iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** XI **Month of publication:** November 2023

DOI: https://doi.org/10.22214/ijraset.2023.57189

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 XI Nov 2023- Available at www.ijraset.com

Emotion Recognition with CNN

Rishikesh Vagmare
Computer Science and Engineering, (Data Science), Swami Vivekanda Institute of Technology, Secunderabad, Telangana, India

Abstract: Emotion is a subjective phenomenon, utilizing knowledge and science behind tagged data and extracting the components that comprise it has been a difficult challenge. With the advancement of deep learning in computer vision, emotion identification has become a popular research topic. This Project presents feature extraction of facial expressions using a neural network combination for the recognition of various facial emotions (sad, happy, neutral, angry, surprised, fear). Convolution Neural Network has been used to achieve a accuracy of 75%, which have excellent recognition of image features. Haar-Cascade has been used to find the region that contains the face, so the model has to only work with the region which has face.
Keywords: Neural Networks, Deep Learning, Convolution Neural Networks (CNN), Haar-Cascade.

I. INTRODUCTION

This Paper helps to accurately recognize an emotion of an individual through his facial expressions. Nonverbal communication is possible through facial expressions. This paper uses convolution neural networks to identify facial features. We will use Haar-Cascade to detect the face in a region which is the input to the convolution neutral networks which will then classify the expressions shown by the individual.

A. Machine Learning

Machine Learning is the technique which allows machines to learn in a way as learnt by humans. Machine Learning uses various algorithms, and these algorithms are divided into three categories namely supervised learning, unsupervised learning and reinforcement learning. In this project supervised learning is used.

B. Supervised Learning

Supervised Learning is the process in which machines are trained with labelled training data which is given as input. This training data contains both the input and its corresponding output. CNN is a supervised learning algorithm that is mainly used for image recognition. CNNs are trained by using large datasets which contain labeled images, during the training phase the network learns to recognize features and patterns which are associated with corresponding classes or objects. Once the CNN is trained it will be able to classify new images, and extract features.

C. Objective

The objectives of this paper is:

- To effectively and efficiently recognize emotions through facial expressions.
- To make use of CNN which helps in classifying the extracted features accurately.

D. Scope

The scope of this paper is given below:

- To recognize emotions accurately based on training data.
- To help various departments and sectors to implement facial recognition to extract emotions.
- This project helps in recognizing facial emotions of criminals, Drivers.

II.SYSTEM ANALYSIS

A. Problem Definition

Emotions can be recognised based on brain activity, brain activity of an individual can be analysed by using Electroencephalography (EEG). We cannot accurately determine emotions solely based on the EEG signals, it depends upon other factors. Various complex electronic devices are used in EEG which are expensive and need specialised crew to make it function properly. Signal receptors are placed on individuals head to transfer brain activity signals to the EEG monitor which analyses the signals. These signal receptors restrict the movement and are prone to electronic interference and disturbance which may effect the accuracy of signals.

B. Problem Analysis

We are using facial expressions which are fed to CNN as input and we obtain features (Emotions) as output. This can be performed on a computer have a webcam without requiring any specialised hardware with accuracy almost similar to the expensive brain activity detectors. It works on TensorFlow framework which allows the program to be platform independent and can be made to work on various platforms such as web apps, android apps, iOS apps.

C. System Requirements

1) Hardware Requirements

- Processor : Intel core i3
- Speed : 2.70 GHz
- RAM : 4GB (minimum)
- Hard Disk Space : 32GB minimum)

2) Software Requirements

- Operating System : Windows 10
- Technology : Python 3.6 version
- IDE : PyCharm

III.SYSTEM DESIGN

A. Modules Used

The system contains three modules:

1) TensorFlow
2) OpenCV
3) Numpy

B. Modules description

1) TensorFlow: Tensorflow is an open-source free software library used for deep learning applications. It is also used in machine learning applications such as neural networks. It allows a program to be deployed across variety of platforms independent of what type of architecture they use. The name TensorFlow is derived from operations that are performed by neural networks on multidimensional data arrays, which are referred to as tensors. A tensor is the data unit, and it consists data in the form of arrays which may be of any dimension. Rank of a tensor can be determined by its number of dimensions.

