IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 4, April 2022

An Analysis on Sentiment identification of Images Using Deep Learning

U D Prasan[1*], P Ramana[2]

[1]Department of Computer Science and Engineering, Aditya Institute of Technology and Management, Tekkali - 532201, India.

[2]Dept of CSE, SRK Institute of Technology

*** Corresponding author: udprasanna@gmail.com**

Abstract:

Human emotion recognition plays an important role in the interpersonal relationship. Emotions are reflected from speech, hand and gestures of the body and through facial expressions. Hence extracting and understanding of emotion has a high importance of the interaction between human and machine communication. He/She expressions are detected and then the dataset is compared to give the desired output on a particular scale. The main objective of this project is to develop the emotion recognition system from facial Expressions. Facial emotion recognition is the process of detecting human emotions from facial expressions. The human brain recognizes emotions automatically, and software has now been developed that can recognize emotions as well. The emotions can be classified into 7 classes — happy, sad, fear, disgust, angry, neutral and surprise. The programming for this will be done in Open CV and Keras using the Haar Cascade library for the detection of facial features.

Keywords: Emotion recognition, Machine communication, Open CV, Keras, Haar Cascade library.

Introduction

Humans interact with each other mainly through speech, but also through body gestures, to emphasize certain parts of their speech and to display emotions. One of the important ways humans display emotions is through facial expressions which are a very important part of communication. There have been several advances in the past few years in terms of face detection, feature extraction mechanisms and the techniques used for expression classification, but development of an automated system that accomplishes this task is difficult. We are Using an approach based on Convolutional Neural Networks (CNN) for facial expression recognition. The input into our system is an image; then, we use CNN to predict the facial expression label which should be one these labels: anger, happiness, fear, sadness, disgust and neutral.

With the development of artificial intelligence and deep learning, numerous FER algorithms have been proposed to deal with the expression information in facial representations, which has improved the accuracy of recognition gradually and achieved better performance than traditional methods. The tasks of FER can be mainly divided into two categories: static images (represented by photographs) and dynamic sequence (represented by videos) that take the dynamic relationship between the continuously changing images into account and therefore pose additional challenges than the former. In addition to the vision-based methods, other biometric techniques can also be adopted to assist the recognition of expression.

Structure of Facial Expression Analysis System

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 4, April 2022

Facial expression analysis includes both measurement of facial motion and recognition of expression. The general approach to automatic facial expression analysis consists of three steps (Fig. 1): face acquisition, facial data extraction and representation, and facial expression recognition. Face acquisition is a processing stage to automatically find the face region for the input images or sequences. It can be a detector to detect faces for each frame or just detect faces in the first frame and then track the face in the remainder of the video sequence. Depending on the different facial feature extraction methods, the effects of in-plane head rotation and different scales of the faces can be eliminated by face normalization before the feature extraction or by feature representation before the step of expression recognition.

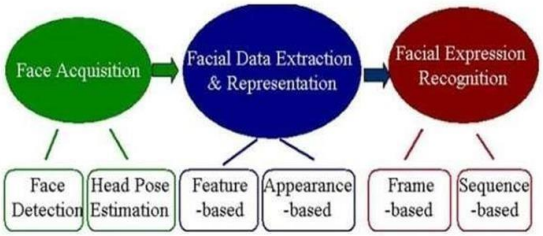


Fig 1: Automatic facial expression analysis

Literature Survey

Recently, Many schools around the world have delayed opening and adopted online education as one of the main teaching methods. However, the efficiency of online classes has long been questioned.

First, the cameras built in the electronic devices are utilized to capture the facial images.Second, the facial expression recognition algorithm trained by the standard facial expression database is employed to detect the faces and classify the facial expressions. Rectified Linear Unit (ReLUs are adopted as the activation function. The CNN consists of two components for Feature extraction and classification. In addition, with a large number of participants in the online courses, we have no way to ensure that everyone keeps the high level of concentration, and then students' expressions may not fully represent their emotions due to these subjective factors. Taking measures like setting thresholds can filter out some invalid information and highlight the main emotions in the image.

The Features which are manually extracted from a given Image. A facial image consists of feature points which represents all the facial parts(such as nose, eyes and lips etc.,). These feature points are then forwarded to a classifier to train the model. Now, The trained model will be used to predict the class labels.

