Emotion Classification and Emoji Mapping using Convolutional Neural Network

Dr. Lutful Islam[1], Arif Ansari[2], Sreejith Nair[3], Azhan Patel[4]

[1]Professor, Dept. of Comp. Engg., M. H. Saboo Siddik College of Engineering, Mumbai, Maharashtra, India
[2-4]Student, Dept. of Comp. Engg., M. H. Saboo Siddik College of Engineering, Mumbai, Maharashtra, India

---

Abstract - *In this paper, we suggest an implementation of a **convolutional neural network (CNN)** to classify emotion and map emoji to the classified emotion. We validate our models through growing a real-time vision machine which accomplishes the responsibilities of **face detection**, **emotion classification** and **emoji mapping** simultaneously in a single mixed step using our proposed **CNN architecture**. After presenting the information of the training procedure setup we maintain to evaluate on standard benchmark sets. We record accuracy of 62% withinside the FER-2013 emotion dataset. We argue that the careful implementation of cutting-edge CNN architectures, using the cutting-edge regularization strategies and the visualization of previously hidden features are important for you to reduce the space amongst slow performances and real-time architectures*

***Key Words*: Convolution Neural Network (CNN), face detection, emotion classification, emoji mapping, CNN architecture**

1. INTRODUCTION

Nonverbal behavior passes on complete feeling and passionate data, to impart thoughts, oversee connections, and disambiguate importance to enhance the effectiveness of discussions[1][2]. One method to illustrate nonverbal activates is through sending emoticon, that are practical symbols (e.g. , ,) oversaw through the Unicode Consortium which might be distinguished through unicode characters and introduced through a systems font bundle.

Emoticons empower people to speak lavishly, and keeping in thoughts that are regarded as display screen designs, they may be managed as textual content structures. Other than Pohl's EmojiZoom[3] who endorse a zooming-based interface, getting into emoticon on cell phone consoles as of now expects customers to make a dedication from extensive records (one rundown for each class of emoticon) (e.g., Apple© iOS 10 emoticon console 2 in Fig. 1). This makes emoticon passage "a linear search task"[3] and given the developing wide variety of emoticons, we anticipate it may result in customer dissatisfaction. While no earlier work expressly addresses this, endeavors, for example, Emojipedia 3 characterize the requirement for higher emoticon search.

To deal with this, we advise a framework and approach to make use of customers' facial emotional expressions as framework input to clear out emoticons with the aid of using emotional classification. In spite of the truth that emoticons can deal with activities, items, nature, and distinctive symbols, the maximum typically applied emoticons are faces with specific feelings[4][5][7]. Additionally, past work has proven that emotions may be placed through assumption (Emoji Sentiment Ranking with the aid of using Novak[6]), literary notification containing emoticons show contrasts in 3-esteemed conclusion throughout stages[8], and for faces, emoticons may be positioned through valence and arousal[7].

Momentum studies examine facilities around emotional recognition. The emotions every on-occasion ease and determine connections among the people. The putting of emotions explicitly attracts out the thoughts boggling and atypical social correspondence. Social correspondence is prominent as the judgment of the alternative people's temperament depending on the emoticon. The acknowledgment of emotions may be diagnosed via different signs by the "frame language, voice intonation" simply as through means of "extra complicated techniques, for example, electroencephalography (EEG)." Nonetheless, the much less difficult, and possible method is to investigate facial expression. By noticing the facial expression, the people's mind-set and behavior are affected. Duncan clarified that "there are seven types of human feeling [that could undoubtedly be unmistakable with an assortment of meanings] across various societies". This examination consists of discovering emotional recognition through the regular collections. The emotions are distinguished as happiness, fear, disgust, anger, sadness, surprise and contempt.

This project objectives to construct a deep learning model to categorize facial expressions from the images. Then we are able to map the labeled emotion to an emoji or an avatar.

