IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 2, June 2023

Light Weight Emotion Recognition System of Facial Images using Convolutional Neural Networks

Mrs. S. Sudha[1] and Dr. R. Karthiyayini[2]

Assistant Professor, Department of IT, Anjalai Ammal Mahalingam Engineering College, Kovilvenni[1]

Assistant Professor Sr. Gr, Department of Computer Applications, BIT Campus, Trichy[2]

Abstract: *Sentiment analysis, also known as sentiment research or sentiment mining, is the process of analyzing consumer generated online content (be it text, images, or video) to determine their mood. Images are a powerful resource where people can share their thoughts and share their experiences via social media. With the increase of social media users who regularly share their memes, videos and pictures, visual content analysis can help users better understand their views on the topic, problem or thing. Real time Emotional Intelligence has been a field of research. in the last few years. Facial expressions convey non-verbal information from facial interactions. The human machine interface plays an important role in automatic face recognition, which has attracted more and more researchers since the early 1990s. Inspired by this challenge, this paper introduces the Light Emotion Recognition (LWER) system using facial images from the FER (Facial Emotion Recognition) dataset provided in the Kaggle Facial Expression Recognition competition. The emotions depicted in the illustrations were divided into five categories: anger, fear, love, sadness, and happiness. Using a deep convolutional neural network (CNN), it recognizes and interprets the semantic information contained in the facial images. The LWER is based on two-part convolutional neural network (CNN): The first-part removes the background from the picture, and the second part concentrates on the facial feature vector extraction. Experimental results show that CNN can be used effectively for image recognition and the proposed method is proven to be efficient for emotion recognition.*

Keywords: convolutional neural network

I. INTRODUCTION

Today, the internet and social media have become a part of people's lives. People share pictures, cartoons, memes and more on social media. This exchange of information gives us a large database of people's thoughts and opinions on a topic. With the popularity of social networks and mobile devices, many photos and videos taken by users are recording their various activities every day and everywhere. Analysis of this theory plays an important role in many business applications such as the stock market, political forecasting, stock analysis, business intelligence, target advertising, e-commerce, and research. Pictures express emotions better than words. We always prefer to express our thoughts with pictures, not words. People from different cultures can easily understand the content of an image or video. Compared to textual opinion analysis, visual analysis is a more difficult task and provides a good insight to show the underlying vision.

Given an image, we try to determine the polarity of the image, whether it is positive, negative or neutral. With CNN, the deep learning framework can learn features from images and categorize them into five emotional categories (anger, fear, love, sadness, and happiness). These groups are available for both the positive and negative poles of the image. For this purpose, we use Progressive CNN (CNN), which learns to use larger data from social media like Twitter, while fine-tunes it through development. The method learns images and selects a sequence and uses this sequence to tune the CNN.

Emotional image analysis is a highly abstract process that deals with emotions conveyed by images and emotions at lower visual levels (eg colour, contour) and bad mood level (eg obsession). To predict the reflection of an image, it is first necessary to remove irrelevant objects from the image and accurately point to the area where the reflection can be captured well. A CNN can be used to generate image vectors for each region of the image, and a collection of these features can be used to train a CNN to analyze visual content. Since there may be different features in the viewer's

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, June 2023

image view, we aim to analyze the representation of different features rather than individual features. The distribution to this perspective is approximated by adjusting the texture and tone of the image. This edited image shows reflections closer than the original image.

Facial expressions are important in human communication and can help us understand the emotions of others. Often people use facial expressions and voice commands to express emotions such as happiness, sadness, and anger. According to various studies, one third of human communication is verbal and two thirds are nonverbal communication. Non-verbal facial expressions, conveying the meaning of the heart, are one of the important tools in personal communication. It is therefore significant that in the last few years the study of facial expressions has received a lot of attention, with its use not only in knowledge and understanding of information, but also in thinking and human Computer interaction.

Most previous work on visual perception analysis has focused on the middle and lower levels of images for emotional recognition. Sentiment classification is done using text definitions such as those found in, SentiWordNet, and WordNet. Previous work in this area includes methods based on maximum relative entropy, binary linear classification, and unsupervised learning. Most of the methods use good features like bag of words, n-grams, tf-idf, ANP, which is considered the simplest. Sentiment analysis is based on handcrafted features and attempts to predict the polarity of images and emotion groups.

