Acta Scientific COMPUTER SCIENCES

Volume 7 Issue 3 June 2025

Research Article

Convolution Neural Network Approaches for Facial Emotion Recognition

Shivamma D[1], Nitheesh Kumar G[2]*, Deepika P[2], Dr. Shreedhara K S[2]

[1]*Assistant Professor, Dept. of CSE(DS), School of Engineering, DSU, India,*

[2]*PG Student, Dept. of CSE UBDT College of Engineering, VTU, Davanagere, India,*

***Corresponding Author:** Nitheesh Kumar G, PG Student, Dept. of CSE UBDT College of Engineering, VTU, Davanagere, India.

Received: May 21, 2025

Published: June 09, 2025

Abstract

A face detection model based on the mapping of behaviors with physical aspects is offered by the facial emotion recognition system described in this study. Geometric structures that have been reconstructed serve as the foundation matching template for the identification system and are linked to the physical characteristics of the human face that correspond to different expressions including happy, sad, fear, angry, surprise, and disgust. These days, face expression recognition is popular due to its broad range of applications. Because of its many uses, emotion recognition is widely used. As the science of learning has advanced, emotion detection has become increasingly important in business. Feeling recognition allows one to perceive a person's feelings. To detect emotional states in pictures, several methods have been developed.

Keywords: CNN (Convolutional Neural Network); LBP (Local Binary Patterns); Emotion Detection; Facial Expression; SVM (Support Vector Machine); KNN (K-Nearest Neighbor

Introduction

Outward appearances are an indication of the passionate mood. Projecting a clear image of an individual's excitement onto the external environment is the goal. Since emotions provide a unique perspective, this viewpoint is closely tied to an individual's interior state. Observers may assess a person's spiritual state in order to validate or respond to their outward appearance [2]. In psychological health care offices, medical staff members search for this to help doctors create therapy applications which are readily available to patients. Therefore, it is essential to comprehend feelings and the causes behind them, as well as the range of emotions that are present in decision- making, healthcare settings, sentiment, educational institutions, and counselling offices. One of the trickiest parts of computer vision difficulties is separating exterior appearances, which is why facial pattern recognition algorithms are used. Since it is hard to design a precise model, the representation should identify the characteristics such as happiness, anger, sadness, calmness, surprise, and so forth that set apart a particular feeling from resentment. Image processing is a type of signal processing in that the input and output signals are images. The most essential uses of image processing is facial expression recognition. Our emotions are revealed through our facial expressions [1]. Facial expressions are vital in interpersonal communication. Facial expression is a nonverbal scientific gesture that we express with our faces based on our emotions. Automatic face expression detection is vital in artificial intelligence and robotics, making it a generational need [3]. This area includes applications such as automated surveillance, forensics, human-computer interaction, videophone and teleconferencing, personal identification and access control, and cosmetology.

Related work

There is no predetermined framework for identifying people's emotional states. We use a variety of biological methods, such checking the body's temperature, pulse, and other vital signs, to determine the state of thought and if it is acceptable. The estimate is essentially based on the many limits of an activity that are a component of the human propensity to evaluate other people's feelings. A few methods for automated recognition of external appearances are shown in article. Angle essential dynamic presence model, is used by remainder of system to identify the 68 face focuses on the basic edge that certain modern frameworks physically name. Pre-

cise feature acquisition might not be feasible in many functional scenarios. Shawn suggested looking at the application of local binary patterns (LBPs) the recognition of external appearance [7]. The articulation check is carried out by articulate support vector machine (SVM) classifiers that support LBP highlights [4]. It is absurd in most functional settings that the writers made it universally beneficial by physically noting the eye locations. The Computer Expression Recognition Toolbox (CERT) proposed by Larry, to detect face highlights, CERT forbids the insertion of Gabor unique letters into facial images. Instead, it employs SVM and multidimensional strategic relapse to identify external appearances. Lyons proposed a method for highlight extraction and class extension from exterior appearances using 2D Gabor wavelets and bunching. These methods specifically call for programming and computational work. For eliminating highlights from photos, Deep Convolution Neural Network (DCNN) Framework is frequently utilized [15]. DCNN uses multiple layers to achieve accurate element learning. The highlights shown here serve as channels, and these channels, as stated by bright and dim across the info picture, provide the highlights used by different layers of the organization. Techniques like Kazoo's models, which are based on convolution neural networks, has been offered as identifying external appearances. Still, they mostly use other facial datasets for model training. For example, although this technique is usually required to create very small picture datasets, Google employed Deep Conventional Activation highlight to extract face highlights for generalized optical recognition, which does not require considerable preparation.

