RESEARCH ARTICLE

RSP Science Hub

Emotica.AI - A Customer feedback system using AI

Ayush Kumar Bar[1], Akankshya Rout[1], Avijit Kumar Chaudhuri[2]

[1]*Department of Computer Science Engineering, Techno Engineering College Banipur, West Bengal, India*

[2]*Assistant Professor, Computer Science and Engineering, Techno Engineering College, Banipur, Kolkata,West Bengal, India*

Email: ayush.kumar.bar.official@gmail.com

Abstract

Our lives are being significantly impacted by the rapid development of wireless technology and mobile gadgets on this day. The digital economy demands that services be developed almost instantly while also paying close attention to client feedback. It becomes difficult to manage and analyse the information gathered about products from customers. Successful businesses typically gather reasonable input on customer behaviour, comprehend their clients, and maintain ongoing contact with them. But it's not an easy task to keep a record of each and every customer' feedback on a daily basis. Also, everyone is not intended to provide clear feedback whether the product was satisfactory or not. It is a very difficult and time-consuming task to analyse the data collected manually. Companies need automation of customer feedback processing in order to quickly use the data that has been collected and analyse consumer feedback. To proceed with the problem and through much research we came across a solution, Emotica.AI, an emotion recognition system which can overcome this situation in real time. Emotion recognition plays an important role in building interpersonal relationships. Speaking, making facial expressions, gesturing, or writing are all ways that people directly or indirectly convey their feelings. Now that AI has mastered the power of learning, it is capable of treating anything just like a human would. The proposed model is built with Haar-Cascade Algorithm and classified with CNN and is able to recognise the emotions of multiple faces in a real time scenario. Accuracy of this model is around 76% is achieved for seven emotions on a real -time basis. Our goal is to develop a real time implementation of an emotion detection system with better accuracy and make it more reliable for businesses and other purposes.

1. Introduction

Facial expressions are visible indications of a person's emotional and mental state, goals, personality, and probable psychiatric disorders. They operate as a channel of communication in social situations. It has seen significant advancement in recent decades after years of study. Because of the diverse and varied nature of expressions, it is still difficult to reliably identify them even after substantial breakthroughs.

Emotica.AI is an AI system that permits a program to "examine" the sentiments on a human face by utilizing sophisticated image dispensation. Businesses are experimenting with employing sophisticated algorithms and image processing methods to analyse films or photos of people's faces in order

to better comprehend their emotional states (X. Li and Huang) . These methods have substantially advanced over the past ten years, which has led to the creation of software that is quite good at identifying emotions. In addition to recognizing basic emotions like happy, sad, surprise, anger, etc. based on a person's facial expressions, emotion detection software may also spot "micro expressions" or small body language clues that researchers say might betray a person's sentiments unknowingly (Haque, X. Li, and X. Li) . By gathering data on consumers' emotions and preferences using cutting-edge face recognition technology, businesses can better understand their target audience. This data may then be utilised to tailor their offers to better match their consumers' requirements and aspirations, while also evaluating the efficiency of their marketing and customer service tactics (Soni and Khanna) . In the end, this may result in higher client satisfaction and corporate success. This may promote creativity, aid in the creation of new products, and cultivate a foundation of devoted customers. The algorithm categories the same person's facial expressions according to their fundamental emotions, which include anger, contempt, fear, happiness, sorrow, and surprise. By utilizing multiple methods, including eye gaze tracking, facial expression detection, and cognitive modeling, this system's primary goal is to enable effective interaction between humans and machines (Sarode and Nimbhorkar) . The system's goal is to improve the human-machine interface by allowing the machine to better comprehend the user's intents, emotions, and preferences and respond in a more natural and intuitive manner. Here, facial emotion recognition and categorization can be a practical technique for promoting organic contact between people and robots. Machines can better grasp human emotions and intentions by analysing facial expressions, enabling more intuitive and individualised interactions. (Y. Li et al.) This can be particularly beneficial in applications such as virtual assistants, online customer service, and other forms of human-machine communication. The intensity of facial expressions varies from person to person and is also influenced by factors such as age, gender, and the size and shape of the face. Additionally, even the expressions of the same person can vary over time. However, recognizing facial expressions is a challenging task due to the inherent variability of facial images caused by factors such as variations in illumination, pose, alignment, and occlusions (Smith) . There have been many surveys on facial feature representations for face recognition and expression analysis that address the challenges and possible solutions in detail.

