Emotion Recognition Methods with the help of Emoji based on Convolutional Neural Networks

Pranoy Dev

School Of Electronics and Electrical Engineering(SENSE), VIT University, TamilNadu, 632014, India

***Abstract*--**We propose a convolutional neural network (CNN) architecture for facial expression recognition. The proposed architecture is independent of any hand-crafted feature extraction and performs better than the earlier proposed convolutional neural network-based approaches. We visualize the automatically extracted features which have been learned by the network in order to provide a better understanding.

Keywords—Introduction, Related Work, Literature Review, Technicalities of Developing Emojis, Background, Neural Network, Library Used, Output, Applications, Future Research, Conclusion, References.

I. INTRODUCTION

Face plays an important role in social communication. Face biometric itself is used in many applications like security, forensic and other commercial applications. Similarly, facial expressions are the fastest means of communication, while conveying any type of information. In 1978, Ekman and Frisen reported that, Happy, Sad, Anger, Fear, Disgust and Surprise are the six basic expressions which are readily recognized across very different cultures. A system designed for analyzing facial actions automatically through a human computer interaction, is called Automatic Facial Expression Recognition System (AFERS). We adopt deep learning technique and propose effective architectures of Convolutional Neural Networks to solve the problem of facial expression recognition. We also apply different loss functions associated with supervised learning and several training tricks in order to learn CNNs with a strong discriminative power.

The AFER system can be applied in many areas of science such as emotion detection, clinical psychology. and pain assessment. There are three major steps in an AFERS; 1. To detect the face from the given input image or video, 2. To extract the facial features like eyes, nose, mouth from the detected face and 3. To classify the facial expressions into different classes like Happy, Angry, Sad, Fear, Disgust, and Surprise. Face detection is a special case of object detection.

In the proposed system, face detection is implemented using skin color detection and segmentation. Also, it involves lighting compensation algorithm and morphological operations to retain the face from the input image. To extract the facial features, Active Appearance Model i.e. AAM method is used. Finally, the expressions are recognized as Happy, Sad, Anger, Fear, Disgust, and Surprise, initially by using simple Euclidean Distance method and then by training the Artificial Neuro-Fuzzy Inference System (ANFIS).

The implementation of an application process is to identify the facial emotion recognition reasons with highlighted emoji indicator for the identification of six expressions. Such expressions are neutral, fear, anger, happy, sad, and surprise. Furthermore, facial expressions are investigated to realize the real impact of the emoji. The expressions are the actual predictor of human behavior. If expressions are good, then it means the person is in a good mood or has a joyful personality. If the person's expressions show anger or sorrow that means the person is not feeling well or his or her personality does not have joyful traits. The expressions are related with the person's behavior and personality. Understanding expressions are important because they give a lot of information regarding behavior and moods of people through expressions; one can know what is going inside the person's mind and how it can be handled.

II. RELATED WORK

With the fast growth of deep learning, the state-of-the-art in many computer vision tasks has been considerably improved. In image classification, there are some well-known deep CNNs can be mentioned as follows. The first network we want to mention is AlexNet. This network has a very similar architecture to LeNet, but is deeper and bigger, and its convolutional layers stack on top of each other. Previously it is common to have only a single convolutional layer followed by a pool layer. The next CNNs are called VGGNet. One important property of VGGNet is that there are many convolutional layers with small filter size 3×3 that stack on top of each other instead of using a single convolutional layer with larger filter size as in previous CNN generations. Recent improved variants of ResNet called Wide ResNet or ResNeXt also demonstrate their impressive power in image classification tasks. Deep learning allows us to extract facial features in an automated manner without requiring manual design of feature descriptors. There have been some studies that employ CNNs to address the problem of facial expression recognition.