2) OpenCV: OpenCV stands for Open Source Computer Vision Library. It was built to provide infrastructure for computer vision applications. It is a collection of various types of algorithms which are mainly used for recognition of faces and objects.

3) NumPy: NumPy is a package which performs all scientific computing operations. It offers several tools for performing operations on arrays.

C. Use-Case Diagram

Use-Case diagram represents the behaviour of system. It provides information about the system functionality by adding use cases, actors, and their relationships.

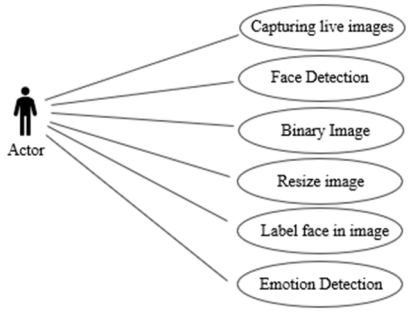


Figure 1

D. Methodology

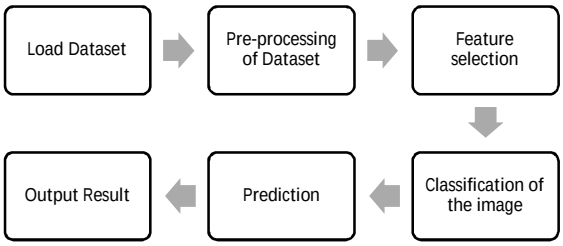


Figure 2

The steps involved in recognizing emotion are

1) Select and load the appropriate Dataset
2) Pre-processing of the Dataset
3) Feature selection
4) Classification of the Image
5) Emotion prediction of the input image
6) Result Generation

E. Implementation

Dataset which is used in this paper is taken from Kaggle. It has various face images of various facial emotions such as sad, happy, surprised, angry, Neutral, Fear. The dataset has thousands of images which will helps us to train the model efficiently. The images present in the Dataset is as follows.

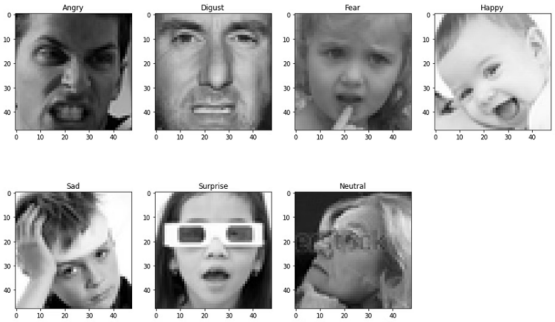


Figure 3

An image is basically a combination of three colour planes Red, Green and Blue. A 4x4 RGB image in matrix representation is as follows.

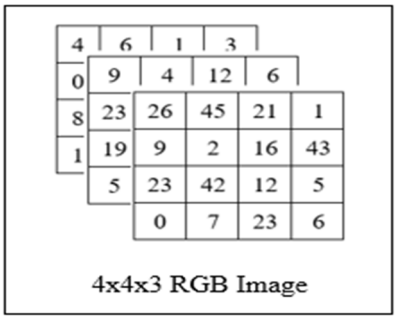


Figure 4

As shown above, the image has width of 4 Units (Pixels) and Height of 4 Units (Pixels) and has 3 colour channels.

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

An image is converted into matrix representation by using OpenCV and Numpy. We can transform an image into various formats such as Grayscale, HSV, CMYK, etc. This is possible when the dimensions of the image are small. But, when the dimensions of the images are huge consider 8k (7680x4320) it will be difficult for the computer to represent it in matrix form and it is an high overhead process. Hence, we use CNN whose role is to reduce the images to a form which is easy to process, without losing features of the image. This can be done by the process called as convolution.

A. Convolution Layer

Convolution is the process in which the image is transformed by applying a kernel over each pixel and its local neighbours across the entire image.

