Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 218 (2023) 375–382

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering

Enhanced Automatic Recognition of Human Emotions Using Machine Learning Techniques

Monisha.G.S[a],Yogashree.G.S[b], Baghyalaksmi.R[c],Haritha.P[d]

[a,b,c,]*Computer Science Department, PanimalarEngineering College.Poonamallee, Chennai-600123, India.*
[d]*Information Technology College. Poonamallee, Chennai-600123, India.*

Abstract

Emotion recognition plays an increasingly significant role in interpersonal relationships. Research on emotion recognition is the most active within the area. Therefore, there are many different techniques for recognizing human emotion. The majority of the time, one's emotions can be expressed through speech, hand movements, or physical gestures. Facial expressions are used to convey most emotions. The recognition of facial expressions is largely dependent on facial expressions. The identification of human emotions has been extensively studied, but there are no effective techniques that can identify them correctly. In order to overcome this problem, the proposed method implements a real-time emotion recognition system by using machine learning.

Keywords: Facial expression, emotion recognition, machine learning, research filed

1. Introduction

Human emotion can be recognized by analyzing the human face or verbal communication. It is crucial [4] for public interaction to know how people are feeling. Facial expressions are essential for recognizing human emotions. Through identifying emotions, we can understand what humans are thinking. Surveys have found that humans interact socially through emotion and universal language. Through facial [10] expressions, people convey information to one another. The mental well-being of a person can be seen through the expression on their face. Hence, automatic recognition [1] of emotion using Image processing, cybersecurity, robotics, psychological studies, and virtual reality applications, to mention a few, can all benefit from high-quality sensors.

10.1016/j.procs.2023.01.020

When a human expression is analysed along with [5,6] multimodal forms of text, physiological audio, or video, its recognition rate is high. The face, body action, gestures, and voice of a person are used to identify a variety of expressions. Based on hybrid approaches, knowledge-based techniques, statistical methods, and knowledge-based techniques existing approaches to emotion recognition have been [2] classified into various types.

In many of the proposed systems, convolutional neural networks (CNN) are used to design real-time CNNs. Currently, face detection is the [3] major problem in existing research. Convolutional neural networks (CNN) are used for both gender and emotion classification. Features of the image are computed [7,8] by convolutional neural networks. There are seven different emotion states: neutral, joy, sorrow, surprise, anger, fear, and disgust. The features of the image were calculated using a three-dimensional face model, and a k-NN classifier was applied to classify them.

In interpersonal communication, facial expressions are one of several important nonverbal components that convey emotional information. The study of facial emotion has [9] therefore gained considerable Over the last few decades, attention has been increasingly important in perceptual and cognitive science, as well as effective [11] computing and computer animations. Interaction between humans and computers could be made easier and more effective with the help of intelligentfacial expression recognition.

The following is a breakdown of the paper's structure. Section 2 delves deeper into the problems that have been raised thus far. Section 3 describes the proposed system in detail, Section 4 shows the results ofthe enhanced system, and Section 5 presents the conclusion.

2. Related Work

Facial Expressions identification because of a clever facial deterioration. At first, seven locales of interest (ROI) signify the space between the brows, nose, and mouth: left brow, right brow, left eye, and right eye. Highlights were eliminated with the help of [1] facial milestones recognized by the connection point calculation. Anyway, unique nearby descriptors, like LBP, CLBP, LTP, and Dynamic LTP, are utilized to separate elements. At long last, the element vector centers around the face picture and is taken care of into a multiclass support vector machine to accomplish the acknowledgment task. Where results on two public datasets show that the proposed strategy is best in class techniques given other facial disintegrations. Exploratory outcomes showed the adequacy of our proposed strategy on completely tried datasets utilizing every single tried descriptor. The proposed facial deterioration beats the best-in-class ones.

Two- and three-dimensional images can be used to determine image features. The K-NN classifier and MLP neural network can be used to extract feature classification. Initially, subject-subordinate - for each client individually - and subject-autonomous - for all clients together - are used to recognize a person's feelings. In all cases, information was arbitrarily split between the showing component (70%) and the testing part (30%) for the 3-NN classifier [2], and into three groups for MLP: instructing (70%), testing (15%), and approval (15%). (15 percent). With the form inclination strategy, the brain network was prepared to use a backpropagation calculation. The expert has performed look recognition using revolutionary deep learning-based to improve the problems of look recognition from photographs. The information information of the [2] face photos is used in a current model to extract hidden nonlinearity, which is essential for identifying the type of feeling and response that individual is transmitting. They improve a profound convolutional brain organisation (DNN) model to create a grid of squares, each with identifying and sub-inspecting convolutional layers. The suggested look acknowledgment organisation (FERNet) restricts existing techniques and model complexity, according to tests on the benchmark FER2013 dataset. In this vein, a model was built using the FER2013 dataset, which is the most demanding of all the datasets available for this task. They achieve a precision of roughly 69.57 percent. Examine the effects of dropout, group standardization, and increase, as well as how they contribute to the reduction of crime.