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 05 | May 2021 www.irjet.net p-ISSN: 2395-0072

Fig -1: Apple© iOS 10 emoji keyboard within iMessage

2. LITERATURE SURVEY

Today, the most widely recognized method of communication among individuals is virtual platforms, regardless of whether utilizing the web or phones (Vissers and Stolle, 2014). The current age utilizes online applications and stages to impart and trade discussions. Be that as it may, imparting emotions is troublesome. Thus, little and straightforward pictures, also called emoticon characters, are utilized to enhance feelings when utilizing composed language (Yeole, Chavan, and Nikose, 2015). They have extraordinary semantic and passionate highlights, but on the other hand are firmly identified with marketing, law, medical care and numerous different zones. The examination on emoticons has become an interesting issue in the scholastic field, and then some and more researchers from the fields of computing, communication, marketing, behavioral science, etc are considering them. Emoticon characters are turning out to be increasingly more promoted hence the variety of these characters has expanded. Be that as it may, the current existing emoticon characters are restricted to predetermined characters. Also, these characters need intricacy and variety. To customize emoticon characters, this examination investigated techniques for clients to "emojify" their photos. This study not just permits individuals to make altered and exceptional methods of imparting emotions, yet in addition makes justification for additional upgrades of emoticon characters.

It is roused by two discoveries from the literature: that an essential capacity of emoticons is to communicate emotions, and that most emoticons utilized are face emoticons. Cramer tracked down that 60% of their dissected messages by US members were emoticons utilized for communicating emotions. In an Instagram emoticon study, faces represented 6 of the main 10 emoticons utilized, giving additional proof that individuals as often as possible use emoticons to communicate emotions. Moreover, as indicated by a 2015 SwiftKey report, faces represented near 60% of emoticon use in their examination of billions of messages. At last, in a subjective report from Lee on emoji sticker utilization, they tracked down that these stickers were utilized fundamentally for communicating emotions.

2.1 Survey of Existing System

Multimodal User Interfaces and Emoji Entry

Identified with our methodology, Filho et al.[9] augmented text chatting in mobile phones by adding automatically identified facial expression responses utilizing computer vision methods, bringing about an emotion upgraded mobile talk. For utilizing the client's face as information, Anand et al.[10] investigated a utilization instance of an eBook reader application wherein the client plays out certain facial expressions normally to control the gadget. Concerning emoticon section, Pohl et al.[3] proposed another zooming keyboard for emoticon passage, EmojiZoom, where clients can see all emoticons immediately. Their strategy, which was tried in a convenience concentrate against the Google keyboard, showed 18% quicker emoticon entry.

Emoji and Emoticon Communication

The smallness of emoticons diminishes the exertion of contribution to communicate feelings, yet in addition serves to change message tone, increment message commitment, oversee discussions and keep up friendly connections. Additionally, emoticons don't have language obstructions, making it feasible for clients across nations and social foundations to impart. In an examination by Barbieri et al.[11], they tracked down that the general semantics of the subset of the emoticons they considered is protected across US English, UK English, Spanish, and Italian. As approval of the convenience of planning emoticons to feelings, primer examinations announced by Jaeger et al.[12] recommend that emoticons may have potential as a strategy for direct estimation of emotional relationship to food varieties and drinks.

2.2 Limitations of Existing system

Emoji Mis-interpretation

As of late, Miller et al.[13] exhibited how a similar emoticon looks diversely across gadgets (iPhone, Android, Samsung) and is in this way contrastingly deciphered across clients. In any event, when members were presented to a similar emoticon delivering, they differed on whether the assumption was positive, unbiased, or negative around 25% of the time. In a connected primer investigation, Tigwell et al.[7] discovered clear contrasts in emoticon valence and arousal evaluations between stage matches because of contrasts in their design, just as varieties in appraisals for emoticons inside a stage. With regards to our work, this features the need to represent different translations, where an emoticon can be delegated having a place with at least one emotion class.

3. PROPOSED SYSTEM

3.1 Details of Hardware & Software

Hardware Requirements:

Laptop (32-bit or 64-bit architecture, 2+ GHz CPU, 4 GB RAM.), Camera (8MP & above)

Software Requirements:

- Operating System: Windows 7/8/8.1/10, Linux
- Database: Firebase/MySQL/MongoDB
- Tools and Framework: OpenCV, Tensor flow, Flask, Bootstrap
- Language Requirement: Python, HTML, CSS.
- Server: Locally hosted

Technology Used:

Image pre-processing, TensorFlow, Keras, OpenCV and Google Collab, Deep Learning, Image Processing.

