Avalokiteshvara Journal of Artificial Intelligence

http://hcapit.org/ajai.html

ISSN - XXXXXXXXXX

Research Article

Enhanced Facial Expression Recognition: A Convolutional Neural Network Approach

Sumit Sanjay Maske[1]

Independent Researcher, Bloomington, Indiana, United States.

sumitmaske8612@gmail.com

Abstract

Within non-verbal communication, the identification of facial expressions stands as a significant and formidable task. The objective of a facial expression recognition system is to categorize real-time facial images into distinct emotional classes, encompassing Anger, Disgust, Fear, Happy, Sad, Surprise, and Neutral. Drawing inspiration from the advancements achieved in image recognition and classification through Convolutional Neural Networks (CNNs), this paper advocates for a CNN-based methodology to tackle the complexities of facial expression recognition. The model presented in this study harnesses various libraries, including OpenCV, Keras, and TensorFlow. Through the utilization of grayscale images from the Face Expression Recognition dataset on Kaggle, the model undergoes training with CNN architectures of varying depths. Notably, the proposed model attains an accuracy of 72.34% for familiar data and 60.54% for previously unseen data.

Keywords: Intrusion Detection System, Deep Learning, Convolutional Neural Network, Re-current Neural Network, Generative Adversarial Network, Long Short-Term Memory

1. Introduction

Facial Expression Recognition serves as a vital system for analysing emotional behaviour through diverse sources, including images and videos. The facial expression of an individual represents a visible manifestation of various aspects, such as personality, internal emotional state, affective state, cognitive activity, intention, and psychopathology. It plays a pivotal role in interpersonal communicative relations, being a form of non-verbal communication that provides insights into an individual's emotional state.

Human expressions, conveyed through intricate facial muscle movements, offer glimpses into internal feelings, making them a subject of interest for researchers in fields such as Human-Computer Interaction (HCI), linguistics, psychology, animation, neurology, medicine, and security. Autonomous facial expression analysis systems contribute to human-computer interaction, presenting a challenging yet essential process. Deep learning and Convolutional Neural Networks (CNNs) prove invaluable for retrieving and analysing facial expression traits for sentimental analysis [1].

The primary objective of this system is to develop a model for facial expression recognition, employing a specific artificial neural network, namely the convolutional neural network. Through the application of a typical CNN with data augmentation, this study classifies facial expressions into emotions like Disgust, Fear, Anger, Surprise, Happy, Sad, and Neutral. The abundance of filters in CNNs makes them superior for image identification tasks. Despite significant progress in recent decades, achieving high accuracy in recognizing individual facial expressions remains challenging due to the complexity and variations in expressions[2].

Nonetheless, recognizing emotions through non-verbal cues, such as facial expressions and gestures, remains a crucial aspect of human communication[3]. This system aims to provide a non-verbal means for individuals to communicate emotions effectively. In daily life, people intuitively identify emotions by observing characteristic features displayed in facial expressions. For instance, happiness is often associated with a smile, while sadness is reflected in a gloomy facial expression. The model presented here categorizes facial expressions into basic human emotions, including anger, disgust, fear, happiness, sadness, surprise, and neutral. However, challenges arise from variations in pose, alignment, illumination, and occlusions, complicating the task of recognizing facial expressions. Surveys on facial feature representations for face recognition and expression analysis address these challenges and offer detailed insights into potential solutions.

2. Methodology

Our proposed methodology centers on employing Convolutional Neural Networks (CNN) for the task of facial expression recognition. The system receives input in the form of an image captured via a webcam. Subsequently, the CNN model is utilized to predict the facial expression label, which falls into one of the following categories: Angry, Happy, Fear, Sad, Disgust, or Neutral. The process of recognizing facial expressions involves three key phases: face detection, feature extraction, and facial expression recognition, as illustrated in Figure 1.

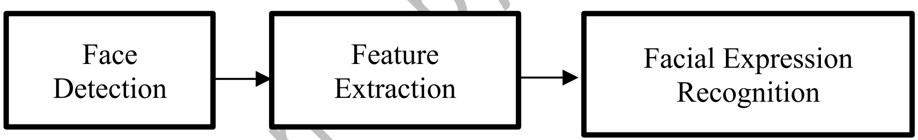


Figure 1: Phases of Facial Expression Recognition System

In the face detection phase, the system identifies the presence of a face or faces within the input image. Following successful face detection, features crucial for facial expression recognition are extracted. These extracted features play a pivotal role in accurately identifying the facial expression depicted in the image. Ultimately, the facial expressions are categorized into distinct labels, including Angry, Happy, Fear, Sad, Disgust, or Neutral. This multi-phase approach ensures a comprehensive analysis of facial expressions, contributing to the robustness and accuracy of our proposed facial expression recognition system.