2. Literature Review

In this section we will discuss the previous work done in this field. As discussed, Emotion recognition systems (ERS) aim to detect and classify the emotional state of a user during interaction with a computer system. To proceed with our model, we went for various researches i.e.,provides an application of feature extraction of facial expressions with a combination of neural networks for the recognition of different facial emotions using Luigi Rosa's Eigen Expressions for Facial Expression Recognition simulator with an accuracy of 97%. (Goshvarpour, Abbasi, and Goshvarpour) (Agrafioti, Hatzinakos, and Anderson) (Rattanyu and Mizukawa) achieved an accuracy of 80-100% using ECG model using different classifiers i.e., KNN, SVM. There are several drawbacks with the proposed systems i.e., first it works on existing dataset not on the real-world problems. Second, it is unable to measure emotions. Third, there is no GUI provided to give a real time outlook. The major aim of Emotica.AI is to work as a feedback system for people dynamically. The GUI solves the real-life emotions dynamically, it is not required to check the stored data (Chowanda) . The proposed is also able to store new emotions by creating a new class and get trained according to that.

3. Problem Statement

Companies get massively charged for the feedback & survey services offered by the third-party companies, which excludes the smaller companies from getting these facilities. Poor user - feedback ratio of products results in lack of reviews and thus lack of proper R & D related to the problems faced by the customers. User Privacy is a major concern while using facial recognition technology.

4. Motivation

Large corporations make huge investments to get feedback and surveys for their product satisfaction. Such corporations spend huge amounts of money.

If we will be able to provide our stakeholders with a system which can track the emotions of their customers automatically, then we can hear the real voice of their customers, whether they are actually satisfied (Majumdar and Avabhrith) . It can easily benefit by monitoring customer behaviour to their products or staff service by using emotion recognition systems (Kokate and Kadam) . This can give them proper data to improve their products and services and optimize their business model accordingly.

5. Objectives

The objectives of the system development are as follows-

- To give businesses a more effective and efficient means to comprehend their consumers' preferences and wants, a facial expression recognition customer feedback system should be created. With the use of this system, businesses may learn more about the opinions of their clients and utilise that knowledge to enhance their products and services.

- The second objective of the system development is to acknowledge the sellers about which product is being chosen by the customers more in a dynamic environment. This information can help sellers to improve their marketing strategies and better understand their customers' preferences.

- Providing real-time feedback without requiring consumers to complete any forms or surveys may be a more effective and easy approach to get customer feedback, which may ultimately save time and effort for both customers and merchants.

- To assist small companies and industries in obtaining insightful client feedback that would enable them to enhance their goods and services and develop in the marketplace.

6. Methodology

The system is designed to recognize human faces and classify facial expressions into seven basic categories. The supervised learning approach involves training the system with a large dataset of images labelled with the corresponding expression categories. During the testing phase, the system is evaluated on new images to assess its accuracy and performance in recognizing facial expressions.

6.1. Video Acquisition

Videos used for the Emotica.AI System are real-time or dynamic and captured using a camera. The

camera resolution can vary. If the frames in the videos are of low resolution, then it gets upscaled if it's of higher resolution then it is downscaled to 1920*1080 pixels.

6.2. Pre-Processing

Image pre-processing is a crucial step in facial expression recognition to ensure that the images are standardized and noise-free, so that the subsequent feature extraction and classification steps can be more accurate This include following steps:Converting frames of videos into greyscale image

1. Converting frames of videos into greyscale image

2. Noise Reduction

3. Image Sharpening

6.3. Face Detection

Face Detection is useful in the detection of facial images. The Viola-Jones face detection algorithm is a popular method for detecting faces in an image. It uses a Haar-like feature cascade classifier to detect faces, which involves the calculation of the difference in intensity between adjacent rectangular regions of an image. To find areas where a face could be present, this procedure is repeated over the whole picture. Once these areas have been located, they are further examined to see if a face actually exists there. This algorithm's implementation is included in OpenCV, a well-liked computer vision library.

