Xmoji (Expression to Emoji)

M Sarath[*1], K.Yaswanth[1], G.Uday[1], S.Harsha[1], Dr M.Padmaja[*2]

1. *Student, Dept. of Computer Science & Engineering, Gitam Institute of Technology, GITAM (Deemed to be University), Visakhapatnam, India.*
2. *Assistant Professor, Dept. of Computer Science & Engineering, Gitam Institute of Technology, GITAM (Deemed to be University), Visakhapatnam, India.*

Email Id: - 1sarathmajji02@gmail.com, 2padmaja.madugula@gmail.com

Abstract: Actions, facial expressions, and poses are considered as a medium for conveying human emotions. This paper proposes a prototype system that automatically detects the human face and generates an emoji corresponding to its features. These emojis have become an essential part of online communication, brand emotion, and many more. It also leads to increasing data-science research dedicated to emoji driven storytelling. Thus a convolutional neural networks based solution is used to classify the expressions and generate emoji and gender, age, and whether the mask is present. This paper will see how we can use convolutional neural networks to develop and generate emojis based on human expressions and certain other features. This results in the development of a system that leads to the development of emojis digitally.

Keywords: Face expression detection, Gender detection, Age Detection, Mask detection, emoji.

1. Introduction

Communication is an elemental part of everyone's life. The Internet and other communication devices make it possible to participate in fast, secure, and dynamic communication. Emojis are used for the visual depiction of human emotions. What is emoji? An expression is a kind of human mental state involving many actions, including meaning, behavior, thoughts, and feelings. Facial expressions can be viewed not only as expressing emotions but also as nonverbal communication. General facial expressions can be divided into six categories: "happy, sad, neutral, disgust, surprise, and fear." The human brain can automatically recognize emotions, and now it has developed software that can recognize emotions. This technology is becoming more and more precise. It has always been able to read emotions like our brains finally. Recently, it has been developed in facial emotion detection, which has been extended to the area of expressions and emoticons based on emotions and expressions. Due to the recent covid-19 pandemic, improvements in human characteristics (such as gender, age) and whether humans wear masks have received additional benefits due to the pandemic situation of covid-19.

2. Related Work

The latest work relevant to the study can be roughly categorized as face detection, facial feature extraction, age and gender detection, mask detection. The research done on these categories is quite familiar and noteworthy.

2.1 Face Detection

Face detection is done using OpenCV, which already contains pre-trained classifiers for face, eyes, smiles, etc. This face classifier from OpenCV pre-trained classifier

will be used to detect the age and classify gender. This face detection works on only grayscale images, so the color images get converted to grayscale images. Faces contain some labeled coordinates where a rectangular region is formed using these coordinates so that face is recognized in these rectangular coordinates. Face detection can also be done in videos in which there will be several frames that are still images, in which face detection will be performed in each frame.

2.2 Facial Feature Extraction

This is the process of extracting facial features such as eyes, mouth, and nose, etc.… The faces detected in the previous stage are removed to identify other features. Facial feature extraction is crucial for the initialization of processing technologies such as facial tracking, facial expression recognition, or facial recognition. Selecting sufficient feature points represents the characteristics of the human face is successful facial feature extraction. All these methods focus on classifying general universal emotions: happiness, sadness, surprise, fear, disgust, anger.

2.3 Age Detection

Age detection is a process of directly describing the age factor of a particular person solely from the photograph containing the face of a person or even a live video feed from a webcam. The convolutional neural network model is implemented to describe the age of a specific person in the image through facial feature extraction. We can use a comparison algorithm to determine the age of a particular person in an image. We need a dataset of images including different ages and different classes for all the age groups for comparison. Then each image is compared with input images, and standard features are selected, and the age is determined through them.

2.4 Gender Detection

Gender Detection is generally essential for any visual tasks for human beings, such as social interactions, or systems such as emotion detectors, emoji generators also critically depend on the gender classification. Generally, whenever we are working with gender or age detection, such image is classifications resizing would be essential because all the images must be fed to the model in the same dimensions, i.e., size. Typically, the real-time datasets for gender classification can be limited in scope, which can cause an over-fitting problem to the model. To avoid this deep convolutional neural network can be used, and more care must be taken to prevent such issues.