2.1 Model Training and Architecture

For model training, we utilized the Face Expression Recognition dataset from Kaggle [5]. Comprehensive details of the training dataset are presented in Table 1, while Table 2 provides specifics about the validation dataset. The combined total of images for training and validation amounted to 35,887, with 28,821 examples in the training set and 7,066 in the public test set. The model's batch size for training and validation was set to 128. The classification task involved seven emotion classes, each representing a distinct human emotion. The model architecture incorporated four CNN layers and two fully connected layers. The initial CNN layer featured 64 filters with a kernel size of (3, 3), while the subsequent layers followed suit, with the second CNN layer containing 128 filters and a (5, 5) kernel size. Figure 2 illustrates the CNN architecture designed for the Facial Expression Recognition System.

Table 1: Details of Training dataset

Type of Emotion	No. of Images for Training
Angry	3993
Disgust	436
Fear	4103
Happy	7164
Neutral	4982
Sad	4938
Surprise	3205

Table 2: Details of Validation dataset

Type of Emotion	No. of Images for Validation
Angry	960
Disgust	111
Fear	1018
Happy	1825
Neutral	1216
Sad	1139
Surprise	797

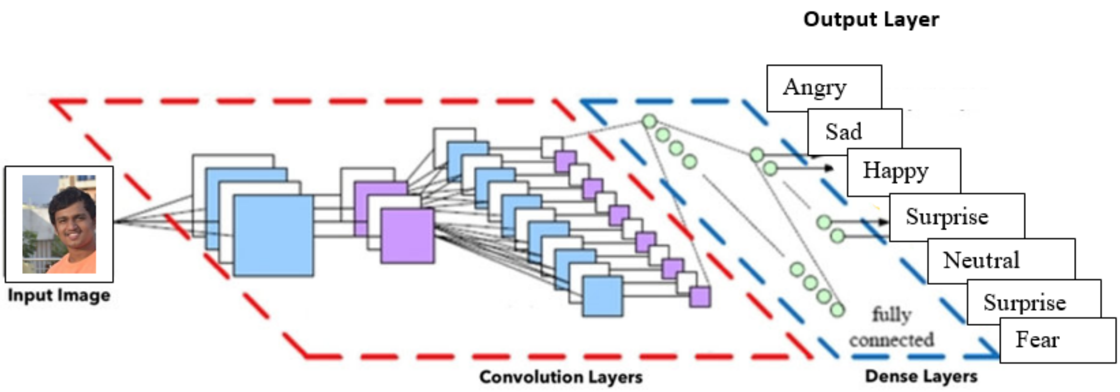


Figure 2: CNN for Facial Expression Recognition System

2.2. Real-time Facial Expression Recognition

In the real-time scenario, input images are captured through a webcam and processed by the trained model to classify facial expressions into one of the seven emotion categories. OpenCV is employed to capture video images through the webcam, and the Haar cascade is utilized for face detection within the image. Subsequently, the image is converted into an array, transformed into grayscale, and resized to a 48*48 megapixel resolution, aligning with the dimensions the model was trained on. The region of interest, representing the face in the image, is identified and converted into an array. Employing a classifier, the image is then categorized into one of the seven emotions discussed in the study, namely Disgust, Fear, Anger, Surprise, Happy, Sad, and Neutral, in real-time. The system's output is visually represented in Figure 3, Figure 4, and Figure 5.