3.2 System

This project will consist of creating two important inter-dependent modules. They can be described as follows:

- Desktop Website (UI)
- Server (Database and Deep Learning model)

Server: (Deep Learning Phase)

- Preprocess the input image from the train dataset collected from sources such as Kaggle as well as creation of the dataset using one's own facial expression.
- Process it in R-CNN/CNN for Facial Expression Detection and Emotion Classification.
- Update Facial Data in Database for corresponding session with output results from DL model
- Send corresponding update to Desktop Website

Server (Database):

- Store User related details such as Name, Google Account ID.
- Store results of submission with classified emoji, Emotion and Expression.
- Database will be implemented using MongoDB/Firebase.

Desktop Website:

- Provide User-Interface to the user.
- Generate and maintain a profile of users for future use in integration.
- Create emoji based on expression.

3.3 Analysis/Framework/Algorithm

Tensorflow

TensorFlow is a free and open-supply software program library for device learning. It may be used throughout various responsibilities however has a selected recognition on schooling and inference of deep neural networks. Tensorflow is a symbolic math library primarily based totally on data flow and differentiable programming. It is the quickest and only manner to do image reputation to your computer or laptop with none GPU due to the fact it's far simply an

API and your CPU is ideally sufficient for this. TensorFlow affords the opportunity to conform a pre-skilled version to new training of records with numerous advantages.

Create Training Data

In Preparation for schooling information we're going to seize video via webcam the usage of a python application along with opencv and imutils additionally enforce the HAAR cascade classifier and create a photograph dataset with the aid of using shooting the frames of specified emotion as a facial expression. Another opportunity is to download pics from Kaggle that have already described a dataset for emotions, including the FER2013 dataset.

OpenCV

OpenCV (Open Source Computer Vision Library) is an open source computer imaginative and prescient and system getting to know software program library. OpenCV became constructed to offer a not unusual place infrastructure for laptop imaginative and prescient programs and to boost up using system notion in business products. Being a BSD-certified product, OpenCV makes it smooth for organizations to make use of and adjust the code. It helps C++, Python, Java, Android SDK, etc. The library has more than 2500 optimized algorithms, which incorporates a complete set of each traditional and state-of-the-art laptop imaginative and prescient and system getting to know algorithms. These algorithms may be used to locate and understand faces, discover objects, classify human movements in videos, music digital digicam movements, music shifting objects, extract 3-d fashions of objects, produce 3-d factor clouds from stereo cameras, sew pix collectively to produce an excessive decision photo of a whole scene, discover comparable pix from a photo database, get rid of purple eyes from pix taken the use of flash, observe eye movements, understand surroundings and set up markers to overlay it with augmented reality, etc. It has C++, Python, Java and MATLAB interfaces and helps Windows, Linux, Android and Mac OS. OpenCV leans in general in the direction of real-time imaginative and prescient programs and takes advantage of MMX and SSE commands whilst available. A full-featured CUDA and OpenCL interface is being actively advanced properly now. There are over 500 algorithms and approximately 10 instances of many features that compose or guide the ones algorithms. OpenCV is written natively in C++ and has a template interface that works seamlessly with STL containers.