Methodology

Machine learning (ML) and other applications of artificial intelligence (AI) rely on algorithms to acquire or analyze data, draw conclusions from it, and then use those findings to direct behavior. The ability of a machine to predict future occurrences using historical data or behavior without explicit programming is known as machine learning. In order to create Artificial Neural Networks (ANN) that can learn on their own and draw logical conclusions, deep learning (DL) employs structural layers. Each layer of an ANN has a large number of neurons or perceptrons. Use deep learning to examine massive volumes of data to provide accurate discoveries with high output performance. Applications that describe pictures employ deep learning. Image description is the process of explaining the information found in a photograph.

Convolutional neural networks

Convolutional neural networks (CNN) categorize images by analyzing them for purposes such as object detection, machine vision, identification and categorization [10]. CNN uses an algorithm to classify images then organizes them into top-to-bottom subcategories to extract important information.

Convolutional neural networks (CNN) are capable of processing input forms that roughly approximate two-dimensional matrices. CNN provides resources for handling images and a 2D matrix is a straightforward way to represent images [11]. CNN is mainly employed for image classification, i.e., identifying if a picture shows superman, a jet, or a bird [5]. Above figure 1 shows working of convolution.

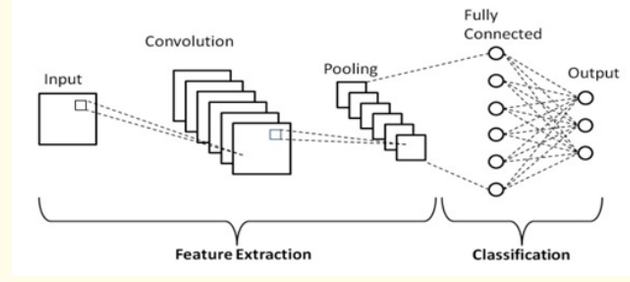

Figure 1: Working of Convolution.

Convolution layers

Convolution layers are the initial layers that are utilized to derive characteristics of images. Convolution Layers picks up feature from a small input data series and maintain the link between pixels [12]. In mathematics, a convolution layer is made up of three parts: two inputs, a photo matrix, and a kernel or filter [8]. The steps listed below are the steps in the matrix multiplication process. The output matrix of the [13]. Convolution Layers for a 5X5 image multiplied by a 3X3 filter is displayed below the feature that results referred as convolved feature or feature map. In figure 2 it explain about Convolution operation.

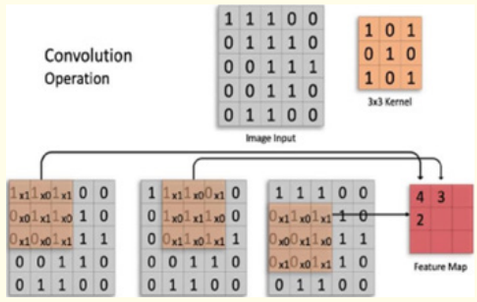

Figure 2: Convolution Operation.

Strides

The transfer of information from a single neuron to another is called stride. During the array creation process, images shift from the matrix that is output to the input matrix, and the filter itself shifts one image at a time when the total amount of Strides is one. When the total amount of Strides hits 2, the filters are also changed to 2 pixels, and so forth. Convolved feature explained in figure 3.

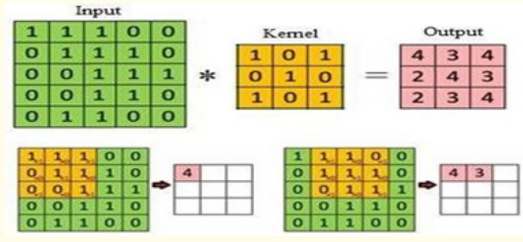

Figure 3: Convolved Feature.

Padding

Convolutional neural network construction requires padding. After the procedure of convolution, the image has shrunk in size. Following each convolution layer stage in an image categorization challenge, the initial image is reduced. Overlapping occurs when the kernel travels through the middle layer more times than the outer layers as it proceeds over the original image. The concept of padding was developed as a solution to the overlapping issue. An additional layer known as padding can expand an image's boundaries without reducing or preserving the dimensions of the original image. Padding represented in figure 4.

