Emotion Classification of Facial Images Using Machine Learning Models

K PRASANNA LAKSHMI
Professor, Department of Information Technology, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, Telangana, India

K SRI DIVYA MUKTHA
Sri Chaitanya Junior College, Hyderabad, Telangana, India

KOMPALLI AVINASH BHARGAV
Graduate Engineer Trainee, Seimens Healthcare Private Limited, India

N RAJASEKHAR
Department of Information Technology, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, Telangana, India

Abstract
Humans find it very easy to detect emotions through facial expressions. It is a herculean task for the machine to replicate the same. But recently Machine Learning (ML) is coming to the aid of computer vision in detecting facial emotions accurately from digital photos. In this work a system that uses machine learning algorithms is designed for identifying the emotions based on captured digital images of faces. A total of eight ML models has been designed to classify the facial images into an emotion. Human emotions mainly include happy, sad, surprise, anger, serious and disappointment. ML algorithms like K-Nearest Neighbour (KNN), Naive Bayes (NB), Decision Tree (DT), Random Forest (RF), Gradient Boosting (GB), AdaBoost Classifier, Histogram Gradient Boosting Classifier (HGB), Linear Support Vector Classification (SVC) were used in designing the models. In this research, the HGB based model attained maximum accuracy of 75.67%. Facial Emotion Classification (FEC) is significant in areas like customer feedback, surveillance, mental disease diagnosis, Human Computer Interaction and social behavioural analysis.

Keywords: Facial expression recognition, Human emotion Classification, Machine Learning, Image Processing, Computer Vision, Human Computer Interaction
DOI Number: 10.14704/NQ.2022.20.11.NQ66774 **NeuroQuantology2022;20(11): 7777-7783**

1. INTRODUCTION

Automatic human emotion recognition systems are need of the day as they find importance in many fields like human computer interface (T.B.Sheridan, 2016), security (X.Zhu et.al, 2015), (T.Chen et.al, 2013),(T.Bai et.al, 2014), behavioural analysis of customer, gaming systems (C.H Chem et. al, 2015), (S.Hickson et.al, 2019) . Human emotions can be classified as speech and linguistics (K.Han et.al, 2014) and facial (T.Abegaz et.al, 2011), (W.Y Zhao et.al, 2000). Moods of students in an online classroom setup can be analysed using facial expression analysis. The effectiveness of the classroom can be gauged by this type of analysis. Security agencies can monitor faces of large crowds and analysis any suspicious activity based upon expression analysis (S.Ghosh et.al, 2018). According to a survey, human beings express majority of their emotions using facial expressions rather than words and tone (A.Mehrabian et.al, 2017). Human-Robot interactions can be improved if a robot can sense the mood of the human being it is interacting with. Robots can detect the positive as well as negative states of mind of a human being with which it is interacting with. It can further decide whether to continue or discontinue its interaction with the human (Samadiani et.al, 2019). Human emotions mainly include happy, sad, surprise, anger, serious and disappointment (Ali et.al, 2020). Facial emotion recognisers play a vital role in designing compact and accurate avatars in the field of augmented reality (Hickson et.al, 2017) and virtual reality (Chen et.al, 2015). Many animation-based applications use facial emotion recognition techniques (Liu.M et.al, 2014) for simulating human emotions in the created characters. Facial emotion recognisers have to deal with two types of human emotions namely, natural & superficial emotions (Gan.Q et.al, 2015). The natural emotions are generated by human beings while executing their day-to-day activities. Superficial emotions are generated by human beings artificially like an actor in a movie. The time an emotion is expressed by a human being can be taken into account for classifying emotions. If an emotion is expressed for less than ½ a second then it is termed as micro expression where as if an expression lasts more than ½ second then it is termed as a macro expression. Digital camera, eye-tracker, electrocardiogram, electromyography and electroencephalograph are utilised by facial emotion classifiers to achieve their tasks. This paper considers images taken by a camera as input to the system. Any Facial emotion Classifier works in three stages viz. image pre-processing, feature extraction and emotion detection. Figure 1 depicts the basic working of any camera based facial emotion classifier.

