AN EFFICIENT DYNAMIC MODEL FOR RECOGNITION AND CLASSIFICATION OF HUMAN EXPRESSION USING NOVEL RANDOM FOREST COMPARED OVER CONVOLUTION NEURAL NETWORK WITH IMPROVED ACCURACY

Shaik.Abhida[1], M.Sandhiya[2*]

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

Aim: To implement the dynamic model for recognition and classification of human expression using Novel Random Forest compared over Convolution Neural Network with improved accuracy.
Materials and Methods: This study contains 2 groups i.e Novel Random Forest (RF) and Convolutional Neural Network (CNN). Each group consists of a sample size of 10 and the study parameters include alpha value 0.05, beta value 0.2, and power value 0.8.
Results: The Novel Random Forest is 77.9% more accurate than the Convolutional Neural Network of 75.8% in classifying the facial expressions of humans with p=0.8.
Conclusion: The Novel Random Forest model is significantly better than the Convolution Neural Network in identifying human expressions.

Keywords: Facial Expression Recognition, Convolution Neural Network, Facial Feature Extraction, Emotion Classification, Novel Random Forest.

[1]Research Scholar, Department of Information Technology, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India. Pincode: 602105.
[2*]Department of Information Technology, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India. Pincode: 602105.

1. Introduction

Humans communicate primarily through speech, but they also use body gestures to emphasize specific parts of their speech and to express emotions. Humans use a variety of methods to express their emotions through facial expressions, which are a crucial part of communication. Despite the fact that nothing is said verbally, there is a lot to learn about the messages we send and receive through nonverbal communication. Automatic facial expression recognition can be a useful feature in natural human-machine interfaces, as well as in behavioral science and clinical practice. Although advances in face detection, facial feature extraction mechanisms, and emotion classification techniques have been made in recent years, developing an automated system that accomplishes this task is difficult. Six basic emotion classification according to Ekman and Friesen, are universal: happiness, surprise, sadness, fear, anger, and disgust. These feelings can be found in people from all walks of life. Preprocessing, feature extraction, and emotion classification are the three steps involved in facial expression recognition. The Facial Feature Extraction feature we expect to minimize the distance between within-class variations of expression while maximizing the distance between between-class variations is a key step in the recognition process. There are two types of feature extraction methods: appearance-based methods and geometric feature-based methods. Geometric characteristics are extracted. Implementation of Facial Expression Recognition system for selecting fashion items Based on like and Dislike Expression (Wang 2012). Facial Expression Recognition in image sequence using Geometric Deformation Features and Support Vector Machines (Kotsia, Nikolaidis, and Pitas 2007). Facial Expression Recognition based on Gabor feature and neural network (Z. Li et al. 2018). The advantage of CNN is that it requires less training data when compared to other models. The Applications are Human behavior understanding, detection of mental disorder and Synthetic human expression.

There are about 25 articles in IEEE explore and 15 Scopus related to this study. In a study by (L. Prevost. 2012) in Dynamics of Facial Expression Recognition of facial actions and their temporal segments from face profile image sequence. In Human Facial Expression Recognition using a 3D morphable model CNN is derived from the usage of sequential data information (Michel F. Valstar. 2013). In Assisting the austic with improved facial expression from mixed expressions (Alessandra Lumini. 2013). It is said that CNN is suitable for learning time series data i.e long term temporal dependencies in Active and dynamic information fusion for Facial Feature Extraction understanding from image sequence (Kaoru Hirota. 2020). Previously there was a rich experience in working on various research projects across multiple disciplines (Diego Ziyong Feng. 2016). Implementation of Facial Expression Recognition system for selecting fashion items Based on like and Dislike Expression (Wadawadagi, Ramesh, and Veerappa Pagi. 2020). Facial Expression Recognition in image sequence using Geometric Deformation Features and Support Vector Machines (Mehrabian, Albert. 2017). Facial Expression Recognition based on Gabor feature and neural network (Zhang and Yu-Jin 2010**).**

Some datasets are aimed at theoretical research rather than processing it as per the real life application. Therefore defining the boundaries between the expression parts is very challenging. Most of the existing standard facial feature extraction processes are for short-term analysis, So researches have made their own feature set, Finally paper is proposed assuming all the limitations. This paper solely focuses on enhancing the facial expressions models to increase the accuracy of human facial expression recognition.