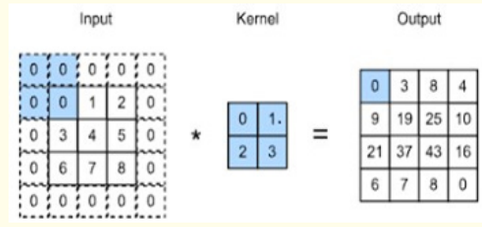

Figure 4: Padding.

Maximum pooling

Maximum Pooling, commonly referred as max pooling, is a rule which prioritizes image elements by using a certain location [9]. Max pooling is a sample-based technique that converts continuous functions to discrete ones [6]. Its main goal is take an input and make assumptions about rejected sub- region data. Max pooling explained in figure 5.

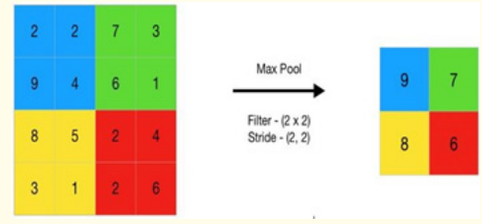

Figure 5: Max pooling.

Average pooling

Max pooling and average pooling are not the same. By dividing the input matrix into rectangular sections and calculating the average figures for each, average pooling reduces the matrix. This preserves data regarding less significant elements. Average pooling represented in figure 6.

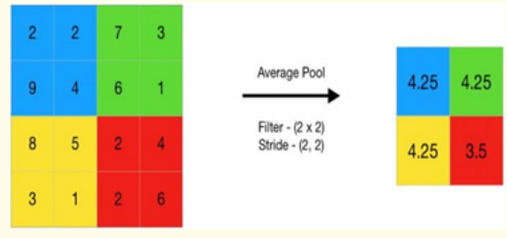

Figure 6: Average pooling.

Neural network

The Neural Network's topology is depicted in figure. It consists of interconnected neurons and each neuron has three characteristics: (i) Weight (ii) Bias and (iii) Activation Function. Below figure 7 shows the structure of neural network.

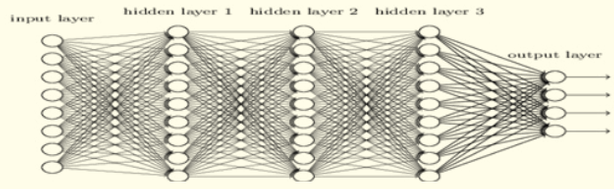

Figure 7: Structure of Neural Network.

After being fed into the input layer, the input is transformed linearly by the neurons using the weights and biases in equation. X = (weight x input) + bias.

Applying an Activation Function to the aforementioned result, which is represented in equation, precedes the linear transformation. Y = Activation (∑ (weight x input) + bias).

Activation function

ReLU (Rectified Linear Unit)

One common irregular activation function used in deep learning is the ReLU. Since ReLU is unable to activate all neurons simultaneously, neurons will only become silent if the outcome of the linear transformation is less than zero. ReLU has this one major benefit above other activation functions.

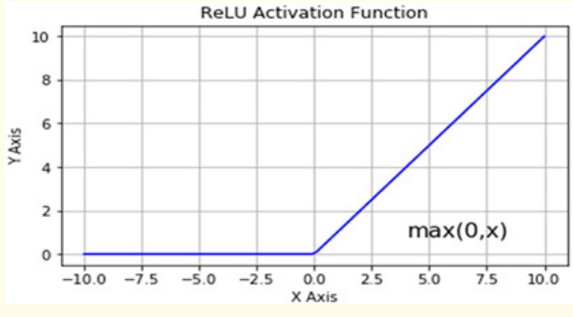

Figure 8: Graphical Representation of ReLU Function.

Dataset

Project requires knowledge of Deep Learning, Python, TensorFlow, Keras, NumPy, OpenCV, and. Install packages TensorFlow, Keras, Pillow, NumPy, and Tqdm to ensure appropriate operation. This experiment used the FER 2013 dataset, which was obtained via Kaggle [14]. Categorize each expression of face into the following seven categories: categories: angry, sadness, happiness, surprised, disgusted, fearful or neutral training set has 28,709 examples, whereas the public test set includes 3,589 examples.

