Irish Interdisciplinary Journal of Science & Research (IIJSR)
Volume 7, Issue 3, Pages 90-97, July-September 2023

Design of Emotion Based Face Recognition System Using Convolutional Neural Networks

Preethiga, E.[1*], Navaneetha Krishnan, P.[2] & Bashirunisha, S.[3]

[1]PG Scholar, [2,3]Assistant Professor, [1-3]Department of Electronics and Communication Engineering, Sir Issac Newton College of Engineering and Technology, Nagapattinam, Tamilnadu-611 102, India.
Corresponding Author Email: preethigasri17@gmail.com*

DOI: https://doi.org/10.46759/IIJSR.2023.7310

Article Received: 09 July 2023 Article Accepted: 20 September 2023 Article Published: 29 September 2023

ABSTRACT

Facial expressions convey non-verbal information between humans in face-to-face interactions. Automatic facial expression recognition, which plays a vital role in human-machine interfaces, has attracted increasing attention from researchers since the early nineties. Classical machine learning approaches often require a complex feature extraction process and produce poor results. In this paper, we apply recent advances in deep learning to propose effective deep Convolutional Neural Networks (CNNs) that can accurately interpret semantic information available in faces in an automated manner without hand-designing of features descriptors. We also apply different loss functions and training tricks in order to learn CNNs with a strong classification power. The experimental results show that our proposed networks outperform state-of-the-art methods on the well-known FERC-2013 dataset provided on the Kaggle facial expression recognition competition. In comparison to the winning model of this competition, the number of parameters in our proposed networks intensively decreases, that accelerates the overall performance speed and makes the proposed networks well suitable for real-time systems.

Keywords: Emotion Recognition; Convolutional Neural Networks; Deep Learning.

1. Introduction

Recognizing emotions is considered to be a chief potential for one's interpersonal skills and plays a vital role in communications. It can be very distinctive and complex task for a human being. We have been trying to recognize human emotions for decades now using our human instincts, however sometimes we failed to successfully identify the emotions and sentiments of the opposite person [1]-[7].

The aim of this research is to successfully recognize the emotions from visual data i.e. images using deep neural networks. The research has attempted to identify and overcome the dead locks in current architecture and systems to identify human emotions. Human emotions can be primarily identified using human facial muscles and facial gestures resulting into expressing their emotions, feelings and opinion of others. The aim of this research is to increases the accuracy of success full prediction of the seven human emotions namely happy, sad, angry, scared, surprise, neutral and disgust while utilizing the Facial Emotion Recognition dataset (FER2013).

Preprocessing, feature extraction, and classification are the three main steps in the Face Expression Recognition (FER) methodologies survey. The various FER technique types are explained in this survey along with their main contributions. The number of expressions identified and the complexity of the algorithms are used to compare the performance of different FER approaches.

In this examination, many face expression databases like JAFFE, CK, and others are discussed. The research on classifiers gathered from recent papers provides research fellows with a more thorough and dependable grasp of the specific traits of classifiers [8-15].

Convolutional Neural Network (CNN) is used to recognize the seven distinct facial expressions of people. The seven categories are surprise, neutral, sad, furious, disgusted, and terror. There were roughly 36,000 gray scale photos in the dataset. Our specifically designed CNN model, which consists of two fully linked layers and four convolutional layers, achieves test data accuracy of 64.3% [2].

In order to categorize children's spontaneous emotion detection, progressive light residual learning is used. As opposed to earlier residual neural networks, ours steadily increases the skip connection as it descends deeper into the network. Limiting the skip connection locally allows the progressive light residual network to explore more feature space. This increases the network's susceptibility to perturbations and aids in resolving the overfitting issue for smaller datasets. The proposed methodology outperformed state-of-the-art methods significantly, according to experimental findings using a benchmark dataset of children's emotions [14-19].

Biometrics are used by facial recognition systems to extract facial traits from images or videos. To discover a match, it compares the data from the known database. Nowadays, a variety of facial recognition methods are employed. One of the most important biometrics is facial recognition. The commonly used facial recognition technology employs machine learning to find, match, and identify the person. When utilized to distinguish between users, the facial recognition technology delivers precise results. Facial recognition and its methods will be briefly discussed in this manuscript [4].

We gain from the veracity of the information we've learned through this channel while making decisions. As a result, numerous investigations in this area are carried out to develop reliable face expression recognition (FER) systems. The informative facial image elements are necessary to develop a strong and trustworthy FER system. In this study, a trustworthy FER strategy was established using the Local Binary Patterns (LBP) and Local Phase Quantization (LPQ) methodologies. The major objective is to ascertain whether or not combining these two approaches contributes. According to experiments, combining these two strategies raises the success rate of categorization from 75% to 88% [20-25].