Recognition of the facial expression using convolution neural network (CNN). because recognizing emotions is a difficult process because their no lank mark to separate the emotion of the face and there is lots of complexity and variability. This proposed model [3] used a machine learning algorithm for the extraction feature of the image. Convolutional neural networks (CNN) are used to enhance the work for the recognition of facial emotion. Rather than computing hand-designed highlights, CNN ascertains highlights by advancing naturally. The suggested technique is unusual in that it uses facial activity units (AUs) of the face, which are initially detected by CNN and then consolidated to perceive the seven basic inclination states. The Cohn-Kanade data set is used to evaluate the model, with the goal of achieving the best precision rate of 97.01 by fusing AU, whereas other efforts in the field used a direct CNN and achieved an exactness rate of 95.75.

The FERC is based on a convolutional brain organization (CNN) with two sections: The foundation is removed from the image in the first segment, and the facial component vector extraction is the emphasis of the second. An expressional vector (EV) is employed in the FERC model [4] to track down the five unique types of ordinary glances. The put away data set of 10,000 photos yielded administrative information (154 people). Using an EV of length 24 characteristics, it was possible to accurately depict the experience with 96 percent correctness. The two-level CNN works in sequence, with the last perceptron layer changing the loads and type values with each emphasis. With single-level CNN, FERC differs from generally used approaches, focusing on precision.Furthermore, a creative foundation ejection technique used prior to the EV period attempts to avoid numerous difficulties that may arise (for instance distance from the camera). FERC was extensively tested on over 750K images using the expanded Cohn-Kanade demeanour, Caltech faces, CMU, and NIST datasets. We think that the FERC emotion recognition will be useful in a variety of applications, including predictive learning, lie detectors, and so on.

3. Proposed Work

A convolutional neural network is used to identify the emotion of the person in order to derive more accurate features. The process of extracting the features of the image is a critical one since the rate of segmentation is low. A training dictionary will then be created based on the extracted features. At first, the dataset was trained with specifications.

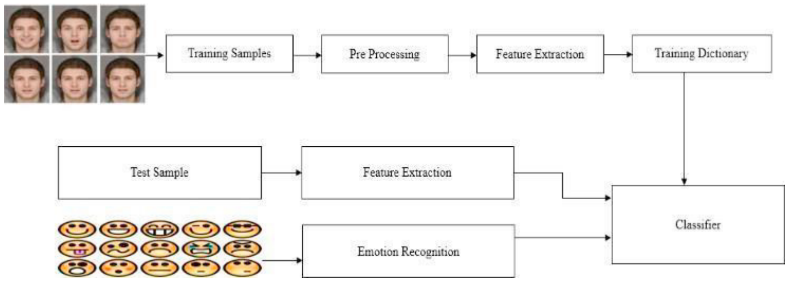

Figure 1. Proposed System Framework

Fig. 1 illustrates the proposed system. At first, an image of the person is captured using a camera or mobile phone. In order to classify an image of a person, it is subjected to various pre-processing steps, segmentation, feature extraction with a convolutional neural network, and, most importantly, a classifier. The proposed method identifies the emotion of the person after classifying the image.

3.1 Image Acquisition

Files Acquisition of images is the first step in the proposed research. The image is captured in the form of pictures using a smartphone camera or a camera. The image is actually the first and most crucial part of any project since, without an image, there is no wayto proceed. Various pre-processing methods are applied after a picture is taken.

3.2 Pre Processing

The image is pre-processed to remove unwanted noise and disruption after acquisition to improve image quality. As the image's intensity rises, the image's noise will decrease. During processing, there is a reduction and normalization of the image. Pre processing can do using a wiener filter using a high range of PSNR values and calculated using MSE.

3.3 Feature Extraction

Feature extraction is one of the major processes in this project. Each image has its own unique character. The process of feature extraction is used to define the shape of the face of people through the formal technique. Feature extraction uses a convolutional neural network (CNN) that classifies data in high dimensions of invariance, scaling, skewing, and other distortions.