II. METHODOLOGY

Light Weight Emotion Recognition System

2.1 Dataset Collection:

FER Kaggle Dataset is used which consists of 28,000 traffic sign images. The dataset is divided into training set and test set. Each sample represents a traffic sign labeled as one of 2 classes. The shape of a traffic sign image is scaled to 256×256 pixels in 3 channel RGB representation.

Figure 1: FER dataset sample

2.2 Preprocessing

Preprocessing data is a common first step in the deep learning workflow to prepare raw data in a format that the network can accept. For example, the image can be resized to match the size of an image input layer. Data can also be preprocessed to enhance desired features or reduce artifacts that can bias the network. Variations that are irrelevant to facial expressions, such as different backgrounds, illuminations and head poses, are fairly common in unconstrained scenarios. Illumination and contrast can vary in different images even from the same person with the same expression, especially in unconstrained environments, which can result in large intra-class variances. Therefore, before training the deep neural network to learn meaningful features, pre-processing is done to align and normalize the visual semantic information conveyed by the facial image.

2.3 CNN Layer Feature Extraction:

Convolutional Neural networks are very successful in image recognition. The key part to understand, which distinguishes CNN from traditional neural networks, is the convolution operation. AlexNet Algorithm is used for Feature Extraction.

Convo2D Layer :

Conv2D is a 2D Convolution Layer, this layer creates a convolution kernel that is wind with layers input which helps to produce a tensor of outputs.

Maxpooling Layer:

Max pooling is a pooling operation that selects the maximum element from the region of the feature map covered by the filter. Thus, the output after max-pooling layer would be a feature map containing the most prominent features of the previousfeature map.

Dense Layer:

Dense Layer is simple layer of neurons in which each neuron receives inputf rom all the neurons of previous layer.

Flatten layer:

Flatten layer is used to make the multidimensional input one-dimensional, commonly used in the transition from the convolution layer to the full connected layer.

Dropout Layer:

Dropout is a technique used to prevent a model from overfitting. Dropout works by randomly setting the outgoing edges of hidden units (neurons that make uphidden layers) to 0 at each update of the training phase.