The main objective of this study is to test and identify various suitable methodologies and proven techniques in order to classify emotions in real-time video stream. The research can be deployed to an online system which can be used for many for various purposes in real-time. This can be beneficial in machine-human interaction in multiple forms. This research aims to use FER 2013 dataset which is available publicly and training our built deep learning model on this dataset and predict or classify the emotions on to the new acquired images. Final system targets to classify the emotions of the live streaming data in a successful manner. The proposed system aims the use of Convolutional Neural Network from the Deep Learning family to achieve success.

Section 2 provides an information about an Emotion Recognition using Machine Learning, Section 3 provides an information about Emotion Recognition using Deep Learning, Section 4 is about Result and Discussion, Section 5 is about Conclusion.

2. Emotion Recognition using Machine Learning

In the existing system, classification is done through simple image processing to classify images only. Existing work includes the application of feature extraction of facial expressions with the combination of neural networks for the recognition of different facial emotions (happy, sad, angry, fear, surprised, neutral, etc..) Humans are capable of producing thousands of facial actions during communication that varies in complexity, intensity, and

Irish Interdisciplinary Journal of Science & Research (IIJSR)
Volume 7, Issue 3, Pages 90-97, July-September 2023

meaning. The existing system is capable of analyzing the limitations of the existing system of Emotion recognition using brain activity. In this work, two machine learning algorithms such as KNN, and Haar Cascade are used to identify and classify facial emotion.

KNN is a simple nonlinear classifier model that classifies data points based on similar points. KNN algorithm is often used in image recognition technology, decision-making models, and simple recommendation systems. KNN is a non-probabilistic learning algorithm used to classify unknown test data based on the majority of similar data among the k-nearest neighbors closest to test/anonymous data. KNN algorithm works on deeply rooted mathematical formulas that are used for classification. When implementing KNN, the foremost step is to transform data points into feature vectors, or a certain mathematical value. Then the algorithm processes it by finding the distance between the mathematical values of these points.

Haar Cascade Detection algorithm is a machine learning-based approach where a cascade function is trained using lots of positive and negative images and then used to detect objects in other images. Haar Cascade is an object detection algorithm to identify faces in an image or real-time video. It uses edge or line detection features.

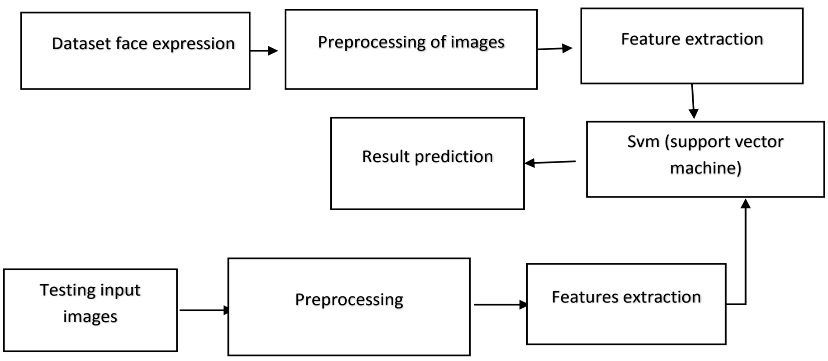

Figure 1. Existing system architecture

3. Emotion Recognition using Deep Learning

In deep learning model multi convolution neural network based algorithm is employed for face expression. Facial expression recognition using a Multi - CNN algorithm involves training a deep learning model that is capable of recognizing and classifying different facial expressions from images or videos of faces. Here is a general approach to implementing this algorithm Collect and prepare a dataset of facial expression images: You'll need a dataset of images that show different facial expressions, such as happiness, sadness, anger, etc. You can use publicly available datasets like the Facial Expression Recognition Challenge dataset, or you can collect your own dataset.

Preprocess the data: You'll need to preprocess the data by resizing the images, normalizing the pixel values, and augmenting the data to increase the size of the dataset. Train the Multi-CNN model: You'll need to train a deep learning model using multiple convolutional neural networks (CNNs) that are combined to recognize different aspects of the facial expression. Each CNN is trained on a different set of features to capture different aspects of the facial expression. The outputs of each CNN are then combined to produce a final prediction. Test the model: Once the model is trained, you'll need to test it on a separate set of facial expression images to evaluate its performance Deploy the model: Finally, you can deploy the model to recognize facial expressions in real-time video streams or images. Overall, the Multi-CNN algorithm is a powerful technique for facial expression recognition, and can be used in a variety of applications, such as emotion recognition for human-computer interaction, facial recognition for security purposes, and more.