3.4 Design Details

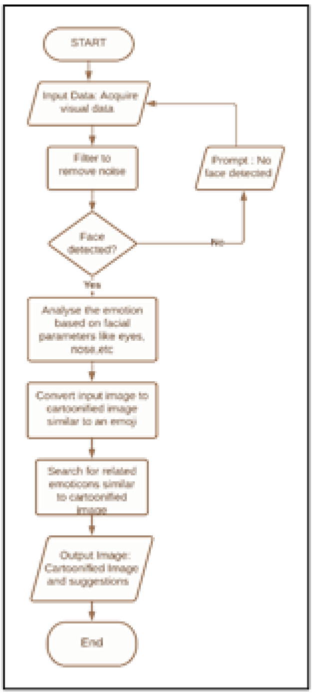

Fig -2: Flowchart

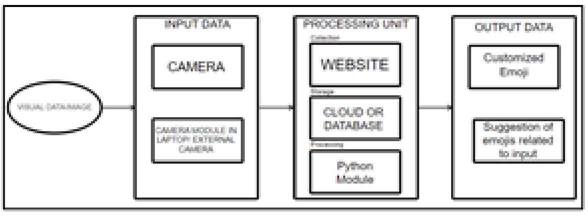

Fig -3: Data Flow Diagram

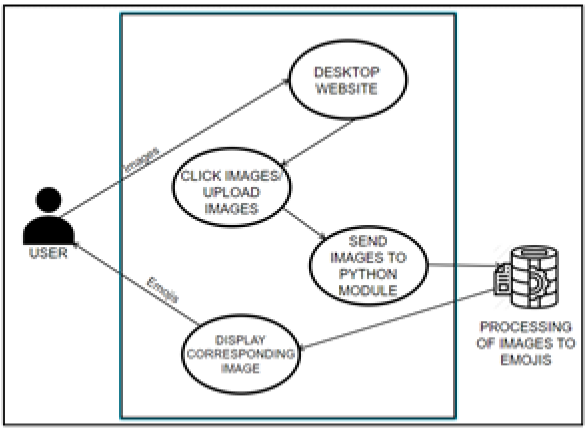

Fig -4: Use Case Diagram

3.5 Steps followed by the proposed system

Step 1: Input the dataset

Below are sample images from the FER 2013 dataset that are used to classify emotions. These Images are categorized based on the emotion shown in the facial expressions such as happiness, neutral, sadness, anger, surprise, disgust, fear.

Fig -5: Sample dataset of emotion "happy"

Step 2: Data pre-processing and applying augmentation Strategies.

Image data augmentation is used to expand the training dataset in order to improve the performance and ability of the model to generalize. Images are rescaled from [0,255] to [0,1] using the ImageDataGenerator python module. Benefits of this are:

- Treat all images in the same manner: some images are high pixel range, some are low pixel range. The images are all sharing the same model, weights and learning rate. The high range image tends to create stronger loss while low range creates weak loss, the sum of them will all contribute to the back propagation update.

- Using typical learning rate: when we reference the learning rate from other's work, we can directly reference their learning rate if both works do the scaling preprocessing over images data set. Otherwise, higher pixel range image results in higher loss and should use a smaller learning rate, lower pixel range image will need a larger learning rate.

Step 3: Neural Network architecture.