6.4. Feature Extraction

Emotica.AI is using CNN for facial feature extraction. CNNs are known to be very effective in computer vision tasks, and particularly in image classification tasks. Here the input images have a resolution of 48x48 pixels, and that there are 7 emotions that are being predicted ie. Angry, Disgust, Fear, Happy, Sad, Surprise, Neutral. Deep learning models are frequently trained with a batch size of 64, which facilitates faster training by processing several pictures concurrently. As it helps to guarantee that all characteristics are on a same size and makes it simpler for the model to learn from the data, standardising the input data is a typical pre-processing step in machine learning. Convolutional, pooling, and fully
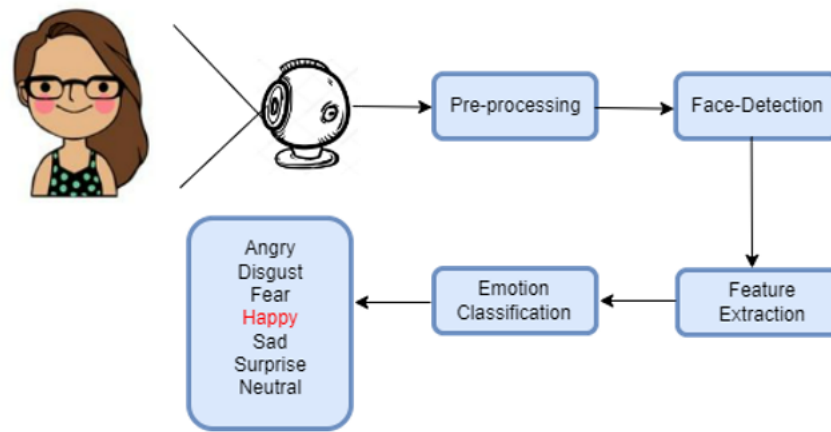
FIGURE 1. Steps for classification

connected layers are some examples of the layers that may be used while building a CNN. Moreover, each layer's characteristics, including the quantity and size of filters, must be supplied. The general architecture of the CNN is built using these layers and parameters, which affects how well it completes the task at hand.

Sequential () - A sequential model in Keras is a stack of layers that is connected to one another and may be added one layer at a time using the add () technique. Up until the final output is created, each layer's output serves as its counterpart's input. While it is the most basic model type in Keras, it is nonetheless capable of handling a wide range of issues, particularly those related to deep learning.

model.add(Conv2D()) - The 2D Convolutional layer performs the convolution operation which involves sliding a small window called a kernel over the input image and multiplying the values in the kernel with the corresponding values in the input image, then summing up these values to produce a single output value. This process is repeated for each location in the input image, producing a new output image. The ReLU activation function sets all negative values to zero, and leaves positive values unchanged. This helps to introduce nonlinearity in the model and is commonly used in deep learning architectures.

model.add(BatchNormalization()) - It conducts the batch normalisation process on the inputs to the following layer so that our inputs are organised into a certain scale, such as 0 to 1, rather than being dispersed across the model.

model.add(MaxPooling2D()) - MaxPooling is a type of pooling operation where the maximum value

of a rectangular neighbourhood is taken and used as the representative value for that neighbourhood. This helps to reduce the spatial dimensionality of the data and extract the most important features while preserving the most prominent patterns. In the current model, MaxPooling is being used with a window size of 2x2 and 2x2 strides to further reduce the size of the features extracted by the convolutional layer.

model.add(Dropout()) - As explained above Dropout is a regularization technique used to prevent overfitting in neural networks. During training, neurons in the network are randomly deactivated, or "dropped out," with a certain probability. This forces the network to learn more robust features and reduces the impact of any single neuron. By doing so, dropout helps prevent the model from relying too much on any particular feature and reduces overfitting.

model.add(Flatten()) - This just flattens the input from ND to 1D and does not affect the batch size.

model.add(Dense()) - The Dense layer in Keras implements the operation: output = activation(dot(input, kernel) + bias), where dot represents the dot product between the input tensor and the kernel (weights), and bias is an optional bias vector. The activation function is applied to the dot product result. This layer is used for the final classification or regression prediction in a neural network model.

