iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IV **Month of publication:** April 2023

DOI: https://doi.org/10.22214/ijraset.2023.50709

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

Emotion Classification of Facial Images Using Machine Learning Models

Aakula Sai Lahari[1], Mula Supraja[2], Cheerneni Akshaya[3], K. Vikram Reddy[4]
[1, 2, 3, 4]*Department of Information Technology, Matrusri Engineering College, Hyderabad, India*

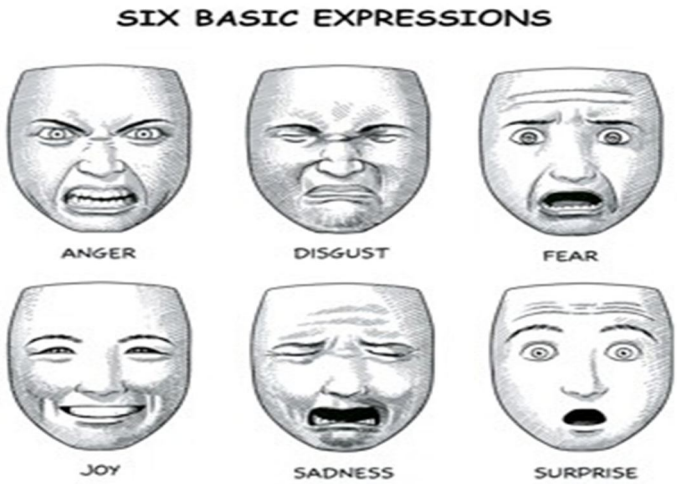
Abstract: ***Visual sentiment analysis, which studies the emotional response of humans on visual stimuli such as images and videos, has been an interesting and challenging problem. It tries to understand the high-level content of visual data. The success of current models can be attributed to the development of robust algorithms from computer vision. Most of the existing models try to solve the problem by proposing either robust features or more complex models. In particular, visual features from the whole image or video are the main proposed inputs. Little attention has been paid to local areas, which we believe is pretty relevant to human's emotional response to the whole image. Application of image recognition to find people in images and analyze their sentiments or emotions. This project uses the CNN algorithm to perform that task. Given an image, it would search for faces, identify them, put a rectangle in their positions and describe the emotion found and emoji is displayed.***
Keywords: face emotion detection, CNN algorithms, Machine Learning, Image recognition, feature extraction

I. INTRODUCTION

The movement of facial muscles beneath the skin is referred to as a facial expression. In nonverbal communication, facial expressions are used. Many different emotions can be expressed on the human face without using words. Additionally, unlike some nonverbal communication techniques, these facial expressions are comprehended by all types of people, unlike some nonverbal communication techniques. People from different cultures all use the same facial expressions to communicate their joy, sorrow, anger, surprise, fear, and disgust.

The study of systems and the creation of tools and gadgets that can identify, comprehend, process, and imitate human feelings is known as affective computing. Through the use of sensors, microphones, and cameras, affective computing systems can detect the user's emotions and respond by carrying out some specific, specified product or service characteristics. Human-computer interaction is one approach to look at efficient computing. In this scenario, a gadget is able to recognise and react to the users' expressed emotions.

Our goal is to analyse photographs taken by a live camera in real time and identify emotions from them. The webcam will now be recording a video, and faces will be recognised in the frames based on facial landmarks like the corners of the mouth, nose, eyes, and brows. Then, from these facial landmarks (dots) faces, the features were extracted that will be used for the detection of facial emotions. After identifying the emotions, we use picture processing tools to check for any discomfort.



©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

II. BACKGROUND

A. Machine Learning: The majority of machine learning algorithms in use are focused on identifying and/or utilising relationships between information. Machine learning is a combination of correlations and relationships. Once Machine Learning Algorithms are able to focus on specific correlations, the model can either generalise the data to highlight intriguing patterns or use these connections to forecast future observations. There are many different types of algorithms used in machine learning, including regression, linear regression, logistic regression, the Bayes theorem, Naive Bayes classifier, decision tree, entropy, ID3, SVM (support vector machines), K-means algorithm, random forest, and others.