III. LITERATURE REVIEW

Artificial intelligence systems to recognize human emotion have attracted much research interest, and potential applications of such systems abound, spanning domains such as customer-attentive marketing, health monitoring, and emotionally intelligent robotic interfaces. In light of the important role that facial expression plays in communicating emotion in humans, there has been substantial research interest in computer vision systems to recognize human emotion. Using facial expressions with ideograms and smileys is the emoji. The Japanese word, "emoji" consist of two parts: "the e means "picture" and moji means "letter". Emoticons were used before emoji as "symbolic representations for facial expressions based on punctuation marks that could be covered using a standard keyboard". Both emojis and emoticons are frequently used in the text messaging, emails, and other electronic forms of communication. Emojis are a part of the life which was first introduced by Japanese mobile phone companies, such as Vodafone and NTT DocoMo. An early ninety was the period

when Japanese companies enabled the use of the emoji in their communication via electronic devices. They were the pioneers in the use of the emoji. Through these companies, the trend enhanced and the other companies also came forward, and used these emojis to make the communication better. Emojis became popular worldwide and are widely being used in the world at an international leve. The emoji was adopted by Apple Inc.; the corporation recognized the use of the smileys and other electronic pictorial symbols to show what the sender is feeling. Besides the text meaning, the pictorial smileys and other expressional symbols were important because they provide the opportunity to show the inner feelings of the sender. After the adaptation of iPhones, the other phones such as Samsung also used these methods. Now, it is used worldwide. The uses of the smileys are important in day-to-day lives to show facial expression. The use of smileys and other pictorial images in every platform is common. These platforms include Android and Windows. The meaning of the word emoji in original form is pictograph. Emoji is now available in colourful forms. For the time being, it has progressed and now multiple forms and types are available through the internet and communication devices. In the beginning, the emoji was only available in the form of black and white shapes and it was also in a basic format. However, at the current time, they have been developed effectively and they are available in variety of shapes. Using emojis has increased the effectiveness of the use of the symbols. The use of symbols is also seen in other communication formats. Besides with electronic devices, the symbols are widely used on the internet. Today is the era of internet as well as communication and information technologies. Thus, the use of this communication innovation is evident. No one can deny its importance and the organizations engaged in the communication are well aware of the importance of all the techniques that are suitable for enchasing the effectiveness of the communication. Therefore, the use of the emoji is quite common and has developed over time. The first international conference on the emoticon was conducted in 2016. This is an important progress that has been noted or seen on the topic. It is expected that, for the time being, a lot of other initiatives must be taken to make sure that the emojis are incorporated with emoticons to enhance their effectiveness. Thus, that was the brief history about the emoji and its use in communication via electronic devices. In addition, emojis are used in publications and social media as well. The importance of the emojis is twofold. Emojis show feelings because they express emotions which make them important. We often times witness in our day to day lives the message conveyance. Emojis suggest that an emotional communication is as significant as using words. By using smileys, readers can understand the sender's sentiments. The use of the smileys and other symbols are imperative in our everyday lives to demonstrate effective communication. That is why more often the smiley signs are used in a time of happiness, an important advancement in communication. For example, in the face-to-face communication, if people do not use gestures or expressions in their conversations, his or her conversation becomes less meaningful. So, it can be said that these are the important considerations in the realm of the development of the symbols and their ultimate use in the communication patterns. Besides this, the emoji are considered as non-verbal tools. The emoji are the most powerful tools to facilitate the communication and allows people to express their linguistic capabilities. An investigation by was also done on the use of the emoji in real time communication to make sure whether the emojis are effectively communicating. For this purpose, the scientists experimented and analysed that when people use the sign of anger in their day to day communication, the receiver perceives that the person to whom they are in talking is angry and they apologize in that context. They also use the sign. of the apology. Thus, this implies that the use of the emoji in relation to strengthen the language is important. The people understand their linguistic cues, and they use emoji in that context. Hence, the emojis are the powerful tools to demonstrate the freedom of expression and the nonverbal cues in our daily textual language. The symbols and the pictures that are used in the emoji have different colors, and these colors provide the best support for the representation of the emotions and the facial expressions. The usage of the appropriate colors and the style is an important indication of the effectiveness of the emoji. The body and skin colors are used in those symbols and the pictures are related to the human body. The use of the skin color is significant and promising because it must match with the reflection of facial expressions. Both the Windows and the Android platforms are used by millions of the people in the world. To ensure an effective communication, these platforms allow users to use symbols. The Windows 8.1 and onward are using smileys and the symbols. Any platform in the world supports different fonts. Even MS Word provides different kinds of the facial expressions. As mentioned earlier, the Japanese organizations have additionally delivered the images and the smileys that are especially identified with the way of life of the Japan, and they are not subject to the impression of the feelings of the entire world. Therefore, they are operable in any platform. Furthermore, the use of the emoji has great cultural influence in terms of the facial expressions that are expressed by the emojis (Figure 1). The Oxford dictionary also considered the year 2015 as the most influential in terms of the development of the emoji and its impact on the culture. The emoji of the year was the smiley with tears of joy. This emoji best illustrated and expressed the emotions of the humans when they express their love and joy for life. The uses of the emoji are worldwide and the representation of the emotions by the people is evident in every culture and the environment. The emoji is also the reflection of the specific culture of the country. Emojis provide the appearance of the facial expressions and emotions that are specific to each culture. For instance, the Japanese companies have also produced the symbols that are particularly related to the culture of the Japanese and they are not subject to the reflection of the emotions of the whole world. Therefore, the role of the diversity of the culture is essential and the emojis have a profound impact on the life of the people

