International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 8 Issue 12 ǁ 2020 ǁ PP. 26-32

A Descriptive Researchon Human Emotion RecognitionUsing CNN Algorithm

VAISHNAVI NAIR
Academic Research Student, Department of Information Technology, B.K. Birla College of Arts, Science and Commerce (Autonomous), Kalyan, Maharashtra, India

ABSTRACT: *In today's trending generation, human- robot interaction has become a challenge. Humans interact through emotions, so if human-robot interaction must be effective, humans must detect and recognize emotions of human beings. Human poses and experience a lot of mixed emotions and their moods change instantly, so to understand and analyze emotions of humans, robots must be emotionally expressive to make human-robot interactions easier. This paper gives a descriptive research on what techniques to be adopted in robots. Some techniques and modal systems are used to detect main human emotions such as anger, sad, surprise, happy but then too there should be some technology which can recognize other emotions too along with the basic human emotions. Thus, the study will give an insight into some advanced technology in identifying and reacting to human emotions.*

KEYWORDS: *FER-2013 Dataset, HRI, Epoch, Accuracy, Emotion Classification.*

Date of Submission: 18-11-2020 Date of acceptance: 03-12-2020

I. INTRODUCTION

Emotions play an important role in human beings when two human beings interact with each other, the key aspect in making the conversation effective and interesting is when the opposite person understands and responds properly. In future, there are predictions that robots are going to take place of a human in each aspect. There are many robots which are made and designed in different countries to perform specific and special tasks. Most of the people say that robots can work faster than humans and can perform all the activities of human-robot interaction, the robots lack behind in understanding the emotions of humans.

It is important to consider the variety of cultural attitudes towards robots and how they influence affective interactions, as people from diverse cultures may interpret robot emotions differently. Therefore, emotion recognition is a key challenge for enabling robots to communicate with people. Human affective cues can be inferred from natural communication modalities such as facial expressions, vocal intonation, and body language. Multimodal inputs provide two key advantages: a combination of modalities providing two diverse complementary information, which increases robustness and performance and when one modality is unavailable, due to occlusion and/or noise, a multimodal recognition system can use the remaining modalities to estimate affect.

There are many aspects to be taken into consideration in human emotion recognition by robots such as body gestures, facial expressions, vocal intonation, etc. In this paper, I have used the FER-2013 dataset from Kaggle and applied an algorithm to detect human emotions giving 91% accuracy. The dataset was trained on some images which were in-built and rest from Google detecting seven basic emotions namely, Angry, Disgust, Fear, Happy, Sad, Surprise, Neutral.Data is an integral part of the existing approaches in emotion recognition and in most cases it is a challenge to obtain annotated data that is necessary to train machine learning algorithms [16]. For the task of classifying different emotion types from multimodal sources in the form of texts, audio, videos or physiological signals, there are many datasets available.Decades of scientific research have been conducted developing and evaluating methods for automated emotion recognition.

There is now an extensive literature proposing and evaluating hundreds of different kinds of methods, leveraging techniques from multiple areas, such as signal processing, machine learning, computer vision, and speech processing. Different methodologies and techniques may be employed to interpret emotion such as Bayesian networks. , Gaussian Mixture models and Hidden Markov Models [16]. There are many models available giving an accuracy between 50% - 70% but I have proposed a model that uses CNN algorithm and this model can be used for further research and can be a contribution for the field of artificial intelligence in robotics to detect human emotions making human-robot interaction interactive as human-human interaction takes place.

A) FER 2013 DATASET

FER 2013 dataset was published on ICML (International Conference on Machine Learning) five years ago, to recognize the facial expression of human. FER 2013 is an open source dataset which is first, created for an ongoing project by Pierre Luc carrier and Aaron Courville, then shared publicly for a Kaggle competition, shortly before ICML 2013.This dataset consist of 35887 grayscale, 48 * 48 sized face images with various emotions 7 emotions namely angry, disgust, fear, happy, sad, surprise, neutral .The dataset consists of 48*48 pixel grayscale images of faces. The faces have been automatically registered so that the face is centered and occupies about the same amount of space in each image. The training set consists of 28,709 examples and the public test set consists of 9,589 examples.

