International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 09 | Sept - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Train Your Own Neural Network for Facial Expression Recognition Using TensorFlow, CNN and Keras

Khushali Pandav[1], Neramaye N Deshpande[2]

[1]Student, Department of Computer and Information Science, Nagindas Khandwala Collage , Mumbai, Maharashtra, India
[2]Assistant Professor, Department of Computer and Information Science, Nagindas Khandwala Collage, Mumbai, Maharashtra, India

Abstract

Facial expression recognition is one of the most important aspects of human-computer interaction, based on which the system needs to perceive and respond to the emotions expressed by a human. This paper presents a facial expression recognition system that makes use of the convolutional neural network, using TensorFlow and Keras for its implementation. It categories emotions included in the FER2013 dataset consisting of 48x48 pixel grayscale images labeled with seven emotional categories: Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral. It basically comprises steps like image preprocessing, designing and training a CNN in order to hierarchically extract features from data, and using some metrics such as accuracy, confusion matrix, and classification report regarding model performance. The results showed a high degree of accuracy in classifying emotions, hence proving the effectiveness of deep learning techniques in this domain. The following findings highlight the model applications that have contributed to affective computing, improving customer experience, and interactive systems; these depict a big move in emotion recognition technologies.

Keywords: Facial Expression Recognition, Convolutional Neural Networks, TensorFlow, Keras, FER2013 Dataset, Emotion Classification.

INTRODUCTION

Facial expression recognition has become one of the important modalities of human–computer interaction, providing systems with an understanding and ability to respond to human emotions. In this paper, a robust facial expression recognition system is described for the classification of emotions from grayscale images of human faces using deep learning methodologies. It uses a very famous dataset of emotional images, all of which are 48 x 48 pixels in size and annotated for one of the following seven emotional categories: Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral. The representation above underlines part of the main steps of the proposed methodology that can be consecutively applied. First, preprocessing for the data is done in order to transform the raw pixel values into a format appropriate for training a model. In this respect, a CNN was designed for learning hierarchical features from facial images and classifying them accurately into predefined emotion categories. The performance of the model was tested rigorously with a myriad of metrics to the goals of accuracy and confusion matrix analysis, which proves the reliability and effectiveness of the model. Moreover, distribution across emotions in the dataset has been analyzed and visualized to give insight into its representativeness and balancing. This can contribute to emotion recognition technologies by demonstrating that deep learning performs well in facial expression interpretation and may further be applied to other applications such as affective computing, enhancement of customer experience, and creation of interactive systems.

RESEARCH OBJECTIVES

- Design and implement a Convolutional Neural Network (CNN) using TensorFlow and Keras to classify facial expressions from grayscale images.
- Execute preprocessing steps to convert raw image data into a format suitable for model training and evaluation.
- Train the CNN model on the prepared dataset and validate its performance using appropriate metrics.
- Examine and visualize the distribution of emotional categories within the dataset.

LITERATURE REVIEW

Duncan. D, et. al. the enthusiasm behind selecting this theme, especially the deceptions in giant investments, bulky corporations indulge in responses and reviews but flop to get reasonable reply on their investments. Emotion Detection through facial movements is a technology that intends to advance product and facilities presentation to customer behavior by nursing certain products or staff by their estimation. "Substantial discussion has mounted previously about the emotion represented in the world famous masterpiece of Mona Lisa". British Weekly "New Scientist‟ has drawn an elaborate profile that she is actually an amalgamation of many discrete emotions, "83%pleased, 9% appalled, 6% afraid, 2% mad". We have also been inspired by the benefits of observing physically handicapped people such as the deaf and the dumb but if any normal human being or an automatic system can understand their desires it develops a lot in perceiving their facial expression informal for them to make the parallel human or automatic system appreciate their wishes.

Piatkowska. E, & Furst. J, (2010) primarily, authors like to outline the elementary impression of the FED system and describe the most vital problems which should be taken under attention in the procedure of system strategy and progress. Afterwards, every FED system phase will be defined in detail, namely: main duty, emblematic difficulties and future approaches. Additionally, the fresh advances in the zone of facial emotion the analysis for detection will be itemized. In summary, some models will reveal applications of FED systems, which show that they are extensively used in many arenas of science as well as they do during everyday natural life.

