International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 06 | June - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Facial Expression Analysis Using Haar-Cascade Face Detection and Convolutional Neural Networks

Sindhushree P V
student, Dept of CSE,
Sea College of Engineering & Technology

Sudha kumari
student, Dept of CSE,
Sea College of Engineering & Technology

Yashaswini G S
student, Dept of CSE,
Sea College of Engineering & Technology

Mrs. Sushma B A
Assistant Professor
Dept of CSE
SEA College of Engineering & Technology

Dr Krishna Kumar P R
Professor Dept of CSE
SEA College of Engineering & Technology

Mrs Jayashri M
Assistant Professor
Dept of CSE
SEA College of Engineering & Technology

Abstract

Facial expression analysis is a crucial aspect of human-computer interaction, enabling machines to recognize and respond to human emotions effectively. This project presents a system that combines Haar-Cascade face detection with Convolutional Neural Networks (CNNs) to identify and classify facial expressions in real-time. The Haar-Cascade classifier is utilized for efficient and accurate face localization in static images or video frames. Once the face is detected, it is processed and fed into a trained CNN model that classifies the expression into predefined categories such as happy, sad, angry, surprised, neutral, etc. The CNN architecture is designed to extract hierarchical features from facial images, capturing subtle variations in muscle movement and facial landmarks. The system is trained and evaluated on benchmark facial expression datasets, demonstrating high accuracy and robustness under various lighting and background conditions. This integrated approach offers a reliable and real-time solution for applications in emotion-aware systems, surveillance, healthcare, and human behavior analysis.

Introduction

Computer vision is an interdisciplinary scientific field that deals with how computers can gain a high- level understanding of digital images or video processing [4]. Computer vision can solve tasks using methods for acquiring, processing, analyzing, and extracting high-dimensional data from images, video, text, and others. Sub-domains of computer vision include scene reconstruction, image compression, image restoration, vehicle detection, and facial emotion recognition. Facial emotion recognition is the central identifier for human expressions. Emotions can be expressed in many varieties such as facial expressions, voices, text, and others.

Currently, this research approaches primarily focuses on the classifier method using *Haar- Cascade* and the deep learning method using Convolution Neural Network. Facial detection using *Haar* feature-based *Cascade* classifier is an effective object detection method [5]. The deep neural network has the same type as CNN in high network depth and algorithm process. The basic concept of CNN almost the same as with Multilayer Perceptron, however, each neuron in CNN will be implemented using a two-dimensional shape. CNN can only be used for image and voice recorder data which has a two-dimensional structure. CNN has many layers of learning which are used to read an input a feature based on the characteristic of the input. The layers are in the form of number vectors. This feature extraction layer consists of a convolution layer and a pooling layer. In the convolution layer, the output of neurons will be calculated, each of which calculates their weight and until the images with small size are connected into input volume.

The manuscript will focus on seven facial expressions based on the training dataset, which are angry, disgust, fear, happy, sad, surprises, neutral. The proposed model is trained using the FER2013 dataset. In the next section will be

presenting about Related Work or Literature Review, Methodology, *Haar-Cascade Classifier,* and CNN. The last section will discuss the results and conclusions.

1. Literature Review

In this section, we will describe and present the latest research about CNN in digital image processing. All humans have facial expression to show their emotional mood. The facial expression has many tools for use in many fields such as robotics, medicine, unmanned aerial vehicle, and lie detectors. Based on knowledge facial expression as follows angry, fear, happy, sad, disgust, neutral, and surprise. Recently, Ninad and their group research present facial emotion recognition using *Convolutional Neural Networks,* and they divide it into two parts of the concept [1]. The first part to removing the background from the images and the second part is focused on the part of vector extraction from the images. They used the supervisory data from the stored database of 10,000 imagesnd 154 faces. The results show if the accuracy is about 96%, using an expressional model with 24 values. In their paper present if used more than 750k dataset of images there are Cohn-Kanade, Caltech, CMU, and NIST face data storage [1].
The other research by Liliana was a present facial expression to detecting the occurrence of facial Action Units (AUs) as a subpart of FACS which presents human emotion [2]. They used a regularization method called "dropout" to prove if CNN was very effective in reducing overfitting. Dataset Cohn Kanade (CK+) was used in their experiment. Then, the results gain an average accuracy rate of 92,81%. The dataset images were classified with eight emotion classes. Furthermore, we will present a recent paper by Zeynab and their group research. The paper proposes CNN method is trained on Cohn-Kanade and RAVDESS dataset and classified the facial expression with five major facial emotions. They mention that the uniqueness of their model it's not trained and performed well for both datasets with few neural layers [3]. The model has outperformed and convolutional layers with the addition of pooling and dropout layers. The results of their research proposed satisfactory in detecting macro facial emotions on the dataset.

