encompass noise reduction, Binary/Grayscale conversion, pixel brightness adjustment, and geometric transformation. Face Registration, a computer technology with versatile applications, identifies human faces in digital images. Initially, faces are located through landmark points known as "face localization" or "face detection." These identified faces are then normalized geometrically to align with a template image, termed "face registration." Subsequently, Facial Features extraction pinpoints regions, landmarks, or contours in a 2D or 3D image. This step generates a numerical feature vector from the registered image, including common features like lips, eyes, eyebrows, and nose tips. Finally, in the classification phase, the algorithm aims to categorize faces into one of the seven basic emotions.

The system flowchart for the proposed system involves training and testing.

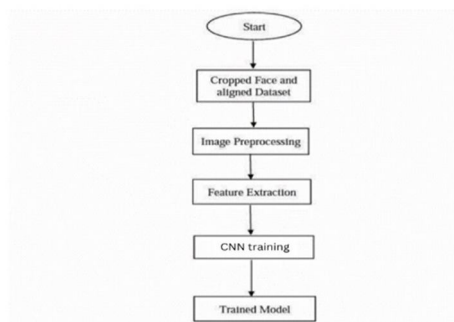


Figure2: Flowchart of Training

- 1) *Cropped Face and Aligned Dataset*: Images that have been trimmed to focus solely on the face, excluding surrounding background; Images in which facial features are adjusted to standardized positions, ensuring consistent facial orientations across images.
- 2) *Image pre-processing*: Image Preprocessing is a set of techniques applied to raw images before they are used for analysis or input into a machine learning model. It involves various operations to enhance image quality, reduce noise, and extract relevant features. Common preprocessing steps include resizing, normalization, noise reduction, and data augmentation. Preprocessing ensures that images are in a suitable format for accurate and effective analysis or training.
- 3) *Feature Extraction*: Feature extraction is the process of converting complex data, like images, into simplified numerical representations that capture important patterns, enabling efficient analysis and machine learning.
- 4) *CNN Training*: Training a Convolutional Neural Network (CNN) involves data preparation, architecture selection, model building, training with optimization, hyperparameter tuning, performance evaluation, adjustments, and utilizing the trained CNN for predictions.
- 5) *Trained CNN Model*: A machine learning model (CNN) that has learned from data and can predict or classify new data based on image patterns.

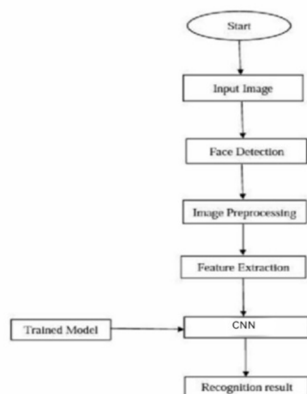


Figure 3: Flowchart of Testing

The testing flowchart for facial expression detection consists of essential steps illustrated in Figure 3:

- a) *Input Image*: Training images for machine learning, processed and labelled for pattern recognition.
- b) *Face Detection*: Identifying human faces using algorithms and bounding boxes.
- c) *Image pre-processing*: Enhancing raw images with resizing, normalization, and noise reduction.
- d) *Feature Extraction*: Capturing key information from raw data to simplify analysis and machine learning.
- e) *CNN*: Convolutional Neural Network utilized for classification and regression tasks.
- f) *Trained Model*: A model learned from data for making predictions on new data.
- g) *Recognition Result*: Outcome of identifying objects or patterns based on input data quality and model accuracy.

IV. RESULT AND DISCUSSION

The project commenced with the meticulous selection and preprocessing of a diverse dataset, containing labelled facial images portraying emotions such as happiness, sadness, anger, disgust, fear, surprise, and neutrality. After resizing, normalization, and augmentation, a specialized Convolutional Neural Network (CNN) architecture was meticulously crafted for precise facial emotion recognition. Comprising convolutional and pooling layers, this architecture effectively extracted intricate facial features, followed by fully connected layers for emotion classification. Training involved splitting the dataset into training and testing sets, using metrics like accuracy, precision, recall, and F1-score for monitoring. The trained CNN model seamlessly integrated into a dynamic pipeline for real-time emotion detection. Face detection algorithms accurately located faces in video frames, followed by resizing, normalization, and emotion prediction using the CNN model. The system achieved real-time emotion display, overlaying emotions like happiness, sadness, anger, disgust, fear, surprise, and neutral on the video stream. Rigorous evaluation through metrics and confusion matrices demonstrated the system's effectiveness, culminating in an 80% accuracy rate and successful recognition of multiple emotions in dynamic scenarios

