Facial Expression Recognition Using Machine Learning: Applications and Future Directions

Md. Sawar Hossain[a], Md. Rifat Hossen[b], Md Uzzal Mia[c]

a. *Associate Professor, Department of Information and Communication Engineering, Pabna University of Science and Technology, Pabna-6600, Bangladesh.*
b. *Department of Information and Communication Engineering, Pabna University of Science and Technology, Pabna-6600, Bangladesh.*
c. *Department of Information and Communication Engineering, Pabna University of Science and Technology, Pabna-6600, Bangladesh.*

Abstract:

Introduction: Facial Expression are a universal Language of emotion, transcending, cultural, and Linguistic barriers. They provide critical cues about a person's internal emotional state, which is pivotal in human communication. In the digital age, the ability to automate facial emotion recognition (FER) has gained immense interest across disciplines such as psychology, human-computer interaction, healthcare, and entertainment, by enabling machines to interact with human behavior and artificial intelligence [1].

Machine learning has revolutionized FER by offering models that learn directly from data, crimination the need for hand-crafted feature extraction [2]. Two commonly used datasets for FER research are FER-2013 and CK+48. The FER-2013 datasets are composed of a large collection of grayscale images categorized into seven emotions: anger, disgust, fear, happiness, sadness, sunrises, and neutral, making it ideal for training robust models [3]. In contrast, CK+48, known for its high-quality images, focuses on posed facial expressions, serving as an excellent benchmark for model performance [3].

Understanding human emotions is crucial for applications like virtual assistants, where FER can make interactions more intuitive and empathetic [4]. It is equally transformative in healthcare, particularly in mental health monitoring, where real-time emotion recognition can assist therapists and patients [5].

FER technology is already being applied in real-time applications, including customer sentiment analysis, smart surveillance, and immersive gaming experiences. With advancements in machine learning, its adoption is expected to grow across industries that demand human-centric Al solutions [6].

This research aims to develop a robust Machine-learning FER model by leveraging the FER-2013 and CK+48 datasets. By addressing challenges like data variability and augmenting and augmenting model training with advanced techniques, this study aspires to push the boundaries of FER accuracy and applicability.

Methodology:

This proposed method for facial expression recognition uses the algorithm, support Vector machine Random Forest, and Decision Tree classifiers. The platform used here is Python. Python is a general-purpose, object-oriented, high-level and powerful modern computer programming language. Python uses English keyword frequently where as other easy to maintain, portable and extendable language. It is processed at run time by the interpreter. You do not used to handle big data and perform complex mathematics. The block diagram for proposed method in show in fig. 1. [8]

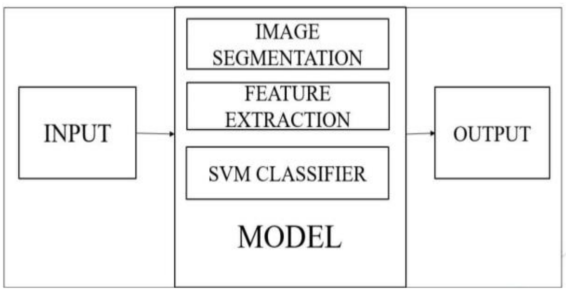

Fig. 1. Block diagram for propose method

The input is the image. Each image consists of dataset values. The input is obtained by using FER -2013 and CK+48 datasets.

Face emotion detection in used to predict the emotion state of the person based on their face expressions. The overview of the detection system I shown in the fig. 2 [9]

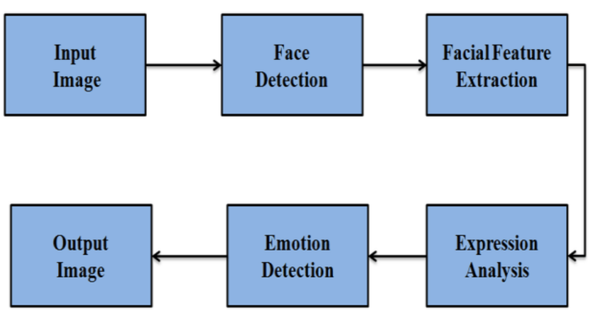

Fig. 2. Block diagram of emotion detection system.