2. Research Methodology

2.1. Data Acquisition

We use the basic emotion in facial emotion recognition and classified into 7 class of expression, namely angry, disgust, fear, happy, sad, surprise, as well as neutral. Furthermore, we use CNN to recognize different expressions or emotions. In fig 1 determine much facial emotion based on dataset FER2013.

Figure 1. Variety Facial Emotion Recognition

2.2. Data Storage

The dataset used in this research is *Facial Emotion Recognition 2013 (FER2013).* The training set consists of 28,709 examples, and the testing set consists of 7179 examples. The data consist of 48 x 48 pixel grayscale images of the face. The faces have been automatically registered so that the face is less or less centered and occupies about the same amount of space in each image. On the other hand, the test data set consists of 70 examples.

2.3. CNN for facial expression detecting

Fig 2 describes the architecture of CNN. CNN has 6 convolutional layers, 2 subsampling layers, 12 convolution layers, and 2 subsampling Neural Network.

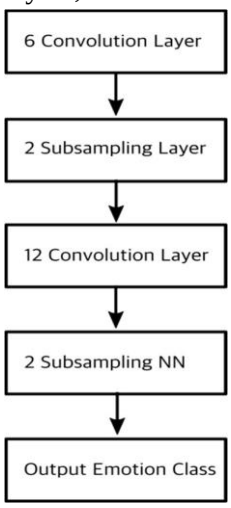

Figure 2. The Diagram of CNN for Facial Expression Detection

2.4. Haar Cascade Classifier

Object detection using Haar feature-based cascade classifier is an object detection method by Paula Viola and Michael Jones. In 2001, they proposed a paper name "Rapid Object Detection using a Boosting Cascade of Simple Features". Haar cascade is a set of Haar Like Features that are combined to form a classifier. The feature is the pixel value in the write subtracted from the pixels value in the blank area. The base of the face detector is 24 x 24. From that base face detector, there around 160k possible Haar-Like Feature. However, not all of these features are used.

3. Experimental Results

We conducted experiments using a training dataset and testing data set in real-time by a video camera. Fer2013 dataset has 28,709 images and we used it all in our experiments. As we can see in table 1, we use varied epochs and get a different result in MSE and Accuracy. From the experiments we can that there a significant decreasing means square error as the epoch of training data raises. Several testing data we are proposed each expression has 10 times to test in real-time.

Table 1. MSE and Model Accuracy of CNN

Epoch	Num of Training Data	Num of Testing Data	MSE	Model Accuracy
30	28,709	70	0,8652	67%
50	28,709	70	0,6754	75%
75	28,709	70	0,5214	81%
100	28,709	70	0,4192	85%
150	28,709	70	0,3356	89%
200	28,709	70	0,2912	92%

Table 1 presents the result of the experiment, we use a variety of epoch to training and testing the data. The results present that the higher the epoch value so the MSE value is getting lower value. Likewise, model accuracy was present more accurately when the epoch higher. These works are showed the CNN method is great for testing and training images.