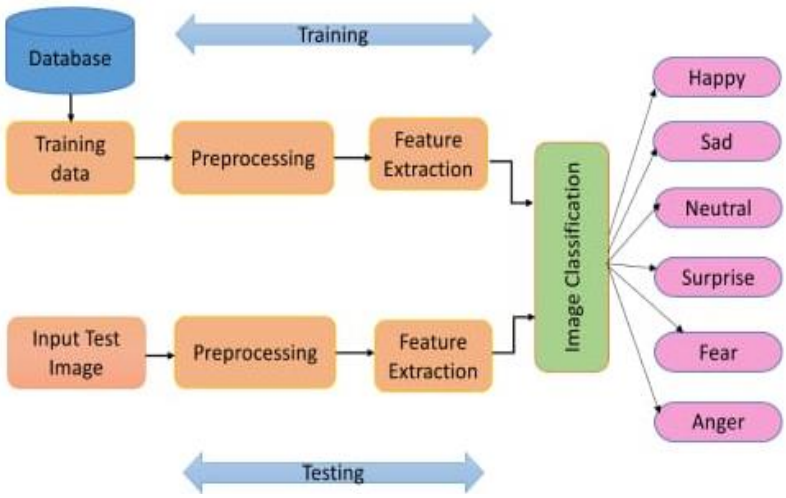

Figure 2. System Architecture

Our initial concept designs would be segregated into 4 main components that are Training, Evaluate, Pre-processing and Testing phases. Within each phase there would be carried out specific asks which can be ranged from classification of training data, calculating accuracy and loss, validating data, resizing the data etc. As per our system design we would initially acquire training dataset and after pre-processing it we would pass it for classification using Convolutional Neural Network while applying the mini-Exception architecture model and storing the model results into file which will be later used for on our test data for prediction and classification. The stored architecture model will then be passed onto our testing data which has been already pre-processed i.e. resized and gray-scaled and now ready for classification from the learned CNN model. The learned CNN model would classify or predict the testing or live data set from one of the 7 categories of the emotions.

3.1. Module list

- Training
- Dataset collection
- Preprocessing
- CNN layer Feature extraction
- Testing with result and analysis

The concept of convolutional neural networks is they are very successful in image recognition. The key part to understand, which distinguishes CNN from traditional neural networks, is the **convolution** operation. Having an image at the input, CNN scans it many times to look for certain **features**. This scanning (convolution) can be set with 2 main parameters: stride and padding type. As we see on below picture, process of the first convolution gives us a set of new frames, shown here in the second column (layer). Each frame contains an information about one

Irish Interdisciplinary Journal of Science & Research (IIJSR)
Volume 7, Issue 3, Pages 90-97, July-September 2023

feature and its presence in scanned image. Resulting frame will have larger values in places where a feature is strongly visible and lower values where there are no or little such features. Afterwards, the process is repeated for each of obtained frames for a chosen number of times. In this study we chose a classic model which contains only two convolution layers. The latter layer we are convolving, the more high-level features are being searched.

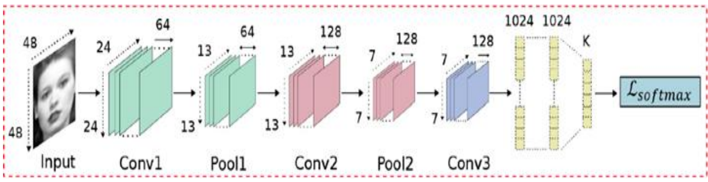

Figure 3. Convolution model architecture

It works similarly to human perception. To give an example, below is a very descriptive picture with features which are searched on different CNN layers. As you can see, the application of this model is face recognition. You may ask how the model knows which features to seek. If you construct the CNN from the beginning, searched features are random. Then, during training process, weights between neurons are being adjusted and slowly CNN starts to find such features which enable to meet predefined goal, i.e. to recognize successfully images from the training set. Different image feature pattern.

4. Results and Discussion

Our systems implies the use of CNN architecture model in order to facial expression detection which can be implemented optimally in real-time detection. Our research aims to fabricate a system to recognize the face and identify the facial emotions. The success rate in recognizing the emotion in test dataset is 62.5% which concludes that our proposed system gives good results. For our real-time system within a live video stream each the system recognizes the emotions for each frame.