Valenti. R, et. al. (2005) subsequently in the first 1970s there have been widely revisited human facial expressions. Author initiates it as an indication towards providing universality in facial terminologies. These 'universal facial expressions' are those on behalf of happiness, sadness, anger, surprise and neutral.

Khan. R, & Sharif. O, (2017) earlier works focus on the production of outputs from unimodal systems. The machines used to forecast emotion by either only facial emotions or only verbal noises. Some time later, multimodal systems that use more than one feature to predict emotion are more active and bounce more truthful Outcomes. Therefore, a mixture of features like audiovisual terminologies, EEG, body signs has been used. Meanwhile, more than one intellectual machine and neural Emotion detection can be instrumented using a network system.

Zhang. H, et. al. (2019) the measurement of this feature is large before and after the achievement of the feature so the dimensionality reduction is sometimes disbursed. The facial expression classification refers to the utilization of specific algorithms to identify the classes of facial expressions in step together with the extracted features, unlike traditional CNN, the convolution core of CNN diverges thanks to the more discriminant differences between feature maps at the identical level, leading to fewer redundant features and more compressed image representation.

Lekhak. D, (2017) in deep learning, a convolutional neural network (CNN) is a type of feedstuff advancing artificial neural network in which the connectivity design between its nerve cell is enthused by the society of the animal pictorial pallium. Distinct cortical Neurons tend to respond to spurs in a limited region of space known as the interested field as the

approachable fields of diverse Neurons partially intersect such that they are stone. pictorial field. The response of an individual neuron to stimuli within its friendly field can be approached precisely by an Operation of Convolution. Basically, convolutional systems were enthused by biological progressions.

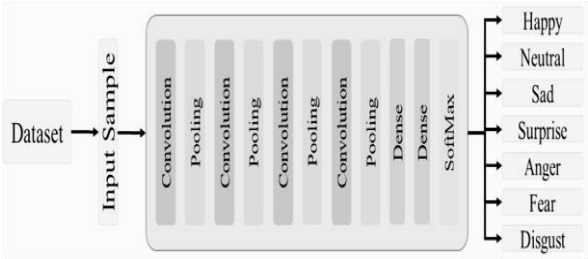
Murugappan. M, et. al. (2020) instinctive human emotion detection has established much courtesy recently with the starter of IOT and smart surroundings at hospitals, smart homes, and smart cities. "Intelligent personal assistants (IPAs), like Siri, Alexia, Cortana and others", exercise natural language processing to connect with humans; when augmented with emotions, it surges up to the neck and neck of effective message and human-level cleverness.

Raval. D, % Sakle. M, (2015) emotion is a very important, very complex and most researched topic in fields of biomedical engineering and psychology, neuroscience and health. Studies in this area focus on how it detects human facial 5 basic emotions. There are diverse methods provided in these works for the detection of emotional states like electroencephalography (EEG), galvanic skin response (GSR), Speech analysis, facial expression, multimodal, visual scanning behavior and so more.

PROPOSED MODEL

In tune with the drawbacks of the Present system, we are going to develop this project. The strategy startswith the preparing CNN model by pleasing an input image (static or dynamic, by stacking a convolution layer, pooling layer, flatten layers, and dense layers. Convolution coats will be added at the end for better accuracy for big datasets. The dataset is collected from the CSV file (in pixel format) and converted into images and then classify emotions with respective expressions. Here emotions are classified as happy, sad, angry, surprise, neutral, disgust, and fear with lots of images for the training dataset and huge images for testing. Each emotion is expressed with different facial features like eyebrows, opening the mouth, Raised cheeks, wrinkles around the nose, wide-open eyelids and many others. So, trained the large dataset for better accuracy and the result is the object class for an input image based on those features. It does the convolution layers and max pooling. In this proposed system, Deep learning is used with the help of Keras, containing several Models. Among these models, the human facial emotion is classified. Here the emotions are categorized as happiness, sadness, anger, surprise, and neutral with plenty of images for the training dataset and huge images for testing. Each emotion is expressed with different facial features such as eyebrows raised, mouth open, cheeks raised, and wrinkles around the nose,wide-open eyelids and many others.Trained the huge dataset for better accuracy and result which is object class for an input image.