Table 2. Facial Emotion Detection Classification

Epoch	Angry	Happy	Disgust	Sad	Neutral	Surprise	Fear
30	60%	70%	50%	60%	60%	60%	60%
50	70%	70%	70%	70%	70%	70%	70%
75	70%	80%	70%	70%	70%	70%	70%
100	80%	90%	80%	70%	70%	80%	70%
150	80%	100%	80%	80%	80%	80%	80%
200	90%	100%	80%	80%	90%	90%	80%

By measuring the facial emotion in seven classes, the accuracy rates for each epoch is different. The problem when we test using video were the facial expression model is still cannot distinguish between fear and sad expression, happy and surprise expression.

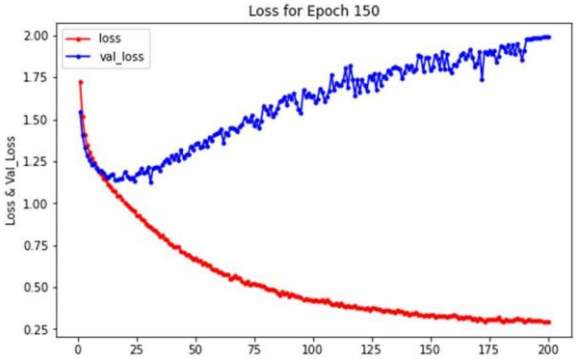

Figure 3. Mean Square Error or Loss when running the program

Figure 3 was described Mean Square Error versus Validation Loss, the more the loss or MSE gets the lower pattern, so the validation loss will be higher graphic.

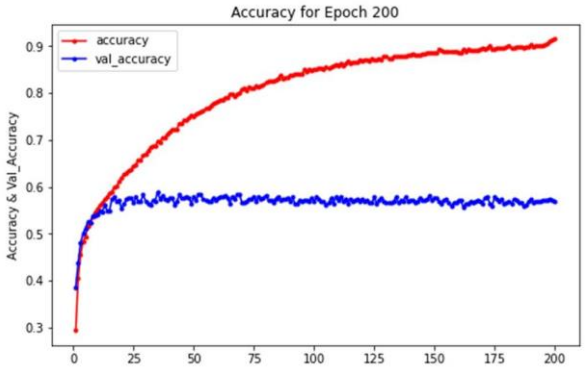

Figure 4. Model Accuracy when running the program

The scenario of the experiment gives us prominent results about the model performance of mean square error value show decreasing process from minimum epoch into maximum epoch in our works. In a different graph in fig 4, the accuracy of model facial emotion detection was proposed increasing value from the beginning of the process, it means the smaller the MSE value, so we can get the greater accuracy value of the model.

4. Conclusion

In this project, we successfully developed a facial expression analysis system that integrates Haar-Cascade face detection with a Convolutional Neural Network (CNN) for accurate emotion classification. The Haar-Cascade algorithm proved to be an effective method for real-time face detection, while the CNN demonstrated strong capability in learning and distinguishing between different facial expressions. The combined approach ensures a fast, reliable, and scalable solution for expression recognition tasks. Experimental results on benchmark datasets validate the system's performance, showing promising accuracy and generalization across varied facial images. This work lays the foundation for future improvements such as incorporating more advanced deep learning models, handling occlusions, and extending the system for multi-face analysis in dynamic environments. The developed system holds significant potential for real-world applications in human-computer interaction, mental health monitoring, security systems, and social robotics

5. References

[1] Mendel, N. 2020. *Facial Emotion Recognition Using Convolutional Neural Networks (FERC).* Springer Nature Switzerland AG

[2] Liliyana DY. 2019. *Emotion Recognition from Facial Expression Using Deep Convolutional Neural Network.* (IOP Conf. Series: Journal of Physics: Conf. Series 1193)

[3] Rzayeva, Z, & Emin Alasgarov. 2020. *Facial Emotion Recognition using Convolutional Neural Networks.* IEEE Xplore (Auckland University of Technology)

[4] Wiley, V, & Thomas Lucas. 2018. *Computer Vision and Image Processing: A Paper Review.* International Journal of Artificial Intelligence Research

[5] Pasadena, A, & Muhammad Habibi. 2019. *Face Detection using Haar Cascades to Filter Selfie Face Image on Instagram.* Conference Paper