IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Facial Emotion Recognition Using Convolutional Neural Networks.

Sumanth Gade[*1], Maneesh Kadavergu[*2], SaiKrishna Mannepalli[*3], A. Kiran kumar[*4] Dr.G.Suwarna Gothane[*5]

[*1]JNTUH, Computer Science and Engineering, CMR Technical Campus, Medchal, Telangana, India

[*2]JNTUH, Computer Science and Engineering, CMR Technical Campus, Medchal, Telangana, India

[*3]JNTUH, Computer Science and Engineering, CMR Technical Campus, Medchal, Telangana, India

[*4]Assistant Professor, Department of Computer Science and Engineering, CMR Technical Campus, Medchal, Telangana, India

[*5]Associate professor, Department of Computer Science and Engineering, CMR Technical Campus, Medchal, Telangana, India

ABSTRACT

Emotion's analysis plays an important role in modern life age. The details about what a consumer is experiencing as a reaction to a product, or to an advertisement, is a very important feature. From response towards advertising to having an improved and customised user interface, needs a solution pertaining to the field of finding the emotional state of the user. This paper aims at predicting the feelings, of a consumer. Without a question, a convolutional neural networks (CNN) model can be prepared to break down pictures and understand face feeling. we make a system that perceives understudies' feelings from their appearances. Our System consists of three phases: face identification using Haar Cascades, normalisation and emotion recognition using CNN with seven types of expressions.

Keywords: Customer facial expression, Convolution Neural Networks, Emotion recognition.

I. INTRODUCTION

The face is the most expressive and communicative aspect of a human being. It's able to express several emotions without saying a sound. Facial expression recognition recognises emotion from face image, it is a representation of the behaviour and personality of an individual. Currently, an instructor uses assessments, questionnaires and observations as sources of input, but these classical approaches also come with low efficiency. Using the facial expression of students, the instructor will adapt their technique and their instructional materials to help promote learning of students.

Facial expression recognition has brought much interest in the past years due to its effects in clinical practise, and education. According to diverse studies, emotion plays an important role in detecting the reaction towards an advertisement or a product.

II. LITERATURE SURVEY

Many researchers are interested in improving the learning environment with Face Emotion Recognition (FER). Tang etal proposed a system which can analyse facial expressions in order to evaluate classroom teaching effect. The system is composed of five phases: data acquisition, face detection, face recognition, face expression recognition and post-preprocessing. The approach uses K-nearest neighbour (KNN) for classification and Uniform Local Garbor Binary Pattern Histogram Sequence (ULGBPHS) for pattern analysis of student's emotions who participating in active face-to-face classroom instruction. The application uses webcams that are installed in classrooms to collect live recordings. In Whitehill et al. suggested an approach that considers commitment from student's facial expressions. The method uses Gabor features and SVM algorithm to define interaction as students interacted with cognitive skills training tools. The authors collected labels from videos annotated by human judges. Then, the writers in used annotated by actual judges. Then, the authors in used computer vision and machine learning techniques to describe the effect of students in a school computer laboratory, where the students were engaging with an educational game intended to demonstrate fundamental concepts of classical mechanics. In the authors proposed a framework that detects and tracks student's emotion and gives feedback in real-time in order to enhance the e-learning environment for a greater content delivery. The system uses moving pattern of eyes and head to deduce relevant information to understand student's mood in an e-learning environment. Ayyaz Etal developed a Facial Emotion Recognition System (FERS), which recognises the emotional statics and motivation of students in video conference style e-learning. The method uses 4 machine learning algorithms (SVM, KNN, Random Forest and Classification & Regression Trees) and the best accuracy rates.

III. PROPOSED SYSTEM

In this section, we explain our proposed framework to analyse facial expressions using a Convolutional Neural Network (CNN) architecture. First, the device detects the face from input image and these detected faces are cropped and scaled to a scale of 48x48. Then, these face images are used as input to CNN. Finally, the performance is the facial expression recognition results (anger, happiness, sadness, disgust, surprise or neutral). Figure 1 Represents the framework of our proposed solution.