OBJECTIVE

Facial expression recognition systems have attracted much research interest within the field of artificial intelligence. Many established facial expression recognition (FER) systems apply standard machine learning to extracted image features, and these methods generalize poorly to previously unseen data. This project objective is to classify images of

human faces into discrete emotion categories like Happy or Sad using convolutional neural networks (CNNs).

METHODOLOGY

We implemented a first-pass CNN with a fixed depth of five convolutional layers. trained using the following characteristics. Parametrized dropout rate, learning rate, and l2 regularization. Batch normalization (optional) after each layer. Weight initialization for using ReLU nonlinearities as presented by He et al. 3x3 convolutional filters with stride 1 and zero padding to preserve spatial size

And for library we are using Keras. Keras is very good for rapid prototyping. Keras was developed to enable deep learning engineers to build and experiment with different models very quickly. Just as TensorFlow is a higher-level framework than Python, Keras is an even higher-level framework and provides additional abstractions. Being able to go from idea to result with the least possible delay is key to finding good models. However, Keras is more restrictive than the lower-level frameworks, so there are some very complex models that you can implement in TensorFlow but not (without more difficulty) in Keras.

Note that Keras uses a different convention with variable names than we've previously used with numpy and TensorFlow. In particular, rather than creating and assigning a new variable on each step of forward propagation such as X, Z1, A1, Z2, A2, etc. for the computations for the different layers, in Keras code each line above just reassigns X to a new value using X = ... In other words, during each step of
forward propagation, we are just writing the latest value in the commputation into the same variable X. The only exception was X_input, which we kept separate and did not overwrite, since we needed it at the end to create the Keras model instance (model = Model(inputs = X_input, ...) above).

We have to built a function to describe the model. Then to train and test this model, there are four steps in Keras:

1. Creating the model by calling the function above

2. Compile the model by calling model.compile(optimizer = "...", loss = "...", metrics = ["accuracy"])

3. Training the model on train data by calling model.fit(x = ..., y = ..., epochs = ..., batch_size = ...)

4. Testing the model on test data by calling model.evaluate(x = ..., y = ...)

To get a point of comparison, our model has to get around 95% test accuracy in 40 epochs (and 99% train accuracy) with a mini batch size of 16 and "adam" optimizer.

TECHNICALITIES OF DEVELOPING EMOJIS

All technicalities were incorporated to develop the emojis. The incorporation of all kinds of coding languages and the binary operations are important. Without the inclusion of these operations, it is quite possible that the emojis are not developed effectively. All programming languages were studied to develop the emojis, and this is an aspect worth noting (Benenson). The basic programming languages which are essential to develop the emoji are given below. These are the most important languages that are widely used in the development of the software and emojis.

- HTML
- JAVA
- PHP
- C++
- JavaScript

The command line procedures are also important for the sake of the development of the emoji-enabled software in the electronic devices. The command lines must be free of any errors and this must be considered in an attempt to develop the bug free software for the emoji. The command line procedures in the software must be incorporated with the programming language which has been used, and which can be used to develop the software and its programming languages. Therefore, these were the basic considerations for the developers. Furthermore, the format of the pictures was also selected for the development of the emoji. The pictures' format, shape, and coloring play a significant role in developing the emoji pictures. Therefore, the correct formatting and the use of the colors and the lining are thus quite important. The use of the correct picture format is important, and the following formats can be used:

- SVG
- PING
- GIF

These are commonly important formats of the pictures which are frequently used in the development of the software. The pictures must be developed and incorporated in the right format in the system to develop the accurate and right emojis. Thus, it is important to notice that the organization of the photos in the emojis is necessary to build, and is likewise chosen for the advancement of the emoji technology in the world.

