Stacked Support Vector Machine Ensembles for Cross-Culture Emotions Classification

Khurshid Asghar[1], Mubbashar Sadddique[2], Inam ul Haq[3], M. Ahmad Nawaz-ul-Ghani[4], Ghulam Ali[5]

[1,2,3,5]Department of Computer Science, University of Okara, Okara, 56300, Pakistan
[4]Department of Computer Science University of Management and Technology, Lahore, 54000, Pakistan

Abstract
Facial expressions play a main role in representing the human's internal emotional state in social communication, but the scope to which they are universal and culture reliant is a subject of discussion. In this paper, we introduced the Stacked Support Vector Machine Ensembles (SVMEs) for cross-culture emotions classification. A pool of SVM ensembles is stacked to learn the cross-culture emotions. The SVM ensemble is a collection of a set of support vector machines. The outcomes of support vector machines were tied to the probability distribution across the support vector machine ensembles. The final decision about the presence of an emotion is made by naive Bayes predictor. The cross-cultural facial images from JAFFE, TFEID, KDEF, CK+ and RadBoud databases are combined to develop the multi-culture dataset. The participants of multi-culture database originate from following geography and ethnicity: Japanese, Taiwanese, Caucasians, Moroccans, Swedish, Asians, Northern Europeans, Euro-American, and Afro-American. The experimental results and inter-expression resemblance analysis demonstrate that the proposed ensemble approach performs significantly better than the stat-of-the-art ensemble techniques.

Key words:
Support Vector Machine Ensemble, facial expression classification, boosted ensemble classifier, universal emotions classification

1. Introduction

All public collaborations judgmentally depend on the common thoughtful of feelings, reached mainly by replacing a set of facial expressions. Due to performing such an essential purpose in human culture, facial expressions have been the basis of attraction and experiential examination between philosophers, biologists and psychologists for over an era. With the rapid growth in globalization and cultural integration within the digital economy, cross-cultural interaction is now becoming important in highly connected digital society. In order to meet these evolving communication needs, understanding the complexities of cross-cultural facial expression representation has recently become very important in various disciplines such as engineering, robotics, e-learning, healthcare, and many more. As a result, it is essential to growing highly sophisticated methods for cross-cultural facial expression recognition. The current facial expression techniques focus only on the culture of exact facial expression recognition. The first core impartial of this paper was the progress of a facial expression recognition system, which could recognize the cross-cultural emotions.

Several techniques for facial expression recognition have been developed. However, no effective effort has been made to recognize the universal emotions. The reason behind that is the human facial structure, which has too much diversity due to the difference of culture, colour, ethnic and geographic region [2]. Due to this diversity, the existing techniques attempt to classify facial expressions using single dataset. The major objectives of this research were to develop a system that recognizes universal emotions while satisfying the following conditions.

1. The classifiers should be trained and tested on universal dataset.
2. The extracted features should be invariant to difference of facial structure and appearance.
3. The availability of enough universal dataset.
4. The system works automatically without human intervention.

The best accuracy rates reported in the literature for [1] expression recognition on Japanese female facial expression (JAFFE) [3] is 98.3% [4]. The expression recognition accuracy on Taiwanese facial expression images database (TFEID) [5] is 96.2% [6]. The expression recognition accuracy on Karolinska directed emotional faces (KDEF) and extended Cohen Kanade (CK+) [7] is 46.17% [8] and 97.3% [9] respectively. However, RadBoud faces database (RaFD) reported in [10] with best recognition accuracy of 98.1% [11].

The HOG and PCA features were employed to construct the SVM classifiers, which are used as a base-level classifier in SVM ensembles. The results of Stacked SVMEs are combined using naive Bayes (NB) classifier to represent the most prominent emotion. The Stacking ensemble approach allows each SVME to focus on the prediction of a single [12] expression. Introduction of Stacked SVM ensembles (SSVME) is an innovative construction that gains the rewards of separate SVM. This technique transforms a multi discussion classification difficult into modest sub-problems.