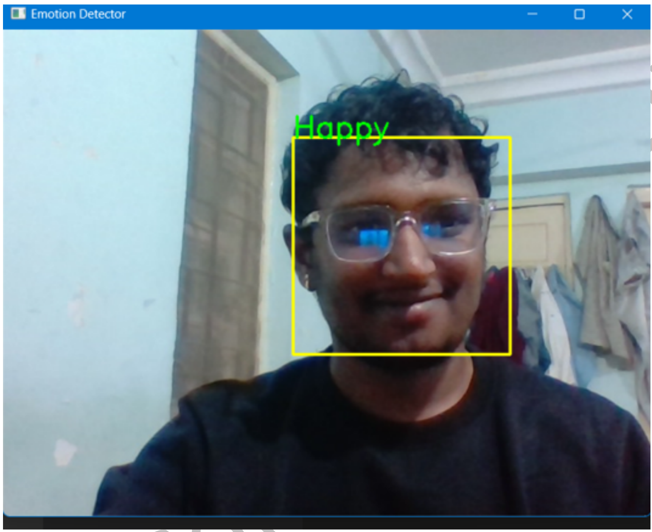


Figure 3: Sample output- 1

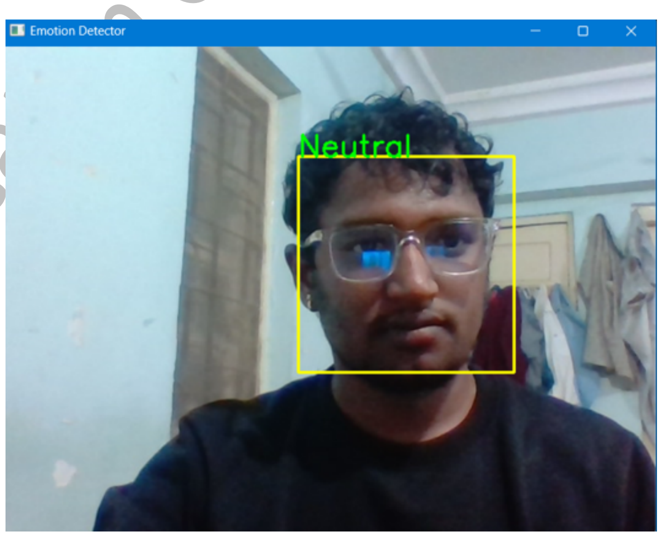


Figure 4: Sample output- 2

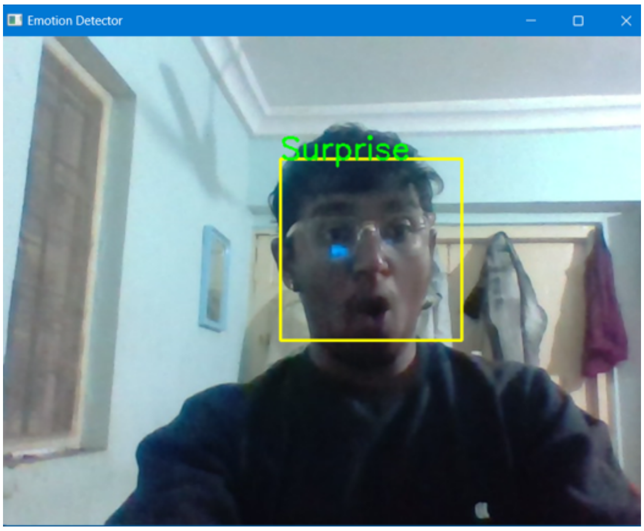


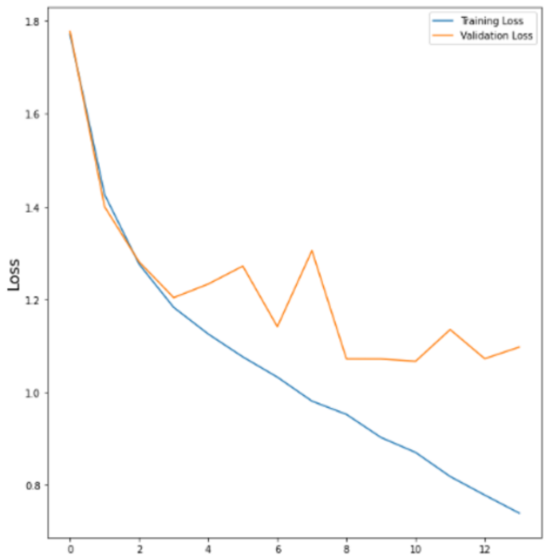
Figure 5: Sample output- 3

3. Results and Discussion

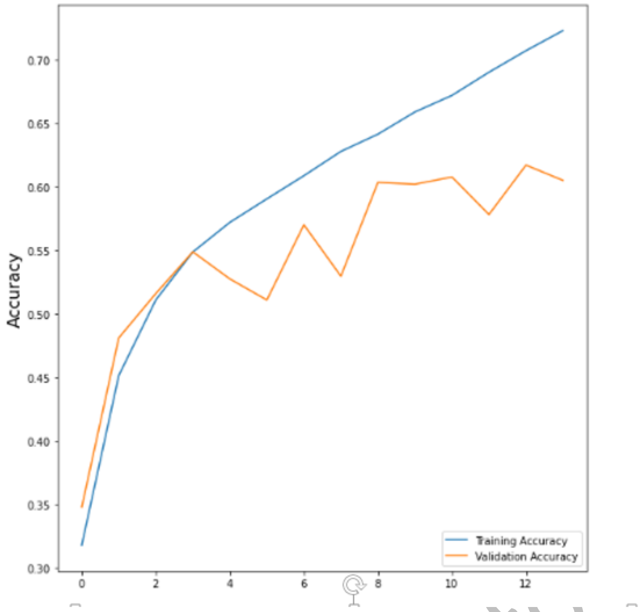
The proposed facial expression recognition system demonstrates commendable accuracy, achieving 72.34% for the training dataset and 60.54% for the validation dataset. To gain deeper insights into the model's performance dynamics, we visualized the training process by plotting loss and accuracy using the Matplotlib library.

Upon examination of the training and validation trends, it is evident that both Training Loss and Validation Loss exhibit a consistent decrease over time. This suggests that the model effectively learns and generalizes from the training data, minimizing its overall loss.

Simultaneously, the accuracy metrics present an encouraging upward trajectory. Both Training Accuracy and Validation Accuracy steadily increase, indicating the model's ability to correctly classify facial expressions. These trends are visually depicted in Graph 1, illustrating the decline in loss, and Graph 2, showcasing the ascending accuracy rates.



Graph 1: Training Loss and Validation Loss



Graph 2: Training Accuracy and Validation Accuracy

The convergence of decreasing loss and increasing accuracy underscores the efficacy of the proposed system in capturing and recognizing facial expressions. These results not only validate the robustness of the model but also pave the way for further discussions on potential enhancements and real-world applications.

4. Conclusions

In this paper, we have explored the intricacies of facial expression recognition, focusing on the classification of facial images into seven distinct categories representing fundamental human emotions. The initial validation of our model yielded an accuracy of approximately 67.27%, with a corresponding validation accuracy of around 53.81%. Subsequent iterations and adjustments were made to refine the model, resulting in a notable improvement. The refined model achieved an accuracy of 72.34%, with a validation accuracy reaching 60.54%. These advancements underscore