BACKGROUND

The techniques of the visual technology for the conversions of the expressions into the art of facial graphical style conversion are making a great contribution and advancement in the digital graphics designing. The initial tools used from the people associated with this field were using the histograms as the mode of representing the pyramids of the facial gradients. The image, which has the static advancements in the input emotions detection, are illustrating in the applications of the visual posture for the body language. The face moods that highlight the EEG are complex stature for the methods for detecting the wide ranges of the pyramids of computer database applications. The networks such as the CNN is also integrated with this small-scale training framework or the models that could relate the facial impressions or the emotions for optimization of the facial defection of emotions in the layer-based technology, such as VGS.

1. Neural network
2. Convolutional neural network
3. General CNN-based face recognition schema
   - Face recognition models based on CNN
   1) Deep-Face model
   2) Web-scaled Deep-Face model

NEURAL NETWORK

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. ANNs, like people, learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. This is true of ANNs as well.

Many important advances have been boosted by the use of inexpensive computer emulations. Following an initial period of enthusiasm, the field survived a period of frustration and disrepute. During this period when funding and professional support was minimal, important advances were made by relatively few researchers. These pioneers were able to develop convincing technology which surpassed the limitations identified by Minsky and Papert. Minsky and Papert, published a book (in 1969) in which they summed up a general feeling of frustration (against neural networks) among researchers, and was thus accepted by most without further analysis. Currently, the neural network field enjoys a resurgence of interest and a corresponding increase in funding. The first artificial neuron was produced in 1943 by the neurophysiologist Warren McCulloch and the logician Walter Pits. But the technology available at that time did not allow them to do too much.

Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyze. This expert can then be used to provide projections given new situations of interest and answer "what if" questions. Other advantages include:

1. Adaptive learning: An ability to learn how to do tasks based on the data given for training or initial experience.
2. Self-Organization: An ANN can create its own organization or representation of the information it receives during learning time.
3. Real Time Operation: ANN computations may be carried out in parallel, and special hardware devices are being designed and manufactured which take advantage of this capability.
4. Fault Tolerance via Redundant Information Coding: Partial destruction of a network leads to the corresponding degradation of performance. However, some network capabilities may be retained even with major network damage.

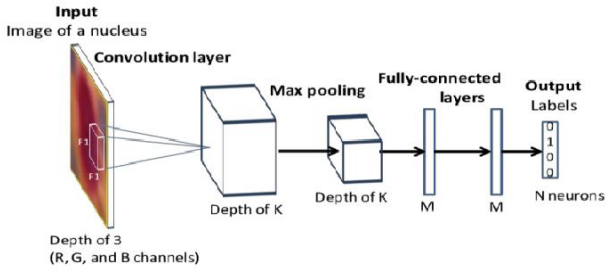
Neural networks take a different approach to problem solving than that of conventional computers. Conventional computers use an algorithmic approach i.e. the computer follows a set of instructions in order to solve a problem. Unless the specific steps that the computer needs to follow are known the computer cannot solve the problem. That restricts the problem-solving capability of conventional computers to problems that we already understand and know how to solve. But computers would be so much more useful if they could do things that we don't exactly know how to do.

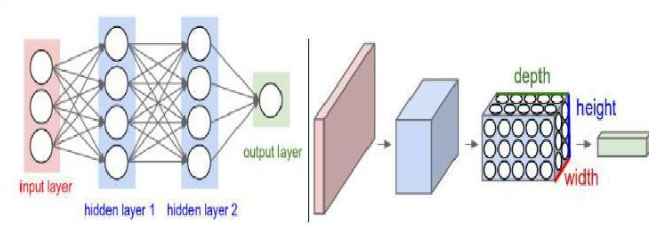
CONVOLUTIONAL NEURAL NETWORK

Convolutional Neural Networks are very similar to ordinary Neural Networks; they are made up of neurons that have learnable weights and biases. Each neuron receives some inputs, performs a dot product and optionally follows it with a non-linearity. The whole network still expresses a single differentiable score function: from the raw image pixels on one end to class scores at the other. And they still have a loss function (e.g. SVM/Softmax) on the last (fully-connected) layer and all the tips/tricks we developed for learning regular Neural Networks still apply.