B) CNN ALGORITHM

A CNN is a deep learning algorithm which can take an input image, assign importance to various aspect/ object in the image and be able to differentiate one from other [17]. Architecture has different convolution layers [3]. Greater, the number of layers in architecture, more is the accuracy obtained. CNN is most widely used technique that can detect, acquire, and interpret features acquired from an image to classify emotions accurately. [5]

The objectives of this descriptive study are as follows:

1. To apply a CNNalgorithm in a dataset giving higher accuracy.
2. Making such a model that can detect human emotion from images.
3. The proposed model can be a small contribution to the researchers ahead in the field of artificial intelligence for human robot interaction.

To achieve the above objectives giving a marginal result the following points as hypotheses can be given through designing the model:

H1:"Most of the emotions of human beings are detected to make human robot interaction effective and interactive."

H2:"The data model set must be applied an algorithm in such a way that emotion of a person is detected with higher accuracy."

II. LITERATURE REVIEW

In a study published by WafaMellouket,[1] in 2020, a review of various databases,architecture present from 2016 to 2019 give an insight about problems and contributions that can be overcome and used for the further work in human robot interaction.

In a study published by Viraj Mavani, et al.[2] in 2017 demonstrated a CNN model for facial emotion recognition using visual saliency of images in their datasets. The comparison must be made between CNN model and the human performance on emotions using different metrics.

In a study published by Ninad Mehendale [3] in 2019 demonstrated FERC which uses CNN and supervised learning for the ability of background removal and only focusing on the face of human making it easier to detect human emotions with high accuracy.

In a study published by Nour Meekiet al. [4] in 2020 to analyze non-verbal sentiment analysis called facial emotion expressions three model of deep learning architecture. It is inferred that more the number of layers in NN model, more accuracy can be gained in detecting human emotions.

In a study published by Michael Noses Thiruvanathan, et al. [5] in 2020, created anEmonet model which is compatible for facial emotion recognition systems, and able to work well with 40 layers. This model can be used in further applications for detecting human emotions.

In a study published by Madhumita A Takalkar, et al. [6] in 2017 inferred that CNNS worked on large datasets yield greater accuracy in detecting human emotion and has also said about the micro – expression recognition giving major contribution to the HRI

In a study published by Wisal Hashim et al. [7] in 2018 gives the survey describing background of facial emotions recognition many researchers have used different techniques for feature extraction and classification and has detected the basic emotions of human using datasets.

In a study published by Arpita Gupta et al. [8] in 2020 has proposed transfer model which attained performance enhancement to the existing state of art models and it is based on deep residual network, pre – training and self - attention module giving an advantage of increased classification accuracy and inductive training.

In a study published by Hao Goa et. al [9] in 2020 have conducted classification and recognition study on facial expressions they have proposed such a model that enhance the future extraction ability from image and reduce the amount of parameters to be detected thus reducing usage of data

In a study published by Raid saabniet al. [10] have presented a novel architecture that simultaneously aggregates three different convents using additional four layers. This can contribute to future work in HRI, increasing the speed and accuracy for detecting emotions from facial expressions.

In a study published by IshaTalehaonkar, et al. [11] presents an approach of FER using CNN to detect facial expressions in real time the model on the FER2013 dataset was experimented giving test accuracy of 0.6012 and validation accuracy of 0.8978.

In a study published by rishi Mittal, et al. [12] in 2020 presented emotion which is a context aware emotion recognition system that incorporates the context interpretations from psychology. They have also introduced a new dataset named Groupwalk which is a collection of videos captured in multiple real worlds setting of people walking.

In a study published by Jose Maria Garcia – Garcia, et al. [13] presented a review of technologies to detect human emotions thus identifying the strength and weak points of current technologies existing for detecting human emotions.

In a study published by Anvita Saxena, et al. [14] in 2020 surveyed and inferred that the best results can be obtained by using stationary wavelet transforms for facial emotion recognition emotion detection through ECG, Electromyography signals can be considered as future work for giving an accuracy in detecting human emotion.