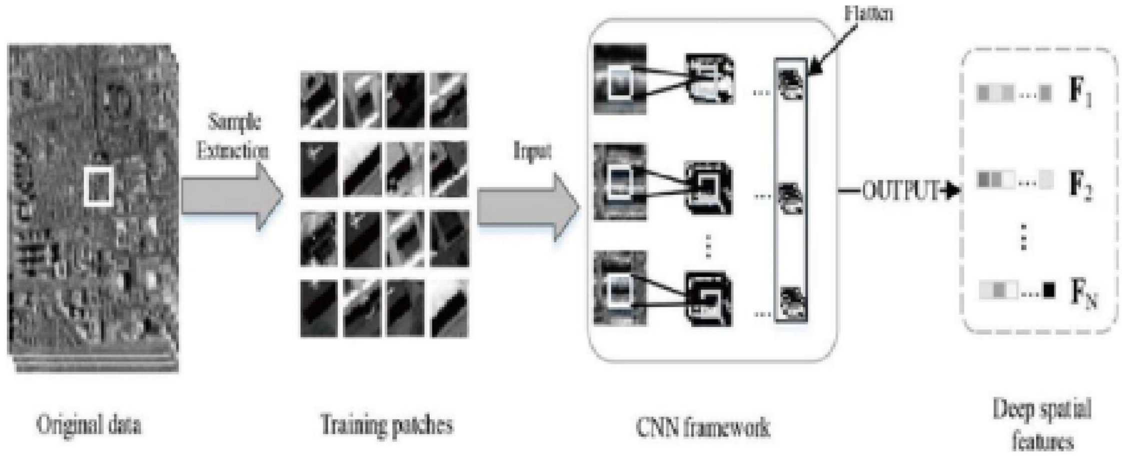

Figure 2 Data Pre-processing

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, June 2023

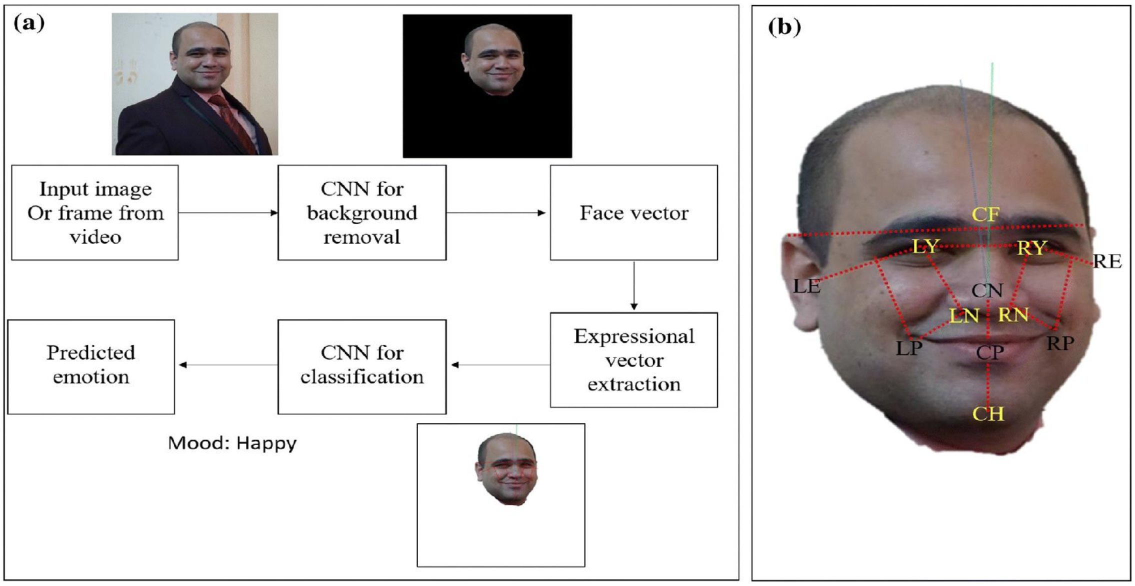

Figure 3a Block Diagram,**3b.** Facial vectors marked on the facial image

2.4 Alex Net :

AlexNet has five layers of CONV, some of which include max pooling layers, three layers, and a final 1000-way softmax. The total number of teach parameters for AlexNet is approximately 60 million. Conversely, LeNet has two layers of coevolution, followed by three full layers and two pooling layers. Total training without LeNet is about 60,000. AlexNet uses the nonlinear function (ReLU), while LeNet uses the sigmoid logistic function. AlexNet uses the last variant, a form of "dropout" constant, to reduce overfitting at all layers. This rule was not used when LeNet was created. Before explaining AlexNet's topology, some terms used in the model should be explained. In the subsections below, the features ofAlexNet are listed in detail.

