Emotions and Hand Gestures Recognition using Computer Vision

Bharti Yogi[1], Divya Godse[2], Gitika Bharadwaj[3], Sanjay Waykar[4]

[1,2,3,4]*Department of Information Technology, Mahatma Gandhi Mission's college of Engineering and Technology, Kamothe, Maharashtra, India -410 209*

---***---

ABSTRACT: *This research would make it possible for people to have effective communication with computers. We would make the computer recognize human emotions and feelings through analyzing human effective state and behavior. Automatic analysis of facial gestures is rapidly becoming an area of great interest in Computer science and human-computer interaction design communities. This paper is designed to focus on different emotions, facial expressions, and hand gestures of humans. In this project we used Tensorflow and Keras libraries to train the deep learning model in the backend and for the front end we used android and OpenCv. Transfer learning is the widely used technology to classify any emotion, there are many types of Emotions but transfer learning briefly classifies the emotions in seven different categories that we have tried to implement in our project. Users will get an appropriate output in seven categories. Seven types of emotions are namely happy, sad, angry, disgust, Surprised, neutral, fear. A gesture is a way of interpreting your message in a nonverbal form. Hand gestures detection and recognition are proposed in this paper using computer vision and image processing techniques. The people who are dumb and Deaf can find this model useful to communicate with the other people. Dataset is the essential thing without this dataset project can not be done. Large amount of datasets are required to train these models and that dataset has been obtained from kaggle.*

Keywords: Emotions, Hand gestures, Android App, Computer Vision, Machine Learning, Artificial Intelligence.

1. INTRODUCTION

Emotion recognition strategies are beneficial anywhere. Like for detecting a person's emotions which would be one of the crucial factors of a conversation. Facial gestures modify our interactions with everyone, they determine whether we have been paying attention(e.g., a person or what has been said) so to clarify the conversation was important, funny, or unpleasant for us. Gesture recognition does not only limit to facial gesture, the other important type of gesture recognition that is essential for communication are Hand Gesture. This is a very natural and innovative way of communication. It can be used in multiple areas in human-computer interaction and sign language.

1.1 Emotion Recognition

Emotion recognition refers in psychology to the attribution of emotional states based on the observation of visual and auditory nonverbal cues. Nonverbal cues include facial, vocal, postural, and gestural cues displayed by a sender, that is, a person displaying an emotional reaction. For understanding these emotions through a computer's point of view a lot of research has been conducted. A way for a computer to understand emotions could be achieved by machine learning. Machine learning (ML) is an application of Artificial Intelligence (AI) that provides the system with the ability to automatically learn and improve from experience rather than explicit programming. This is possible because today a large amount of data is available which lets machines be trained rather than programmed. It is considered a major technological revolution that can analyze a massive amount of data. Thus, in recent years, there has been a tremendous interest in automating facial gesture analysis. Most approaches to automatic facial gesture analysis in face image sequences attempt to recognize a set of prototypic emotional facial expressions, i.e., happiness, sadness, fear, surprise, anger, and disgust.

1.2 Hand Gesture Recognition

The primary goal of HCI is to improve the interaction between users as well as computers by making the computer more receptive to user needs. Human-Computer Interaction with a personal computer, it is not just limited to keyboard or mouse interaction. A hand gesture recognition system to be able to successfully work, needs to be able to precisely detect each finger and hand orientation in real-time and should be robust to various changes in hand measurements, rotation, color, and lighting. This is a very complex concept and requires advanced image processing and computer vision concepts. Interaction between humans takes place with the help of different sensory modes like gesture, speech, facial and body expressions. Being able to interact with the system naturally is becoming ever more important in many fields of Human-Computer Interaction.

In hand Gesture recognition users need to know what sign language is so they can use this application to convey their message. Any alphabet can be recognized by this application.

1.3 Objective

The proposed System will help detect the user's facial expressions. If someone is angry, happy, disgusted, surprised, scared that would be detected. It can be used

by physically challenged people to express their feelings. They also can use hand gestures to convey their thoughts to other people. In this project we are going to make an android application which would use artificial intelligence and Machine learning. We have trained the model for that and we have used some libraries from python. This application will help the people to share their feelings through our app. It is a very useful application, can be used in various places like, we can use facial and emotion recognition in security systems, like ATMs, for banks , for gaming we can use this system.

This app can detect whether someone is truthful or not. We've seen that security is the main reason for identifying someone. If we can Identify the intent of the person we will be able to avoid that threat. This may be helpful in attack prone areas such as major public gatherings like politician rallies, concerts, cricket matches and airports, which have seen many breaches in recent years.

1.4 Related Works

Our objective is to efficiently integrate emotions recognized from facial expressions and upper body pose of humans using images. The scanned image (testing dataset) is being compared to the training dataset and thus emotion and gestures are predicted.

A lot of work and Research has been done on Emotion and Hand Gesture Recognition using different approaches. All the related work has been done before they have made this into a different project/application but in our case, we have merged these two topics together in a group and one more thing we have done this project for android users so that people get to know about this project. Because android makes it easier to use.