3.4 CNN Architecture

Face Recognition acknowledgment general CNN design with preprocessing, the info layer can be predefined to a proper size, which can then be placed into the following layer. In this progression, the venture utilized OpenCV, a well-known Computer vision library that remembers pre-prepared channels for identifying faces for pictures. The venture additionally utilized Ada lift to find and trim appearances. By doing this, the picture's size is essentially diminished. Then, the info layer is moved to the Convolution2D layer, where the quantity of channels is indicated as a super-boundary.

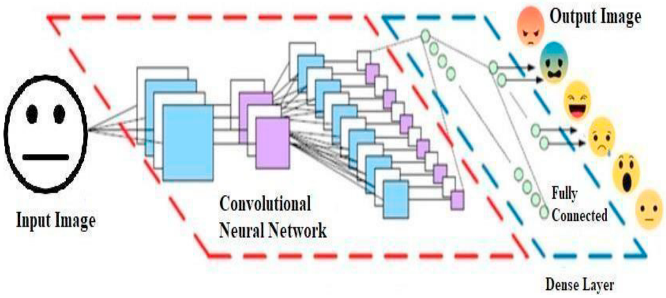

Figure 2. Architecture of Convolutional Neural Network

In Figure 2, we explain how a Convolutional Neural Network recognizes facial emotions. The kernel, for example, is composed of randomly generated weights. The portion, for instance, is made out of arbitrarily produced loads. Each channel, for example, a sliding window, navigates the whole picture to develop a component diagram with shared loads. A convolutional layer produces a component map that shows how pixels are raised, like edges, light, and examples. To fabricate a conventional CNN design, pooling to decrease the [5] aspect behind the convolution layer is fundamental. This is on the grounds that adding more convolution layers builds the computational expense. MaxPooling2D is a famous pooling method that utilizes a 2x2 window to navigate highlight maps, holding just the pixels with the most noteworthy worth. Whenever pixels are combined, the picture size is diminished by four. As an option in contrast to the sigma layer returns the likelihood of every look class as its enactment work. Consequently, the CNN [7] model can assess the likelihood of every inclination and select the feeling with the most elevated prescient score as the acknowledgment result. Figure 8 shows the CNN engineering that was at last intended to distinguish faces and perceive explicit looks for each face.

4. Facial Expression Classification

Following feature extraction, image classification can be performed using a convolutional neural network, allowing the test sample to detect the features and expressions of the person in an efficient way. Classifiers classify emotions based on the facial expressions of a particular person. Using feature extraction and emotional recognition, facial expressions can be categorized into different [6] categories. The extracted feature image goes through various processes of sample testing, and the tested samples are compared with the emotions and displayed from the facial expressions. This operation is performed with three kernels. Input pixels are divided into eight neighboring pixels, and each neighboring pixel is multiplied by the kernel matrix's associated value Finally, the final output value is calculated by adding all multiplied values together.

5. Result and Discussion

MATLAB is used to implement the proposed system. An image is used as an input for emotion recognition of a person and also for recognizing and diagnosing their emotions. It is then subjected to a variety of processing methods that allow it to be segmented.

`

Table 1. Shows the Comparisons of Different Filters

Input Image	Images	Median Filter	Gaussian Filter	Wiener Filter
Image1				
Image 2				
Image 3				

Table 1. Comparison of the three different filters' accuracy. To compare image compression quality then use the peak signal-to-noise ratio (PSNR) and mean-square error (MSE) computed from the various filters. The peak error is measured by PSNR, and the squared error between the compressed and original image is measured by MSE. The error is low when the MSE is low. See Table 2 for information on the Median, Gaussian, and Wiener Filters, as well as the PSNR.

Table 2. Comparison of filter performance using PSNR value

Types of Filters	PSNR VALUE		
	Median Filter	Gaussian Filter	Wiener Filter
Image 1	64.974	87.56	96.89
Image 2	83.64	72.356	90.67
Image 3	74.02	64.34	86.90