In a study published by Chevl young park,et al. [15] in 2020 presented K- Emocon dataset which is the first publicly available emotion dataset accommodating the multi-perspective assessment of emotions.

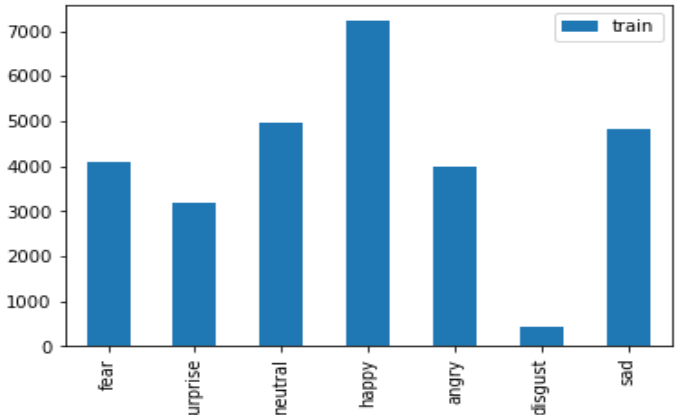
III. METHODOLOGY

A) FER2013 dataset is trained using CNN algorithm which involved various steps as follows: -

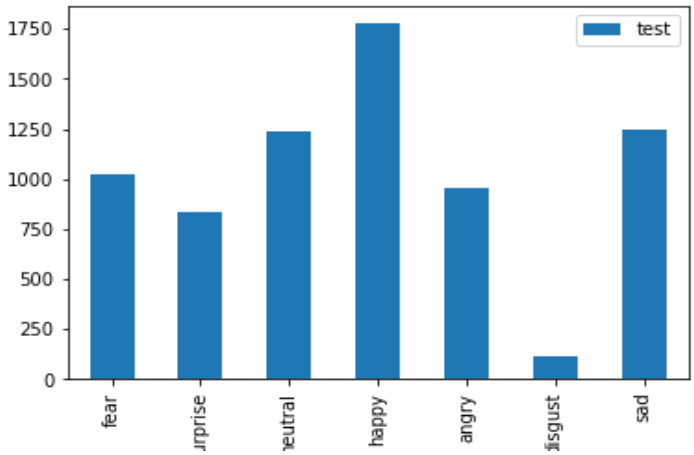
a. Importing libraries:The necessary libraries involved for testing, training, and giving the accuracy of dataset are imported

b. Importing Dataset:The FER2013 dataset is used so importing this dataset is the second major step.

c. Plot the number of images in training set:The images in the training are plotted using the necessary library and represented by a bar graph.



d. Plot the number of images in test set:The images in the test set are plotted using the necessary library and represented by a bar graph.