ConvNets architectures make the explicit assumption that the inputs are images, which allows us to encode certain properties into the architecture. These then make the forward function more efficient to implement and vastly reduce the number of parameters in the network.

Convolutional Neural Networks take advantage of the fact that the input consists of images and they constrain the architecture in a more sensible way. In particular, unlike a regular Neural Network, the layers of a ConvNet have neurons arranged in 3 dimensions: width, height, depth. It should be noted that the word depth here refers to the third dimension of an activation volume, not to the depth of a full Neural Network, which can refer to the total number of layers in a network. For example, the input images in CIFAR-10 are an input volume of activations, and the volume has dimensions 32x32x3 (width, height, depth respectively). As we will soon see, the neurons in a layer will only be connected to a small region of the layer before it, instead of all of the neurons in a fully-connected manner. Moreover, the final output layer would for CIFAR-10 have dimensions 1x1x10, because by the end of the ConvNet architecture we will reduce the full image into a single vector of class scores arranged along the depth dimension.





LAYERS TO BUILD COVNETS

A simple ConvNet is a sequence of layers, and every layer of a ConvNet transforms one volume of activations to another through a differentiable function. We use three main types of layers to build ConvNets architectures: Convolutional Layer, Pooling-Layer, and Fully Connected Layer (exactly as seen in regular Neural Networks). We will stack these layers to form a full ConvNet **architecture**. A simple ConvNet for CIFAR-10 classification could have the architecture [INPUT - CONV - RELU - POOL - FC]:

1) INPUT [32x32x3] will hold the raw pixel values of the image, in this case an image of width 32, height 32, and with three color channels R,G,B.

2) CONV layer will compute the output of neurons that are connected to local regions in the input, each computing a dot product between their weights and a small region they are connected to in the input volume. This may result in volume such as [32x32x12] if we decided to use 12 filters.

3) RELU layer will apply an elementwise activation function, such as the max(0,x)max(0,x) thresholding at zero. This leaves the size of the volume unchanged ([32x32x12]).

4) POOL layer will perform a down-sampling operation along the spatial dimensions (width, height), resulting in volume such as [16x16x12].

5) FC (i.e. fully-connected) layer will compute the class scores, resulting in volume of size [1x1x10], where each of the 10 numbers correspond to a class score, such as among the 10 categories of CIFAR-10. As with ordinary Neural Networks and as the name implies, each neuron in this layer will be connected to all the numbers in the previous volume.

In this way, ConvNets transform the original image layer by layer from the original pixel values to the final class scores. Some layers contain parameters and other don't. In particular, the CONV/FC layers perform transformations that are a function of not only the activations in the input volume, but also of the parameters (the weights and biases of the neurons). On the other hand, the RELU/POOL layers will implement a fixed function. The parameters in the CONV/FC layers will be trained with gradient descent so that the class scores that the ConvNets computes are consistent with the labels in the training set for each image.

1) General CNN-based face recognition schema

2) Common steps:

Face detection

Viola-Jones, Cascade CNN

Pre-processing

Geometric & lighting normalization

CNN training

Supervised vs. unsupervised

Face identification

Classification problem

Metric learning

Joint-Bayesian, Cosine similarity, Triplet Similarity, Energy-based similarity

3) Face verification

4) Face recognition models based on CNN

5) CNN based models are different:

1. Architecture of CNN

2. Depth of neural network

3. Number of parameters

4. Scale of training dataset

5. Similarity metric

6. Alignment vs. non-alignment pre-processing

LIBRARY USED

KERAS LIBRARY

Keras (κέρας) means horn in Greek. It is a reference to a literary image from ancient Greek and Latin literature, first found in the Odyssey, where dream spirits (Oneiroi, singular Oneiros) are divided between those who deceive men with false visions, who arrive to Earth through a gate of ivory, and those who announce a future that will come to pass, who arrive through a gate of horn. Keras is a high-level neural networks API, written in Python and capable of running on top of TensorFlow, CNTK, or Theano. It was developed with a focus on enabling fast experimentation. Keras is used if there is a need of a deep learning library that:

1) Allows for easy and fast prototyping (through user friendliness, modularity, and extensibility).
2) Supports both convolutional networks and recurrent networks, as well as combinations of the two.
3) Runs seamlessly on CPU and GPU.