2. Materials and Methods

This work was carried out at the Data Analytics Lab of Information Technology in Saveetha School of Engineering. The study consists of two sample groups i.e Random Forest and Convolution Neural Network. Each group consists of 10 samples with pre-test power of 0.18. The sample size was collected from (Wadawadagi and Pagi 2020) by keeping the threshold at 0.05, G power of 80%, confidence interval at 95%, and enrolment ratio as 1.

The dataset used for classification is taken from the Berlin Database of Facial Expression Recognition through bilderbar©, an open-source data for Facial Expression. The dataset contains 2 columns. In the first column we’ll have test size and in the second column it represents the accuracy values. The dataset was split into training and testing parts accordingly using a test size of 0.2.

For training of the Novel Random Forest, the test set size is about 20% of the total dataset and the remaining 80% is used for the training set. The Novel Random Forest training set consists in determining a hyperplane to separate the training data belonging to two classes, whereas the CNN model uses backpropagation for training. The whole dataset is fitted for training the Novel Random Forest and CNN model. Accuracies of both models are tested with a sample size of 10 using Python 2.7.

Novel Random Forest

A Novel Random Forest (RF) algorithm for regression is proposed in this paper. Group 1 is Novel Random Forest. Novel Random forest tries to find the hyperplane in a N-dimensional space that distinctly classifies the data points. The data points here are called random forests which are obtained from extraction and labeling. X-axis is time and the Y-axis is frequency of facial expression taken from a dataset. Using the past labeled data, the Random Forest tries to identify and predict the new data done by model training.

First the variables are taken from the dataset. Then perform facial feature extraction using Mel-frequency cepstral coefficients and classify them into extraction labels i.e Angry, Neutral, Happy, Sad. Now data is split into a training set and test set using test size=0.2. Scaling the sets using the standard scaler. Fitting the training set with Novel Random Forest Classifier (RFC) and fitting the test set to the model to get final predictions done. Checking the accuracy of the predictions. Accuracy is calculated using following equation 1.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

where,
TP= no of true positive classified by the model
FP= no of false positive classified by the model
TN= no of true negative classified by the model
FN= no of false negative classified by the model

Convolution Neural Network

Convolution Neural Networks (CNN) have improved the state of the art in many applications, especially the face recognition area. In this work, we present a review on the latest face verification techniques based on Convolutional Neural Networks. In addition, we give a comparison on these techniques regarding their architecture, depth level, number of parameters in the network, and the obtained accuracy in identification and/or verification. Furthermore, as the availability of large-scale training dataset has significantly affected the performance of Convolution Neural Network based recognition methods, we present a preface to the most common large-scale face datasets, and then we describe some of the successful automatic data collection procedures.

The step by step process of CNN algorithm is first loading the Berlin Emo-DB dataset. Splitting the dataset into 80% training set and 20% test set. Truncate and pad the input sequences so that they all have the same length for modeling. The model will learn that the zero values have no value in them. Compile and fit LSTD. Adding the CNN layers such as Embedded layer, LSTD layer. Dense output layer with a sigmoid activation function is used and a single neutron to make 0 or 1 predictions for the two classes (good and bad) which is shown in equation 2.

$$= [\mathbf{i} \times (\mathbf{f}\times\mathbf{f}) \times \mathbf{o}] + \mathbf{o} \quad \mathbf{(2)}$$

Where,
i= no. of input maps (or channels)
f= filter size (just the length)
o= no. of output maps (or channels. this is also defined by how many filters are used)

The minimum requirement to run the softwares used here are intel core I3 dual core cpu@3.2 GHz, 4GB RAM, 64 bit OS, 1TB Hard disk Space Personal Computer and Software specification includes Windows 8 , 10 , 11 , Python 3.8, and MS-Office.

Statistical Analysis

Statistical Package for the Social Sciences Version 26 software tool was used for statistical analysis. An independent sample T-test was conducted for accuracy. Standard deviation, standard mean errors were also calculated using the SPSS Software tool. The significance values of proposed and existing algorithms are shown in Table 5. Table 6 contains group statistical values of proposed and existing algorithms. The independent variable is image files and the dependent variable is recognized emotions, natural emotions from the dataset. The dependent variables are accuracy and precision.