Kin Yun Lum has also Used mobileNetV2 to train their model but they have used this model to train the sign language recognition and in our project we also have used this model to train our Emotion recognition model their work is also good and their model is as accurate as our model[8]. We have used different models to train aur hand gesture Recognition model in our project.

Samta Jain has used a vision-based multi modal analyzer in their model and the analyzer is used to combine the face and hand gestures together in one model. In this model the user's hand and emotions both are required to classify the user's emotions. It detects the user's emotion using their hand Gestures surrounding the user's face [3].

Md Abdur Rahim has done optimal segmentation in hand Gesture recognition which focuses on human computer interactions[5].

author	Name	Accuracy
Kin Yun Lum* , Yeh Huann Goh, Yi Bin Lee	American Sign Language Recognition Based on MobileNetV2	95%
Samta Jain Goyal, Arvind Kumar Upadhyay, Rakesh Singh Jadon	Facial Emotion Recognition Through Hand Gesture and Its Position Surrounding The Face. Feb, 2019	90%
Md Abdur Rahim 1, Abu Saleh Musa Miah2, Abu Sayeed 3, Jungpil Shin4	Hand Gesture Recognition Based on Optimal Segmentation in Human-Computer Interaction. 2020	95%
Bharti Yogi1, Divya Godse2, Gitika Bhardwaj3	Emotion and Gesture recognition using computer vision	98%

Table-1: comparison of this work and previous work

2. PROPOSED SYSTEM

The system we proposed has uncovered a couple of Deep Learning models. The models we have developed can be used to identify anyone's emotions [reaction] and hand gesture models can be used to recognize sign language. There two different models namely Emotion recognition and Hand gesture recognition, emotion recognition model is used when someone wants to check another person's emotions. This application can identify a couple of emotions namely Sadness, happiness, anger, surprise and few more. This application provides real time face detection whenever a user wants he/she can check anyone's reactions. Transfer learning is a key concept used to develop this model and the custom layers are added to improve the performance of the models, so that users can get better and higher accuracy. The result of the live image scanned can be shown to the user in a text format. The benchmark we use is accurateness and achieved a validation accuracy of **a** minimum of 98% for the emotions model and a validation accuracy of over 95% for the gestures model.

Human emotions are very precise, they cannot be recognized accurately by a person and also it is difficult to recognize correct emotions. Also expressions of same person and different person may vary because everyone has different facial structure. To accomplish recognition and classification tasks Neural networks and machine learning can be used and can give satisfactory results.

These machine learning models provide effective and useful results in terms of pattern recognition and image recognition.

Emotion classification may be a task where a computer will predict the classification of the image. Before deep learning starts growing, tasks like image classification, speech recognition, text recognition cannot achieve human-level performance. Because of the facility of deep learning, image classification tasks can reach an individual's level of performance employing a model called Convolutional Neural Network (CNN). CNN could be a variety of deep learning model that learns representation from a picture. This model can learn from low to high-level features without human involvement. Although deep learning can do human-level performance, it needs an oversized amount of information. In our application we've got lakhs of images to get good results.

2.1 System diagram

Here is a system diagram. This explains the inner working flow of the project, and how the emotions and the hand gestures are recognized through this application. First we will see the diagram and the deep explanation about the project. This project includes data pre-processing and facial expressions are extracted.
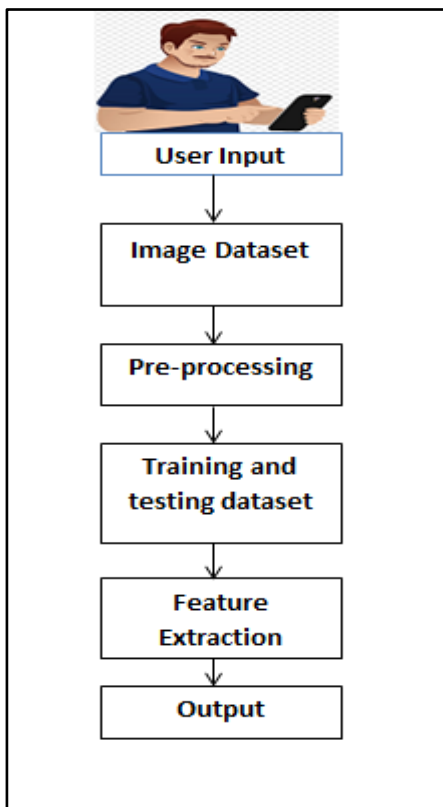

Figure-1: System Design

This model will first capture the user's facial expressions, after that the model will pre-process the image.

Dataset we have used to train models consists of images more than 1 lakh which we have obtained from k

Kaggle and used to train models. This dataset is divided into categories like angry, sad, happy etc. dataset is the essential thing in terms of training the models. They can be used to classify and validate different emotions.

When we talk about data pre-processing we usually think about large datasets with large amounts of rows and columns, but this is not the case- data can be in any format like text data, images, structured tables, audio files, videos, etc. Machines do not understand free text, it only understands the data in 0s and 1s format. So by putting lakhs of images you cannot expect your machine should understand it all and the model will be trained.