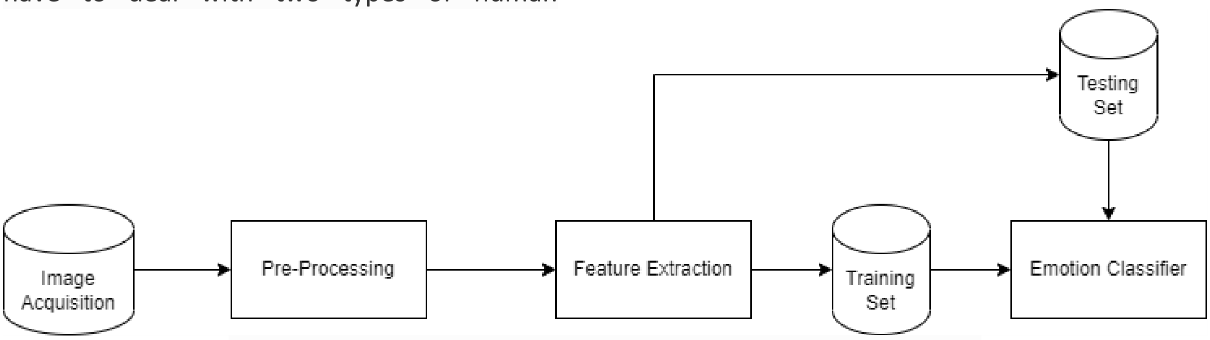

Fig 1. Stages in a general camera based facial emotion classifier

This paper proposes eight models to classify facial emotions that are designed using ML algorithms like KNN, NB, DT, RF, GB, AdaBoost Classifier, HGB, and SVC respectively. The rest of the paper is organised as follows. Section two presents the review of literature related to FEC's. Section 3 throws light on data set used in this research. Section 4 explains about the proposed methodology and discusses broadly on the achieved results. Finally, section 5 concludes the paper with some insights to future prospects.

2. RELATED WORK

An FEC faces a number of challenges which researchers have tried to address through their investigations. The main challenges arise due to illumination of the image, diverse poses of the facial images and subject-dependence. Facial expressions are very complex to capture in runtime due to noise, movement of people, lighting conditions and occlusions (Tian et.al, 2001). This section throws light on the work done so far in this area. Most of the FEC's developed so far prove to be effective in simulated environments and fail to achieve their objectives in a real world scenario.

Feature extraction is an important phase in any facial emotion classification framework. The effectiveness of an FEC mostly depends on the quality & quantity of features extracted from the facial image. Geometric and appearance-based techniques are popular feature extractors with respect to images (Calvo et.al, 2010), (Yan.H et.al, 2014). Geometric FEC's construct feature vectors from ear, eye, mouth and nose using geometry. The shape, position and angles of the facial organs are considered for constructing the feature vectors. Appearance based FEC's are popular among researchers due to their simplicity and accuracy. In the real world, as it is near to impossible to collect geometrical features therefore most of the researchers favour appearance based FEC's (Wang.N et.al, 2017).

Classification of the emotions is the second important phase in the working of a FEC. After extracting facial features the next task is to reduce their dimensions and protecting their significance. Researchers in the past have employed classifiers like SVM, RF, NB, AdaBoost, GB, for accurately classifying the emotions present in the human images/videos. FEC resistant to illumination disparity was developed by a group of researchers (Liu.M et.al, 2014) that could acquire the facial features. The proposed FEC used kernel SVM, Logistic Regression (LG) and partial least squares as classifiers in combination with Riemannian manifold kernels. Some FEC's were designed using Principal Component Analysis (PCA) (Zhang et.al, 2015), (Garg.A et.al, 2015), (Patil.M.N et.al, 2016). Researchers even experimented with Local Binary Patterns (LBP) to build FEC's (Levi.G et.al, 2015), (Chao.w et.al, 2015). LBP based FEC's performed effectively when compared to PCA based FEC's when illumination was the challenge. A parallel research (Chengeta.K et.al, 2018) implemented Local Directional Patterns (LDP) for building an FEC that worked effectively even when there was unstable illumination. Another group of researchers (Ding.Y et.al, 2017) used LBP and Taylor expansion to construct a robust FEC promised to work effectively even with unstable illumination. A group of researchers implemented (Gupta.O et.al, 2017) deep neural networks to recognise facial expressions. Deep learning was also used by other set of researchers (Uddin.M.Z et.al, 2018) to build a FEC. This FEC recognised facial expressions from videos. Some researchers also used mathematical concepts like Fourier transforms integrated with Contrast Limited Adaptive Histogram Equalization (Munir.A et.al, 2018) for recognising facial expressions even in dim light. Deep evolutional spatial-temporal networks were also used by some researchers (Zhang.K et.al, 2017) to perform facial expression and emotion classification. Another study (Uddin.M.Z et.al, 2017) utilised deep belief networks with local direction feature extractor for analysing the emotions in facial images of humans. A method (Chen.J et.al, 2018) was proposed using Warp transformation combined with Histogram of Oriented Gradients for building a facial emotion classifier. A novel machine learning technique (Yan.H et.al, 2018) that uses multiple metrics for discriminating the features collaboratively was proposed by some researchers.