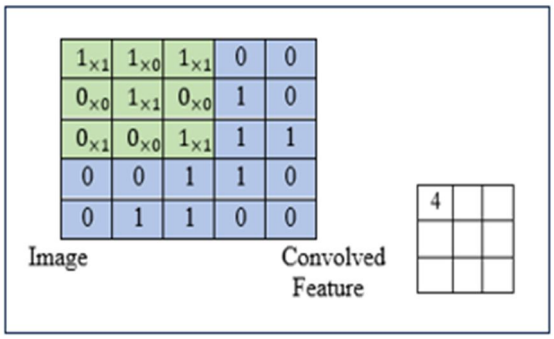


Figure 5

In the above a 5x5x1 image is convoluted with a 3x3x1 kernel to get a 3x3x1 convolved feature.
In the above figure, the blue represents or 5x5x1 input image. The element responsible for convolution operation is called as Kernel/Filter, represented in green colour. We have selected our kernel as 3x3x1 matrix and is represented using green colour.

$$\text{Kernel/Filter K} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Same Padding: 5x5x1 image is padded with 0s to create a 6x6x1 image. When we convert a 5x5x1 image into a 6x6x1 image and then apply 3x3x1 kernel over it, we find that the convolved matrix obtained is of dimensions 5x5x1. If we perform the same operation without passing, we obtain a matrix which has the dimensions of the kernel itself, as shown before.

B. Pooling Layer

It is used for reducing the size of the obtained convolved feature. This is done to decrease the power required to perform computations. This layer is also useful for extracting features which is used for training the model effectively. Max Pooling and Average Pooling are two possible types of pooling.
Max Pooling returns the maximum value and Average pooling returns the average of all values. This can be illustrated as follows

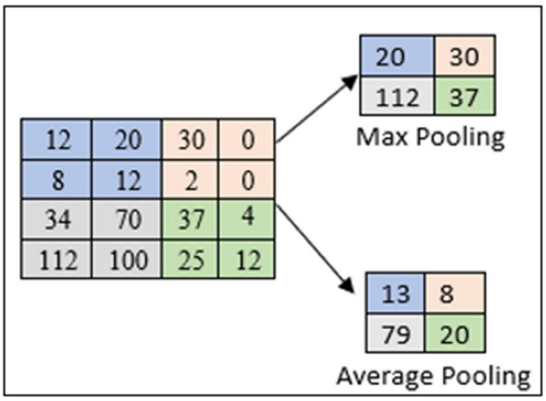


Figure 6

The convolutional layer and the pooling layer, both together form the i-th layer of CNN. The number of layers may be increased depending on the image complexities.

By performing the above process, we have successfully allowed the model to recognize facial expressions and extract features from it. Final step is to flatten the output and give it as input to a regular neural network for classification purposes. The image is flattened into a column vector and is feed as input to a feed-forward neural network and back propagation technique is applied to every iteration of training. After training the model with thousands of images, it is able to classify the image which is given as input from a webcam by using Softmax Classification technique and is able to distinguish among various features of the image and it is able to output the expected results.

IV. RESULTS AND OUTPUTS

Figure 7: (a)

Figure 7: (b)

Figure 7: (c)

Figure 7: (d)

Figure 7: (e)