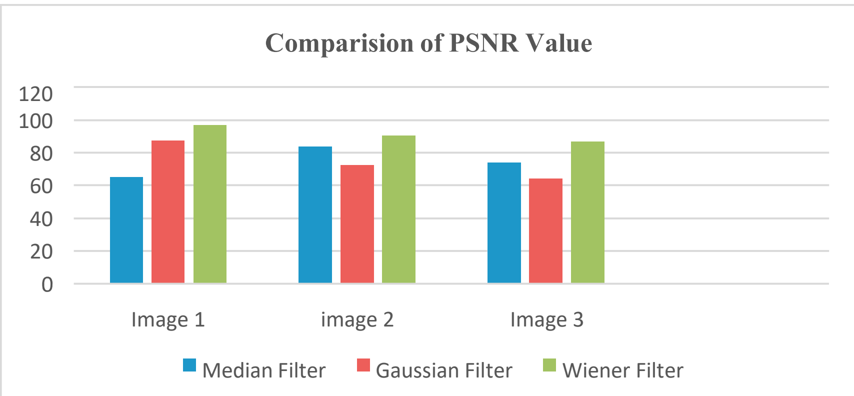

Figure 3. Comparisons of PSNR Value

Figure 3. shows Compare the performance of the three filters. The processed image uses a Wiener filter because of the range of PSNR values and [8] the Wiener filter works efficiently in comparison to another filter. To segment features of

the human face for emotion analysis, feature extraction and semantic segmentation technique is used with three different inputs: image 1, image 2, image. K When compared with other segmentation processes, K Means Algorithm segmentation performs better.

Table 3. Performance of segmentation.

Images	Input Images	Pre processed Image	Segmented Image	Specificity	Sensitivity	Accuracy
1				94.6%	86.78%	97.23%
2				85.01%	92.90%	97.21%
3				96.90%	93.76%	98.23%

Table 3 shows the three classes of input images with their segmented portion undergoes the segmentation to find and computed using specificity, sensitivity and accuracy. There are 95.6% specificity, 86.78 % sensitivity, and 97.23 % accuracy in image 1. Image 2 has a specificity of 85.01%, a sensitivity of 92.90%, and an accuracy of 97.21%. Its specificity is 96.90%, its sensitivity is 93.76%, and its accuracy is 98.23%.

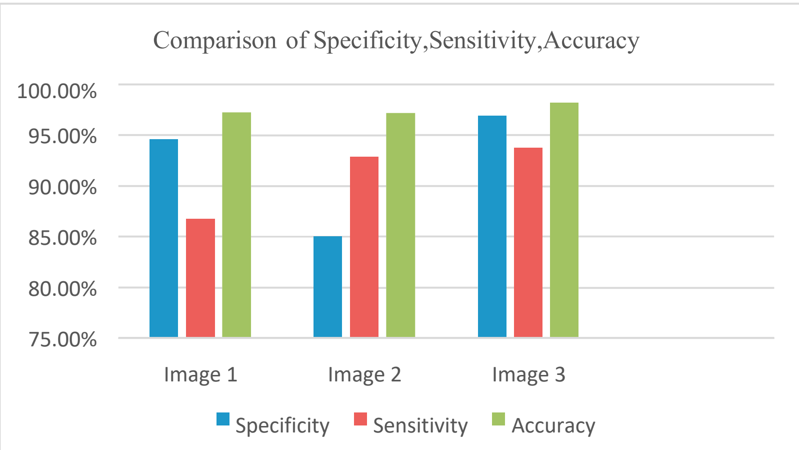

Figure 4. shows the comparisons of Specificity, Sensitivity and Accuracy

Figure 4. shows the comparisons of Specificity, Sensitivity and Accuracy of the segmented Image for the Feature Extraction using Convolutional Neural Network.

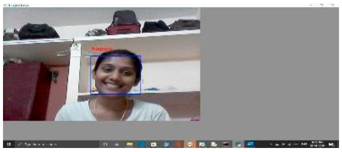

Figure 5. Happy Emotion

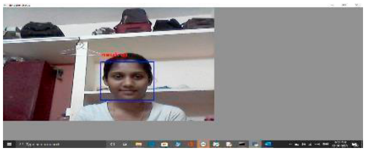

Figure 6. Neutral Emotion

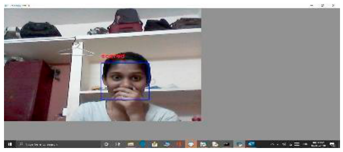

Figure 7. Scared Emotion

Figure 5, 6 and 7. Shows the output of the Proposed project and also shows the Major accuracy rate. The emotion of the human can be identified in an efficient manner by using the Wiener filter to remove the unwanted noise which is presented in the image of the human. After removing the unwanted noise preprocessed image goes through a segmentation process to provide a high range of accuracy in the given outcome.