3. DATA SET DESCRIPTION

The proposed work utilises the FER-2013 dataset (Kaggle, 2013) for training and testing the models. FER-2013 is a repository of 32,298 images out of which 28,709 images are used for training and 3,589 images are used testing the models. All images are facial images which are gray scale in nature. Each image is of size 48X48 pixels. The mapping of the images with respect to emotions can be done as follows- 0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral. The dataset contains 958 angry, 111 disgust, 1024 fear, 1174 happy, 1223 neutral, 1247 sad and 831 surprise images. Figure 2. Depicts the various classes of facial emotions.

Fig 2. Various classes of facial emotions (Source: FER-2013 dataset/Kaggle)

4. CLASSIFICATION METHODOLOGY AND RESULT ANALYSIS

The main objective of an FEC is to classify an emotion expressed in a human face present in a digital image. The first step in this process is to acquire image using a camera. The image is converted to gray scale. This image is then pre-processed using various normalisation techniques. Noise is also removed from the image. Later the dimensions of the image to are reduced to 48X48 pixels. Significant features are extracted from the image to construct a feature vector of the captured image. This discriminate features acquired from the feature extraction phase is fed to the classifiers of the various ML models for classification. Seven ML models are designed as part of this research. Each model classifies the facial emotion present in the image as either angry or disgust or fear or happy or sad or surprise or neutral. Model 1 uses KNN as its classifier, model 2 uses NB as its classifier, model 3 uses DT as its classifier, model 4 uses RF as its classifier, model 5 uses GB as its classifier, model 6 uses AdaBoost as its classifier and model 7 uses HGB as its classifier. Figure 3 represents the flow diagram of the classification process.

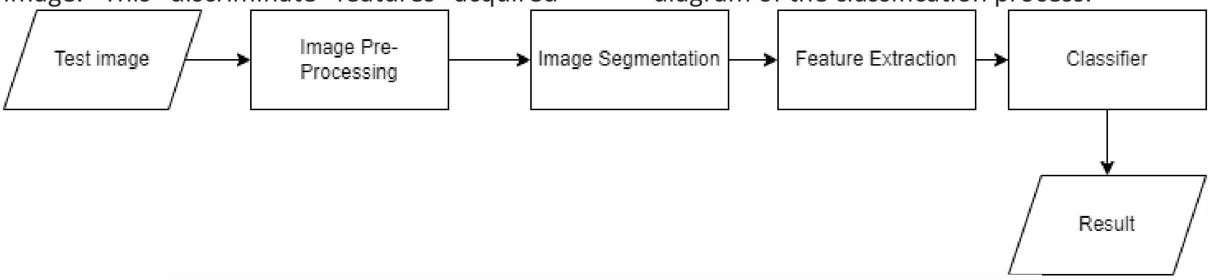

Fig 3. Flow diagram of the classification process of a facial emotion image

The preprocessing is performed using the StandardScaler package imported from SKLearn python library. Figure 4 depicts the pre-processed images. Python3 programming language was used for conducting the experiments. Numpy and Pandas packages were also imported in to the code. OpenCV python library was also imported in to the notebook for performing the required tasks. Kaggle cloud was used to run the experiments. CPU, Storage and RAMs were provided by Kaggle.com for running the investigations. Some of the classifiers require fine tuning of the hyper parameters to avoid over fitting. Table 1 lists the hyper parameters used by the classifiers in the designed models.