3. Results

The group statistical analysis on the two groups shows Novel Random Forest (group 1) has more mean accuracy than Convolution Neural Network (group 2) and the standard error mean is slightly less than Novel Random Forest. Fig. 1 represents the bar chart of accuracies with standard deviation error is plotted for both the algorithms. Multiple images in Fig. 2 show various expressions of human beings like Happiness, Sadness, Disgust, Surprise, Contempt, Anger and fear. The above images are collected from an independent dataset.

The Novel Random Forest algorithm scored an accuracy of 77.9% as shown in Table 1 and Convolution Neural Network has scored 75.8% as shown in Table 2. The accuracies are recorded by testing the algorithms with 10 different sample sizes and the average accuracy is calculated for each algorithm.

4. Discussion

From the results of this study, Random Forest is proved to be having better accuracy than the Convolutional Neural Network model. RF has an accuracy of 77.9% whereas CNN has an accuracy of 75.8%. In Table 3 and Table 4, the group statistical analysis on the two groups shows that Random Forest (group 1) has more mean accuracy than Convolutional Neural Network (group 2) and the standard error mean including standard deviation mean is slightly less than Novel Random Forest.

In the study by Dynamics of facial expression recognition of facial actions and their temporal segments from face profile image sequence. In Human Facial expression recognition using a 3D morphable model (Yang et al. 2016). In Assisting the austic with improved facial expression from mixed expressions (Matsugu et al. 2003). In Active and dynamic information fusion for facial expression understanding from image sequence (Sugie and Sugie 2016). In Affective facial expressions recognition for human-robot interaction (Taejin Lee. 2016). They have summarized the results for various face based Facial classification systems reported in the literature (Corneanu et al. 2016). The performance of a neural network depends on the type of parameters extracted from the facial image (Zeng et al. 2009).

The limitation in this model is that the accuracy of RF may get affected due to the inconsistent data and difficulty in getting the right datasets for analysis. Most of the data is simulated from nature which is far from reality. As with any technology, there are potential drawbacks to using facial recognition, such as threats to privacy, violations of rights and personal freedoms, potential data theft and other crimes. There's also the risk of errors due to flaws in the technology. The Future work can be concentrated on effective data processing techniques and usage of ensemble machine learning algorithms can be focussed.

5. Conclusion

Based on the experimental results, the Novel Random Forest (RF) has been proved to recognize expression more significantly than the Convolutional Neural Network (CNN). Further research is needed on how to recognize facial expressions under these extreme conditions. In addition, for the convolutional neural network, it is necessary to collect as many samples as possible and make the trained network have a good generalization performance.

Declarations

Conflicts of Interest

No conflicts of interest in this manuscript.

Authors Contribution

Author SA was involved in data collection, data analysis, data extraction, manuscript writing. Author MS was involved in conceptualization, data validation, and critical review of the manuscript.

Acknowledgement

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.

Funding

We thank the following organizations for providing financial support that enabled us to complete the study.

1. Vee Eee Technologies Solution Pvt. Ltd., Chennai.
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of Engineering.

6. References

Almaev, Timur R., and Michel F. Valstar. 2013. “Local Gabor Binary Patterns from Three Orthogonal Planes for Automatic Facial Expression Recognition.” 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction. https://doi.org/10.1109/acii.2013.65.

Anas, Lutfiyatul Fatjriyati, Nana Ramadijanti, and Achmad Basuki. 2018. “Implementation of Facial Expression Recognition System for Selecting Fashion Items Based on Like and Dislike Expression.” 2018 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC). https://doi.org/10.1109/kcic.2018.8628516.

Anbarjafari, Gholamreza, Jelena Gorbova, Rain Eric Hammer, Pejman Rasti, and Fatemeh Noroozi. 2018. Machine Learning for Face, Emotion, and Pain Recognition.

Branham, Sheryl, Lakhmi C. Jain, Loris Nanni, and Alessandra Lumini. 2013. Local Binary Patterns: New Variants and Applications. Springer.