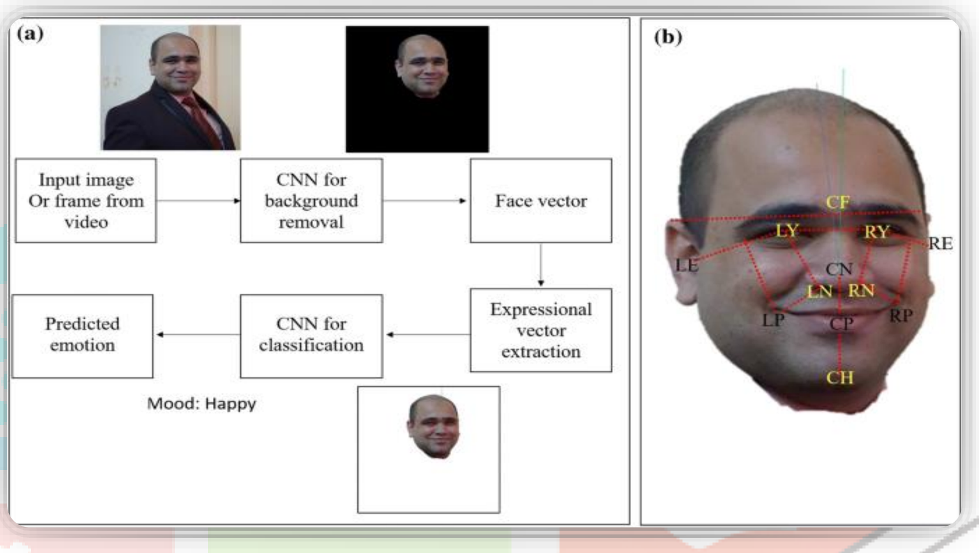


Fig: Structure of Facial recognition System.

A Convolutional Neural Network is a deep artificial neural network that can classify visual patterns from input image with minimal pre-processing compared to other image classification algorithms. This means that the network learns the filters that in conventional algorithms were hand-engineered. The critical unit within a CNN is a neuron. They are bound together, in order that the output of the neurons at the next layer.

In order to compute the partial derivatives of the cost function the backpropagation algorithm is used. The term Convolutional refers to the use of a filter or kemel on the input image to generate a function map. In reality, CNN model contains 3 types of layers as shown in fig.

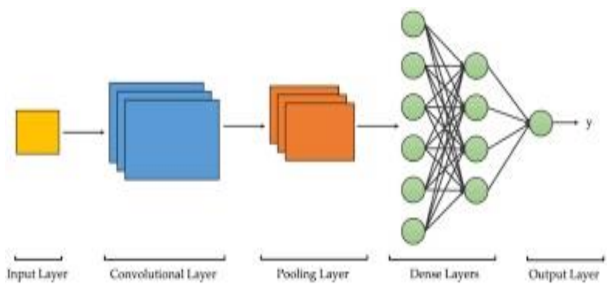
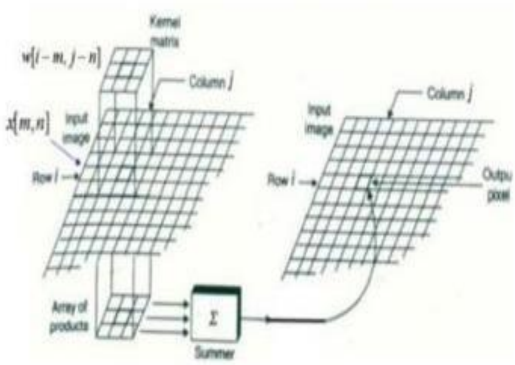


Fig: CNN Architecture

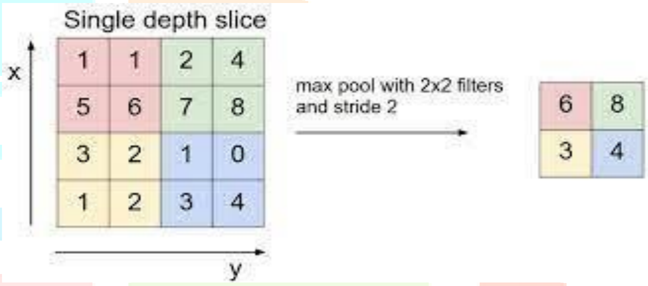
Convolutional Layer: The CNN layer is the first to extract features from an image input. In the case of ConvNet, the primary goal of Convolution is to extract features from the input image. By learning image features using small squares of input data, convolution preserves spatial relationships between pixels. It does a dot product on two matrices, one of which is the image and the other the kernel. Equation 1 represents the convolutional formula.