Manuscript received July 5, 2019
Manuscript revised July 20, 2019

The NB classifier [7] was introduced as a final predictor, which consumes the output of SVM ensembles and predicts the sample image as one of the facial expression.
Rest of the paper is planned as follows; related work is elaborated in Section II, proposed stacked SVM ensembles explained in section III. A thorough description of experimental results is represented in section IV. Finally, concluding remarks and future directions are presented in section V.

2. Related Work

Prolonged and at time impassioned controversy has failed to settle the question of whether the facial expressions are universal. Since Darwin's significant effort on emotions in 1872, the representation of universal facial expressions has considered a debatable topic in psychological and social sciences. Until early 2009, the issue of cross-cultural facial expression recognition was comparatively ignored. In the facial expression recognition literature, the use of universal facial expression dataset is rare; and comparatively less attention has been given to the problem of universal facial expression recognition. Matsumoto performed cross-culture emotions analysis on spontaneously produced facial expressions by American, Japanese and British students. The signal clarity of each expression was similar across all cultures. Matthew et al. [13] introduced the neural network based machine learning technique for cross-cultural facial expression recognition. Presented that subtle variation in facial structure and expression representation is enough to confuse the classifier. It provides the directions to develop efficient classification techniques to deal with the cultural variations in facial expression representation. Silva and Pedrini [14] performed a similar experiment. The authors proposed that there are six basic facial expressions, which are innate and universal with delicate changes. A dataset consists of multi-culture facial expressions has been used in experiments. This dataset consists of four different facial expression databases. To outline boundaries and face surfaces, Gabor filters, HOG descriptors, and LBP have been used. ANN, SVM, and KNN have been used as classifiers to train the expression recognition model. However, the reported approach is unable to achieve hopeful results on the multi-culture dataset. Similarly, a method for cross-cultural facial expression recognition has been proposed by Zia and Jaffar [15], which uses an incremental learning model. Incremental learning models allow a classifier to accommodate the cross-cultural expression representation variations. JAFFE, CK, MUG, and FEEDTUM databases have been used to train the learned model. Experiments showed a recognition rate in terms of accuracy as 86.3%. As seen in the methods reported in [13-16], learned model did not outperform. Therefore, development of robust classifiers that can learn cross-cultural expression variations is still demanding.
Support vector machine has prospective applications in the field of machine learning. It is believed that a combination of SVM with HOG, LBP, and PCA will obtain best expression recognition rates. Monasor et al. [17] implemented a dynamic emotion detector using the active shape model and support vector machine. Active shape models are applied on Cohn-Kanade, JAFFE and MMI databases for feature extraction and SVM for expression classification. Moeini and Moeini [11], introduced SVM for facial expression recognition form the grouping of 2D and 3D feature vectors. Mlakar and Potocnik [18], presented a combination of SVM and optimal HOG descriptor for facial expression recognition. This combination of HOG descriptor and SVM classifier outperform the state-of-the-art techniques on CK dataset. In a recent work reported by [19] a novel approach based on active shape model has been proposed. The reported approach has been proposed to locate facial features from dynamic face region. To predict the presence of an expression, multi-class SVM has been trained.
The performance of SVM is remarkably better than traditional techniques of pattern recognition. Zavaschi et al. [20], employs the combination of Gabor wavelet transform and local binary pattern in an ensemble approach. A pool of support vector machines was trained using these two feature sets. The prediction about the existence of facial expression was made by combining the predictions of an optimal set of classifiers. Zavaschi et al. [21], represented a support vector machine-based ensemble approach for facial expression recognition using Gabor and LBP descriptors to represent the features of a facial image. A set of SVM classifiers was constructed using Gabor and LBP filters, and the multi-objective algorithm was employed to find the optimal combination of SVM classifiers.
Zheng [24] proposed a grid-based facial feature representation technique, where a sample image is divided into a set of sub-regions. Each region was independently processed for facial feature representation. The prediction about the presence of expression was made by the group sparse reduced-rank regression technique. This approach enables the system to automatically select the optimal sub-regions, which contribute most to the expression representation.
Recently, Renda et al. [22] compared several ensemble generation and aggregation techniques with respect to number of base level classifiers in the context of facial expression recognition. In order to add variability in base classifiers different pre-processing and pre-training procedures were adopted. Experiments showed that increasing the number of base classifiers beyond a certain limit becomes not significant, thus suggesting using boosting with average voting as an appropriate ensemble approach. Similarly, Chen et al. [26] proposed a two stage