GUIDING PRINCIPLES

User friendliness: Keras is an API designed for human beings, not machines. It puts user experience front and centre. Keras follows best practices for reducing cognitive load: it offers consistent & simple APIs, it minimizes the number of user actions required for common use cases, and it provides clear and actionable feedback upon user error.

Modularity: A model is understood as a sequence or a graph of standalone, fully-configurable modules that can be plugged together with as few restrictions as possible. In particular, neural layers, cost functions, optimizers, initialization schemes, activation functions, regularization schemes are all standalone modules that you can combine to create new models.

Easy extensibility: New modules are simple to add (as new classes and functions), and existing modules provide ample examples. To be able to easily create new modules allows for total expressiveness, making Keras suitable for advanced research.

Work with Python: No separate models configuration files in a declarative format. Models are described in Python code, which is compact, easier to debug, and allows for ease of extensibility.

Here is the sequential model:

```
from keras.models import Sequential

model = Sequential()
```

Stacking layers is as easy as .add() :

```
from keras.layers import Dense

model.add(Dense(units=64, activation='relu',
input_dim=100))
model.add(Dense(units=10, activation='softmax'))
```

Once your model looks good, configure its learning process with .compile() :

```
model.compile(loss='categorical_crossentropy',
        optimizer='sgd',
        metrics=['accuracy'])
```

TENSORFLOW LIBRARY

TensorFlow is a Python library for fast numerical computing created and released by Google. It is a foundation library that can be used to create Deep Learning models directly or by using wrapper libraries that simplify the process built on top of TensorFlow. TensorFlow is an open source library for fast numerical computing.

It was created and is maintained by Google and released under the Apache 2.0 open source license. The API is nominally for the Python programming language, although there is access to the underlying C++ API.

Unlike other numerical libraries intended for use in Deep Learning like Theano, TensorFlow was designed for use both in research and development and in production systems.

It can run on single CPU systems, GPUs as well as mobile devices and large-scale distributed systems of hundreds of machines.

Nodes: Nodes perform computation and have zero or more inputs and outputs. Data that moves between nodes are known as tensors, which are multi-dimensional arrays of real values.

Edges: The graph defines the flow of data, branching, looping and updates to state. Special edges can be used to synchronize behavior within the graph, for example waiting for computation on a number of inputs to complete.

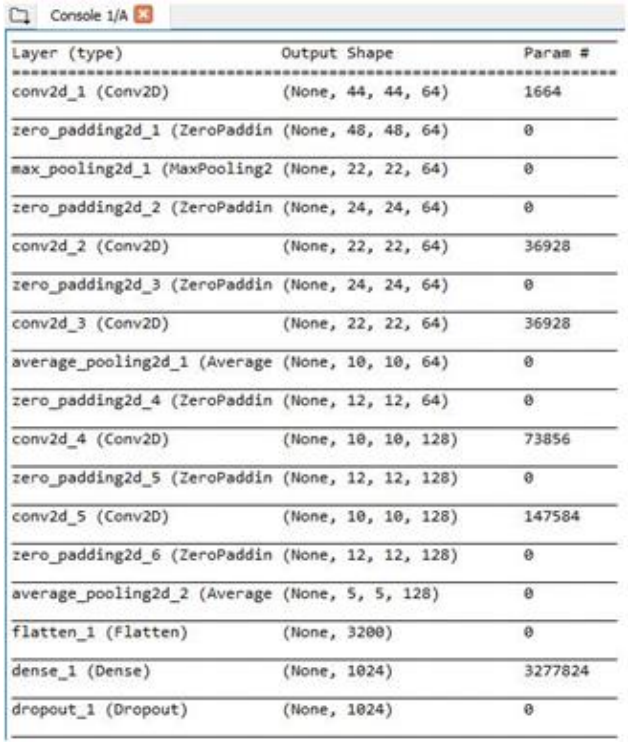
Operation: An operation is a named abstract computation which can take input attributes and produce output attributes. For example, you could define an add or multiply operation.