Chen, Chi Hau, Louis-Francois Pau, and Patrick S. L. Prevost. 2012. Handbook Of Pattern

Recognition And Computer Vision (2nd Edition). World Scientific.

Chen, Lu Feng, Min Wu, Witold Pedrycz, and Kaoru Hirota. 2020. Emotion Recognition and Understanding for Emotional Human-Robot Interaction Systems. Springer Nature.

Davis, Martha. 1975. Recognition of Facial Expression.

Kotsia, Irene, Nikolaos Nikolaidis, and Ioannis Pitas. 2007. "Facial Expression Recognition in Videos Using a Novel Multi-Class Support Vector Machines Variant." 2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07. https://doi.org/10.1109/icassp.2007.366303.

Li, Stan Z., and Anil K. Jain. 2011. Handbook of Face Recognition. Springer Science & Business Media.

Li, Zhuyi, Ming Lei, Minjian Zhao, and Min Li. 2018. "Joint Channel Estimation and Signal Detection for FBMC Based on Artificial Neural Network." 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall). https://doi.org/10.1109/vtcfall.2018.8690845.

Matsugu, Masakazu, Katsuhiko Mori, Yusuke Mitari, and Yuji Kaneda. 2003. "Subject Independent Facial Expression Recognition with Robust Face Detection Using a Convolutional Neural Network." Neural Networks: The Official Journal of the International Neural Network Society 16 (5-6): 555–59.

Senechal, T., V. Rapp, H. Salam, R. Seguier, K. Bailly, and L. Prevost. 2012. "Facial Action Recognition Combining Heterogeneous Features via Multi Kernel Learning." IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics: A Publication of the IEEE Systems, Man, and Cybernetics Society 42 (4): 993–1005.

Shan, Caifeng, Shaogang Gong, and Peter W. McOwan. 2009. "Facial Expression Recognition Based on Local Binary Patterns: A Comprehensive Study." Image and Vision Computing. https://doi.org/10.1016/j.imavis.2008.08.005.

Slezak, Dominik, Sankar Pal, Byeong-Ho Kang, Junzhong Gu, Hideo Kuroda, and Tai-Hoon Kim. 2010. Signal Processing, Image Processing and Pattern Recognition: International Conference, SIP 2009, Held as Part of the Future Generation Information Technology Conference, GIT 2009, Jeju Island, Korea, December 10-12, 2009. Proceedings. Springer.

Sugie, Miho, and Kazuma Sugie. 2016. "Characteristics of Risk-Factor Profiles Associated with Stroke in Patients with Myotonic Dystrophy Type 1." Journal of Rare Disorders: Diagnosis & Therapy. https://doi.org/10.21767/2380-7245.100048.(Sugie and Sugie 2016)

Wadawadagi, Ramesh, and Veerappa Pagi. 2020. "Sentiment Analysis with Deep Neural Networks: Comparative Study and Performance Assessment." Artificial Intelligence Review. https://doi.org/10.1007/s10462-020-09845-2.

Wang, Patrick S. P. 2012. Pattern Recognition, Machine Intelligence and Biometrics. Springer Science & Business Media.

Yang, Weixin, Lianwen Jin, Dacheng Tao, Zecheng Xie, and Zhiyong Feng. 2016. "DropSample : A New Training Method to Enhance Deep Convolutional Neural Networks for Large-Scale Unconstrained Handwritten Chinese Character Recognition." Pattern Recognition. https://doi.org/10.1016/j.patcog.2016.04.007. (Matsugu et al. 2003)

Zhao, Guoying, and Matti Pietikäinen. 2007. "Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions." IEEE Transactions on Pattern Analysis and Machine Intelligence 29 (6): 915–28.

(Shan, Gong, and McOwan 2009; Almaev and Valstar 2013; Senechal et al. 2012; Anbarjafari et al. 2018; C. H. Chen, Pau, and Wang 1999; Brahnam et al. 2013)(Davis 1975) (Zhao and Pietikäinen 2007) (L. Chen et al. 2020; S. Z. Li and Jain 2011; Slezak et al. 2010) (Matsugu et al. 2003(Matsugu et al. 2003)https://paperpile.com/c/dUe94w/

(Yang et al. 2016) Konar, Amit, and Aruna Chakraborty. 2014. Emotion Recognition: A Pattern Analysis Approach. John Wiley & Sons.