CNN model: in the first stage expression frames are picked from image sequences, then in the second stage a DCNN recognize the six basic facial expression from picked frames. Achieved facial expression recognition accuracy of 95.4% and 77.4% on CK+ and BU-4DFE respectively.

3. Proposed Stacked SVM Ensemble

A support vector machine ensemble is a classifier containing a set of support vector machines {SVM (d, Vi), i=1, ..., N}. Where the Vn are independent identically distributed random vectors and each support vector machine contribute to determining the presence or absence of expression at the input (d, VI). Algorithm 1 presents the detailed description of the proposed SVM ensemble training procedure [24]. To apply the boosting approach to SVME in order to [22] construct the Stacked SVM ensembles we use the boosting technique presented in [23]. SVME rely on the binary output of SVM to compute the

Algorithm 1:

1. **Initialize:** **Training dataset {d1, y1}, …, {*dN*, *yN*}[1];**
   d_i is the sample dataset and y_i is the expression;
2. **Repeat** $y = 1$ to Y **do**
   Randomly select subset d_E from training set D
   Repeat $n = 1$ to N **do** (Number of SVMEs)
   Repeat k = 1 to K **do** (SVM count)
   Select sample set S from d_E
   Train SVM
   End
   Combine the predictions of SVM in SVME
   $$P(x) = \frac{1}{K}\sum_{i=1}^{K} h_i$$
   End
   End
3. Combine the decisions of SVMEs
4. Naive Bayes Predictor
5. Construct meta-data using training set D
   $$Y(d) = \max\ p(y)\prod_{d} p_{yx}^{d_x}(1 - p_{yx})^{1-dx}$$
   End

probability value in order to determine the presence of emotion. The probability to tie the outputs of individual support vector machines in SVME construction was computed using (1).

$$P(x) = \frac{1}{K}\sum_{i=1}^{K} h_i \qquad (1)$$

Where K is the number of SVM classifiers in an individual SVM ensemble. The prior probabilities were also calculated in the training process of support vector machine ensembles, which is further used to calculate the posterior probabilities in naïve Bayes classifier. The posterior probabilities were used to estimate the presence of facial expression from an unknown facial image. Consequently, an SVM ensemble dependent probability value was attained to compute the intensity of a possible facial expression over SVM ensemble. Subsequently, the process of construction of SVM ensembles encompasses three steps: first is the training of base-level SVM classifiers, second is to combine the predictions of SVM classifiers to construct SVM ensemble, finally stack the SVM ensembles. The decisions of individual SVM ensembles (SVMEs) were combined to detect the presence or absence of facial expression. In this research, we use the naïve Bayes classifier as a final predictor to combine the decisions of SVM ensembles (see Equation 2).