A. Accuracy and Loss

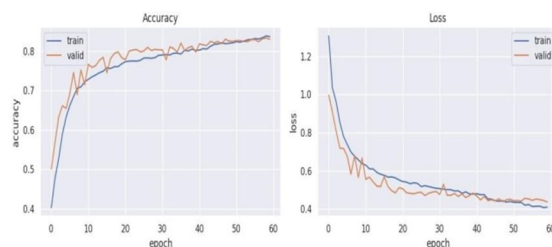


Figure 4: Accuracy and Loss

It can be observed from figure 4 that the loss decreases, and the accuracy increases with each epoch. The training versus testing curve for accuracy remains ideal over the first five epochs, after which it begins to deviate from the ideal values. The training and validation accuracy obtained for the FER2013 dataset using CNN are given in table 1.

epoch	Training accuracy	Validation accuracy
1	40.16	52.99
2	48.81	56.93
3	56.11	57.92
4	62.14	70.05
10	72.62	73.25
20	76.67	78.04
60	84.30	81.86

Table 1: Accuracy table

B. Confusion Matrix

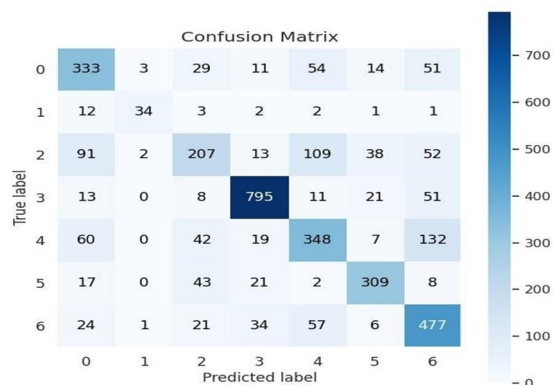


Figure 5: Confusion matrix

The confusion matrix generated over the test data is shown in following figure 5. The dark blocks along the diagonal show that the test data has been classified well. It can be observed that the number of correct classifications is low for disgust, followed by fear. The numbers on either side of the diagonal represent the number of wrongly classified images. As these numbers are lower compared to the numbers on the diagonal, it can be concluded that the algorithm has worked correctly and achieved state of the art results.

V. CONCLUSION AND FUTURE SCOPE

In this project an Emotion/Facial Recognition model has been trained and saved. It can recognize/detect the facial expressions of an individual on a real time basis that whether the individual is Neutral, Angry, Disgust, Fear, Happy, Sad, Surprised. The system has been evaluated using accuracy. The model achieved accuracy of 64.3%. This project also includes a user interface which will detect emotions in real time. Emotion recognition provides benefits to many institutions and aspects of life. It is useful and important for security and healthcare purposes. Also, it is crucial for easy and simple detection of human feelings at a specific moment without actually asking them. Experiments results on the dataset FER. Facial emotion recognition is a very challenging problem. More efforts should be made to improve the classification performance for important applications. Our future work will focus on improving the performance of the system and deriving more appropriate classifications which may be useful in many real-world applications.

REFERENCES

- [1] S. Jain, Changbo Hu and J. K. Aggarwal, "Facial expression recognition with temporal modelling of shapes," 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops) Barcelona, Spain, 2011, pp. 1642-1649, doi: 10.1109/ICCVW.2011.6130446.
- [2] Georgia Sandbach, Stefanos Zafeiriou, Maja Pantic, Lijun Yin, Static and dynamic 3D facial expression recognition: A comprehensive survey, Image and Vision Computing, Volume 30, Issue 10, 2012, Pages 683-697, ISSN 0262-8856.
- [3] A. Geetha, V. Ramalingam, S. Palanivel, B. Palaniappan, Facial expression recognition- A real time approach, Expert Systems with Applications, Volume 36, Issue 1, 2009, Pages 303-308, ISSN 0957-4174.
- [4] M. F. Valstar, B. Jiang, M. Mehu, M. Pantic and K. Scherer, "The first facial expression recognition and analysis challenge," 2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG), Santa Barbara, CA, USA, 2011, pp. 921-926, doi: 10.1109/FG.2011.5771374.
- [5] Q. Chen, X. Jing, F. Zhang and J. Mu, "Facial Expression Recognition Based on A Lightweight CNN Model," 2022 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Bilbao, Spain, 2022, pp. 1-5, doi:10.1109/BMSB55706.2022.9828739.
- [6] N. Zhou, R. Liang and W. Shi, "A Lightweight Convolutional Neural Network for Real-Time Facial Expression Detection," in IEEE Access, vol. 9, pp. 5573-5584, 2021, doi:10.1109/ACCESS.2020.3046715.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)