TFLearn features include:

1) Easy-to-use and understand high-level API for implementing deep neural networks, with tutorial and examples.
2) Fast prototyping through highly modular built-in neural network layers, regularizes, optimizers, metrics...
3) Full transparency over Tensorflow. All functions are built over tensors and can be used independently of TFLearn.
4) Powerful helper functions to train any TensorFlow graph, with support of multiple inputs, outputs and optimizers.
5) Easy and beautiful graph visualization, with details about weights, gradients, activations and more...
6) Effortless device placement for using multiple CPU/GPU

Software used- Spyder (Python 3.6)

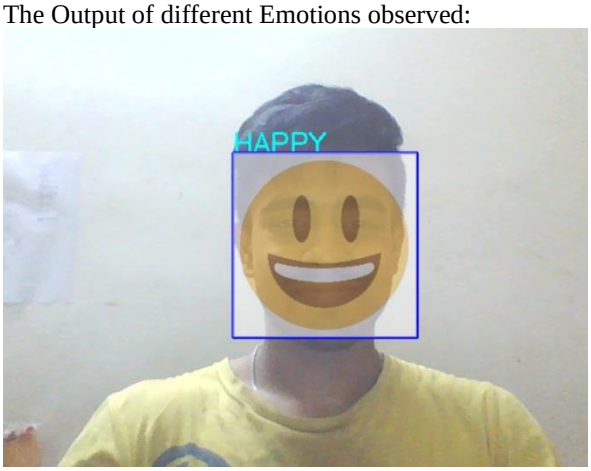
OUTPUT:

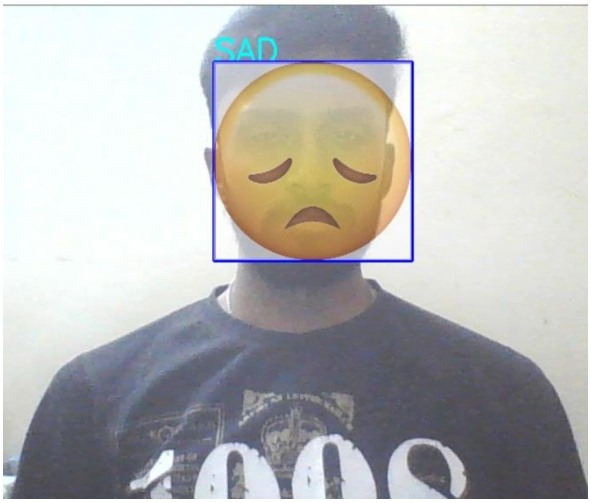


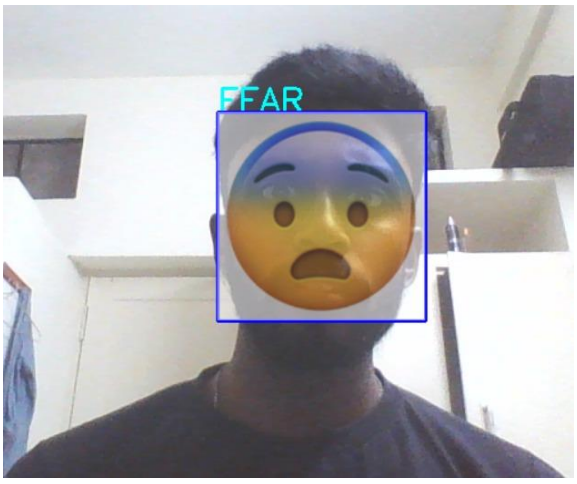
dropout_1 (Dropout)	(None, 1024)	0
dense_2 (Dense)	(None, 1024)	1049600
dropout_2 (Dropout)	(None, 1024)	0
dense_3 (Dense)	(None, 7)	7175

Total params: 4,631,559
Trainable params: 4,631,559
Non-trainable params: 0

The Output of different Emotions observed:







APPLICATIONS

1) It can be used widely to trace highly threatning terrorist organizations and personnel.

2) It can be used in medicine, e-learning, monitoring, entertainment, law and marketing

3) It can be used in home automation system. For eg. Face Recognition door lock.

4) It can help in rehabilitation and counselling.

5) It can be used in school to observe the behavior of students and in hospitals to detect various expressions of patients and especially Autistic people and take actions based on that.

6) This device can also be used to provide security in other areas provided face recognition capabilities are added to this project.