Lim, Wootaek, Daeyoung Jang, and Taejin Lee. 2016. "Speech Emotion Recognition Using Convolutional and Recurrent Neural Networks." 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA). https://doi.org/10.1109/apsipa.2016.7820699.

Milton, A., S. Sharmy Roy, and S. Tamil Selvi. 2013. "SVM Scheme for Speech Emotion Recognition Using MFCC Feature." International Journal of Computer Applications. https://doi.org/10.5120/11872-7667.

Mirsamadi, Seyedmahdad, Emad Barsoum, and Cha Zhang. 2017. "Automatic Speech Emotion Recognition Using Recurrent Neural Networks with Local Attention." 2017 IEEE International Conference on

Acoustics, Speech and Signal Processing (ICASSP). https://doi.org/10.1109/icassp.2017.7952552.

TABLES AND FIGURES

Table 1. Pseudocode for Novel Random Forest

// I : Input dataset records
1. Import the required packages.
2. Convert the image files into numerical values after the extraction feature.
3. Assign the data to X_train, y_train, X_test and y_test variables.
4. Using train_test_split() function, pass the training and testing variables.
5. Give test_size and the random_state as parameters for splitting the data using RF training.
6. Importing the Classifier from sklearn library.
7. Using Classifier, predict the output of the testing data.
8. Calculate the accuracy of the model.
OUTPUT //Accuracy

Table 2. Pseudocode for Convolutional Neural Network

// I : Input dataset records
1. Import the required packages.
2. Convert the image files and truncate the input sequences so that they are all the same length for modeling.
3. Assign the data to X_train, y_train, X_test and y_test variables.
4. Using train_test_split() function, pass the training and testing variables.
5. Give test_size and the random_state as parameters for splitting the data.
6. Adding Embedding layer,LSTM layer,Dense Layer to the model
7. Compiling the model using metrics as accuracy.
7. Evaluate the output using X_test and y_test function
8. Get the accuracy of the model.
OUTPUT //Accuracy

Table 3. Accuracy of Facial expression recognition using Novel Random Forest for 10 samples out of 30 (Accuracy= 77.9%)

Test size	**Accuracy**
Test 1	74.6
Test 2	80.3

Test 3	76.4
Test 4	81.6
Test 5	78.5
Test 6	75.4
Test 7	69.9
Test 8	80.7
Test 9	78.7
Test 10	77.9

Table 4. Accuracy of Facial expression recognition using Convolutional Neural Network for 10 samples out of 30 samples (Accuracy= 75.80%)

Test size	**Accuracy**
Test 1	75.3
Test 2	78.3
Test 3	81.3
Test 4	74.6
Test 5	77.6
Test 6	79.6
Test 7	80.6
Test 8	76.4
Test 9	78.4
Test 10	75.8

Table 5. Group Statistic analysis, representing Novel Random Forest (mean accuracy 77.9%, standard deviation 3.75306) and Convolutional Neural Network (mean accuracy 75.8%, standard deviation 2.53835)

Algorithm	**N**	**Mean**	**Std. Deviation**	**Std.Error Mean**
Accuracy NRF	10	77.9700	3.75306	1.18682
Accuracy CNN	10	75.8000	2.53835	.80270

Table 6. Independent Sample Tests results with confidence interval as 95% and level of significance as 0.05 (Novel Random Forest appears to perform significantly better than Convolutional Neural Network with the value of p=0.8).