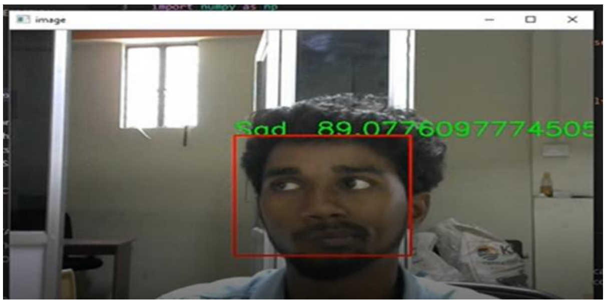

Figure 4. Sad Emotion

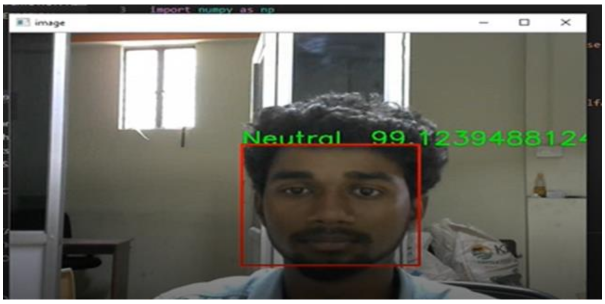

Figure 5. Neutral Emotion

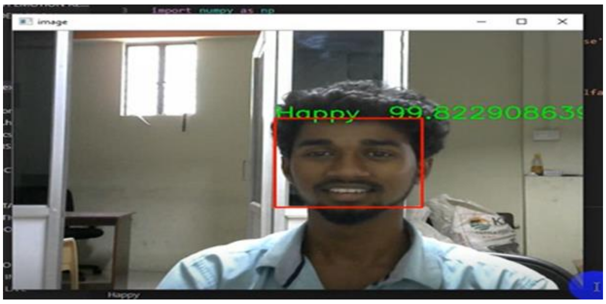

Figure 6. Happy Emotion

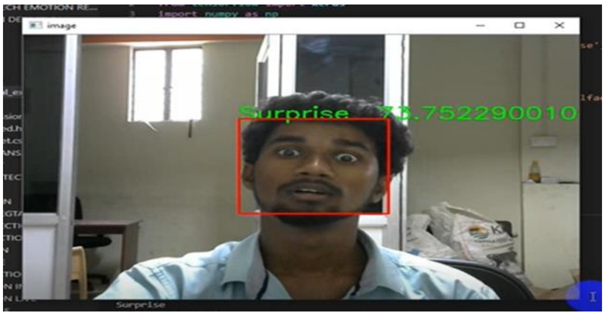

Figure 7. Surprise Emotion

5. Conclusion

This Study Presents a complete and fully automated approach for facial expression identification by simultaneously utilizing the face surface and face subsurface features. We presented a new algorithm for the face identification and recognition, which can more reliably extract the face features and achieve much higher accuracy than previously proposed facial identification approaches. The proposed approach presents a very low degree of complexity, which makes it suitable for real-time applications. Depending upon the selected features and the measured region properties of the human face, the different expression of the human was further classified using CNN (convolution layers). The proposed method is superior compared with other state-of-the-art approaches and that the analysis of the general image quality of the face images reveals highly valuable information that may be very efficiently used to discriminate them from fake traits.

Declarations

Source of Funding

This study has not received any funds from any organization.

Conflict of Interest

The authors declare that they have no conflict of interest.

Consent for Publication

The authors declare that they consented to the publication of this study.

Authors' Contribution

All the authors took part in literature review, research, and manuscript writing equally.

References

[1] M.S. Bilkhu, S. Gupta & V.K. Srivastava (2019). Emotion classification from facial expressions using cascaded regression trees and SVM. Computational Intelligence: Theories, Applications and Future Directions, Springer, Pages 585–594.

[2] Awan M.J., Raza A., & Yasin A. (2021). The Customized Convolutional Neural Network of Face Emotion Expression Classification. Journal of Mobile Embedded and Distributed Systems, Volume 3, Issue 4.

[3] G.N. Foley & J.P. Gentile (2010). Nonverbal communication in psychotherapy. Psychiatry, Edgemont, 7(6): 38–44. doi: 10.4324/9781315134925-16.

[4] Durmusoglu A., & Kahraman Y. (2022). Facial Expression Recognition Using a Combination of Local Binary Patterns and Local Phase Quantization. IEEE Trans. Multimedia, 12(6): 477–480.