a. Angry b. Disgust c. Fear d. Happy e. Neutral f. Sad g. Surprise

Fig 4. Sample images after Pre-processing (Source: FER-2013 dataset/Kaggle)

Accuracy is the metric used here to measure the quality of the model. Table 2 represents the accuracy obtained for various models recognising the facial emotions. The model 7 attains a maximum accuracy of 75.67% followed by model 5(69.97%). Model 4 attains an accuracy of 69.01%, which is followed by model 6 (68.91%). Linear SVC, KNN and NB achieved an accuracy of 65.70%, 65.21% and 59.81% respectively. The analysis also proves that boosting algorithms are outperforming traditional ML algorithms. A comparative analysis of the accuracy achieved by the models is depicted as a bar graph in figure 4.

Table 1 lists the hyper parameters used by the classifiers in the designed models.

Model Number	Classifier	Hyper Parameters
1.	KNN	n_neighbors=k
2.	NB	-----
3.	DT	-----
4.	RF	n_estimators = 10, criterion = 'entropy', random_state = 0
5.	GB	n_estimators=100, learning_rate=1.0,max_depth=1, random_state=0
6.	AdaBoost	n_estimators=100, random_state=0
7.	HGB	max_iter=100

Table 2 Accuracy of the various models designed

Model No	Classifier	Accuracy
1.	K-Nearest Neighbour	65.21
2.	Naive Bayes	59.81
3.	Decision Tree	64.48
4.	Random Forest	69.01
5.	Gradient Boosting	69.97
6.	Adaptive Boosting Classifier	68.91
7.	Histogram Gradient Boosting Classifier	75.67
8.	Linear Support Vector Classifier	65.70

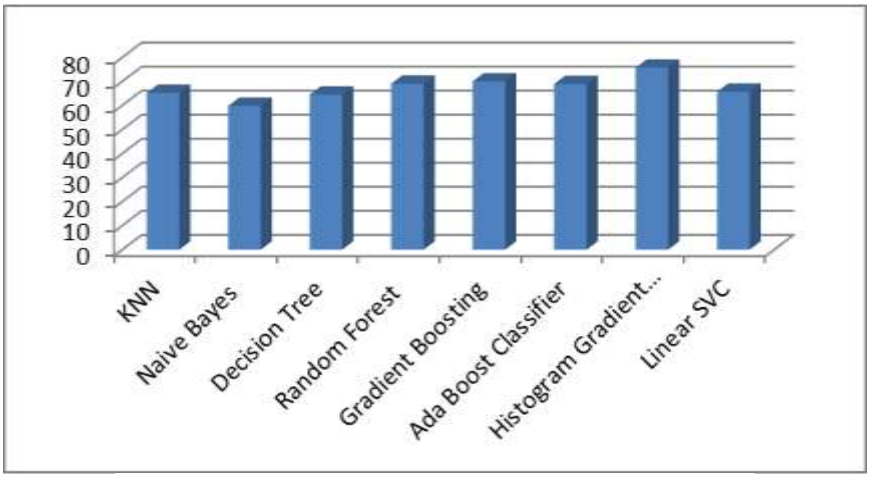

Fig 4. Comparative Analysis of the various FEC models

5. CONCLUSIONS AND FUTURE SCOPE

This paper tried to throw light on the various challenges faced by a Facial Emotion Classifier in real time. This work provided a solution to the problem of facial emotion recognition from images that house human facial expressions. The proposed methodology contains 4 major phases' image acquisition, image pre-processing, feature extraction, classification. Eight ML models were created using the proposed methodology as a skeleton. The features extracted were fed in to eight classifiers namely, K-Nearest Neighbour, Naive Bayes, Decision Tree, Random Forest, Gradient Boosting, AdaBoost Classifier, Histogram Gradient Boosting Classifier and Linear Support Vector Classification. The model based on HGB outperformed the remaining models. The research founded that ML based approaches achieve optimum performance levels in a short time with less computational costs. In future, Deep Learning techniques can be used in place of ML approaches to increase the accuracy of the system. Emotions from voice, text, videos can also be classified in future using the same models. The models can also be exposed to other datasets for improving their performance and reliability.