In any machine learning state data pre-processing is the state in which the data gets transformed, encoded, to bring it to such a state that a machine can easily parse it.

After that process features of every data is extracted and categorized into a suitable category. i.e happy, sad, angry etc.

Our trained model will be working while the user is using this application. The numpy, pandas, matplotlib, keras, tensorflow libraries are used to train this model. After that to train the model raw data(images) are added. Images are a crucial element of any recognition model which can be used to train the model. After that those images will be divided in the batches. To train the model MobileNetv2 model of tensorFlow is imported.

Hand Gesture Recognition

In hand gesture recognition we have done the same, we have imported large amounts of images and used those images to train the hand gesture model which can be used by the users. The spoken language binds a population and a way to communicate with the other person, the people who have listening impairment and cannot even hear or say any word. Those people may find this model useful. Sign language comes to help these deaf and physically Challenged people. Sign language using hand gestures, Facial expression and their body posture can be the medium to communicate with other people.

3. RESULT ANALYSIS

3.1 Emotion Recognition

If a user wants to check someone's expression whether he is happy, sad, disgusted, surprised, angry etc.

If you find someone with this expression, what would you recognize from this, you can use our application and check.

Figure-2: Emotion Recognition Output

We can say this person is angry with someone and our application recognizes it correctly, this is how you can check anyone's feelings and expressions through this application.

This seems interesting, so as we already said our application recognizes 7 different emotions so this was one of those 7 emotions. The image is recognized correctly. we will check it on another emotion also. The accuracy of the emotion and hand gesture recognition is accurate, so it is very useful. We will try this application on one another with different emotions.

Figure-3: Emotion Recognition Output

This is how our application will show output to the user if you check expressions of this baby, As depicted in figure 3 below.

3.2 Hand Gesture Recognition

Same for the hand recognition: if a user is not able to talk or express his feelings or want any help from the other people he can use this app to convey his/her message. Hand Gestures can be used to recognize the alphabets.

Figure-4: Input Image for Hand Recognition

Figure-5: Hand Recognition Output

Hand gestures work like this: if you do some gestures with your fingers it will show you the letter of the alphabet for that gesture. For example, this image showing 3 fingers will be the alphabet W in sign language. If you know sign language then you can use this application.

4. CONCLUSIONS

This application is very useful in every sector to provide security. The Gesture Recognition System applies to normal people and physically challenged people for non-verbal communication between human to human and computer and human. It helps and aids dumb and deaf people to live independently. It eliminates the gap among people and leads to a better society. Here the system takes the hand gestures as input and gives output in the form of text.

The physiological characteristics of the human face with relevance to various expressions such as happiness, sadness, fear, anger, surprise, and disgust are associated with geometrical structures which are restored as base matching templates for the recognition system. In

conclusion, the Detailed Design Report of the Controlling Mobile Phone via Gesture Recognition gives the definition, purpose, and scope of the project. The detailed design is explained and the constraints that it possibly faced are clarified. Data structures and the architecture of the system are explained and corresponding design issues are stated with UML diagrams that the Detailed Design will be more understandable. After the user interface is visually illustrated, the goal of the final product is more apparent.

5. REFERENCES

[1] X. Huang, J. Kortelainen, G. Zhao, X. Li, A. Moilanen, T. Seppänen, M. Pietikäinen Multi-modal emotion analysis from facial expressions and electroencephalogram Comput Vis Image Understand, 147 , pp. 114-124, feb 2021.

[2] D. Keltiner, P. EkrmanM. Lewis, J.M. Haviland Jones (Eds.), Facial expression of emotion, handbook of emotions, Gilford Press, New York, pp. 236-249.

[3] Samta Jain Goyal, Arvind Kumar Upadhyay, Rakesh Singh Jadon Facial Emotion Recognition through Hand Gesture and Its Position Surrounding The Face. Feb, 2019

[4]Meenakshi PanwarHand Gesture recognition based on shape parameters.

[5] Md Abdur Rahim 1, Abu Saleh Musa Miah2, Abu Sayeed 3, Jungpil Shin4 Hand Gesture Recognition Based on Optimal Segmentation in Human-Computer Interaction. 2020

[6] Zhou Ren, Jingjing Meng, Junsong Yuan School of EEE, Nanyang Technological University Depth Camera Based Hand Gesture Recognition and its Applications in Human-Computer-Interaction

[7] Rishabh Agrawal, Nikita Gupta Real Time Hand Gesture Recognition for Human Computer Interaction, 2016.

[8]Kin Yun Lum, Yeh Huann Goh, Yi Bin Lee American Sign Language Recognition Based on MobileNetV2

[9] Supaporn Bunrit, Nittaya Kerdprasop, and Kittisak Kerdprasop Evaluating on the Transfer Learning of CNN Architectures to a Construction Material Image Classification Task 2019

[10] Vishwanath C. Burkapalli1 and Priyadarshini C. Patil2

TRANSFER LEARNING: INCEPTION-V3 BASED CUSTOM CLASSIFICATION APPROACH FOR FOOD IMAGES 2018