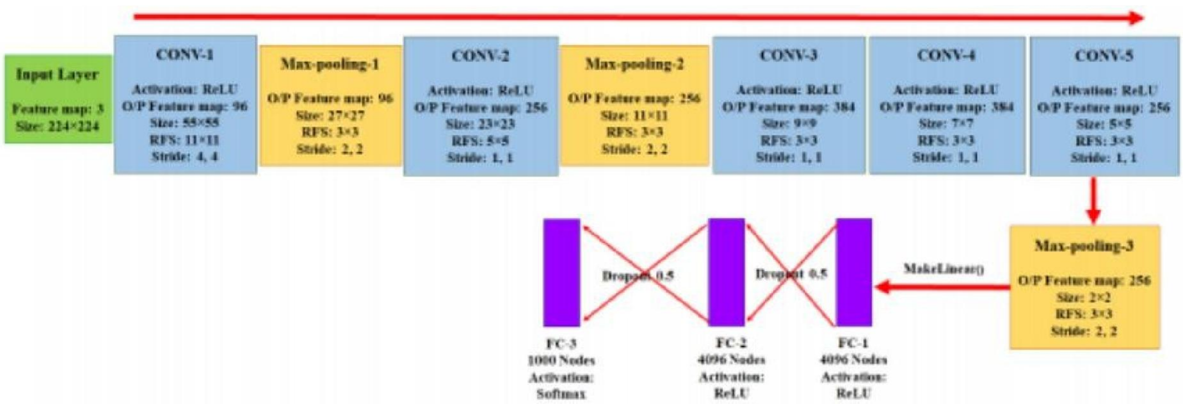

Figure 4 Alexnet Layers

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 2, June 2023

2.5 Results:

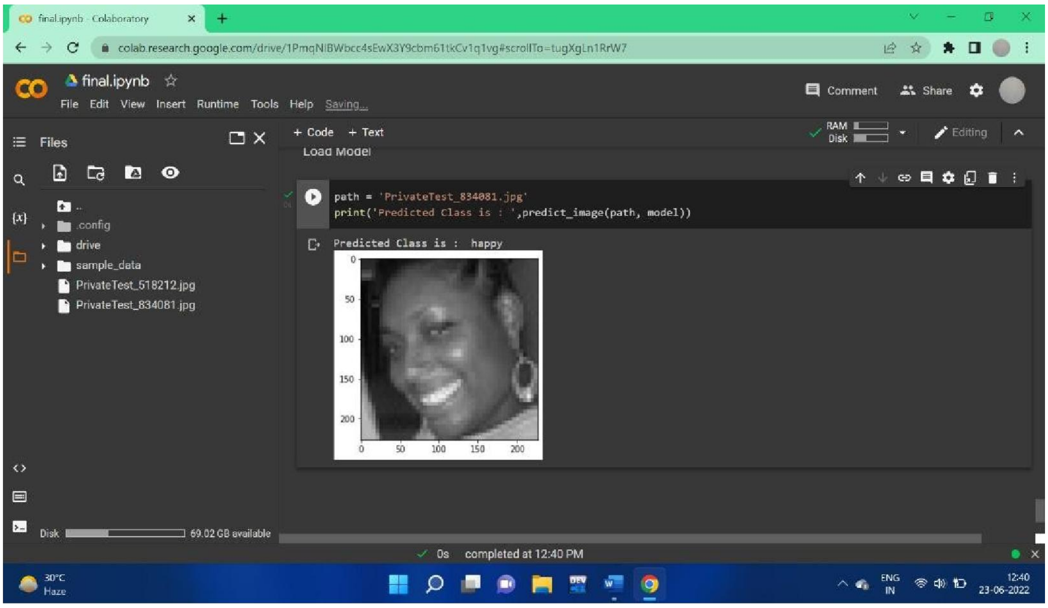

Figure 5: Happy emotion

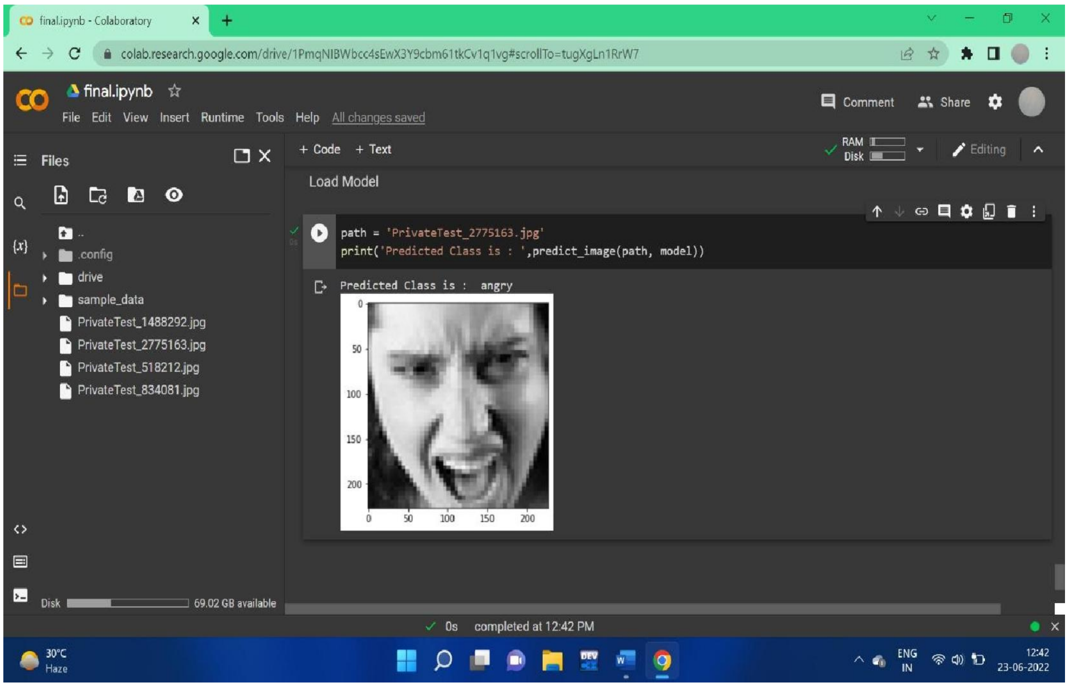

Figure 6: Angry emotion

III. CONCLUSION

Till now the sentiment analysis were mainly focussed on text and visual sentiment analysis is a most challenging and interesting domain where many researchers focus on. Most of the previous work on Visual emotion analysis were based on hand crafted features and try to predict the polarity of images and emotion categories. Motivated by the needs and challenges, we develop a cognitive model that uses a deep convolutional neural network and a good strategy to tune the CNN on a small set of images. This proposed method applies to large-scale content created by customers, and we use advanced learning and transfer learning of pre-trained images with high scores. Experimental results show that the deep

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, June 2023

learning neural network can accurately predict emotions from images and classify them into one of 5 emotional groups. Well-trained CNNs can outperform classifiers for visual recognition analysis tasks that use low- and intermediate-level features of images.

REFERENCES

[1] A.C.RafaelandD.Sidney,IEEETrans.AffectiveComput,vol.1,no.1,pp.18– 34, Jun. 2010. "Affect detection: An interdisciplinary review of models, methods, and theirapplications".
[2] C. Darwin, John Murray, 1872, "The Expression of Emotions in Man and Animals. (reprinted by the University of Chicago Press, Chicago, IL,1965)".
[3] Daniel Georgescu, vol. III, no. 4, 2011 "A Real-Time Face Recognition System Using Eigenfaces, Journal of Mobile Embedded and DistributedSystems".
[4] G. Guo and C. R. Dyer, IEEE Trans. Syst., Man, Cybern. B, vol. 35, no. 3, pp. 477–488, Jun. 2005. "Learning from examples in the small sample case: Face expression recognition."