2.5 Mask Detection

This pandemic veil has become a fundamental part, so this turned into a thought where we utilized the present circumstance and added mask detection as an additional element to this model while at the same time creating. We gather mask dataset with two classes of pictures marked as with_mask and without_mask organizers with their comparing pictures. The essential face mask identifier is set up by stacking the dataset, setting up the model using Tensor Flow, and serializes the face mask classifier toward the end. Another, we stack they confront mask classifier from the disk, recognize the appearances through video/live transfer through webcam, find the area of interest/region of interest(ROI) on the face, and after that apply the face mask identifier at the particular ROI and conclude whether there's mask or no_mask.

3. Proposed System

Here we will explain how the proposed system will work and the implementation and design of the system. As mentioned so far, the system is going to have the capability to detect the universal emotions shown in the human face and present the emotion in the form of an emoji along with the attributes such as gender, age. The emoji will be generated based on all the attributes.

4. Proposed Algorithm

Here the algorithm for the proposed system is determined, and its process is also being explained. The algorithm used here is Convolutional Neural Networks. A convolutional neural network is a deep learning algorithm that takes images as input and assigns labels, weights, and biases to them and predicts the image's output labels. The pre-processing required in this algorithm is less compared to other image classification algorithms. Hence, this was used in the figure 1.

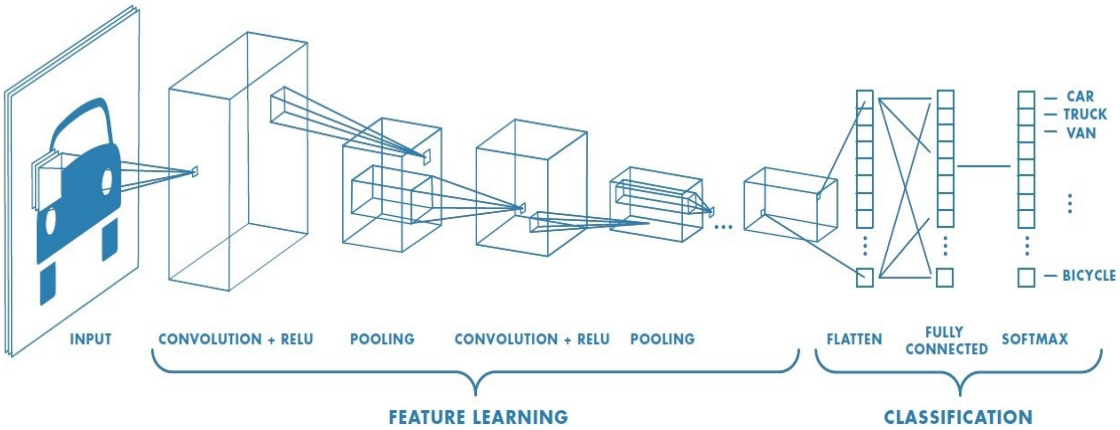

Figure 1. ARCHITECTURE OF CONVOLUTIONAL NEURAL NETWORKS

In CNN, numerous hidden layers are shown between the input and yield layers. Convolutional Neural Networks require a huge sum of data to avoid overfitting. CNN comprises of numerous covered up layers such as Convolutional Layer, ReLu Layer, Max Pooling Layer, Completely Associated Layer, etc. Utilizing these layers, the input picture is changed over to weights, and the weights are utilized to predict the unknown image.

Let's get into the information layers then-

Convolution Layer — the Kernel:-

The convolution layer is used to play out a convolutional activity of the input pictures using kernel features to highlight and to separate the picture's primary features. For instance, for figuring a 2D image, the acceptable filter size is $f \times f \times 2$, where $f = 3, 5, 7$, etc. Yet, the filter size should be smaller contrasted with that of the input image. The channel cover slides over the whole input image bit by bit. It gauges the dot product between the weights of kernel filters with the input images value, which brings about creating a 2D activation map.