output = activation (dot(input, kernel) which means the output layer takes the features learned by the previous layers and produces the final output, which in this case is a probability distribution over the possible classes (i.e., the seven emotions in this case). The activation function used in the output

layer depends on the task being solved

The model is trained with categorical cross-entropy which is a common loss function used in multiclass classification problems with Adam optimizer which is an optimization algorithm that adapts the learning rate during training, and is often used as a good general-purpose optimizer. Accuracy is a commonly used metric to evaluate the performance of the model during validation, as it provides the percentage of correct predictions out of all predictions made.

6.5. Emotion Classification

Emotion classification is as important as feedback meant to you. We came across many emotion classification mechanisms and we chose the best possible mechanism for this problem. Hence, we are using a model trained in CNN (Convolutional Neural Network) for this purpose:

```python
model = Sequential()

#1st CNN layer
model.add(Conv2D(128,(24,24),padding = 'same',input_shape = (48,48,1)))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size = (2,2)))
model.add(Dropout(0.25))

#2nd CNN layer
model.add(Conv2D(512,(12,12),padding = 'same'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size = (2,2)))
model.add(Dropout (0.25))

#3rd CNN layer
model.add(Conv2D(512,(6,6),padding = 'same'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size = (2,2)))
model.add(Dropout (0.25))

#4th CNN layer
model.add(Conv2D(512,(3,3), padding='same'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Dropout(0.25))
```

FIGURE 2. CNN Layers

6.6. Framework

The Tkinter and Pillow libraries from Python served as the foundation for this application's user interface, and Keras was utilised for image processing. CV2 for camera modules with real-time AI processing of the frames. Specifically for computer vision, CV2 is a library for OpenCV.

The interface is very handy and easy to understand. The application processes the output as the

```python
model.add(Flatten())

#Fully connected 1st layer
model.add(Dense(512))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(Dropout(0.25))

# Fully connected layer 2nd layer
model.add(Dense(512))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(Dropout(0.25))
```

**FIGURE 3. Fully Connected layers**
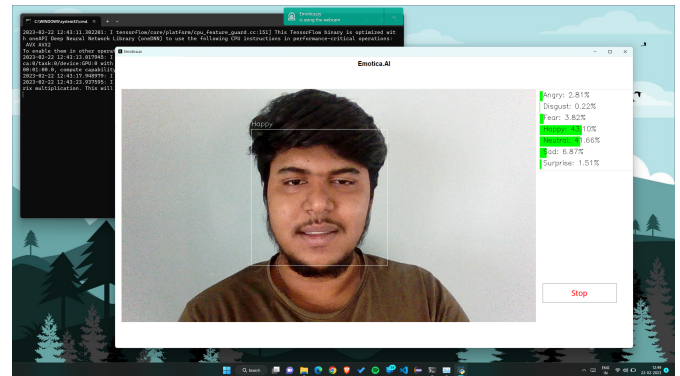
FIGURE 4. Software Interface

emotion class which has highest probability among them. The emotion gauge dynamically changes the values for different instances of time.