Proposed System:

Haar Cascade face detection algorithm is used to detect the face in the images and given as input to Haar Cascade Face detection algorithm. Once the face is detected, Haar Cascade Face detection algorithm is used to extract the Face region from the facial images and given as input to CNN. CNN with eight convolutional layers are used to extract the deep features and those features are passed to fully connected layer. Soft max layer in CNN classify the images and to identify the emotions of a person. The main objective of this project is to develop the emotion recognition system from facial Expressions. Facial emotion recognition is the process of detecting human emotions from facial expressions. The human brain recognizes emotions automatically, and software has now been developed that canrecognize emotions as well. The emotions can be classified into 7 Emotions — happy, sad, fear, disgust, angry, neutral and surprise. The programming for this will be done in

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 4, April 2022

Open CV, Keras and TensorFlow using the Haar Cascade library for the detection of facial features of the given input image . The proposed architecture of Emotion Recognition system From Facial Expressions is given in Figure 2.

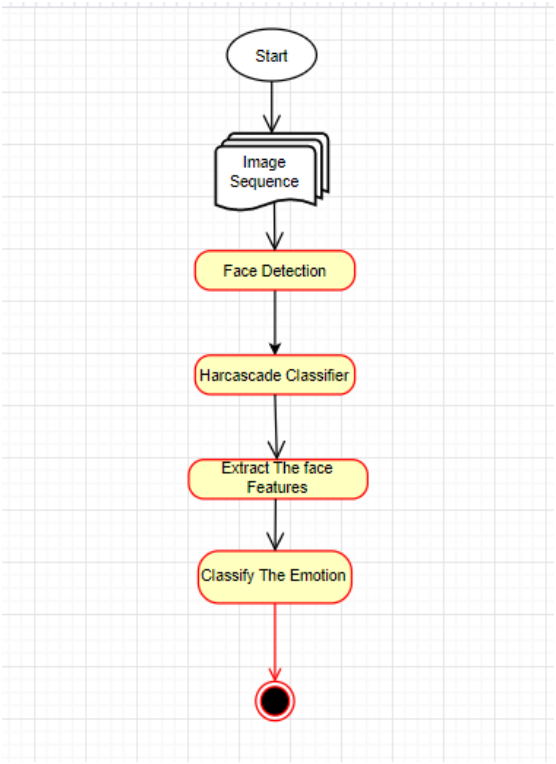


Fig 2: System Architecture

Requirements And Technical Description

Python is a high-level, general-purpose and a very popular programming language. Python programming language (latest Python 3) is being used in web development used in this research.

Keras:

Keras is a deep learning framework. Keras is a central part of the tightly-connected TensorFlow 2.0 ecosystem, covering every step of the machine learning workflow, from data management to hyper parameter training to deployment solutions. Keras is a high-level library intended to stream-line the process of building deep learning networks.

Inside Convnets:

There exists a filter or neuron or kernel which lays over some of the pixels of the input image depending on the dimensions of the Kernel size. The Kernel actually slides over the input image; thus, it is multiplying the values in the filter with the original pixel values of the image.

Kernel:

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 4, April 2022

The kernel is nothing but a filter that is used to extract the features from the images. The kernel is a matrix that moves over the input data, performs the dot product with the sub region of input data, and gets the output as the matrix of dot products. Kernel moves on the input data by the stride value.

Stride in Convnets:

stride denotes how many steps we are moving in each step in convolution. By default it is one. Stride controls how the filter convolves around the input volume. Stride is normally set in a way so that the output volume is an integer and not a fraction.

Tensor flow:

TensorFlow is a popular framework of machine learning and deep learning. It is a free and opensource library which is released on 9 November 2015 and developed by Google Brain Team. It is entirely based on Python programming language and use for numerical computation and data flow, which makes machine learning faster and easier. TensorFlow can train and run the deep neural networks for image recognition, handwritten digit classification, recurrent neural network, word embedding, natural language processing, video detection, and many more. TensorFlow is run on multiple CPUs or GPUs and also mobile operating systems. The word TensorFlow is made by two words, i.e., Tensor and Flow Tensor is a multidimensional array and Flow is used to define the flow of data in operation.

Methodology

Image Acquisition:

Image acquisition is the creation of a digitally encoded representation of the visual characteristics of an object, such as a physical scene or the interior structure of an object. The general aim of Image Acquisition is to transform an optical image (Real World Data) into an array of numerical data which could be later manipulated on a computer. Images used for facial expression recognition are static images or image sequences.

Face Detection:

Generally, Face Detection can be carried out Haar Cascade Classifiers using OpenCv and Convolution Neural Networks.

Haar Cascade Classifiers:

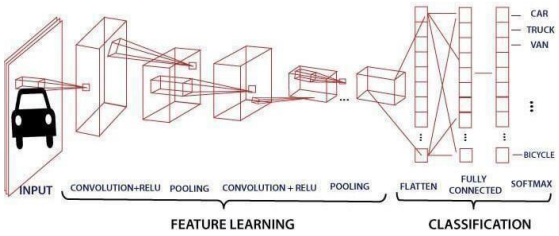
There are some common features that we find on most common human faces:
A dark eye region compared to upper-cheeks.A bright nose bridge region compared to the eyes some specific location of eyes, mouth, nose.