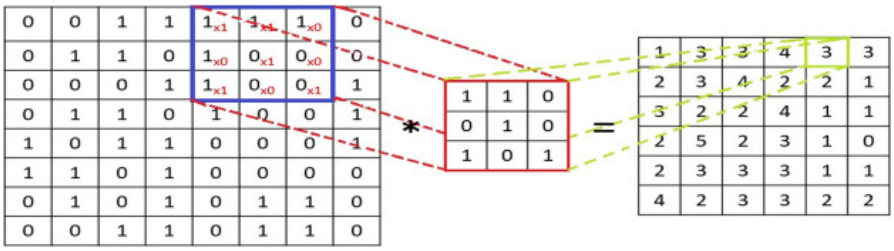

Figure 2. CONVOLUTIONAL LAYER

Pooling Layer-

This layer is the one used for reducing the spatial size of the image from the first layer. This is used in reducing the computational power for processing the data through dimensionality reduction. There are two types of pooling layers:-Max Pooling and Average Pooling. Max pooling layer is utilized to calculate the greatest or maximum value from the piece of the picture covered by the kernel. We use the max-pooling layer when training the layer in this xmoji model.

Fully Connected Layer-

The fully connected layer is formed in the last few layers of the network—the input to the fully connected layer taken from the pooling layer or convolution layer. The input taken is used to generate the labels for the images as the input.

5. Experimental Results

We have used convolutional neural networks for this emoji generator in which by giving inputs and getting outputs that redress to a great extent, a few of the results appear underneath in figure 3 and figure 4.

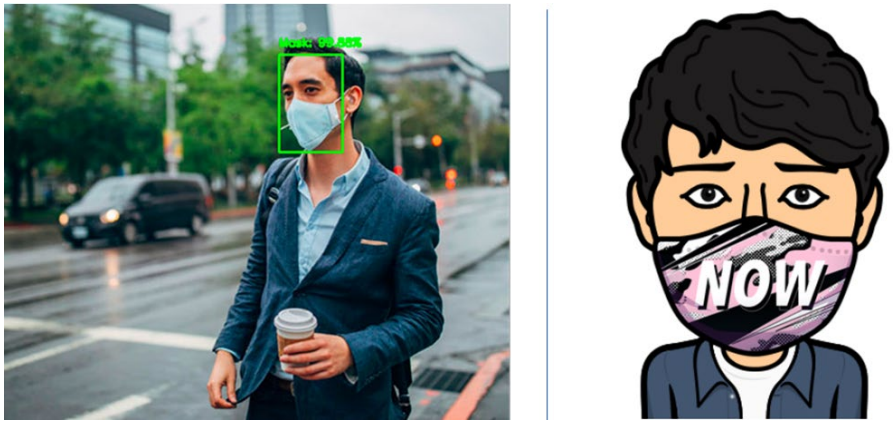

Figure 3. SAMPLE EXPERIMENTAL IMAGE - 1

ANGRY

Figure 4. SAMPLE EXPERIMENTAL IMAGE - 2

As mentioned above our system is capable of detecting all the universal emotions in the human face and generates the desired emoji. The algorithm developed performs various steps such as inputting the image, detecting the emotion along with attributes such as gender, age, and if a mask is present, no attributes are detected, directly the mask emoji is generated.

6. Developed System

In this section, we defined the design of the system, namely its appearance or what way the input will be provided. The design has been created as visual fundamental and has been actualized utilizing HTML, CSS, and JavaScript. The figure 5 and figure 6 below shows the system we thought and implemented.