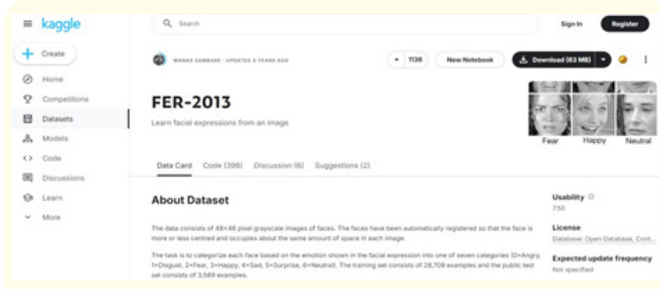

Figure 9: Fer-2013 Dataset from Kaggle.

Comparative study

Sl. no	Year	Classifier	Emotions
1	[1] 2023	CNN	Happy, Angry, Scared
2	[2] 2021	CNN	Happy, Sad, Surprise
3	[3] 2019	LBP	Happy, Angry, Sad
4	[4] 2022	SVM, KNN	Happy, Scared, Neutral
5	[5] 2022	SVM, KNN	Happy, Sad, Surprise
6	[6] 2021	CNN, KNN	Happy, Neutral, Sad,
7	[7] 2021	LBP, CNN	Happy, Sad, Angry
8	[8] 2020	CNN	Happy, Sad, Angry
9	2024 This paper	CNN	Happy, Sad, Neutral, Angry, Surprise, Fear

Table 1: Comparative analysis between the papers**.**

Results

Our Convolutional Neural Network model was trained using the FER 2013 database, which includes seven emotions: happy, angry, sad, disgust, neutral, fear, and surprise. The recognized face pictures were input into the CNN model after being resized to 48x48 pixels and converted to grayscale. The predicted emotion label

is shown by red letters, and the red bar displays the feeling. We achieved a 62% accuracy rate at the 30 epochs. Our method accurately predicts faces that are happy, angry, neutral, fear, surprised. However, it predicts faces pretty poorly and blends with emotions of disgust and sad.