$$\text{net}(t, f) = (x * w)[t, f] = \sum^{m}\sum^{n} x[m, n]w[t-m, f-n] \quad (1)$$

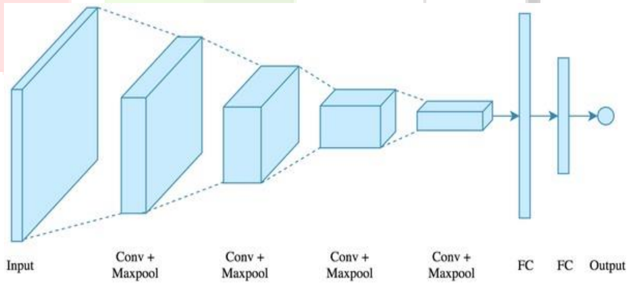
Where net(t, f) is the output in the next layer, x is the input image, w is the filter matrix and $*$ is the convolution operation. Figure 3, shows how the convolution works.



Pooling Layer: Each function map's dimensionality is reduced by the pooling layer, but the most important information is retained. There are three forms of pooling: maximum pooling, average pooling, and sum pooling. Pooling has the effect of gradually shrinking the input representation's spatial size and making the network insensitive to minor transformations, distortions, and translations in the input picture. In our research, we used the block's limit as the single output to the pooling sheet.



Fully Connected Layer: A typical Multi-Layer Perceptron with an activation feature in the output layer is referred to as a fully connected layer. Every neuron in the previous layer is connected to every neuron in the next layer, according to the word "Fully Connected." The completely connected layer's aim is to use the performance of the convolution and pooling layers to classify input images into different groups based on the training dataset. As a result, the convolution, and pooling layers extract features from the input image, while the Fully Connected layer service as a classifier.



It has four Convolutional layers, four pooling layers for extracting features, two completely linked layers, and finally a soft max layer with seven emotion classes. The input image is a grayscale face image with 48*48-pixel resolution. We used 3*3 filters with a stride of 2 for each convolution sheet. We used max pooling layer and 2*2 kernels with stride 2 for the pooling layers. As a result, we used the rectified Linear unit (ReLU), Which is the most commonly used activation function recently, to integrate minerality into our model.

$$R(z)=\max(0, z)$$

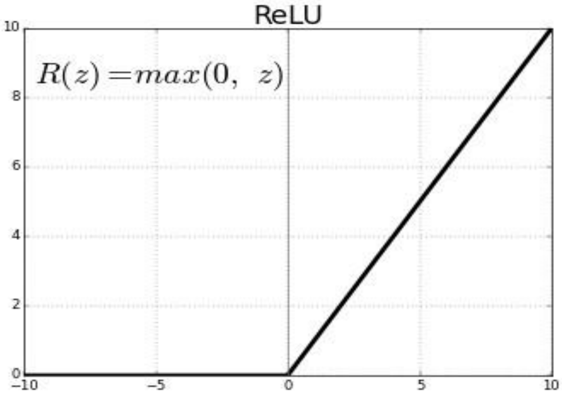


Fig: ReLu Function

TABLE I. CNN CONFIGURATION

Layer type	Size	Stride
Data	48x48	-
Convolution 1	3x3	2
Max Pooling 1	2x2	2
Convolution 2	3x3	2
Max Pooling 2	2x2	2
Convolution 3	3x3	2
Max Pooling 3	2x2	2
Convolution 4	3x3	2
Max Pooling 4	2x2	2
Fully Connected	-	-
Fully Connected	-	-

IV. IMPLEMENTAION AND PROPOSED WORK

Data Acquisition:

It was created with the help of the Google image search API and displayed at the ICML 2013 challenges. Faces in the database have been normalised to 48x48 pixels automatically. The FER2013 database includes 35887 images with seven expression marks (28709 training images, 3589 validation images, and 3589 test images). Table II shows the total number of images for each emotion

Fig: Samples from FER Database.

Table I: Number of Images for each Emotion

Emotion label	Emotion	Number of image
0	Angry	4593
1	Disgust	547
2	Fear	5121
3	Happy	8989
4	Sad	6077
5	Surprise	4002
6	Neutral	6198

We used the OpenCV library to capture live frames from a web camera and to detect students' faces using the Haar Cascades process. Haar Cascades is based on Freund Etal's Adaboost learning algorithm, which won the Godel Prize in 2003. In order to provide an efficient result of classifiers, the Adaboost learning algorithm selected a small number of significant features from a large collection. Using TensorFlow and Keras' high-level API, we built a Convolutional Neural Network model.