Here input images are classified into two types,

- Training images
- Test images

Training images are used for training of classifier. Testing images are used to verify the algorithm by prediction the different emotions of the face. Expression analysis is the major part of the emotion detection, the schematic of expression analysis for classifying different emotions is shown in fig. 3.

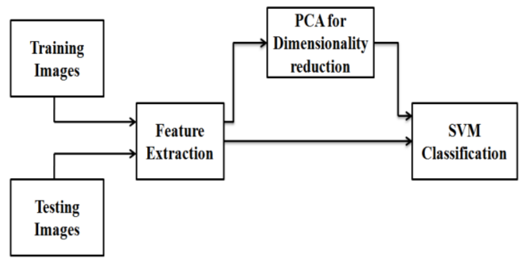

Fig.3. Block diagram of Emotion detection system.

PCA is applied to training images to reduce the dimensionality. Because training images are more compared to testing and if the dimension is high then the time taken for processing will more. Support Vector Machine classification is done for classifying different emotions namely, anger, disgust, fear, happiness, sadness, sunrises, and neutral. The emotion detection system detailed flow diagram shown in fig. 4.

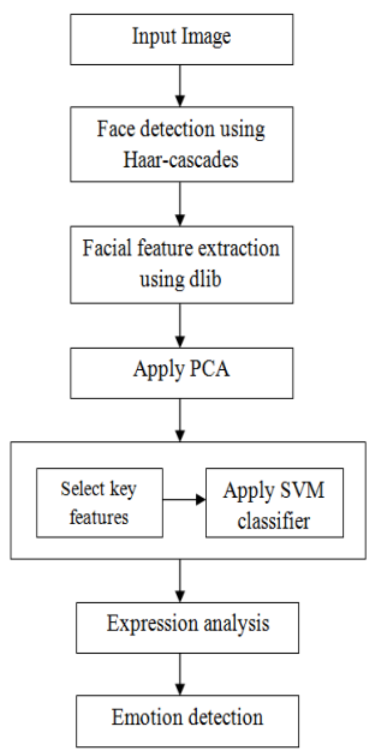

Fig. 4. Workflow of emotion detection system.

Support Vector Machine (SVM): A support vector machine is a supervised machine learning algorithm used for both classification and regression tasks. While it can be applied to regression problems, SVM is the best suited for classification tasks. The primary objective of the SVM is to identify the optimal hyperplane in an N-dimensional space that can effectively separate data points into different classes in the feature space. The algorithm ensures that the margin between the closest points of different classes, known as support vectors, is maximized. The classification of image can be done in linearly and nonlinearly separable data. Classification in linearly and nonlinearly separable data is shown in fig. 5. And fig. 6.

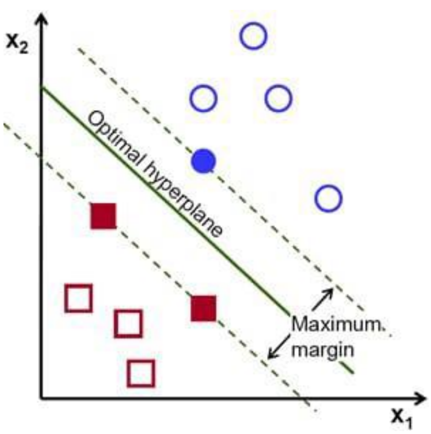

Fig. 5. Classification of Linearly separable data.

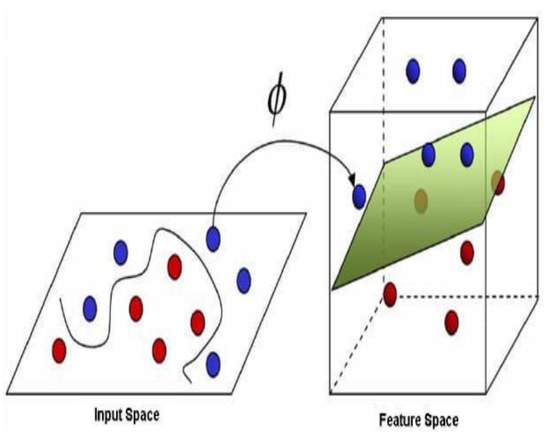

Fig. 6. Classification of non-linear separable data.

SVM have some popular kernel functions. The SVM is function that takes low-dimensional input space and transforms it into higher-dimensional space, i.e. it converts non separable problems to separable problems. It is mostly useful in non-linear separation problems. Simply put the kernel, does some extremely complex data transformations and then finds out the process to separate the data based on the labels or outputs defined.