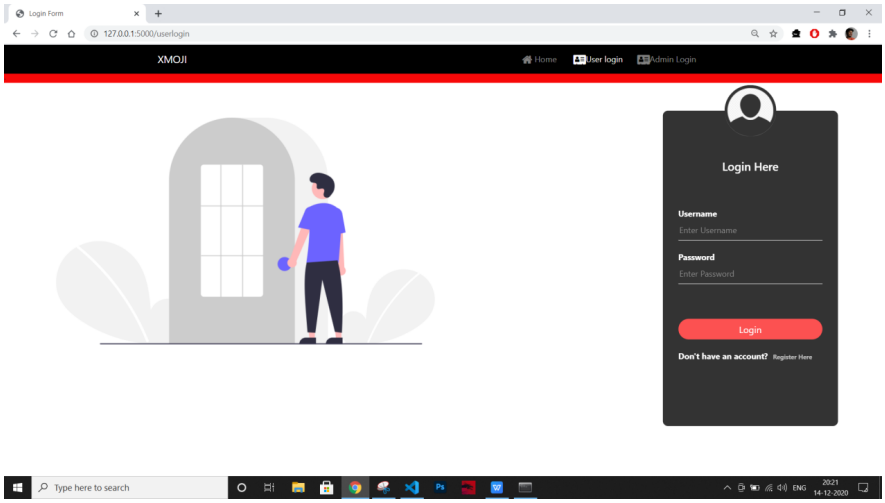

Figure 5. SYSTEM DESIGN WITH LOGIN PAGE

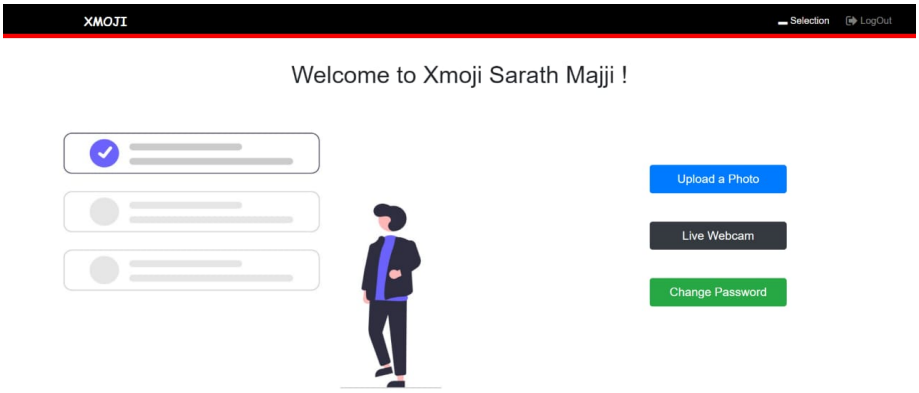

Figure 6. SYSTEM DESIGN – WELCOME PAGE

We have added three buttons to this system:-

1) Upload photo button, which allows the user who has logged in to browse an image from his system and select that image and allow the model to predict the image and generate the emoji according to the features.

2) Live webcam button, which allows the user to capture the human in front of the webcam which has been accessed using OpenCV, and live feeds the human in-front of and records the expressions of humans while they demonstrate them and then generate the emoji according to that.

3) Change password button is related to the user's interest, whether he/she is interested in changing the password for the account which he/she registered and used the same credentials for the login portal.

7. Performance Results

After implementing the convolutional neural networks for facial expressions detection, we have compared the experimental results with the original results and have given an evident performance for the datasets collected and implemented the algorithm on it. We also develop adaptive strategies to detect different features.