CONFLICT OF INTEREST
NONE.
FUNDING SOURCE
NONE.
REFERENCES
T. B. Sheridan, "Human-robot interaction," Human Factors: The Journal of the Human Factors and Ergonomics Society, vol. 58, no. 4, pp. 525–532, 2016.

X. Zhu, X. Li, and S. Zhang, "Block-row sparse multiview multilabel learning for image classification," IEEE Transactions on Cybernetics, vol. 46, no. 2, pp. 450–461, 2015.

T. Chen and K.-H. Yap, "Discriminative BoW framework for mobile landmark recognition," IEEE Transactions on Cybernetics, vol. 44, no. 5, pp. 695–706, 2013.

T. Bai, Y.-F. Li, and X. Zhou, "Learning local appearances with sparse representation for robust and fast visual tracking," IEEE Transactions on Cybernetics, vol. 45, no. 4, pp. 663–675, 2014.

C.-H. Chen, I.-J. Lee, and L.-Y. Lin, "Augmented reality-based self-facial modeling to promote the emotional expression and social skills of adolescents with autism spectrum disorders," Research in Developmental Disabilities, vol. 36, pp. 396–403, 2015.

S. Hickson, N. Dufour, A. Sud, V. Kwatra, and I. Essa, "Eyemotion: classifying facial expressions in VR using eyetracking cameras," in Proceedings of the 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1626–1635, Waikoloa, HI, USA, January 2019.

K. Han, D. Yu, and I. Tashev, "Speech emotion recognition using deep neural network and extreme learning machine," in Proceedings of the Fifteenth Annual Conference of the International Speech Communication Association, Singapore, September 2014.

T. Abegaz, G. Dozier, K. Bryant et al., "Hybrid GAs for Eigenbased facial recognition," in Proceedings of the 2011 IEEE Workshop on Computational Intelligence in Biometrics and Identity Management (CIBIM), pp. 127–130, Paris, France, April 2011.

W. Y. Zhao and R. Chellappa, "SFS based view synthesis for robust face recognition," in Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 285–292, Grenoble, France, March 2000.

S. Ghosh, A. Dhall, and N. Sebe, "Automatic group affect analysis in images via visual attribute and feature networks," in Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), pp. 1967–1971, Athens, Greece, October 2018.

A. Mehrabian, "Communication without words," in Communication Deory, pp. 193–200, Routledge, UK, London, 2017.

Samadiani, N., Huang, G., Cai, B., Luo, W., Chi, C. H., Xiang, Y., & He, J. (2019). A Review on Automatic Facial Expression Recognition Systems Assisted by Multimodal Sensor Data. Sensors (Basel, Switzerland), 19(8), 1863. https://doi.org/10.3390/s19081863

Ali, M. F., Khatun, M., & Turzo, N. A. (2020). Facial emotion detection using neural network. the international journal of scientific and engineering research.

Hickson S., Dufour N., Sud A., Kwatra V., Essa I. Eyemotion: Classifying facial expressions in VR using eye-tracking cameras. arXiv. 20171707.07204

Chen C.-H., Lee I.-J., Lin L.-Y. Augmented reality-based self-facial modeling to promote the emotional expression and social skills of adolescents with autism spectrum disorders. Res. Dev. Disabil. 2015;36:396–403. doi: 10.1016/j.ridd.2014.10.015.

Liu M., Li S., Shan S., Wang R., Chen X. Deeply learning deformable facial action parts model for dynamic expression analysis; Proceedings of the Asian Conference on Computer Vision; Singapore. 1–5 November 2014; pp. 143–157.

Gan Q., Wu C., Wang S., Ji Q. Posed and spontaneous facial expression differentiation using deep Boltzmann machines; Proceedings of the 2015 International Conference on Affective Computing and Intelligent Interaction (ACII); Xi'an, China. 21–24 September 2015; pp. 643–648.

Tian Y.-I., Kanade T., Cohn J.F. Recognizing action units for facial expression analysis. IEEE Trans. Pattern Anal. Mach. Intell. 2001;23:97–115. doi: 10.1109/34.908962.