Convolution Neural Network

It is one of the main categories to do image classification and image recognition in neural networks. Scene labelling, objects detections, and face recognition. are some of the areas where convolutional neural networks are widelyused.

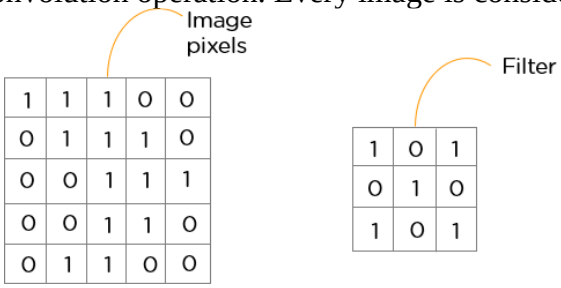
ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 4, April 2022



CONVOLUTION LAYER

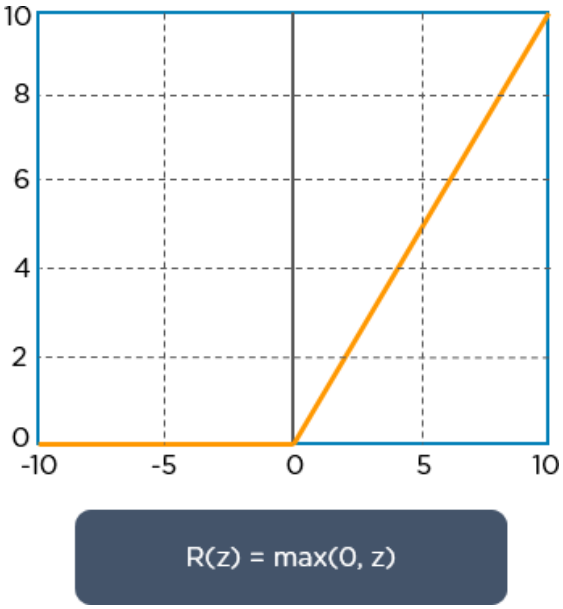
This is the first step in the process of extracting valuable features from an image. A convolution layer has several filters that perform the convolution operation. Every image is considered as a matrix of pixel values.



ReLU LAYER

ReLU stands for the rectified linear unit. Once the feature maps are extracted, the next step is to move them to a ReLU layer.

ReLU performs an element-wise operation and sets all the negative pixels to 0. It introduces non-linearity to the network, and the generated output is a rectified feature map. Below is the graph of a ReLU function.



Pooling

It progressively reduces the size of the input representation. It makes it possible to detect objects in an image no matter where they're located. Pooling helps to reduce the number of required parameters and the amount

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 4, April 2022

of computation required. It also helps control overfitting. If pooling is not done periodically then the size of output will be increased exponentially. There are two types of poolings that can be applied in convnets. They are Global Average pooling, Max pooling. In global average pooling, the given matrix will be replaced by its average and in Max pooling It will be replaced by the maximum value.

FULLY CONNECTED LAYER

A fully connected layer that utilizes the output from the convolution process and predicts the class of the image based on the features extracted in previous stages.The fully connected layer is a layer in which the input from the other layers will be flattened into a vector and sent. It will transform the output into the desired number of classes by the network.

Softmax Layer:

A Softmax function is a type of squashing function. Squashing functions limit the output of the function into the range 0 to 1. This allows the output to be interpreted directly as a probability. Similarly, SoftMax functions are multi-class sigmoid, meaning they are used in determining probability of multiple classes at once. Since the outputs of a SoftMax function can be interpreted as a probability (i.e., They must sum to 1), a SoftMax layer is typically the final layer used in neural network functions.

Datasets

FER Datasets:

A facial expression database is a collection of images or video clips with facial expressions of a range of emotions. Well-annotated (emotion-tagged) media content of facial behavior is essential for training, testing, and validation of algorithms for the development of expression recognition systems. The emotion annotation can be done in discrete emotion labels or on a continuous scale.

Extended Cohn-Kanade Dataset (CK+)

The Extended Cohn-Kanade Dataset (CK+) is a public benchmark dataset for action units and emotion recognition. The dataset comprises a total of 5,876 labelled images of 123 individuals, where the sequences range from neutral to peak expression. Images in the CK+ dataset are all posed with similar backgrounds, mostly grayscale, and 640×490 pixels. The dataset contains 593 video sequences from a total of 123 different subjects, ranging from 18 to 50 years of age with a variety of genders and heritage.