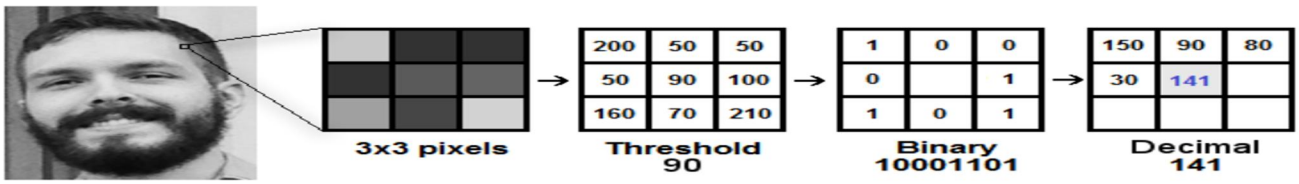
B. Image Recognition: Image recognition refers to the field of computer science that analyses photographs to recognize items, places, people, logos, objects, buildings, and other things. Image recognition, a procedure that can recognize and recognize an object in a digital video or image, is a subset of computer vision. In computer vision, techniques for acquiring, processing, and analyzing data from video or static images taken in the real world are included. These sources produce high-dimensional data that can be used to make decisions that are either numerical or symbolic. Computer vision also encompasses object recognition, learning, event detection, video tracking, and picture reconstruction in addition to image recognition.

A picture is perceived by the human eye as a collection of impulses that the visual cortex in our brain will process. The perfect replication of this mechanism is the goal of image recognition. A computer can distinguish between raster and vector images. Raster images are made up of a series of discrete numerically valued pixels, whereas vector images are made up of a collection of polygons with color annotations. Geometric encoding is converted into constructs that represent physical characteristics and objects in order to analyze images. The computer then does a logical analysis of these constructs. Classification and feature extraction are involved in data organization. Making an image simple by removing all but the most necessary details and leaving the rest out is the first stage in image classification. Building a prediction model is the second phase. This can be classified using an algorithm. We must train the classification algorithm before it can function by displaying tens of thousands of project-related and irrelevant photos. We utilize neural networks to create a predictive model. The neural network is a system that combines hardware and software to mimic the activities of the human brain. It can estimate functions based on a vast number of unknowable inputs. Support vector machines (SVM), face landmark estimation, K-nearest neighbors (KNN), logistic regression, and other image classification techniques are only a few examples. C. Feature Extraction: A starting collection of raw data is reduced in dimension through the process of feature extraction in order to serve some sort of processing need. An image's behavior is determined by its features. In essence, a feature is a pattern in an image, such a point or an edge. When you need to use fewer resources for processing while keeping the crucial and pertinent information, the feature extraction procedure might be helpful. Feature extraction can help to reduce the amount of redundant data. The sampled image is subjected to image preprocessing techniques including thresholding, scaling, normalization, binarization, etc., after which the features are extracted. Techniques for feature extraction are used to get features for picture classification and recognition.

III. LEARNING METHODS

A. CNN Implementation

We used OpenCV library to capture live frames from web camera and to detect students' faces based on Haar Cascades method. Haar Cascades uses the Adab Oost learning algorithm invented, who won the 2003 Gödel Prize for their work. The Adab Oost learning algorithm chose a few numbers of significant features from a large set-in order to provide an effective result of classifiers. In Keras, we used Image Data Generator class to perform image augmentation as shown in Figure 9. This class allowed us to transform the training images by rotation, shifts, shear, zoom and flip. The configuration used is: rotation range=10,width_shift_range=0.1,zoom_range=0.1, height_shift_range=0.1 and horizontal flip=True. Then we defined our CNN model with 4 convolutional layers, 4 pooling layers and 2 fully connected layers. After that, to provide nonlinearity in our CNN model we applied the ReLU function, and we used batch normalization to normalize the activation of the precedent layer at each batch and L2 regularization to apply penalties on the different parameters of the model.