Calvo R.A., D'Mello S. Affect detection: An interdisciplinary review of models, methods, and their applications. IEEE Trans. Affect. Comput. 2010;1:18–37. doi: 10.1109/T-AFFC.2010.1.

Yan H., Lu J., Zhou X. Prototype-based discriminative feature learning for kinship verification. IEEE Trans. Cybern. 2015;45:2535–2545. doi: 10.1109/TCYB.2014.2376934.

Wang N., Gao X., Tao D., Yang H., Li X. Facial feature point detection: A comprehensive survey. Neurocomputing. 2018;275:50–65. doi: 10.1016/j.neucom.2017.05.013.

Liu M., Wang R., Li S., Shan S., Huang Z., Chen X. Combining Multiple Kernel Methods on Riemannian Manifold for Emotional expression recognition in the Wild; Proceedings of the 16th International Conference on Multimodal Interaction; Istanbul, Turkey. 12–16 November 2014; pp. 494–501.

Zhang D., Ding D., Li J., Liu Q. Transactions on Engineering Technologies. Springer; Berlin/Heidelberg, Germany: 2015. Pca based extracting feature using fast fourier transform for facial expression recognition

Garg A., Bajaj R. Facial expression recognition & classification using hybridization of ica, ga, and neural network for human–computer interaction. J. Netw. Commun. Emerg. Technol. 2015;2:49–57.

Patil M.N., Iyer B., Arya R. Performance Evaluation of PCA and ICA Algorithm for Facial Expression Recognition Application; Proceedings of the Fifth International Conference on Soft Computing for Problem Solving; Uttarakhand, India. 15 March 2016; pp. 965–976.

Levi G., Hassner T. Emotional expression recognition in the Wild via Convolutional Neural Networks and Mapped Binary Patterns; Proceedings of the 2015 ACM on International Conference on Multimodal Interaction; Seattle, WA, USA. 9–13 November 2015; pp. 503–510.

Chao W.-L., Ding J.-J., Liu J.-Z. Facial expression recognition based on improved local binary pattern and class-regularized locality preserving projection. Signal Process. 2015;117:1–10.

Chengeta K., Viriri S. A survey on facial recognition based on local directional and local binary patterns; Proceedings of the 2018 Conference on Information Communications Technology and Society (ICTAS); Durban, South Africa. 8–9 March 2018; pp. 1–6.

Ding Y., Zhao Q., Li B., Yuan X. Facial expression recognition from image sequence based on lbp and taylor expansion. IEEE Access. 2017;5:19409–19419. doi: 10.1109/ACCESS.2017.2737821.

Gupta O., Raviv D., Raskar R. Illumination invariants in deep video expression recognition. Pattern Recognit. 2018;76:25–35. doi: 10.1016/j.patcog.2017.10.017.

Uddin M.Z., Hassan M.M., Almogren A., Zuair M., Fortino G., Torresen J. A facial expression recognition system using robust face features from depth videos and deep learning. Comput. Elect. Eng. 2017;63:114–125. doi: 10.1016/j.compeleceng.2017.04.019.

Munir A., Hussain A., Khan S.A., Nadeem M., Arshid S. Illumination invariant facial expression recognition using selected merged binary patterns for real world images. Optik. 2018;158:1016–1025. doi: 10.1016/j.ijleo.2018.01.003.

Zhang K., Huang Y., Du Y., Wang L. Facial expression recognition based on deep evolutional spatial-temporal networks. IEEE Trans. Image Procss. 2017;26:4193–4203. doi: 10.1109/TIP.2017.2689999.

Uddin M.Z., Hassan M.M., Almogren A., Alamri A., Alrubaian M., Fortino G. Facial expression recognition utilizing local direction-based robust features and deep belief network. IEEE Access. 2017;5:4525–4536. doi: 10.1109/ACCESS.2017.2676238

Chen J., Chen Z., Chi Z., Fu H. Facial expression recognition in video with multiple feature fusion. IEEE Trans. Affect. Comput. 2018;9:38–50. doi: 10.1109/TAFFC.2016.2593719.

Yan H. Collaborative discriminative multi-metric learning for facial expression recognition in video. Pattern Recognit. 2018;75:33–40.

https://www.kaggle.com/datasets/msambare/fer2013