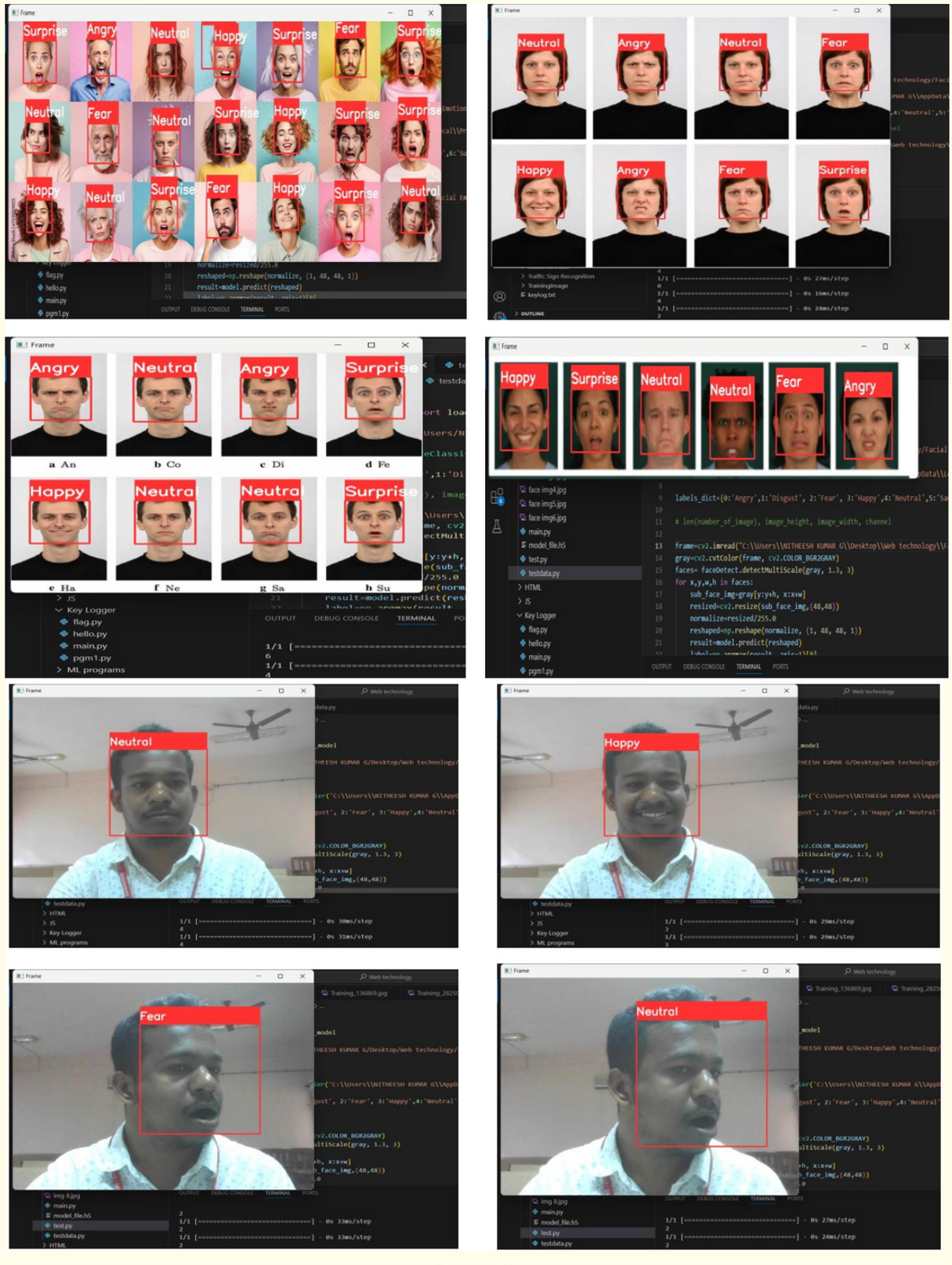

Figure 10

Conclusion

A flexible facial recognition model that links behavioral and physical features is used in this study's facial emotion recognition using convolutional networks technique. The human face represents a wide range of emotions, such as surprise, disgust, fear, angry, sad, happy, and neutral. In this study, face expressions were identified using the Fer-2013 dataset and Python source code. After being trained on photographs, the model is able to recognize the expressions in the images in real time. The system under research for facial expression analysis using convolutional networks is covered in this article.

Bibliography

1. Dr. V Suresh. "Facial emotion recognition using cnn". *International Journal of Creative Research Thoughts* 11 (2023): 2320-2882.

2. Apeksha Khopkar. "Facial expression recognition using cnn with keras". *Research Gate* 14 (2021): 47- 50.

3. Shivam Singh. "Face recognition system". *International Journal of Creative Research Thoughts* 8 (2019): 2278-0181.

4. Imane Lasri. "Facial emotion recognition of students using convolutional neural network". *Research Gate* (2022).

5. Rishab Kumar. "Facial emotion detection". *International Journal of Creative Research Thoughts* 10 (2022): 2320-2882.

6. Nur Alia Syahirah Badrulhisham. "Emotion recognition using convolutional neural network". *Iop* (2021).

7. Dr. V Umarani. "Face emotion recognition". *International Journal of Creative Research Thoughts* 9 (2021): 2320-2882.

8. Ninad Mehendale. "Facial emotion recognition using convolutional neural networks". *Springer Nature* (2020).

9. Suci Dwijayanti. "Facial Expression Recognition and Face Recognition Using a Convolutional Neural Network". *Isriti* (2020): 978-1-7281-8406-7.

10. Mohammad A Haghpanah. "Real-Time Facial Expression Recognition using Facial Landmarks and Neural Networks". (2021).

11. Sabrina Begaj. "Emotion Recognition Based on Facial Expressions Using Convolutional Neural Network (CNN)". IEEE Xplore (2021).

12. Zi-Yu Huang. "A study on computer vision for facial emotion recognition". *Scientific Reports* (2021): 41598-02335446-4.

13. Pengyun Hu. "LCANet: amodelforanalysis of students real-time sentiment by integrating attention mechanism and joint loss function". Springer (2024).

14. Ning Zhou. "A Lightweight Convolutional Neural Network for Real-Time Facial Expression Detection". IEEE Access (2021).

15. MAH Akhand. "Facial Emotion Recognition Using Transfer Learning in the Deep CNN". *Electronics* 10 (2021): 1036.