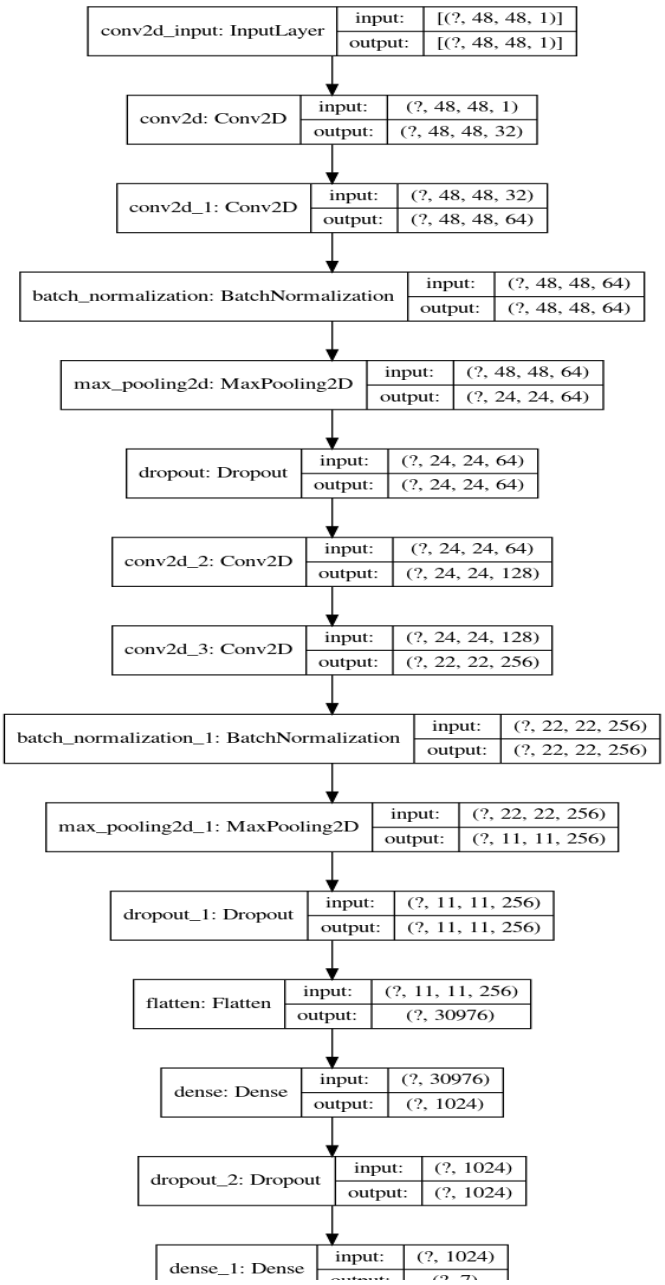
e. Labeling the image:The images in the dataset are labeled according to the respective emotions.

f. Creating training and datasets:The images in the dataset are classified as training set and test set by creating two different such sets.

g. Giving indices in training set:The emotions to be detected were given indices such as O for angry 1 for disgust, 2 for fear, 3 for happy, 4 for neutral, 5 for sad and 6 for surprise.

h. Defining model:The model is defined which is the highlight and important step of our research.

i. Plotting the model: The model is plotted using flow chat giving an easier view of the processes involved in defining the model.

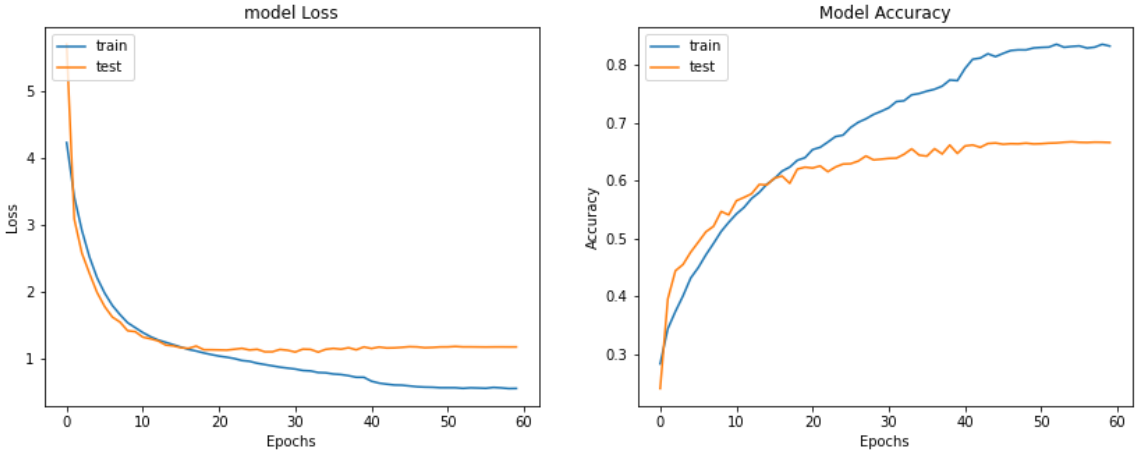


j. Training the model:The model is trained with 60 epochs with each epoch giving an accuracy value, Val- loss and Val-accuracy