the effectiveness of the proposed facial expression recognition system, demonstrating its adaptability and potential for accurately identifying and categorizing diverse facial expressions. The ongoing pursuit of refining and enhancing such systems holds promise for applications in various fields, from human-computer interaction to emotion-aware technologies.

References

[1] selection and ensemble classifier. Comput Netw. 2020;174:107247.
[2] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-444.
[3] Roy S, Cheung H, Ding Z. A hybrid deep learning approach for intrusion detection. IEEE Access. 2020;8:136431-136441.
[4] Siva Shankar S, Bui Thanh Hung, Prasun Chakrabarti, Tulika Chakrabarti, Gayatri Parasa. A novel optimization-based deep learning with artificial intelligence approach to detect intrusion attacks in network systems. Educ Inf Technol. 2024;29(4):3859-3883.
[5] Tang T, Chen J, Luo H. Insider threat detection based on deep learning. IEEE Access. 2018;6:11074-11083.
[6] Hettich S, Bay SD. The UCI KDD Archive. Univ Calif Dept Inf Comput Sci. 1999.
[7] Tavallaee M, Bagheri E, Lu W, Ghorbani AA. A Detailed Analysis of the KDD CUP 99 Data Set. Proc Second IEEE Symp Comput Intell Secur Def Appl (CISDA). 2009.

[8] Moustafa N, Slay J. UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems (UNSW-NB15 Network Data Set). Mil Commun Inf Syst Conf (MilCIS). 2015.

[9] Sharafaldin I, Lashkari AH, Ghorbani AA. Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. ICISSP. 2018.

[10] Lashkari AH, Draper Gil G, Mamun MSI, Ghorbani AA. Characterization of Tor Traffic Using Time Based Features. Proc 3rd Int Conf Inf Syst Secur Privacy (ICISSP). 2018.

[11] MAWI Working Group Traffic Archive. Samplepoint-G. 2006.

Author Details

***Sumit Sanjay Maske**, M.Tech (Pursuing)*
Pursuing Master's Degree in Data Science at Indiana University Bloomington, Indiana, United States.

***Research Interest:** Machine Learning, Artificial Intelligence, Data Science*

***Brief Bio:** Sumit S. Maske is a student at Indiana University Bloomington, pursuing his masters degree in Data Science. A representative for Data Science department at Graduate and Professional Student Government. His research focuses on machine learning and artificial intelligence.*