- Linear: $K(w,b) = w^T x + b$
- Polynomial : $K(w,x) = (\gamma w^T x+b)^N$
- RBF: $K(w,x) = \exp(\gamma||x_i-x_j)^n$
- Sigmoid: $K(w,x) = \tanh(\alpha x_i^T x_j + b)$

Random Forest: Random Forest algorithm is a powerful tree learning technique in machine learning. It works by creating a number of Decision trees during the training phase. Each tree is constructed using a random subset of the data set to measure a random subset of features in each partition. This randomness introduces variability among individual trees, reducing the risk of overfitting and improving overall prediction performance. The below diagram explains the working of the Random Forest Algorithm fig. 7.

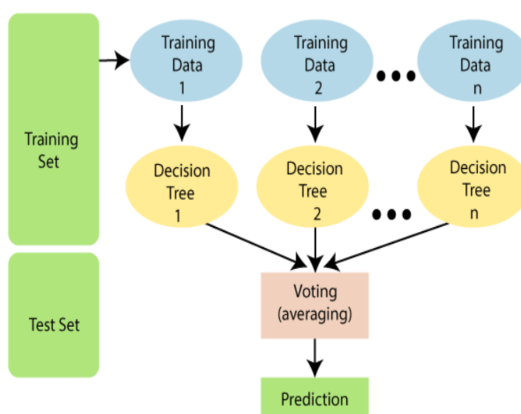

Fig. 7. Random Forest Algorithm Workflow for Classification and Prediction.

Decision Tree: A decision tree is a type of supervised learning algorithm that is commonly used in machine learning to model and predict outcomes based on input data. It is a tree-like structure where each internal node tests on attribute, each branch corresponds to attribute value and each leaf node represents the final decision or prediction. The decision tree algorithm falls under the category of supervised leaning. They can be used to solve the both regression and classification problems. We have two popular attribute selection measures:

- Information Gain: Gain(S,A)=Entropy(S) $+\sum_{V}^{A}\frac{|Sv|}{|S|}$.Entropy ($S_v$)
- Gain(S,A)=Entropy(S)- $\sum_{v\in Values(A)}^{A}\frac{|Sv|}{|S|}$. Entropy ($S_v$)

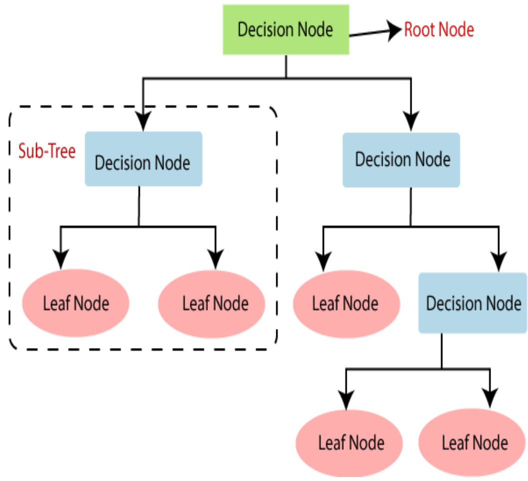

Fig. 8. Decision Algorithm Workflow for Classification and Prediction.

Data Visualization: The data collected from [10], contains two main folders named train and test. Train folder contains 7 subfolders (FER-2013) and CK+48 folder contains 7 main folders. Each of which contain around 1024 (FER-2013) and 150 (CK+48). Each image is in RGB format which size 48x48 fig. 8 and fig. 9. Shows sample database.

Fig. 8. FER-2013 Sample database.

Fig. 9. CK+48 Sample Database

Result:

The accuracy of linear, polynomial, and RBF kernel of SVM, Random Forest and Decision tree for classification report and confusion matrix. Both datasets calculate for different training and test ratio

For CK+48:

Table 1: Classification report for SVM

	Precision	Recall	F1-score	Support
Anger	1.00	1.00	1.00	27
Contempt	1.00	1.00	1.00	11
Disgust	1.00	1.00	1.00	35
Fear	1.00	1.00	1.00	15
Happy	1.00	1.00	1.00	42
Sadness	1.00	1.00	1.00	17
Surprise	1.00	1.00	1.00	50
Accuracy			1.00	197
Macro avg	1.00	1.00	1.00	197
Weighted avg	1.00	1.00	1.00	197