IV. EXPERIMENT

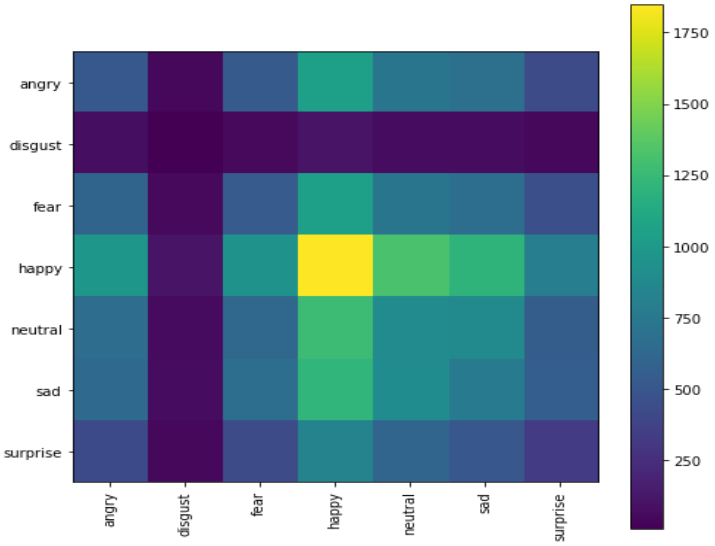
The proposed model is defined and trained as mentioned in methodology and the output obtained from the model is expressed using the following steps:

a. Plotting the accuracy –There are two graphs namely model loss which is showed two parameters train model and test model against loss and epoch inferring that the model loss in both doesn't show huge margin and model accuracy which showed the same above two parameters against accuracy and epochs inferring that training the model gave higher accuracy than testing the model.

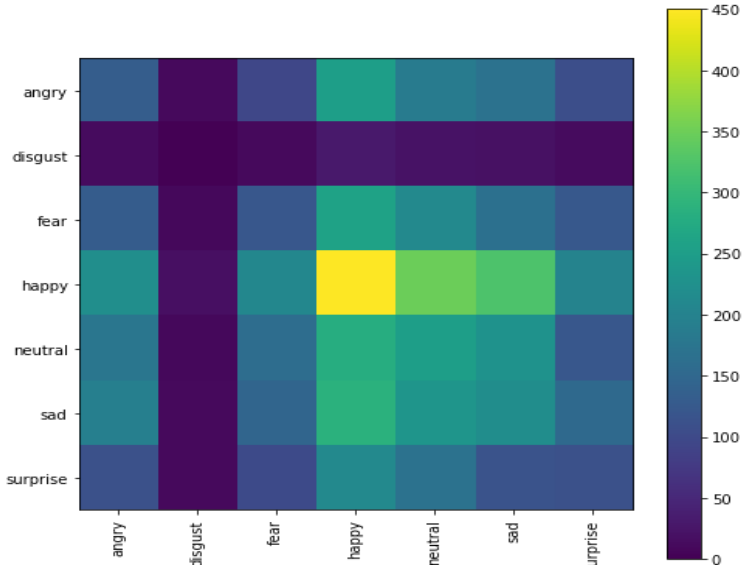


b. Model evaluation: The training set and test set was evaluated giving the accuracy of 91% and 66% respectively. The final trained accuracy is 91% and validation accuracy is 66%.

c. Confusion matrix and classification on training set: The output of the training set is represented using confusion matrix giving a precision of 0.17.



d. Confusion matrix and classification on test set: The output of the test set is represented using confusion matrix. Giving a precision of 0.15.



V. RESULTS

The proposed model can detect seven basic human emotions. This model can be effective when used to have an interaction between human and robots. Human in general go through these seven basic emotions so if a robot can detect these emotions then HRI can be easier just like human- human interaction. The proposed model was applied a CNN algorithm giving higher accuracy as compared to the previous proposed models according to my study. This level of accuracy in detecting human emotions from images can be a small contribution in HRI.

VI. CONCLUSION

The robots are machines designed to perform a particular task but interacting with humans is an emerging field of research. When it comes for robots to interact with humans, detecting the emotions of human can make the interaction interactive. If this model is used in robots, then it can give a small contribution for making the HRI easier. Thus, this model with advanced research in this model can make greater contribution in the field of artificial intelligence.