After pre-processing the dataset, the next step is to build a convolutional neural network. The convolution layer consists of the input layers, the hidden layers and the output layer. Depending on the architecture that is built in the neural network we add convolutional layers with filters. Below is the architecture of the neural network.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 48, 48, 64)        640
_________________________________________________________________
batch_normalization (BatchNo (None, 48, 48, 64)        256
_________________________________________________________________
activation (Activation)      (None, 48, 48, 64)        0
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 24, 24, 64)        0
_________________________________________________________________
dropout (Dropout)            (None, 24, 24, 64)        0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 24, 24, 128)       204928
_________________________________________________________________
batch_normalization_1 (Batch (None, 24, 24, 128)       512
_________________________________________________________________
activation_1 (Activation)    (None, 24, 24, 128)       0
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 12, 12, 128)       0
_________________________________________________________________
dropout_1 (Dropout)          (None, 12, 12, 128)       0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 12, 12, 512)       590336
_________________________________________________________________
batch_normalization_2 (Batch (None, 12, 12, 512)       2048
_________________________________________________________________
activation_2 (Activation)    (None, 12, 12, 512)       0
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 6, 6, 512)         0
_________________________________________________________________
dropout_2 (Dropout)          (None, 6, 6, 512)         0
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 6, 6, 512)         2359808
_________________________________________________________________
batch_normalization_3 (Batch (None, 6, 6, 512)         2048
_________________________________________________________________
activation_3 (Activation)    (None, 6, 6, 512)         0
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 3, 3, 512)         0
_________________________________________________________________
dropout_3 (Dropout)          (None, 3, 3, 512)         0
_________________________________________________________________
flatten (Flatten)            (None, 4608)              0
_________________________________________________________________
dense (Dense)                (None, 256)               1179904
_________________________________________________________________
batch_normalization_4 (Batch (None, 256)               1024
_________________________________________________________________
activation_4 (Activation)    (None, 256)               0
_________________________________________________________________
dropout_4 (Dropout)          (None, 256)               0
_________________________________________________________________
dense_1 (Dense)              (None, 512)               131584
_________________________________________________________________
batch_normalization_5 (Batch (None, 512)               2048
_________________________________________________________________
activation_5 (Activation)    (None, 512)               0
_________________________________________________________________
dropout_5 (Dropout)          (None, 512)               0
_________________________________________________________________
dense_2 (Dense)              (None, 7)                 3591
=================================================================
Total params: 4,478,727
Trainable params: 4,474,759
Non-trainable params: 3,968
_________________________________________________________________
```

Fig -6: summary of the model

Step 4: Accuracy and loss.

With the implementation of the above neural network, we got accuracy of 77% and loss of 0.36 on training data and accuracy of 62% on validation data.

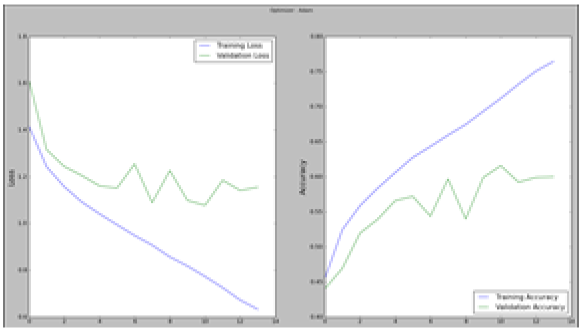

Fig -7: training accuracy and loss against validation accuracy and loss.

3.6 Output of the Implemented Model.