[5] I. Cohen, N. Sebe, Y. Sun, M. S. Lew, and T. S. Huang, Jul. 2003, pp. 184–195 "Evaluation of expression recognition techniques," in Proc. Int. Conf. Image Video Retrieval".
[6] I.Yugashini, S. Vidhyasri, K.Gayathri Devi, Volume-1, Issue-12, November 2013"DesignAndImplementationOfAutomatedDoorAccsessingSystemWith Face Recognition, International Journal of Science and Modern Engineering(IJISME)".
[7] J. F. Cohn, IEEE Signal Process. Mag., vol. 27, no. 6, pp. 128–133, Nov. 2010 "Advances in behavioral science using automated facial image analysis and synthesis".
[8] I. K. I. Chang, K. W. Bowyer, and P. J. Flynn, IEEE Trans. Pattern Anal. Mach. Intell.,vol.28,no.10,pp.1695–1700,Oct.2006"Multiplenoseregionmatching for 3-D face recognition under varying facialexpression".
[9] I. Liton Chandra Paul, Abdulla Al Sumam. Volume 1, Issue 9, November 2012 "Face Recognition Using Principal Component Analysis Method. International Journal of Advanced Research in Computer Engineering &Technology".
[10] L. D. Silva and S. C. Hui, in Proc. 4th Int. Conf. Inform. Commun. Signal Process, vol. 3. Dec. 2003, pp. 1310–1314"Real-time facial feature extraction and emotionrecognition".
[11] M. J. Lyons, J. Budynek, A. Plante, and S. Akamatsu, Mar. 2000, pp. 202–207 "Classifying facial attributes using a 2-D Gabor wavelet representation and discriminant analysis," in Proc. 4th Int. Conf. Automatic Face Gesture Recognit".
[12] M. Pantic and I. Patras, vol. 36, no. 2, pp. 433–449, Apr. 2006 "Dynamics of facialexpression:Recognitionoffacialactionsandtheirtemporalsegmentsfrom face profile image sequences," IEEE Trans. Syst., Man, Cybern.B".
[13] N. Sebe, H. Aghajan, T. Huang, N. M. Thalmann, and C. Shan, IEEE Trans. Multimedia,vol.12,no.6,pp.477–480,Oct.2010"Specialissueonmultimodal affectiveinteraction".
[14] Ole Helvig Jensen, San Francisco, CA: Consulting Psychologist, 2002, "Implementing the Viola-Jones Face DetectionAlgorithm".
[15] P. Ekman, W. V. Friesen, and J. C. Hager, San Francisco, CA: Consulting Psychologist, 2002"The Facial Action Coding System: A Technique for the Measurement of FacialMovement".