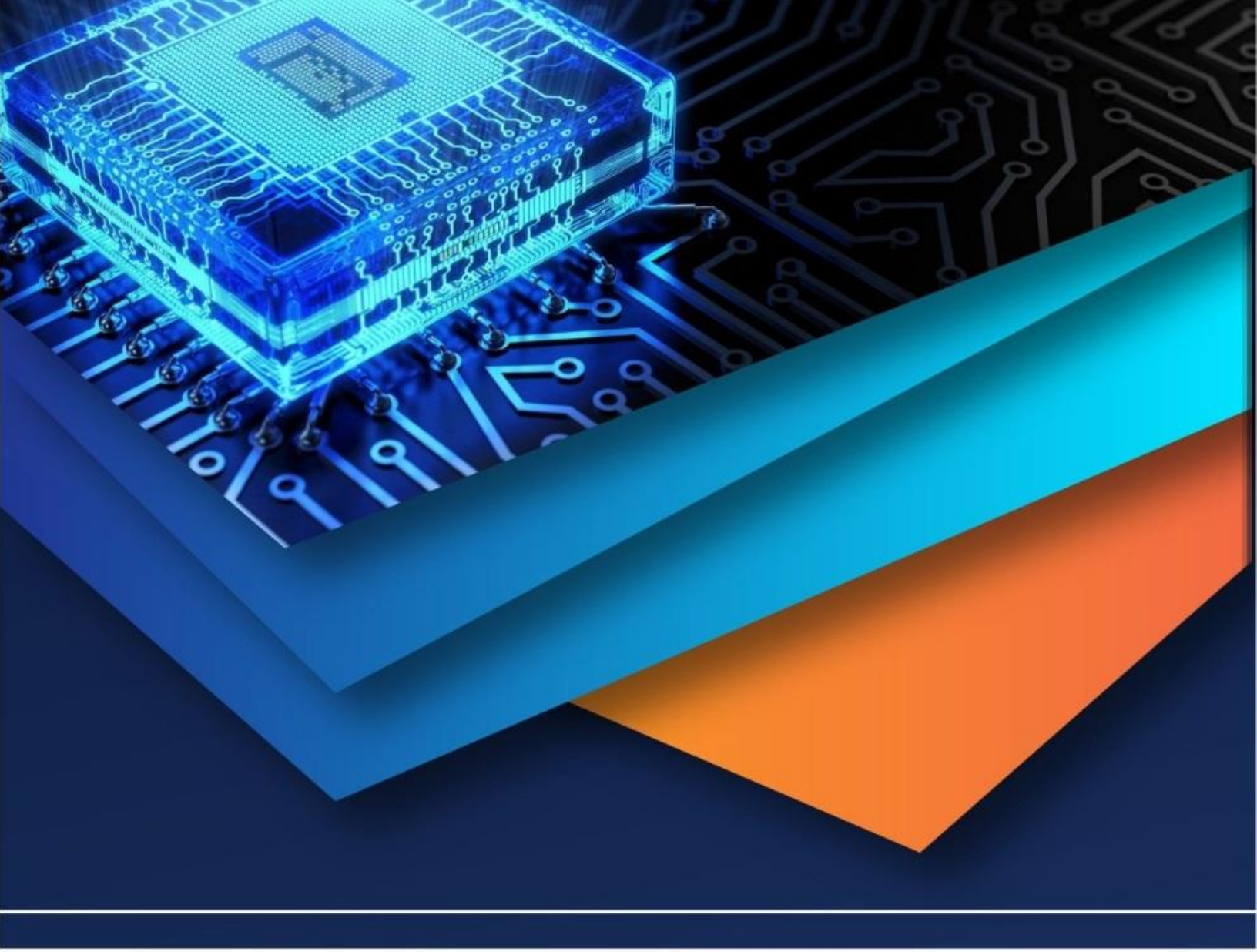
Figure 7: (f)

V. CONCLUSION

Emotion Recognition can have various applications in various sectors and departments. In this CNN were used to perform emotion detection. Using this we have achieved an accuracy about 75%. This can be implemented by anyone having a PC and a webcam without the need of any specialised hardware. This works on TensorFlow framework which allows the program to be ported to different platforms and applications. In future further study can be made in the direction of allele of gene matching to the geometric factors of the facial expressions. The genetic evolution can also be studied using facial expression recognition.

REFERENCES

[1] Andre Teixeira Lopes et al. “A Facial Expression Recognition with Convolutional Neural Networks”, Vol. 00, pg. 273-280,2015.
[2] Siyue Xie and Haifeng Hu, “Facial expression recognition with FRR – CNN”, Electronic Letters, 2017, Vol.53 (4), pg. 235-237.
[3] Razavian et al. “CNN Features off-the-shelf: an Astounding Baseline for Recognition”, arXiv: 1403.6382v3[cs.CV], May 2014.
[4] Francois Chollet, Keras, GitHub, https://github.com/fchollet/keras, 2015.
[5] Alex Krizhevsky et al. “ImageNet Classification with Deep Convolutional Neural Networks,” Neural Information Processing Systems (NIPS), 2012.
[6] Nima Mousavi et al. “Understanding how deep neural networks learn face expressions”, International Joint Conference on Neural Networks, July 2016.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY