Generate Avatar Using Facial Expressions

Kshitij Antani
Computer Science and Engineering
Devang Patel Institute of Advance Technology and Research
Charotar university of Science and Technology
Rajkot, Gujarat, India
18dcs005@charusat.edu.in

Rutika Mehta
Computer Science and Engineering
Devang Patel Institute of Advance Technology and Research
Charotar university of Science and Technology
Rajkot, Gujarat, India
18dcs045@charusat.edu.in

***Abstract*— Nowadays, each and every person is willing to express emotions graphically. With such a good environment that the social interactive sessions become more interactive and effective in day-to-day life. Thus, this approach helps to improvise a person's bonding skills in terms of innovativeness, transformation and refactoring. Here, we use communication over a computer network, with the help of abstract faces such as computer graphics avatars. These avatars convey users' emotions, expressions and enrich their communications. Seeing some avatars showing expressions the same as users, we feel that they are more vivid and lively than photorealistic faces. Some recent advancements in facial tracking systems provide accurate facial movements tracking so that to get accurate data. However, the existing systems are lacking in context of converting or expressing the data acquired from face tracking to other sources. In this paper we are trying an approach of development in an immensely innovative way to turn facial data into graphical representation via showing the avatar nearly matching the expression of a person using Convolutional Neural Network (CNN). CNN helped to achieve the accuracy of 92% which is higher than the traditional approaches. Thus, results in development of a system lead to development of tools to express emotions digitally.**

Keywords*— *Avtar, CNN, Emotions, Facial Expression, Emoji

I. INTRODUCTION

Communication is an important part of everyday life whether it is Verbal or non-verbal, it allows one to engage in conversations with the world. With the help of communication technologies that is the internet and other communication devices have made it possible to engage in fast, dynamic, and effective communication. Emojis are being used for the visual depictions of human emotions through their expressions. In this research paper, emotions of a person are detected and generated in a real time-based avatar using facial expressions. This identification of facial expressions plays an important key role in identifying patterns and image processing [3]. Also helps identifying facial expressions and generates the emoji that almost matches the person's emotions. Some of the application, related to face and its expressions include person identification and access control, video call and teleconferencing, forensic applications, human-computer interaction, automated surveillance, cosmetology, and so on [4]. Also, they are used as an important tool in psychological studies and in medical rehabilitation.

This proposed system helps to express the expressions of humans using real time avatars or emojis. The role of the system is to, pictographic forms of facial expressions, objects, and symbols. Emoji can potentially serve to portray personality traits of its users namely "emotional stability, extraversion, and agreeableness". In the older version of emoji generator, emoticon's use aims at enriching the feature of polarity classification. In fact, emoticons demonstrate interpersonal functions such as personal expressions and one's emotions/feelings in distinct virtual platforms. This system proposes various emojis that try to match exact feelings or emotions the user is showing. Emotions play different roles in human life. Considering this, the system gives basic ideas and accurate avatars to the expression or emotions of a person. Our proposed system dynamically generates avatars; by using Machine Learning, it processes the input taken by camera and an alternative avatar is generated dynamically compared with input data. The generated avatars consist of 7 different emotions. This system brings different emotions to learn or sense the human feelings or mood at a time. Different types of emotions in this system are happy, sad, disgusted, angry, fearful, crying, shocked [3].

II. RELATED WORKS

A. Maker- Free Personal Animated Emojis

Phone supports and enables users to use a new type of emojis called Animated Emoji. It also supports characters like fox, pig, dog, cat and different animals. This app allows users to send and receive animated emojis through smartphones. One can express himself with 3D animated stickers and emojis featuring yourself as a cartoon avatar maker [3].

B. Bitmoji – Your Personal Emoji

People can create an expressive cartoon avatar a lot like themselves. One can choose from a huge and various library of stickers all featuring your looks. One can use this application wherever he wants like third party apps.

C. GIFs, Facemoji Keyboard-Emoji Keyboard, and Sticker:

The First ever Emoji-Centric Android keyboard. It holds over 3600+ Emoji, emoticons, GIFs, stickers on this Emoji keyboard. This application helps us to spice up and make chat more interesting on Facebook, Twitter, Instagram, Messenger, WhatsApp and for many more apps.

D. Elite Emoji

This application has sensational emojis and attractive stickers to express emotions, moods, messages and ideas beautifully in your chat conversations. It contains almost 2000 exclusive high definition emojis and stickers to choose from. Also, with one more feature as it supports emojis with a single touch in Social networking sites. By using this application, we get access to a huge selection of gifs as well and even create your own animated images to make communication describe more perfectly [3].

III. Motivation

Understanding the human facial expressions and the study of expressions has many aspects, from computer analysis, emotion recognition, lie detectors, airport security, nonverbal communication and even the role of expressions in art.

Improving the skills of reading expressions is an important step towards successful relations.

IV. Proposed Approach

To build our Facial Expression Recognition module, we used Convolutional Neural Networks (CNNs). Deep learning-based approaches, mainly which comprises CNNs have a high success rate for tasks which are related to images as they are proficient in extracting adequate representations from the data.[6] Convolutional Neural Networks is a special type of neural network that is very effectively used for image recognition and classification. It is highly proficient in areas like identification of objects, faces, traffic signs. They are also used in self driving cars and robots. Convolutional is a simple application of a filter to an input that results in an activation. So, there is some sort of input and there are some certain sorts of thresholds, and when the input meets those thresholds then there is an activation. Now, repeated application of the same filter to an input results in a map of activations which is called a feature map. If there is a certain type of input that goes in and that input repeats itself in a certain way then we can see there is a feature map forming that indicates their locations and strengths of a detected feature in input. We can then train the model with help of those features. Facial expressions are arrangements of facial muscles which are used for conveying emotional state to the one who is observing. Emotions can be divided into seven broad categories—Anger, Disgust, Fear, Happiness, Neutral, Sad and Surprised.[4] We will build a deep learning model to classify facial expressions from the images. Then we will map the classified emotion to a predefined avatar.

The neural network needs sets of training data and teaching signals for training.[2] In the proposed method, we selected 7 facial expressions (neutral, happy, angry, surprised, sad) as the recognition target. We have adopted a 5-layer neural network. The learning rate was set to 0.0001, epoch was set to 75 and decay was set 1^{e}-6. We have used two activation functions one Relu and Softmax. Relu is better activation function then other [1].

A. Dataset

We have used the FER-2013 facial expression dataset for training and validation.[5] The data consists of 48x48 pixel grayscale images of faces. The faces in the dataset have been automatically recorded so that the face occupies the same amount of space in each image. The dataset is categorized into each face based on the emotion shown in the facial expression into one of seven categories (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral).

B. Data Preprocessing

Data preprocessing refers to a process of developing the raw data and making it befitting for a machine learning model. It is the first and foremost step while building the machine learning model. When building machine learning projects, using a clean and well formatted dataset is a predominant factor. The data gathered from real-world generally consists of missing values, noises or can be in format which cannot be used directly. Hence, data preprocessing plays a crucial role by cleaning the data and making it appropriate for machine learning models. A well formatted dataset also increases the efficiency and accuracy of the model. In our project, we have used the dataset FER- 2013 in which the training set consists of 28,709 examples and the public test set consists of 3,589 examples.

C. Model Architecture

We will generate datasets for training data and validation data from image files in our directory from scale of 1 to 255, since we define our images in pixels which is an array consisting of values from 1 to 255. Then we'll create our generator from the imported data. The generator will be specified with target size, batch size, color mode and class mode. The target size refers to the pixel size of the data, batch size is a term used in machine learning and refers to the number of training examples utilized in one iteration, the color mode specifies the mode of the image dataset, and here the class mode will be categorical since we have categories of angry, disgust, fearful, happiness, sadness, surprised and neutral.

We will use Keras library to create a model which uses Sequential Convolutional Network. The Sequential Convolutional Network allows to create layers of a model in a linear stack of layers i.e., layer by layer, which is the preferred way for most problems.[4] The convolutional network is divided into following components:

a) Convolutional Layer

It is of two dimensions. It creates a convolutional kernel that is convolved with the layer input to produce a tensor of outputs. Here we will provide input shape, height and width of convolutional network, activation function and input shape of the image.

b) Activation functions

In the neural network, activation function decides whether a neuron should be activated or not by calculating the weighted sum and adding it further with bias. Bias is an additional parameter used to adjust the weighted sum of inputs along with outputs to the neuron. In our project, we will use two activation functions, Relu and SoftMax. Relu stands for rectifier which is an activation function used in a neural network. It is one of the most commonly used activation functions. Here, Softmax function is used for normalizing the outputs by converting them from weighted sum values into probabilities that aggregates to one.

c) Pooling Layer

The pooling layers down sample the data. Down sampling means training on a disproportionately low subset of the classic samples. We are down sampling to include the lower subset of the data. In other words, pooling is the process of merging data. In this project, max pooling is used where only maximum value is selected from the input.

d) Dropout Layers

The dropout layers randomly set input units to zero with the frequency rate at each during the training time which helps to prevent overfitting. Overfitting occurs when a

model has learned the noise instead of signal that is considered to be considered to be the required signal for learning. Thus, overfitting means training the model more on the noise than the actual data we need.

e) Flatten Layer

Here, the data is converted into a 1-dimensional array so that it can be used as input for the next layer. The output of the convolutional layer is flattened to create a single long feature vector. We need to flatten the values before we turn all those neurons into the next layer.

f) Dense Layer

Dense layer is the hidden layer. In this project, the first dense layer is going to contain 1024 neurons. And the second dense layer contains seven, so this layer is going to contain the exact number of possibilities that we can have. Since in our dataset, we have expressions divided into seven categories. So, we have to convert 1024 neurons from the first hidden layer into those seven categories.