METHODOLOGY



1.THE DATASET

FER2013 is the dataset used in training the model. The dataset is from a Kaggle Facial Expression Recognition contest a few years ago. The data consists of 48x48 pixel grayscale pictures of faces. The face features have been mechanized recorded so the face is comparatively positioned and dwells in roughly the same amount of space for each image. The mission is categorizing each face according to the emotion displayed in the facial expression into one of seven categories (1.Angry, 2.Disgust, 3.Fear, 4.Happy, 5.Sad, 6.Surprise, 7.Neutral). Used here, the FER2013 dataset consists of 30,219 images which is further divided into training sets of 24,282 and validation (5937) images respectively.

2. TRAINING NEURAL NETWORK

- A good amount of images with different emotions will be collected.
- These datasets are in training the Convolutional Neural Network(CNN) classifier.

- This would be by feeding the coordinates of features extracted from the images.
- Each emotion would produce specific examples after Training.
- These instances would be used to detect the emotions of facial expression.
- 90% of the images would act as the training dataset.

3. TESTING NEURAL NETWORK

- The trained classifiers are then tested with a dataset of images which were not used for the classifier training.
- Image dataset: 10% would be used as a testing dataset.
- Then used in finding the accuracy of the System

4. DEPENDENCIES

OpenCV : OpenCV (Open Source Computer Vision Library) is an open source computer vision machine learning software library. This is the library that we will use for image conversion functions like changing the image to grayscale.

Keras : Keras is anAPIs offerings to model deep learning model.It is a high-level neural link API, written in python. It is an API designed for human beings, not machines.

TensorFlow : TensorFlow(TF)is an open source software library in Python and C++ for Machine Learning. It was released some months earlier (Nov 15) with a strong press exposure. These include speech recognition in Google Now , search features in Google Images, and the smart reply feature in Gmail's Inbox.

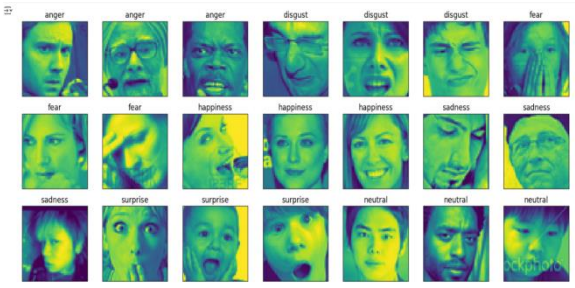
NumPy : NumPy or Numerical Python is an open-source Python library that leads easy to complex numerical operations.

Pandas : Pandas is a Python library, built for Python, which efficiently provides the language with data manipulation and analysis, having core structures like Series and DataFrame. It makes it quite easier to handle and process structured data.

Matplotlib : Matplotlib is a cross-platform Python library for mathematics, mainly used to create static, animated, or interactive visualizations in any format. It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits.