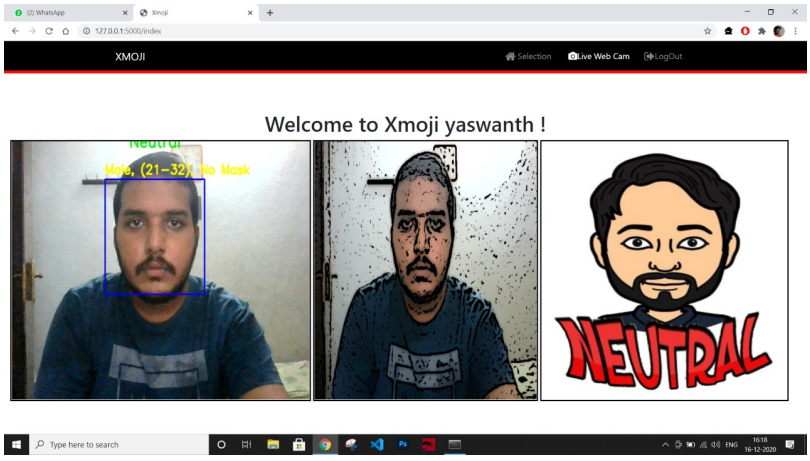

Figure 7.OUTPUT OBTAINED

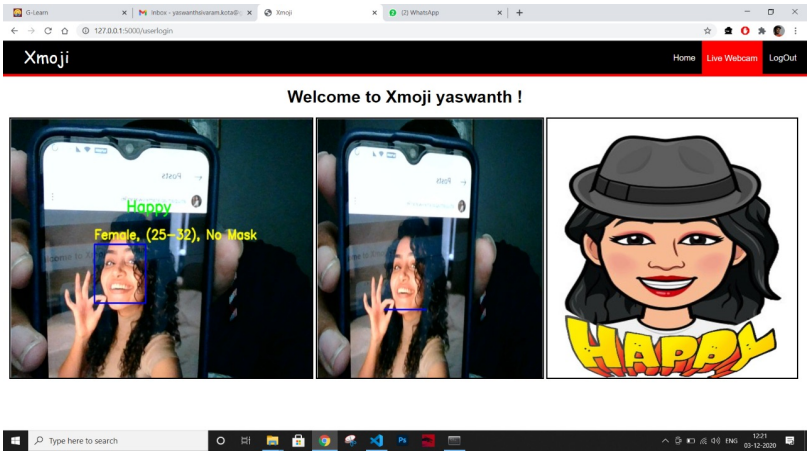

Figure8. OUTPUT OBTAINED

8. Scope of This Work

1) This research just arrangements with the outward facial appearances of people and no different pieces of the body.

2) As there are some predefined cases that do the same work, some of them are also considered.

3) Different types of test cases will be implemented using the proposed work/strategy.

4) Visualizing the experimental results and checking the performance evaluation.

5) Based upon the appropriate performance conclusion will be made and penned.

6) Future work will be drawn from the suitable working condition by considering the limitation beforehand.

Throughout the entire research, the emphasis has been made on the use of technologies and software thoroughly.

9. Conclusion

In this project, we have analyzed the limitation and scope of the project/research thoroughly and understood it up to a great extent. This paper provided a proposed model of the xmoji (expression to emoji) in virtual learning environments. The accuracy and efficiency are also considered up to a great extent. Using Haar-cascades also made us achieve the combination of accuracy and efficiency. According to sources, expression detection is the more researched topic, and also complex patterns make detecting faces hard. So it is required to study, improve and research the problem for further improvements.

10. Future Scope

To make the hypothesis and innovation of the xmoji completely meet the functional prerequisite. Our future work will focus on improving the presentation of the framework and inferring more suitable orders that might be valuable in some real time applications. Because of an absence of involvement with this examination field, we have invested the greater part of our energy understanding fundamental ideas and existing methods than building our own strategies. Nonetheless, we have

seen that we have great outcomes from convolutional neural networks. Yet, it has a few impediments like once in a while in some cases the recognition of different ascribes, for example, gender and age changes which causes to exhibit some unacceptable emoticon. We will attempt to work and diminish this change. Later on, we would likewise focus in on playing tunes utilizing neural networks dependent on the mood by the identification of expression as an additional improvement for this model as a continuation.

REFERENCES

[1] *https://www.kaggle.com/c/challenges-inrepresentation-learning-facial expression recognition challenge/data*

[2] *Lonare, Ashish, and Shweta V. Jain (2013). A Survey on Facial Expression Analysis for Emotion Recognition. International Journal of Advanced Research in Computer and Communication Engineering*

[3] *Ito, K., et al.: Age and gender prediction from face images using convolutional neural network. In: 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). IEEE (2018)*

[4] *Emotion recognition from facial expression using deep convolutional neural network D Y Liliana 2019 J. Phys.: Conf. Ser. 1193 012004*

[5] *M. B. Blaschkozoand C. T.H.Lampertor. Learning to localize objects with the help of structured output regression. In Computer Vision–ECCV 2008*

[6] *https://www.pyimagesearch.com/2020/05/04/covid-19-face-mask-detector-with-opencv-keras-tensorflow-and-deep-learning/*