FER-2013

The FER-2013 dataset consists of 28,000 labelled images in the training set, 3,500 labelled images in the development set, and 3,500 images in the test set. The dataset was created by gathering the results of a Google image search of each emotion and synonyms of the emotions. Each image in FER-2013 is labelled as one of seven emotions, such as happy, sad, angry, afraid, surprise, disgust, and neutral, with happy.

Results and Discussions:

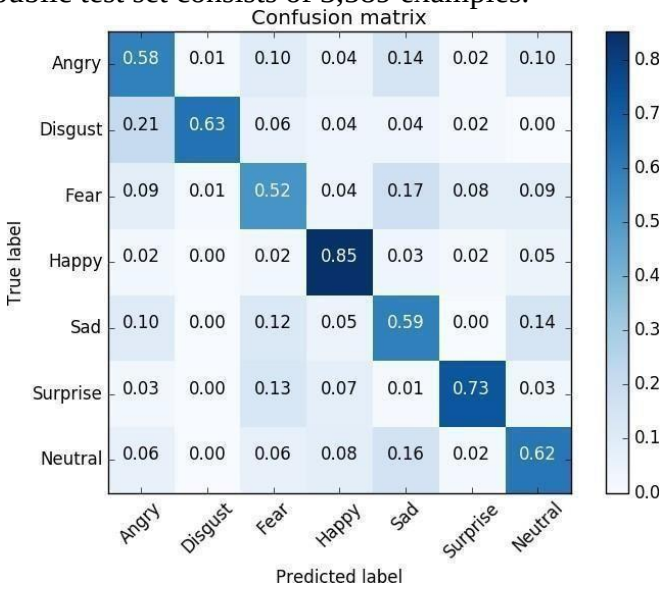
Results obtained for emotions using proposed architecture:

In The dataset the data consists of 48x48 pixel grayscale images of faces. The faces have been automatically registered so that the face is more or less centered and occupies about the same amount of space in each image. The task is to categorize each face based on the emotion shown in the facial expression into one of seven

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 4, April 2022**

categories (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral). The training set consists of 28,709 examples and the public test set consists of 3,589 examples.



See how Anger and Disgust were confused with each other as they are very similar negative emotions. Something similar happened with Fear and Sadness.

Conclusion and Future Scope

In this project, we have proposed an Emotion Recognition of Images Using Deep Learning. The face detection and emotion recognition are very challenging problems. They require a heavy effort for enhancing the performance measure of face detection and emotion recognition. This area of emotion recognition is gaining attention owing to its applications in various domains such as gaming, software engineering, and education. In This We Proposed Using Facial Expressions We Can Classify the Emotion of a Person.

Future Scope:

In future, this work can be extended to propose an application which will be placed at client side to monitor the mood of the learner continuously. Based on the mood of the user the course contents will be developed. This work can also be extended to analyze the mood of the user based on their interaction (such as., Assessments completed, Posted Feedback and active participation) along with the visual features. This may increase the accuracy of the proposed system.

References

[1] J. Shao and Y. Qian, "Three convolutional neural network models for facial expression recognition in the wild," *Neurocomputing*, vol. 355, pp. 82–92, 2019.42

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved,** UGC CARE Listed (Group -I) Journal Volume 11, Iss 4, April 2022

[2] A. Agrawal and N. Mittal, "Using CNN for facial expression recognition: a study of the effects of kernel size and number of filters on accuracy,"26
[3] S. Bayrakdar, D. Akgün, and İ. Yücedağ, "A survey on automatic analysis of facial expressions," *Sakarya University Journal of Science*, vol. 20, no. 2, pp. 383–398, 2016.24
[4] Xiao-Xiao Niu, Ching Y. Suen,A novel hybrid CNN–SVM classifier for recognizing handwritten digits, Pattern Recognition
[5] Pramerdorfer, Christopher &Kampel, Martin. (2016). Facial Expression Recognition using Convolutional Neural Networks: State of the Art.
[6] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, "A survey of affect recognition methods: audio, visual, and spontaneous expressions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 1, pp. 39–58, 2009.
[7] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended Cohn-Kanade dataset (CK+): a complete dataset for action unit and emotion-specified expression," in *Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*, pp. 94–101, San Francisco, CA, USA, July 2010.
[8] MaëlFabien , A guide to Face Detection in Python.
[9] A. Agrawal and N. Mittal, "Using CNN for facial expression recognition: a study of the effects of kernel size and number of filters on accuracy," *The Visual Computer*, vol. 36, no. 2, pp. 405–412, 2020.