RESULTS

1. Sample Dataset



2. Four Convolution Layer

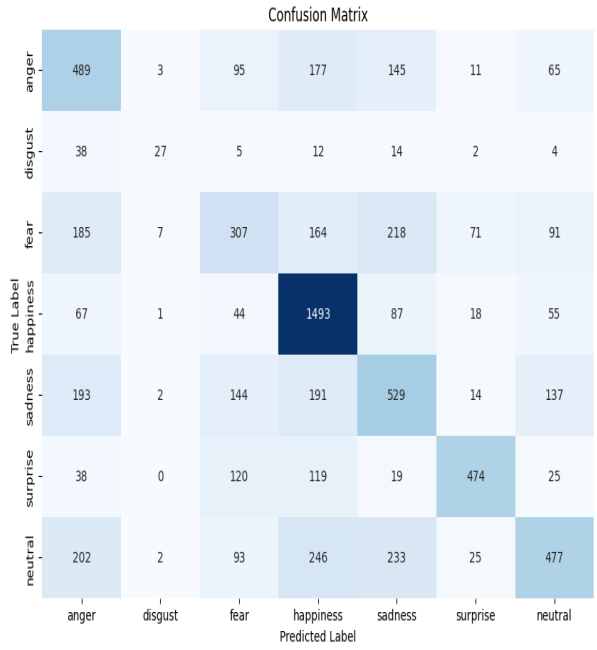
```
basemodel = tf.keras.models.Sequential([tf.keras.layers.Conv2D(32,(3,3),activation='relu',input_shape = (48,48,1)),
                                        tf.keras.layers.MaxPool2D(2,2),
                                        tf.keras.layers.BatchNormalization(),
                                        #
                                        tf.keras.layers.Conv2D(64,(3,3),activation='relu',input_shape = (48,48,1)),
                                        tf.keras.layers.MaxPool2D(2,2),
                                        tf.keras.layers.BatchNormalization(),
                                        #
                                        tf.keras.layers.Conv2D(128,(3,3),activation='relu',input_shape=(48,48,1)),
                                        tf.keras.layers.MaxPool2D(2,2),
                                        tf.keras.layers.BatchNormalization(),
                                        #
                                        tf.keras.layers.Conv2D(256,(3,3),activation='relu',input_shape = (48,48,1)),
                                        tf.keras.layers.MaxPool2D(2,2),
                                        tf.keras.layers.Flatten(),
                                        tf.keras.layers.Dense(1000,activation='relu'),
                                        tf.keras.layers.Dense(7,activation = 'softmax')
                                        ])
```

Facial expression recognition CNN first takes an input image of size 48x48 pixels. It is composed of four convolutional blocks, each comprising a Conv2D layer for feature extraction, a MaxPooling layer for reducing the dimensionality, and BatchNormalization for the purpose of stability. Later on, these features are flattened and passed through a Dense layer consisting of 1000 units with ReLU activation. This last Dense layer, of 7 units with Softmax activation, classifies the images into seven categories of emotion: Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral.

3. Actual label VS Predicted label

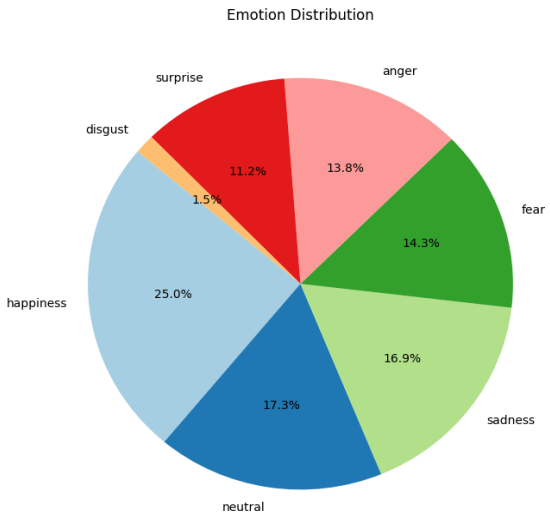
The predictions of a trained TensorFlow/Keras model on test images and loads the model from a checkpoint file. Then, for the first 40 test samples, it prints the actual label of the image, predicts the label using the model, and prints the predicted label. A test image is plotted using Matplotlib, with a 5-second pause between each image to see each one of them. This clears the output at every turn with the latest result of the process at hand. The following way, it will be useful to provide some idea about the result of the model performance by visually comparing the predictions of the model against the true labels.

4. Confusion Matrix



It calculates the confusion matrix to measure the performance of the trained model and plots it. It first predicts classes on the test set, and with the true and predicted labels, it calculates the confusion matrix. This is converted into a DataFrame for clear emotion labels. It plots the confusion matrix as a heatmap using Seaborn, which is then annotated to show the count with a color map to visualize better for assessing how well the model has performed for each class of emotion.