7. Experimental Results

After generating the model, we then further did a validation and accuracy check if the model is trained perfectly or not. Through this rigorous training and testing we received an accuracy of 76% which is better for the purpose this application is serving.
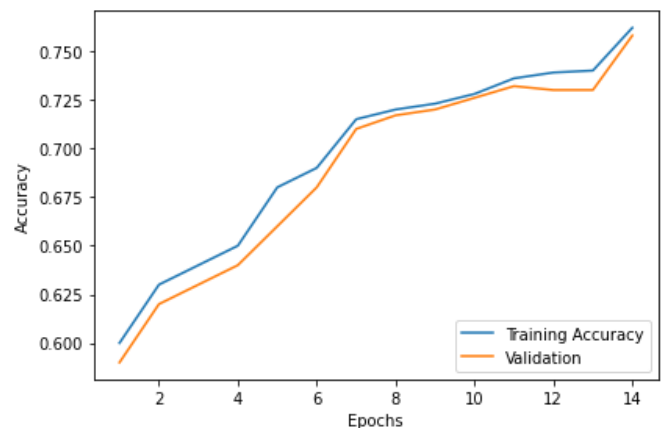
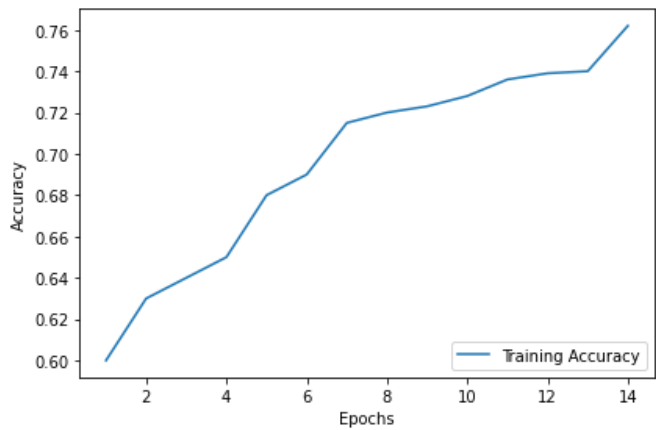**FIGURE 5. Accuracy and Validation per epoch**

FIGURE 6. Accuracy per epochs

This result is good for commercial purposes as we don't need the highest accuracy but enough accuracy to recognise the emotions to perform the purpose of the application.

```
              precision    recall  f1-score   support

       Angry       0.84      0.77      0.80       100
     Disgust       0.76      0.63      0.69       100
        Fear       0.69      0.73      0.71       100
       Happy       0.82      0.84      0.83       100
     Neutral       0.74      0.83      0.78       100
         Sad       0.71      0.75      0.73       100
    Surprise       0.80      0.79      0.79       100

    accuracy                           0.76       700
   macro avg       0.76      0.76      0.76       700
weighted avg       0.76      0.76      0.76       700
```

FIGURE 7. Overall report

Here, the result shows that each of the emotions are well trained and performed well during testing and validation. Now, we have an idea how will it perform against same classes of emotion, we want to analyse how one class performs against other classes.

With this confusion matrix we can determine how it predicted one single emotion against other emotions in terms of similarities and differences or how many times the model gets confused with the emotions.

8. Conclusions

Emotica.AI has the potential to greatly improve customer feedback systems by providing real-time emotion recognition. A 76% accuracy is quite impressive for a facial emotion recognition model, especially when dealing with real-time video. With
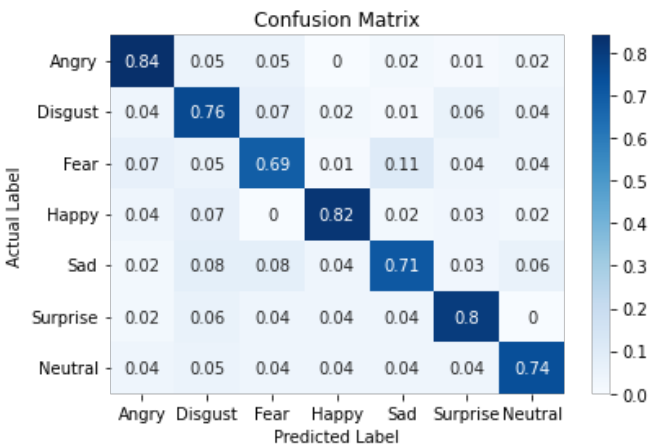

FIGURE 8. Confusion Matrix

further refinement, this technology could be an invaluable tool for businesses looking to gauge customer sentiment and improve their overall customer experience.