V. Experimental Analysis

A. Testing and Training of Dataset

Furthermore, we have compiled, trained and saved the model. In order to compile, we have to specify certain things. We have specified loss which is the degree of error. The model doesn't maximize the accuracy instead it tries to minimize the loss. We have defined loss as categorical cross-entropy. It is the most common one, which is a type of metric for how we are specifying the losses. Here, Adam is used as an optimizer. The optimizers shape and cast our model in such a way that it is in the most accurate possible form. Adam denotes Adaptive Moment Estimation. It works on theory of momentum i.e., by adding fractions of previous gradients to the current gradients. For tracking accuracy, we have specified accuracy as metric.

For training the model, we have specified training generator and validation generator, epochs and steps per epoch. An epoch indicates the number of passes of the entire training dataset the machine learning algorithm has completed. The steps per epoch is defined by the total number of training dataset examples divided by the batch size. After the model is trained, the model weights have been saved. The training dataset accuracy is 92%.

B. Results

In the below figure, i.e., the sample output window the user expressions are captured on the left-hand side of the window through the webcam.

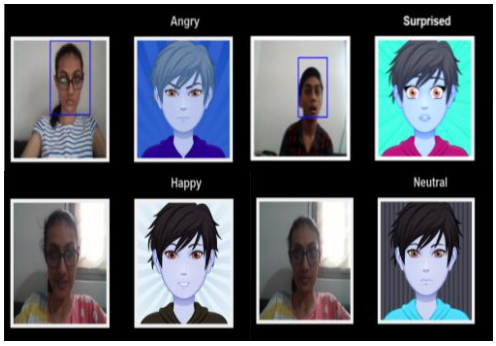

Fig 1: The Output, while user gives any expression

On the right-hand side of the output window, the corresponding avatar is displayed on the basis of the user's facial expression.

On top of the avatar, the respective emotion text is displayed. The avatar changes when the expression of the user changes and that is displayed on the output screen. This real time application can be used in the fields of computer science, neuroscience, psychology and various other fields.

VI. Conclusion and Future Work

The system detects human emotions via their expressions using machine learning, python to predict emotions and mood of the people and represent them using different avatars. These include image acquisition, preprocessing of an image, detection of face, feature extraction, classification. That emotion is classified from the detected face expression through a webcam.[4] The emotions used for the experiments include happy, sad, surprise, fear, disgust and anger emotions.

This project shows how the use of technology can immensely help in knowing the mood of the person. The model named as "Avtar" developed by us shows the high accuracy and therefore be used in real life. This model also show that CNN layers help in improving the result and to further increase the accuracy we can increase the layers, and also, we can further add more type of emotions in the model to be more helpful.

References

[1] Guifang Lin, WeiShen. Research on convolutional neural networks based on improved Relu piecewise activation function. Procedia Computer Science, volume = 131, doi: https://doi.org/10.1016/j.procs.2018.04.239

[2] K. Suzuki et al., "Recognition and mapping of facial expressions to avatar by embedded photo reflective sensors in head mounted display," 2017 IEEE Virtual Reality (VR), Los Angeles, CA, USA, 2017, pp. 177-185, doi: 10.1109/VR.2017.7892245

[3] Pradnya Bhor, Ashwini Raut, Sachin Kate, Anurag Salve, Chandan Prasad. "Research paper on Dynamic Emotion Recognition and Emoji Generation", 2019 IRJET.

[4] N. Swapna Goud, K. Revanth Reddy, G. Alekhya, G. S. Sucheta. "Research Paper on Facial Emoji Recognition", volume = 3, doi: https://doi.org/10.31142/ijtsrd23166

[5] Abdallah El Ali, Torben Wallbaum, Merlin Wasmann, Wilko Heuten, "Face2Emoji: Using Facial Emotional Expressions to Filter Emojis", doi: http://dx.doi.org/10.1145/3027063.3053086

[6] https://www.kaggle.com/msambare/fer2013/metadata