VII. LIMITATIONS AND FUTURESCOPE

Since the dataset is fed with many images, so the images with similar emotions are studied. High accuracy is achieved on training set, but accuracy of validation is stuck at 66% and no over fitting can be seen in the dataset hence the inefficiency may be due to unbalanced dataset.Multimodal approach by combining two high accuracy giving datasets can be an interested field of research. Humans go through a lot of mood swings, changing their emotions in seconds, so in future such a model must be made that can detect this change and act accordingly.

ACKNOWLEDGEMENT

Sincere thanks to Prof. Swapna Augustine Nikale, Department of Information Technology, B. K. Birla College of Arts, Science and Commerce, Kalyan, and my colleagues for supporting me with the work of this paper.

REFERENCES

[1]. Mellouk, W., &Handouzi, W. (2020). Facial emotion recognition using deep learning: review and insights. *Procedia Computer Science*, *175*, 689–694. https://doi.org/10.1016/j.procs.2020.07.101

[2]. Mavani, V, Raman, S, Miyapuram, K, (2020), L. D. College of Engineering, Indian Institute of Technology.

[3]. Mehendale, N. (2020). Facial emotion recognition using convolutional neural networks (FERC). *SN AppliedSciences*, *2*(3), 0. https://doi.org/10.1007/s42452-020-2234-1

[4]. MEEKI, N, AMINE, A, BOUDIA, M, HAMOU, R, (2020), Deep Learning for Non Verbal Sentiment Analysis: Facial Emotional Expressions, GeCoDe Laboratory, Department of Computer Science, Tahar Moulay University of Saida.

[5]. Thiruthuvanathan, M., & Krishnan, B. (2020). EMONET: A Cross Database Progressive Deep Network for Facial Expression Recognition. *International Journal of Intelligent Engineering and Systems*, *13*(6), 31–41. https://doi.org/10.22266/ijies2020.1231.04

[6]. Gao, H., & Ma, B. (2020). A Robust Improved Network for Facial Expression Recognition. *Frontiers in Signal Processing*, *4*(4), 4. https://doi.org/10.22606/fsp.2020.44001

[7]. Gupta, A. (2020). TRANSFER:- DEEP INDUCTIVE NETWORK FOR FACIAL EMOTION RECOGNITION. *JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES*, *15*(7), 352. https://doi.org/10.26782/jmcms.2020.07.00029

[8]. Saabni, R., &Schclar, A. (2020). Facial Expression Recognition Using Combined Pre-Trained Convnets. *Computer Science & Information Technology*, 95. https://doi.org/10.5121/csit.2020.100908

[9]. Talegaonkar, I, &Joshi, K, Valunj, S, Kohok, R, Kulkarni, A. (2019). Real Time Facial Expression Recognition using Deep Learning, MKSSS's Cummins College of Engineering for Women.
[10]. Mittal,T, Guhan,P , Bhattacharya, U, Chandra R,Bera A, Manocha D. (2020). EmotiCon: Context-Aware Multimodal Emotion Recognition using Frege's Principle, Department of Computer Science, University of Maryland, Department of Electrical and Computer Engineering, University of Maryland.
[11]. Garcia-Garcia, J, Penichet, V, Lozano, M, (2017). Emotion Detection: A Technology review, Computer Science Research Institute, University of Castilla-La Mancha, Albacete.
[12]. Saxena, A., Khanna, A., & Gupta, D. (2020). Emotion Recognition and Detection Methods: A Comprehensive Survey. *Journal of Artificial Intelligence and Systems*, *2*(1), 53–79. https://doi.org/10.33969/ais.2020.21005
[13]. Park, C. Y., Cha, N., Kang, S., Kim, A., Khandoker, A. H., Hadjileontiadis, L., Oh, A., Jeong, Y., & Lee, U. (2020). K-EmoCon, a multimodal sensor dataset for continuous emotion recognition in naturalistic conversations. *Scientific Data*, *7*(1), 0. https://doi.org/10.1038/s41597-020-00630-y
[14]. https://analyticsindiamag.com/top-8-datasets-available-for-emotion-detection/.
[15]. https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53/