Future Scope

The market for Facial Expression Recognition (FER) technologies is estimated to grow from $19.5 billion in 2020 to $56 billion by 2024. With increase in deep fake technology and videos, spread of misinformation is becoming rampant. In 2019, the Computer Vision Foundation partnered with UC Berkeley, Google, and DARPA to produce a system claimed to identify deepfake manipulations by analysing facial expressions in the targeted subjects. Those with ASD and other illnesses that impair their capacity to perceive facial expressions can benefit from using it. There have been additional research and initiatives that employ machine learning to provide tools and interventions to help with emotion recognition in addition to the one you provided. For those who have trouble understanding facial expressions, for instance, researchers have utilised machine learning to teach computers to discern emotions from speech patterns. Applications that employ AI to sense emotions and offer individualised assistance and feedback to people with ASD or other problems are also under development. Automotive is another industry where emotion detection and recognition technologies are in high demand. A number of cars trained by machine learning already have emotion recognition included. Such systems can understand if a driver is not looking at the road, is making a hands-on phone call or if the driver is falling asleep and can give appropriate alerts/warnings and make changes to

the autonomous driving system. Emotica.AI can be a useful tool for HR departments in various ways. In addition to helping with candidate selection, it can also be used to assess employee morale and engagement, and to design policies that better align with employees' needs and preferences. By analysing facial expressions and other nonverbal cues, Emotica.AI can help HR professionals gain a deeper understanding of employees' attitudes and emotions, which can inform decisions about training, promotions, and other workplace initiatives. Overall, Emotica.AI has the potential to improve HR practices and create a more productive and engaged workforce.

References

Agrafioti, F, D Hatzinakos, and A K Anderson. "ECG Pattern Analysis for Emotion Detection". *IEEE Transactions on Affective Computing* 3.1 (2012): 102–115. 10.1109/t-affc.2011.28.

Chowanda, Andry. "Separable convolutional neural networks for facial expressions recognition". *International Journal of Advanced Computer Science and Applications* 10.10 (2019): 110–116. 10.14569/IJACSA.2019.0101014.

Goshvarpour, Atefeh, Ataollah Abbasi, and Ateke Goshvarpour. "An accurate emotion recognition system using ECG and GSR signals and matching pursuit method". *Biomedical Journal* 40.6 (2017): 355–368. 10.1016/j.bj.2017.11.001.

Haque, M. A, X. Li, and X. Li. "Facial expression recognition using a dynamic region-based approach." 13.11 (2013): 15549–15568. 10.3390/s131115549;%20https://www.mdpi.com/1424-8220/13/11/15549.

Kokate, P. M and M. B. Kadam. "Emotion detection from speech using deep learning: A review. " *European Journal of Molecular & Clinical Medicine* 7.7 (2020): 4136–4142. https://ejmcm.com/article_2727_0ee2ff3a947a883d78c6e1132b510e0f.pdf.

Li, X. and Z. Huang. " Facial Expression Recognition Based on Edge Detection and Neural Network." *In Proceedings of the 10th International Conference on Computer Vision Theory and Applications (VISAPP 2015)* (2015): 537–541. 10.5220/0005480505370541;%20https://www.scitepress.org/Papers/2015/54805/.

Li, Y., et al. "Human emotion recognition system". *International Journal of Machine Learning and Cybernetics* 4.2 (2013): 91–98. 10.1007/s13042-012-0081-6;%20https://www.researchgate.net/publication/267229317_Human_Emotion_Recognition_System.

Majumdar, A and S Avabhrith. "Human face expression recognition". *International Journal of Research and Analytical Reviews* 4.1 (2017): 415–421. https://www.semanticscholar.org/paper/Human-Face-Expression-Recognition-Majumdar-Avabhrith/0c6ae7bf7a82e216d8ac55b01a0ca9fbdbf8cfc5/figure/1.

Rattanyu, Kanlaya and Makoto Mizukawa. "Emotion Recognition Based on ECG Signals for Service Robots in the Intelligent Space During Daily Life". *Journal of Advanced Computational Intelligence and Intelligent Informatics* 15.5 (2011): 582–591. 10.20965/jaciii.2011.p0582.

Sarode, V. V. and P. R Nimbhorkar. "Facial expression recognition using SVM classification". *Proceedings of the International Conference on Industrial and Information Systems* (2007): 145–148. 10.1109/ICIINFS.2007.4412815;%20https://ieeexplore.ieee.org/abstract/document/4412815.

Soni, D and Khanna. "Human face expression recognition". *International Journal of Advanced Research in Computer Science and Software Engineering* 5.9 (2015): 539–541. https://www.semanticscholar.org/paper/Human-Face-Expression-Recognition-.