6. Conclusion

Facial expression recognition is a novel way to express the emotion of humans beginning using a convolution neural network (CNN). By recognizing the expression of human feelings and also, we can identify the mental health of the person. Convolution neural networks work more affectedly manner in the feature extraction of the Face of the human. It works less in the 30° orientation. The ability to accurately determine emotions was greatly enhanced by the removal of the background. The calculation time for this approach to recognize the emotions on a human face is 15.3 seconds. For many emotion-based applications, such as lie detectors and mood-based learning for students, facial emotional detection could be the first step. Fetching the features of the human face in a more advanced manner, as well as determining the mood of the individual. As a result of the studies conducted, we were able to classify emotions with an accuracy of 97.23 % for data divided randomly and 75 % for data divided in a 'natural' way. The MLP classifier and "natural" data division (subject-independent) led to the same outcome for all users. All experiments used the same settings and the same position of the Kinect unit. Certain facial expressions were correlated to classification accuracy depending on how they were acted out by participants. There are many factors that can influence classification accuracy in real-life situations, actions. When you are experiencing true emotions, your facial expression may be more or less noticeable.

References

[1] A Lekdioui, Khadija; Ruichek, Yassine; Messoussi, Rochdi; Chaabi, Youness; Touahni, Raja (2017). [IEEE 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP) - Fez, Morocco (2017.5.22-2017.5.24)] 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP) - Facial expression recognition using face-regions., (), 1–6. doi:10.1109/ATSIP.2017.8075517.

[2] Mostafa Mohammadpour, Seyyed Mohammad. R Hashemi,Shahrood, Hossein Khaliliardali , Mohammad. M AlyanNezhadi (2017) “Facial emotion recognition using deep convolutional networks”.EEE 4th I International Conference on Knowledge-Based Engineering and Innovation (KBEI).doi: 10.1109/KBEI.2017.8324974.

[3] Tarnowski, Paweł; Kołodziej, Marcin; Majkowski, Andrzej; Rak, Remigiusz J. (2017). Emotion recognition using facial expressions. Procedia Computer Science, 108(), 1175–1184. doi: 10.1016/j.procs.2017.05.025.

[4] Mehendale, Ninad (2020). Facial emotion recognition using convolutional neural networks (FERC). SN Applied Sciences, 2(3), 446 –

[5] . doi:10.1007/s42452-020-2234-1

[6] Monisha G.S., Malathi S. (2021).” Effective Survey on Handwriting Character Recognition”. In: Singh V., Asari V.K., Kumar S., Patel R.B. (eds) Computational Methods and Data Engineering. Advances in Intelligent Systems and Computing, vol 1257. Springer, Singapor e. https://doi.org/10.1007/978-981-15-7907-3_9.

[7] R. K. Madupu, C. Kothapalli, V. Yarra, S. Harika And C. Z. Basha, "Automatic Human Emotion Recognition System using Facial Expressions with Convolution Neural Network," 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2020, pp . 1179-1183, doi: 10.1109/ICECA49313.2020.9297483.

[8] Daniel C. Tozadore, Caetano M. Ranieri, Guilherme V. Nardari, Roseli A.F. Romero, Vitor C. Guizilini, "Effects of Emotion Grouping for Recognitionin Human-Robot Interactions", Intelligent Systems (BRACIS) 2018 7th Brazilian Conference on, pp. 438-443, 2018.

[9] Adam H.M. Pinto, Caetano M. Ranieri, Guilherme Nardari, Daniel C. Tozadore, Roseli A.F. Romero, "Users' Perception Variance in Emotional Embodied Robots for Domestic Tasks", Robotic Symposium 2018 Brazilian Symposium on Robotics (SBR) and 2018 Workshop on Robotics in Education

[10] (WRE) 2018 Latin American, pp. 476-482, 2018.
[11] Yahaya Isah Shehu, Ariel Ruiz-Garcia, Vasile Palade, Anne James, "Detailed Identification of Fingerprints Using Convolutional NeuralNetworks", Machine Learning and Applications (ICMLA) 2018 17th IEEE International Conference on, pp. 1161-1165, 2018.
[12] Inthiyaz Bahsa Kattubadi, Rama Murthy Garimella, "Emotion Classification: Novel Deep Learning Architectures", Advanced Computing &Communication Systems (ICACCS) 2019 5th International Conference on, pp. 285-290,2019.
[13] Ruiz-Garcia, M. Elshaw, A. Altahhan and V. Palade, "Stacked deep convolutional auto-encoders for emotionrecognition from facial expressions," 2017 International Joint Conference on Neural Networks (IJCNN), 2017, pp. 1586-1593, doi: 10.1109/IJCNN.2017.7966040.