Accuracy	Levene's Test for Equality of Variances		T-test for Equality of Means						
	F	**Sig.**	**t**	**df**	**Sig**	**Mean Difference**	**Std. Error Difference**	**95% Conf. Interval Lower**	**95% Conf. Interval Upper**
Equal Variances assumed	1.402	.252	.140	18	.445	.20000	1.43278	-2.81017	3.21017
Equal Variances not assumed	1.402	.252	.140	15.809	.445	.20000	1.43278	-2.84035	3.24035

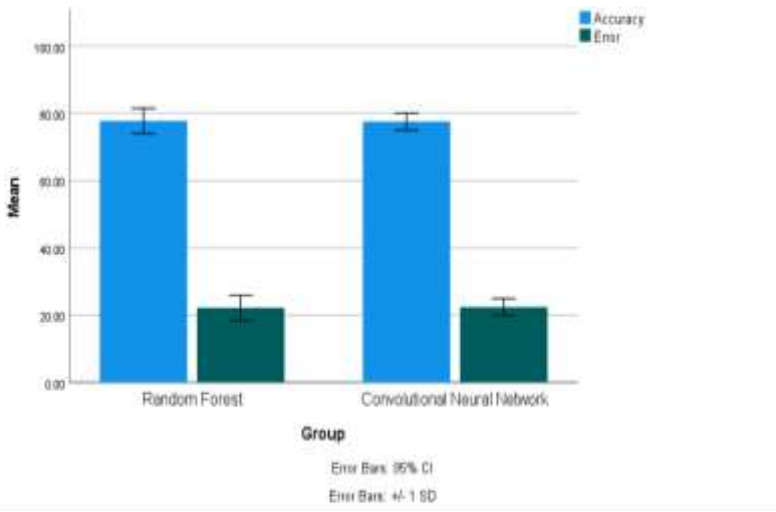

Fig. 1. Comparison of Novel Random Forest and Convolutional Neural Network in terms of accuracy. The mean accuracy of Novel Random Forest is greater than Convolutional Neural Network and the standard deviation is also slightly higher than Convolutional Neural Network. X-axis: Novel Random Forest vs Convolutional Neural Network. Y-axis: Mean accuracy of detection + 1 SD

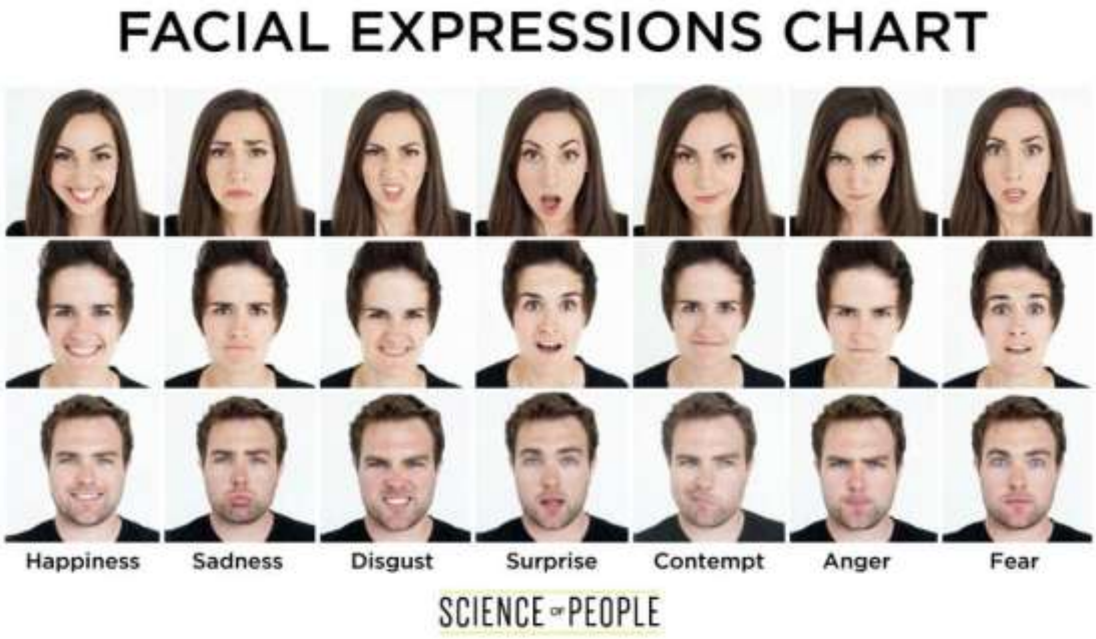

Fig. 2. Multiple images showing various expressions of human beings like Happiness, Sadness, Disgust, Surprise, Contempt, Anger and fear. Those images are collected from an independent dataset.