[5] N. Libbrecht, F. Lievens, B. Carette & S. Côté (2014). Emotional intelligence predicts success in medical school. Emotion, 14(1): 64–73. doi: 10.1037/a0034392.

[6] F. Noroozi, M. Marjanovic, A. Njegus, S. Escalera & G. Anbarjafari (2019). Audio-visual emotion recognition in video clips. IEEE Trans. Affect. Comput., 10(1): 60–75. doi: 10.1109/TAFFC.2017.2713783.

[7] E. Sariyanidi, H. Gunes & A. Cavallaro (2015). Automatic analysis of facial affect: A survey of registration, representation, and recognition. IEEE Trans. Pattern Anal. Mach. Intell., 37(6): 1113–1133.

[8] K. Han, D. Yu & I. Tashev (2014). Speech emotion recognition using deep neural network and extreme learning machine. In Proc. Annu. Conf. Int. Speech Commun. Assoc. (INTERSPEECH), Pages 223–227. doi: 10.21437/interspeech.2014-57.

[9] A. Brogi et al. (2018). Survey on AI-based multimodal methods for emotion detection. Futur. Gener. Comput. Syst., 29(1): 1–18. doi: 10.1007/978-3-030-16272-6.

[10] S. Singh, V. Sharma, K. Jain & R. Bhall (2015). EDBL–algorithm for detection and analysis of emotion using body language. In Proc. 1st Int. Conf. Next Gener. Comput. Technol. (NGCT), Pages 820–823. doi: 10.1109/NGCT.2015.7375234.

[11] C. Ma, Q. Liu & Y. Dang (2021). Multimodal art pose recognition and interaction with human intelligence enhancement. Frontiers Psychol., 12: 1–13. doi: 10.3389/fpsyg.2021.769509.

[12] K. Schindler, L. Van Gool & B. de Gelder (2008). Recognizing emotions expressed by body pose: A biologically inspired neural model. Neural Netw., 21(9): 1238–1246. doi: 10.1016/j.neunet.2008.05.003.

[13] M. Abdulrahman & A. Eleyan (2015). Facial expression recognition using support vector machines. In Proc. 23rd Signal Process. Commun. Appl. Conf. (SIU), Pages 276–279.

[14] Qayyum A., & Razzak I. (2021). Deep Residual Neural Network for Child's Spontaneous Facial Expressions Recognition Structural. Syntactic, and Statistical Pattern Recognition, Pages 282–291.

[15] Revina I.M., & Emmanuel W.R.S. (2021). A Survey on Human Face Expression Recognition Technique. International Journal of Computer Vision, Volume 56, Number 2.

[16] T. Jabid, M.H. Kabir & O. Chae (2010). Robust facial expression recognition based on local directional pattern. ETRI J., 32(5): 784–794. doi: 10.4218/etrij.10.1510.0132.

[17] C. Shan, S. Gong & P.W. Mc Owan (2009). Facial expression recognition based on local binary patterns: A comprehensive study. Image Vis. Comput., 27(6): 803–816. doi: 10.1016/j.imavis.2008. 08.005.

[18] C. Wang (2018). Human emotional facial expression recognition. ArXiv:1803.10864, Pages 1–8.

[19] S. Kang, B. Choi & D. Jo (2016). Faces detection method based on skin color modeling. J. Syst. Archit., 64: 100–109. doi: 10.1016/j.sysarc.2015.11.009.

[20] S.L. Happy & A. Routray (2015). Robust facial expression classification using shape and appearance features. In Proc. 8th Int. Conf. Adv. Pattern Recognit. (ICAPR), Pages 1–5. doi: 10.1109/ICAPR.2015. 7050661.

[21] P. Lucey, J.F. Cohn, T. Kanade, J. Saragih, Z. Ambadar & I. Matthews (2010). The extended Cohn–Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression. In Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPRW Workshops), Pages 94–101. doi: 10.1109/CVPRW.2010.5543262.

[22] P. Ekman & W.V. Friesen (2002). Facial action coding system.

[23] P. Ekman (1970). Universal facial expressions of emotion. California Mental Health, 8(4): 151–158.

[24] N.C. Ebner, M.K. Johnson & H. Fischer (2012). Neural mechanisms of reading facial emotions in young and older adults. Frontiers Psychol., 3: 1–19. doi: 10.3389/fpsyg.2012.00223.

[25] Ranjani S. (2022). A Study on Human Face Recognition Techniques. Computational Intelligence and Neuro Science, Volume 15, Issue 9.