5. Emotion Distribution in Pie Chart



A pie chart to show the distribution of emotions in the dataset. First, it counts the occurrences of every emotion in the said DataFrame and creates a pie chart where it labels each segment for each emotion. The chart is configured to display percentage shares, and everything is colored differently for clarity of view on the distribution of the emotions.

6.Classification Report

```
225/225 ━━━━━━━━━━━━━━━━━━━━ 6s 29ms/step
Classification Report:
              precision    recall  f1-score   support

       anger       0.40      0.50      0.45       985
     disgust       0.64      0.26      0.37       102
        fear       0.38      0.29      0.33      1043
   happiness       0.62      0.85      0.72      1765
     sadness       0.42      0.44      0.43      1210
    surprise       0.77      0.60      0.67       795
     neutral       0.56      0.37      0.45      1278

    accuracy                           0.53      7178
   macro avg       0.54      0.47      0.49      7178
weighted avg       0.53      0.53      0.52      7178
```

Classification report for the performance of the trained model on the test set: predict class probabilities, convert these probabilities into class labels, and then compute the classification report. It

prints this report in order to provide detailed insight into the model's performance regarding the accuracy of the classification of different emotions.

CONCLUSION

A facial expression recognition system based on CNN shows the most robust and effective way to classify emotions from grayscale facial images. The approach used here includes the FER2013 dataset with images in 48x48 pixels annotated into seven categories of emotion, which include Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral. The CNN model was designed and trained in such a way that it makes use of TensorFlow and Keras, multiple convolutional and pooling layers to extract hierarchical features from the images. After training, different performance metrics have been used, including a confusion matrix and a classification report. Confusion matrix-It gives the number of correctly and incorrectly classified images for each emotion. The classification report gives insight into various metrics like precision, recall, F1-score for each kind of emotion. From these visual and quantitative results, it can be concluded that the model has a high degree of accuracy to identify and distinguish between emotions, making it very reliable for applications requiring emotional intelligence. The fruition of such a system points toward deep learning techniques for the development and improvement of human-computer interaction-based systems and opens new horizons toward the advancement of affective computing, improving customer experience, and the development of interactive systems.

REFERENCES

[1] Dan Duncan, Gautam Shine, Chris English “Facial Emotion Recognition in Real Time”.https://cs231n.stanford.edu › pdfs › 022_Report

[2] Ewa Piątkowska, Prof. Dr. Jacob Furst“Facial Expression Recognition System”, DePaul University College of Computing and Digital Media.

[3] Aitor Azcarate, Felix Hageloh, Koen van de Sande, Roberto Valenti“Automatic facial emotion recognition” Universiteit van Amsterdam June 2005.

[4] Reeshad Khan, Omar Sharif “A Literature Review on Emotion Recognition using Various Methods”, Global Journal of Computer Science andTechnology: FGraphics & vision Volume 17 Issue 1 Version 1.0 Year 2017 Publisher: Global Journals Inc. Reviewed International Research Journal By American International University.

[5] Hongli Zhang , Alireza Jolfaei , and Mamoun Alazab “A Face Emotion Recognition Method Using Convolutional Neural Network and Image Edge Computing” Hongli Zhang: Preparation of Papers for IEEE Access (October 2019).

[6] Deepesh Lekhak, “A Facial Expression Recognition System using Convolutional Neural Network”, Department of Electronics and Computer Engineering (072/MSCS/654) May 2017.

[7] Aya Hassouneh a, A.M. Mutawa a , M. Murugappan, “Development of a Real-Time Emotion Recognition System Using Facial Expressions and EEG based on machine learning and deep neural network methods” Informatics in Medicine Unlocked 20 (2020) 100372.

[8] Dipika Raval , Mukesh Sakle “A Literature review on Emotion Recognition System Using Various Facial Expression”, Vol-1 Issue-2 2015 IJARIIE-ISSN(O)-2395- 4396 1177 www.ijariie.com 326