Four concepts in CNN are:

- Convolution
- ReLu
- Pooling
- Full connectedness

1) *Convolution for Feature Extraction:* Using a filter or kernel, CNN performs convolution on an input image. Convolution and filtering involve scanning the screen, which begins at the top left corner and moves down after completing the breadth of the screen. This process is repeated until the entire screen has been scanned. The individual's face feature lines up with the illustration. The appropriate feature pixel is multiplied by the picture pixel. The values are multiplied by the feature's overall pixel count and then put together.
2) *Extraction of Non-Linearity as a Feature:* The output that we get after applying our filter to the original image is then passed through another An activation function is a type of mathematical function. Rectified Linear Unit, also known as ReLu, is the activation function that is frequently employed in CNN feature extraction. This maintains the positive values while converting all of the negative values to zero. The convolution must be cleared of all negative values. The negative values change to zero while all the positive values stay the same.
3) *Feature Extraction of Pooling:* Pooling After a convolution layer once you get the feature maps, we need to add a pooling or a sub-sampling layer in CNN layers. Similar to the Convolutional Layer, the Pooling layer is responsible for reducing the spatial size of the Convolved Feature. This is to decrease the computational power required to process the data through dimensionality reduction. Further, it is useful for extracting dominant features which are rotational and positional invariant, thus maintaining the process of effectively training of the model. Pooling shortens the training time and controls overfitting.
4) *Classification of Fully Connected Layer:* We will now flatten our input image into a column vector after converting it into a suitable format for classification. This is known as the Fully Connected Layer (FC Layer). A feedforward neural network receives the flattened output, and back propagation is used for each training cycle. The Softmax Classification approach is used by the model to distinguish between dominant and specific low-level features in images.

IV. CONCLUSION

We presented a Convolutional Neural Network model for students' facial expression recognition. The proposed model includes 4 convolutional layers, 4 max pooling and 2 fully connected layers. The system recognizes faces from students' input images using Haar-like detector and classifies them into seven facial expressions: surprise, fear, disgust, sad, happy, angry and neutral. The proposed model achieved an accuracy rate of 70% on FER 2013 database. Our facial expression recognition system can help the teacher to recognize students' comprehension towards his presentation. Thus, in our future work we will focus on applying Convolutional Neural Network model on 3D students' face image in order to extract their emotions and show emoji.

REFERENCES

[1] R. G. Harper, A. N. Wiens, and J. D. Matarazzo, Nonverbal communication: the state of the art. New York: Wiley, 1978.
[2] P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion," Journal of Personality and Social Psychology, vol. 17, no 2, p. 124-129, 1971.
[3] C. Tang, P. Xu, Z. Luo, G. Zhao, and T. Zou, "Automatic Facial Expression Analysis of Students in Teaching Environments," in Biometric Recognition, vol. 9428, J. Yang, J. Yang, Z. Sun, S. Shan, W. Zheng, et J. Feng, Éd. Cham: Springer International Publishing, 2015, p. 439-447.
[4] A. Savva, V. Stylianou, K. Kyriacou, and F. Domenach, "Recognizing student facial expressions: A web application," in 2018 IEEE Global Engineering Education Conference (EDUCON), Tenerife, 2018, p. 1459-1462.
[5] J. Whitehill, Z. Serpell, Y.-C. Lin, A. Foster, and J. R. Movellan, "The Faces of Engagement: Automatic Recognition of Student Engagementfrom Facial Expressions," IEEE Transactions on Affective Computing, vol. 5, no 1, p. 86-98, janv. 2014