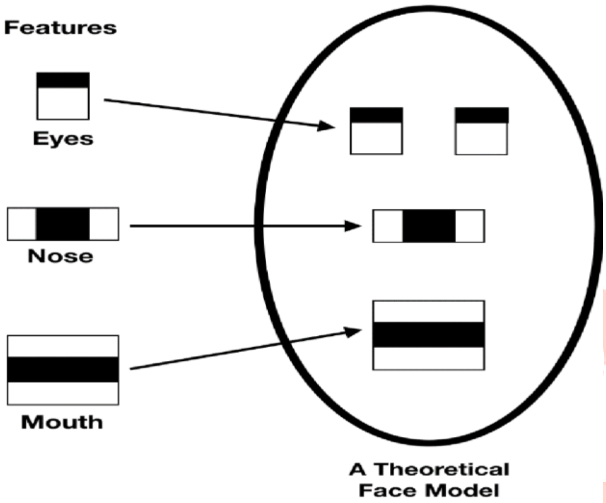


Fig: Face detection using Haar Cascades.

To perform image augmentation in Keras, we used the Image Data Generator class, as shown in Fig.9. We were able to rotate, turn, shear, zoom, and flip the training images. using this class. Rotation shift-10, width shiftrange-0.1, zoom range 0.1, height shift-range-0.1, and horizontal flip-True are the settings used.

Fig: Image augmentation using Keras

Then we added four Convolutional layers to our CNN model. There are four pooling layers in total, in addition to fully connected layers. After that, we used the ReLU function to provide nonlinearity in our CNN model, as well as batch normalisation to normalise the activation of

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{k} e^{z_k}} \; for\; j = 1, \dots, k$$

We split the database into 80 percent training data and 20 percent test data to train our CNN model, and then used the stochastic gradient descent (SGD) optimizer to compile the model. Keras tests whether our model performed better than the models from previous epochs at each epoch. The latest best model weights are saved into a file if this is the case. This allows us to load the weights directly without having to retrain it if we want to use it in a different situation.

V. EXPERMENTAL RESULTS

We used the FER database to train our Convolutional Neural Network model which consists of seven emotions (happiness, anger, sadness, disgust, neutral, fear and surprise). Face images are detected and resized to 48x48 pixels.

The images were then converted to grayscale and used as inputs to the CNN model. Thus, there are six users, two of whom are wearing glasses. Figure 11 depicts the emotions of six members. The expected emotion mark is represented by red text, and the likelihood of the emotion is represented by a red bar.

At the 106 epochs, we had a 70 percent accuracy rate. We measured the confusion matrix, accuracy, recall, and F1- score to assess the efficiency and consistency of our proposed process. Our model is very good in predicting happy and surprised faces.

Confusion matrix of the proposed method. add penalties to the model's various parameters. As a result, we choose SoftMax as our final activation function.it accepts a vector of z K-numbers as input and normalises it into a probability distribution.