Below is the real time result of the implemented model. Model classifies the emotion based on facial expression and maps the emotion to the emoji or avatar.

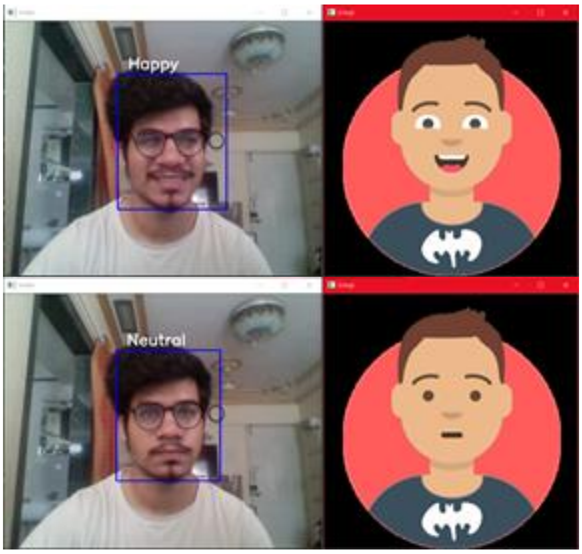

Fig -8: Real-time result of emotion classification.

G. Future Implementation.

Implement this emotion classification and emoji mapping using CNN into a website so people can access and create their own emoji. Customization of emoji based on user requirement as well as extra emotion customization.

4. CONCLUSION

Emojis are approaches to signify nonverbal cues. These cues have come to be a crucial part of on-line chatting, product review, logo emotion, and plenty of more. It additionally led to growing information technology studies devoted to emoji-pushed storytelling. We construct a convolution neural community structure and educate the version on FER2013 dataset for emotion popularity from photographs additionally with improvements in computer –vision and deep learning, it's far viable to come across human feelings from photographs. In this project, we are able to classify human facial expressions to clear out and map corresponding emojis.

REFERENCES

[1] Mark L. Knapp and Judith A. Hall. 2010. Nonver-bal Communication in Human Interaction (7 ed.).Wadsworth: Cengage Learning, Bosten, USA

[2] Jessica L Tracy, Daniel Randles, and Conor M Steck-ler. 2015. The nonverbal communication of emotions.Current Opinion in Behavioral Sciences 3 (2015), 25– 30. DOI:http://dx.doi.org/10.1016/j.cobeha.2015.01.001 Social behavior.

[3] Henning Pohl, Dennis Stanke, and Michael Rohs.2016. EmojiZoom: Emoji Entry via Large OverviewMaps 😄🔍. In Proc. MobileHCI'16. ACM, New York, NY, USA, 510–517. DOI:http://dx.doi.org/10.1145/2935334.2935382

[4] Henriette Cramer, Paloma de Juan, and Joel Tetreault.2016. Sender-intended Functions of Emojis in USMessaging. In Proc. MobileHCI '16. ACM, New York,NY, USA, 504–509. DOI:http://dx.doi.org/10.1145/2935334.2935370

[5] Joon Young Lee, Nahi Hong, Soomin Kim, JonghwanOh, and Joonhwan Lee. 2016. Smiley Face: Why WeUse Emoticon Stickers in Mobile Messaging. In Proc.MobileHCI '16. ACM, New York, NY, USA, 760–766.DOI:http://dx.doi.org/10.1145/2957265.2961858

[6] Petra Kralj Novak, Jasmina Smailovi´c, Borut Sluban,and Igor Mozetiˇc. 2015. Sentiment of Emojis. PLOSONE 10, 12 (12 2015), 1–22. DOI:http://dx.doi.org/10.1371/journal.pone.0144296

[7] Garreth W. Tigwell and David R. Flatla. 2016. OhThat's What You Meant!: Reducing Emoji Misunder-standing. In Proc. MobileHCI '16. ACM, New York, NY,USA, 859–866. DOI:http://dx.doi.org/10.1145/2957265.2961844

[8] Channary Tauch and Eiman Kanjo. 2016. The Role Of Emojis in Mobile Phone Notifications. In Proceed-ings of the 2016 ACM International Joint Conferenceon Pervasive and Ubiquitous Computing: Adjunct (Ubi-Comp '16). ACM, New York, NY, USA, 1560–1565. DOI:http://dx.doi.org/10.1145/2968219.2968549

[9] Jackson Feijó Filho, Thiago Valle, and Wilson Prata. 2014. Non-verbal Communications in Mobile Text Chat: Emotion-enhanced Mobile Chat. In Proc. MobileHCI '14. ACM, New York, NY, USA, 443–446. DOI: http://dx.doi.org/10.1145/2628363.2633576

[10] B. Anand, B. B. Navathe, S. Velusamy, H. Kannan, A. Sharma, and V. Gopalakrishnan. 2012. Beyond touch: Natural interactions using facial expressions. In proc. CCNC ' 12. 255–259. DOI: http://dx.doi.org/10.1109/CCNC.2012.6181097

[11] Francesco Barbieri, German Kruszewski, Francesco Ronzano, and Horacio Saggion. 2016. How Cosmopolitan Are Emojis?: Exploring Emojis Usage and Meaning over Different Languages with Distributional Semantics. In Proc. MM '16. ACM, New York, NY, USA, 531–535. DOI:http://dx.doi.org/10.1145/2964284.2967278

[12] Sara R. Jaeger, Leticia Vidal, Karrie Kam, and GastÃ3n Ares. 2017. Can emoji be used as a direct method to measure emotional associations to food names? Preliminary investigations with consumers in {USA} and China. Food Quality and Preference 56, Part A (2017), 38 – 48. DOI:http://dx.doi.org/10.1016/j.foodqual.2016.09.005

Hannah Miller, Jacob Thebault-Spieker, Shuo Chang, Isaac Johnson, Loren Terveen, and Brent Hecht. 2016. "blissfully happy" or "ready to fight": Varying interpretations of emoji. In Proc. ICWSM 2016. AAAI press, 259–268