FUTURE RESEARCH

1) We can identify the following avenues for improving performance.

2) More precise normalization of the images to account for translation, rotation, and scale changes. Any normalization would be limited by the desired recognition speed.

3) The various facial features could be ranked according to their importance in recognizing faces and separate modules could be introduced for various parts of the face, e.g., the eye region, the nose region, and the mouth region (Brunelli and Poggio obtain very good performance using a simple template matching strategy on precisely these regions).

4) An ensemble of recognizers could be used. These could be combined via simple methods such as a linear combination based on the performance of each network, or via a gating network and the expectation-maximization algorithm. Examination of the errors made by networks trained with different random seeds and by networks trained with the SOM data versus networks trained with the KL data shows that a combination of networks should improve performance (the set of common errors between the recognizers is often much smaller than the total number of errors).

5) Invariance to a group of desired transformations could be enhanced with the addition of pseudo-data to the training database—i.e., the addition of new examples created from the current examples using translation, etc.

CONCLUSION

An automated Facial Expression Recognition System has a wide range of applications in psychological research and human-computer interaction applications. The system plays a communicative role in interpersonal relations because they can reveal the affective state, cognitive activity, personality, intention and psychological state of a person. In this system, the static images as well as video can be given as input and tested for the different expressions. Also, the system can work accurately for person-independent database. The accuracy of facial expression recognition varies with number of training samples. The system gives the recognition rate close to 100% for large number of training samples. A neuro-fuzzy approach for facial expression recognition is applicable for real time applications such as human emotion analysis, human computer interaction, surveillance and online conferencing and for entertainments.

The method is capable of rapid classification, requires only fast approximate normalization and pre-processing, and consistently exhibits better classification performance than the eigenfaces approach on the database considered as the number of images per person in the training database is varied from one to five. This project is developed to create an application for emoji. Emojis are basically pictorial depiction of the human emotions such as fear, happiness, angriness, and related emotions. These emotions are important in day to day conversations, which take place in our daily lives. The use of the emoji was first introduced by the Japanese companies. After then, emojis were present in almost every electronic communication form, from e-mail to SMS and through related means of communication. In addition, the uses of the emojis were first introduced in mobile phones such as Vodafone. The emoji has power to reflect the same level of the human expression as it happens in the real life. Therefore, the roles of emojis are important in our lives. Thus, they are vital in our lives for application purpose as well like we did in this project.

ACKNOWLEDGMENT

This report would not have been possible without the essential and gracious support of Prof. Yogesh Kumar Choukiker. His willingness to motivate us contributed tremendously to our report. We also would like to thank him for showing some examples related to the topic of report. He encouraged us even when we had any sort of obstacles or difficulties and guided through it. It couldn’t be possible without him. He inspired all of us and gave thoughtful ideas and innovative hints which paid dividends to the project.

Besides, we would like to thank the authority of VIT UNIVERSITY for providing us the good environment and facilities to complete this report. Also, we would like to thank them for offering this subject. It gave all of us an opportunity to participate and learn which could be fruitful for our future and carrier. Finally, we would like to thank my family and friends for their support and understanding towards my project as without their help we couldn’t even possibly have had been able to finish our assignment.

Thank you

REFERENCES

1) P. Ekman, T. Huang, T. Sejnowski, J. Hager, "Final Report to NSF of the Planning Workshop on Facial Expression Understanding", 1992.

2) Ewa Piatkowska "Facial Expression Recognition System", Master's Thesis Technical Report.

3) H.-U. Bauer and K. R. Pawelzik, “Quantifying the neighbourhood preservation of self-organizing feature maps,” IEEE Trans. Neural Networks, vol. 3, pp. 570–579, 1992
.

4) https://ieeexplore.ieee.org/abstract/document/6892834

5) http://cs231n.github.io/convolutional-networks/

6) http://krasserm.github.io/2018/02/07/deep-face-recognition/

7) https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html

8) https://arxiv.org/abs/1404.2188

9) https://arxiv.org/abs/1708.00524

10) http://www.lrecconf.org/proceedings/lrec2016/pdf/735_Paper.pdf

11) https://patents.google.com/patent/US20100125811A1/en

12) https://ieeexplore.ieee.org/abstract/document/7427425