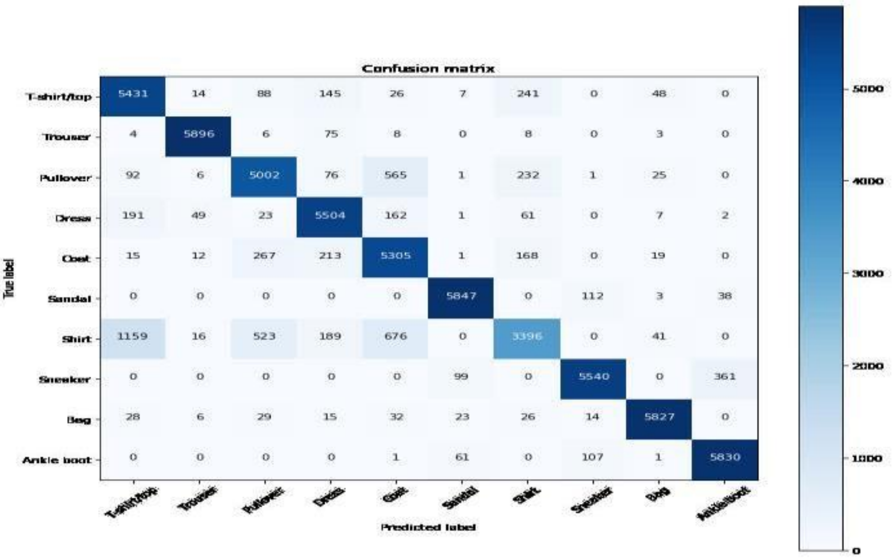


Fig: Confusion Matrix

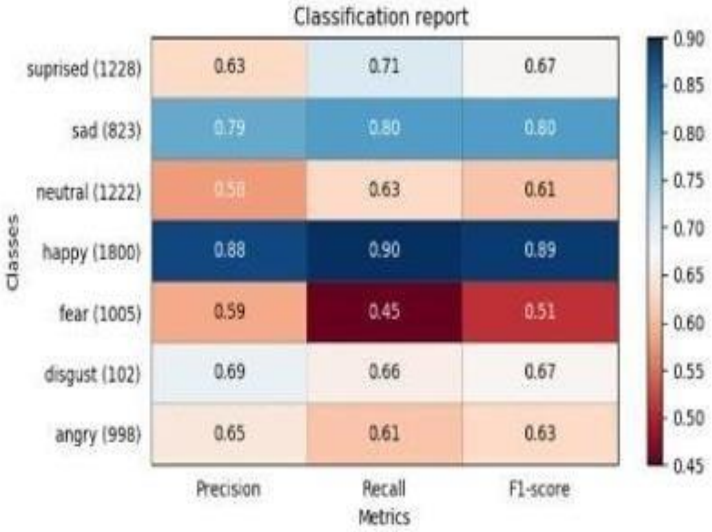


Fig: Classification report of the proposed method.

VI. CONCLUSION

We proposed a facial expression recognition system that extracts facial features effectively using a CNN model. Unlike conventional approaches, the proposed method will learn pattern features automatically and minimise incompleteness induced by artificial design features.

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. The proposed approach uses training sample image data to directly input the image pixel value. Autonomous learning can gain more abstract image function expressions implicitly.

VII. ACKNOWLEDGEMENT

Apart from the efforts of us, the success of paper depends largely on encouragement and guidelines of many others. We take this opportunity to express our profound gratitude to CMR Technical Campus College Management for motivating us and for providing all the facilities required for this work. We are deeply indebted to Chairman Shri C. Gopal Reddy, Secretary Smt. C. Vasantha Latha , Director Dr . A. Raji Reddy, HOD CSE Dr. K. Srujan Raju, PRC Co-ordinator Dr. Suvarna Gothane, Project Guide A. Kiran Kumar who always has been a constant source of inspiration for us.

VIII. REFERENCES

[1] T. Mohana Priya, Dr. M. Punithavalli & Dr. R. Rajesh Kanna, Machine Learning Algorithm for Development of Enhanced Support Vector Machine Technique to Predict Stress, Global Journal of Computer Science and Technology: C Software & Data Engineering, Volume 20, Issue 2, No. 2020, pp 12-20

[2] Ganesh Kumar and P.Vasanth Sena, "Novel Artificial Neural Networks and Logistic Approach for Detecting Credit Card Deceit," International Journal of Computer Science and Network Security, Vol. 15, issue 9, Sep. 2015, pp. 222-234

[3] Gyusoo Kim and Seulgi Lee, "2014 Payment Research", Bank of Korea, Vol. 2015, No. 1, Jan. 2015.

[4] Chengwei Liu, Yixiang Chan, Syed Hasnain Alam Kazmi, Hao Fu, "Financial Fraud Detection Model: Based on Random Forest," International Journal of Economics and Finance, Vol. 7, Issue. 7, pp. 178-188, 2015.

[5] Hitesh D. Bambhava, Prof. Jayeshkumar Pitroda, Prof. Jaydev J. Bhavsar (2013), "A Comparative Study on Bamboo Scaffolding And Metal Scaffolding in Construction Industry Using Statistical Methods", International Journal of Engineering Trends and Technology (IJETT) – Volume 4, Issue 6, June 2013, Pg.2330-2337.

[6] P. Ganesh Prabhu, D. Ambika, "Study on Behaviour of Workers in Construction Industry to Improve Production Efficiency", International Journal of Civil, Structural, Environmental and Infrastructure Engineering Research and Development (IJCSEIERD), Vol. 3, Issue 1, Mar 2013, 59-66