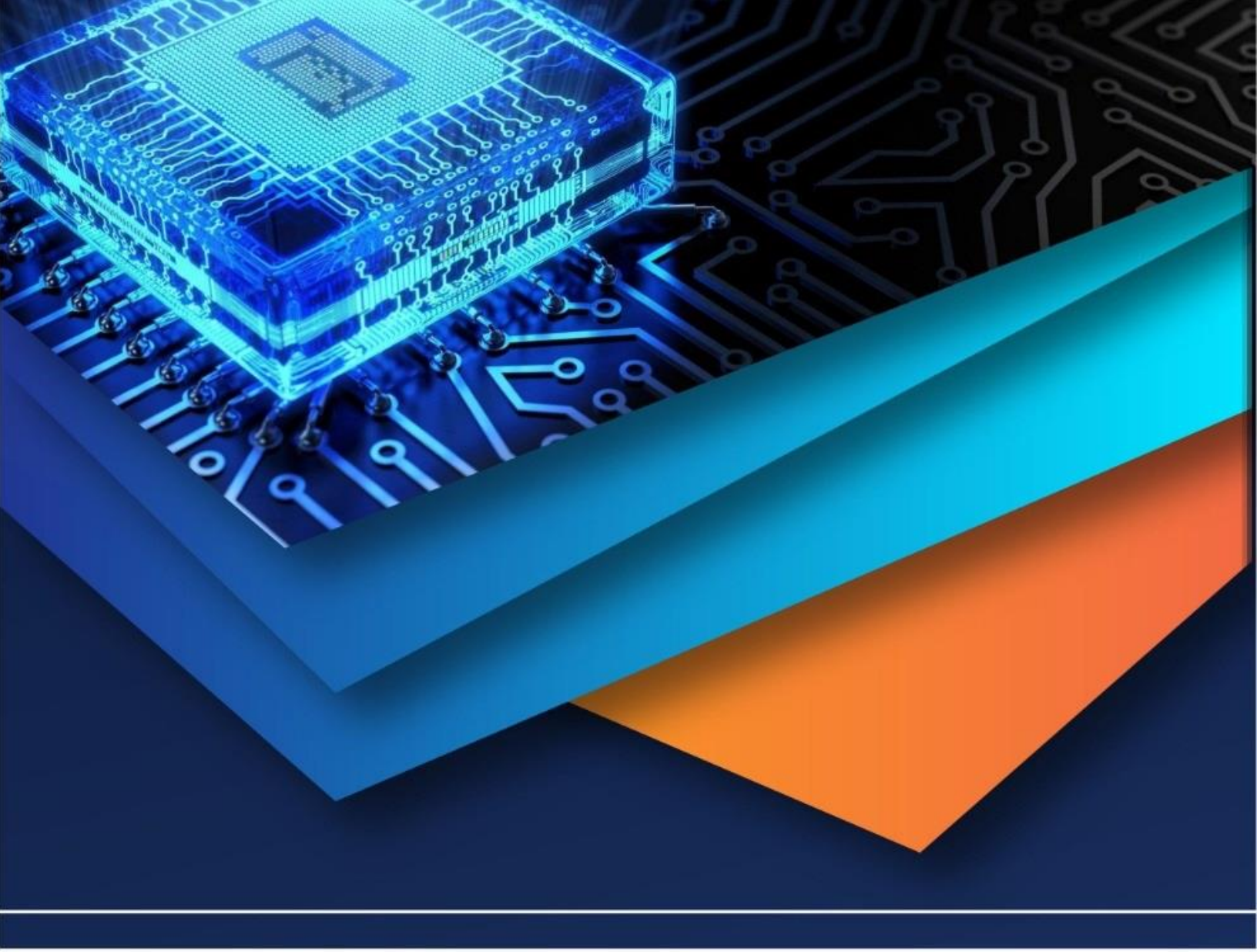
©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

AUTHORS

First Author - Aakula Sai Lahari, Department of Information Technology, Matrusri Engineering College, Telangana, India. (lahariii177@gmail.com).
Second Author - Mula Supraja, Department of Information Technology, Matrusri Engineering College, Telangana, India (mulasupraja13@gmail.com).
Third Author – Cheerneni Akshaya, Department of Information Technology, Matrusri Engineering College, India. (akshayach05@gmail.com).
Correspondence Author - K.Vikram Reddy, Faculty of Information Technology, Matrusri Engineering College, Telangana, India. (vikramreddy@matrusri.edu.in).

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)