Table 2: Confusion Matrix for SVM

	Anger	Contempt	Disgust	fear	Happy	Sadness	Surprise
Anger	27	0	0	0	0	0	0
Contempt	0	11	0	0	0	0	0
Disgust	0	0	35	0	0	0	0
Fear	0	0	0	15	0	0	0
Happy	0	0	0	0	42	0	0
Sadness	0	0	0	0	0	17	0
Surprise	0	0	0	0	0	0	50

Table 3: Accuracy Results for SVM

Train-Test Ratio	Accuracy		
	Linear Kernel	Ploy kernel	RBF kernel
10% Test Set	100%	81.81%	90.90%
20% Test Set	100%	77.15%	85.27%
30% Test Set	98.98%	72.20%	87.79%
40% Test Set	96.18%	71.24%	85.24%
50% Test Set	92.87%	69.24%	83.50%

Table 4: Classification report for Decision Tree

	Precision	Recall	F1-score	Support
Anger	0.75	0.78	0.76	27
Contempt	0.86	0.55	0.67	11
Disgust	0.71	0.86	0.78	35
Fear	0.91	0.67	0.77	15
Happy	0.73	0.79	0.76	42
Sadness	0.83	0.59	0.69	17
Surprise	0.87	0.90	0.88	50
Accuracy			0.79	197
Macro avg	0.81	0.73	0.76	197
Weighted avg	0.79	0.79	0.78	197

Table 5: Confusion Matrix for Decision Tree

	Anger	Contempt	Disgust	fear	Happy	Sadness	Surprise
Anger	21	0	4	0	0	1	1
Contempt	2	6	2	0	1	0	0
Disgust	2	0	30	0	3	0	0
Fear	1	0	1	10	2	0	1
Happy	0	1	3	0	33	0	5
Sadness	1	0	0	1	5	10	0
Surprise	1	0	2	0	1	1	45

Table 6: Classification report for Random Forest

	Precision	Recall	F1-score	Support
Anger	1.00	0.96	0.98	27
Contempt	1.00	1.00	1.00	11
Disgust	1.00	0.91	0.96	35
Fear	1.00	1.00	1.00	15
Happy	0.93	1.00	0.97	42
Sadness	0.94	0.94	0.94	17
Surprise	0.98	1.00	0.99	50
Accuracy			0.97	197
Macro avg	0.98	0.97	0.98	197
Weighted avg	0.98	0.97	0.97	197

Table 7: Confusion matrix for Random Forest

	Anger	Contempt	Disgust	fear	Happy	Sadness	Surprise
Anger	26	0	0	0	0	1	0
Contempt	0	11	0	0	0	0	0
Disgust	0	0	32	0	0	0	0
Fear	0	0	0	15	0	0	0
Happy	0	0	0	0	42	0	0
Sadness	0	0	0	0	0	16	1
Surprise	0	0	0	0	0	0	50

For FER 2013:

Table 8: Classification report for SVM

	Precision	Recall	F1-score	Support
Anger	0.42	0.45	0.44	958
Disgust	0.83	0.54	0.66	111
Fear	0.44	0.41	0.43	1024
Happy	0.69	0.77	0.73	1774
Neutral	0.51	0.52	0.51	1233
Sad	0.46	0.43	0.45	1247
Surprise	0.77	0.69	0.73	831
Accuracy			0.56	7178
Macro avg	0.59	0.54	0.56	7178
Weighted avg	0.56	0.56	0.56	7178

Table 9: Confusion Matrix for SVM

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	433	5	107	129	126	142	16
Disgust	26	60	8	6	2	8	1
Fear	141	3	419	98	112	179	72
Happy	81	1	60	1368	129	100	35
Neutral	133	0	116	152	638	164	30
Sad	169	2	156	152	207	539	22
Surprise	45	1	76	64	40	32	573

For CK+48:

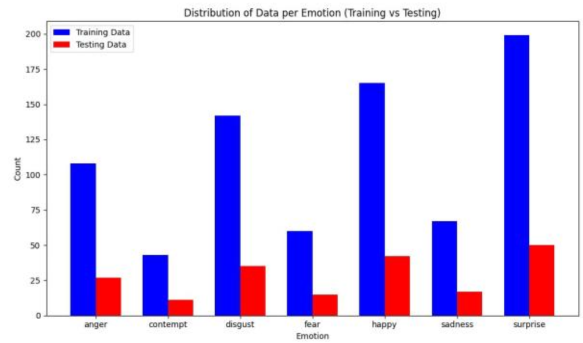

Fig.10. Training data vs Testing data

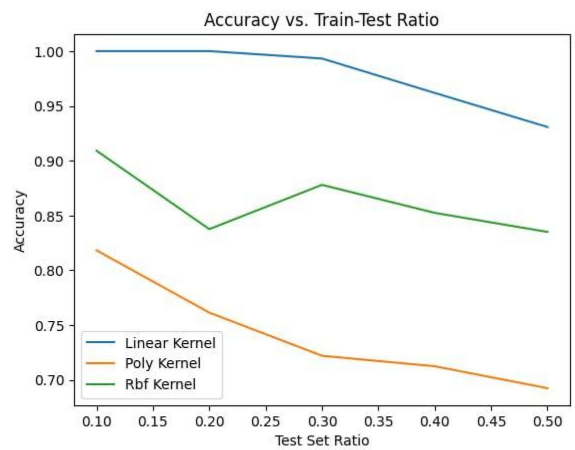

Fig.11. Accuracy for different Kernel Train-Test Ratio

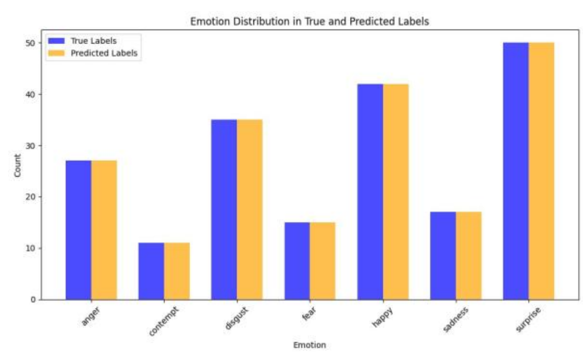

Fig. 12. True Labels vs Prediction Labels

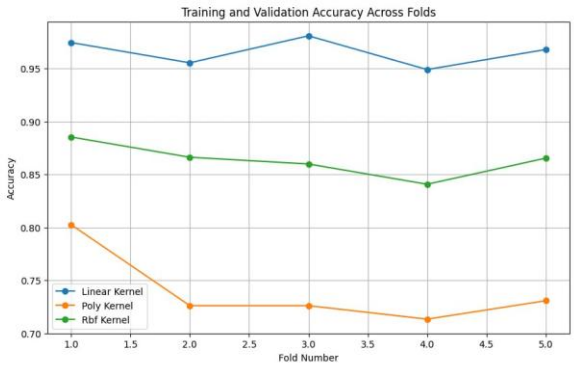

Fig. 13. Training and Validation Accuracy Across Fold

Fig. 14. Actual and Predicted Output

For FER 2013:

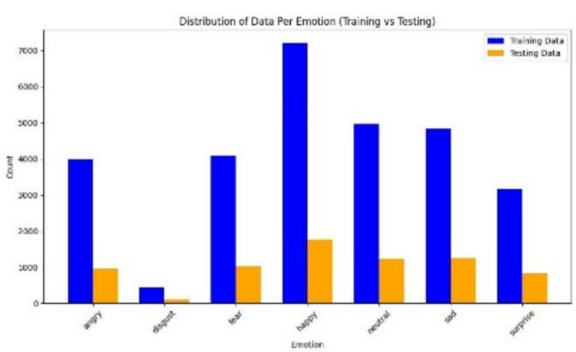

Fig. 15. Training vs Testing

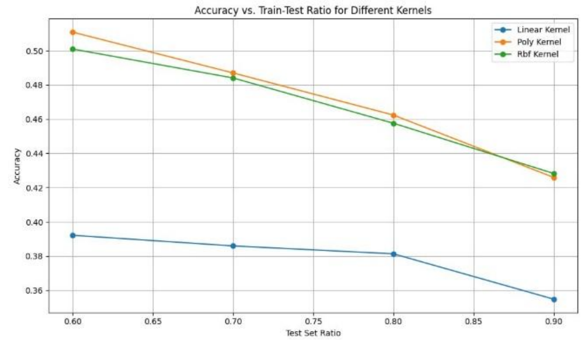

Fig. 16. Accuracy vs Train-Test for Different Kernels

Fig. 17. Actual and Predicted Output

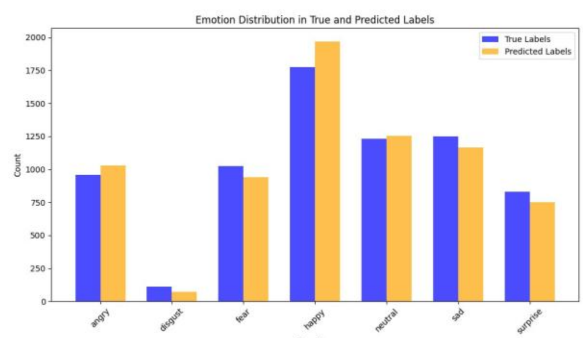

Fig. 18. True Labels vs Prediction Labels

For FER-2013, Random Forest and Decision Tree outer performs less than SVM. We can't include these models for these datasets. For CK+48, SVM archives the highest accuracy due to dataset's-controlled nature, While Random

Forest provides Competitive and robust performance. Decision Trees, though simpler, are better suited for preliminary experiments or interpretability-focused applications.

Conclusion

Facial Expression Recognition was performed using three machine learning models-Support Vector Machine (SVM), Random Forest, and Decision Tree- on two widely used datasets: RER-2013 and CK+48. The results highlight the strengths and limitations of each model in handling these datasets.

For the CK+48 dataset, the SVM with a linear kernel achieved the highest accuracy of 99.99% demonstrating its effectiveness in dealing with high-quality, controlled images of posed expression. Similarly, Random Forest and Decision Tree Models performed well, though sightly lower than SVM, owing to the small and balanced nature of the dataset, which suits all three algorithms.

In contrast, the FER-2013 dataset, known for its real-word complexity and class imbalances, posed significant challenges. The SVM with linear kernel achieved and accuracy of 56.65%, reflecting the dataset's noisy and imbalanced nature. Random Forest, with its ensemble approach, performed batter in handling class imbalance and noise compared to the Decision Tree, which often overfitting on the training data.

Overall, the study demonstrates that SVM is highly effective for controlled environment like CK+48, while Random Forest offers more robust performance for challenging real-world datasets like FER-2013.

REFERENCES

[1] Goodfellow, I., Bengio., & Courville, A. (20016). *Deep Learning*. MIT Press.

[2] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. *Nature*, 521(7553), 436-444

[3] Goodfellow, l., et al. (2013). Challenges in Representation Learning: A Report on Three Machine Learning Contests. *Proceedings of the IEEE International Conference on Machine Learning (ICML).*

[4] Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). The Extended Cohn-Kanade Dataset (CK+): A Complete Dataset for Action Unit and Emotion-Specified Expression. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*.

[5] Mollahosseini, A., Chan, D., & Mahoor, M. H. (2016). AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild. *Proceedings of the IEEE International Conference on Affective Computing and Intelligent Interaction (ACII)*.

[6] Dhall, A., Goecke, R., Joshi, J., Wagner, M., & Gedeon, T. (2012). Emotion Recognition in The Wild Challenge 2012: A Dataset and Comparison. *Proceedings of the ACM International Conference on Multimodal Interaction (ICMI).*

[7] Zhao, X., et al. (2021). Robust Emotion Recognition with Deep Learning Models: A Review. *Frontiers in Robotics and AI*.

[8] Afeefa Muhammed, Ramsi Mol, L. Revathy Vijay, S. S. Ajith, A. R. Shamna. Facial Expression Recognition using Support Vector Machine (SVM) and Convolutional Neural Network (CNN) (2020). *International Journal of Research in Engineering, Science and Management Volume-3, Issue-8, August-2020 journals.resaim.com/ijresm | ISSN (Online): 2581-5792 | RESAIM Publishing*

[9] Rajesh K M, Naveen kumar M. A robust method for face recognition and face emotion detection system using support vector machines (2016). *International Conference on Electrical,*

Electronics, Communication, Computer and Optimization Techniques (ICEECCOT)

[10] CK+48 and FER-2013 dataset's download site. *https://www.kaggle.com/datasets/hamzalahlal/ck48-data* *https://www.kaggle.com/datasets/msambare/fer2013*