$$Y(d) = \max\ p(y)\prod_{d} p_{yx}^{d_x}(1 - p_{yx})^{1-dx} \qquad (2)$$

3.1 Boosted SVM Ensembles Training

The proposed ensemble construction approach based on multiple SVM ensembles. It detains over-fitting in the training process by randomly selecting the sample data for training of base-level classifiers. In order to train the individual SVM ensemble the bagging feature selection technique was applied, which randomly selects a subset of samples from the training samples. This subset training sample guides the SVM classifier, to learn the variations in the representation of cross-cultural facial expressions. The number of SVMs per expression in each SVME was varied to be 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100. The final decision about the number of SVMEs in the proposed ensemble approach was decided by analysing the performance of the classifiers. We adapted the similar boosting technique proposed in [27] and [28], where each individual is a complement to one another. The SVMEs were trained iteratively, where each new SVME is affected by the performance of previously trained SVMEs. The boosting approach encourages new SVMEs to efficiently recognize those subjects misclassified by earlier ones. It randomly selects a subset of samples from training data to train the base-level support vector machines. Subsequently, the decisions of base-level SVMs were combining to construct SVM ensembles. The weights of training data were modified according to the performance of the current SVM ensemble on training data. The weights of misclassified samples were increased, and the weights of correctly classified samples were increased. This weight modification process divides the training samples into two different sets, one with lower weights and the other with higher weights. Each new SVM ensemble in successive repetitions innately concentrate on those samples, which have higher weights. As a result, the weights of training samples are modified with respect to the expression recognition accuracy of current SVM ensemble. It could be probable that some samples may have very high weights and some samples have very low weights. The weights of training samples emulate how frequently the sample images have been correctly recognized by the SVM ensembles trained so far. By retaining the ratio of modification in weights, this technique gives a smooth method of training SVM ensembles that complement one another. The ratio of modification in weights depends on the performance of current SVME. However, Stacked SVM ensembles provide high generalization because of a random selection of training samples for boosted SVM ensembles generation.

4. Experiments

In this research, the experiments were performed using the multi-cultural facial expression database. This dataset contains 2006 facial images of female and male objects. Whereas, 838 images from RaFD, 150 images from JAFFE, 398 images from TFEID, 244 images from CK+ and 376 images from KDEF. For all the experiments the 1404 images of cross-cultural database have been used for training and 602 images used to evaluate the performance of Stacked SVMEs. Before representing the detailed experimental results of proposed ensemble technique, we used the various combinations of three facial feature representation techniques (PCA, LBP, HOG) with NB and KNN classifiers that provide the best expression recognition accuracy. In this regard, we performed several experiments by changing the SVMEs count for each expression from 10 to 20, 30, 40, 50, 60, 70, 80, 90 and 100 with 20 SVM in each SVME. The combination of HOG and naive Bayes with best and most consistent expression recognition accuracy was considered as the best combination of facial features representation technique and final predictor.

Table 1 demonstrates the performance of boosted SVMEs, where NB and KNN classifiers were used as a final predictor and PCA, LBP and HOG filters for facial features representation. The experimental results indicate that the accuracy of 100 SVMEs with the combination of HOG and NB is superior as compared to other SVME count. The best emotion recognition accuracy achieved is 90.7%. These experiments indicate that the expression recognition accuracy of SVMEs with NB and KNN as a final predictor and PCA feature vectors is a little bit lower as compare to HOG feature vectors. The expression recognition accuracy of proposed ensemble classifier is poor on LBP descriptors as illustrated in Table 1. From these interpretations, we can say that LBP is not a good choice for universal emotions classification. The similar work presented in [19] also supports these arguments.

Table 1: Expression recognition accuracy with NB and KNN final predictors

	PCA	**LBP**	**HOG**
NB			
$SVME_{50}$	88.04	84.72	90.03
$SVME_{100}$	87.38	84.72	**90.70**
$SVME_{150}$	87.04	84.72	90.03
$SVME_{200}$	88.04	84.39	89.87
$SVME_{250}$	87.54	84.72	90.37
$SVME_{300}$	87.38	84.22	90.03
$SVME_{350}$	**88.04**	84.88	90.37
$SVME_{400}$	87.21	**85.22**	90.03
$SVME_{450}$	87.54	84.55	90.03
$SVME_{500}$	87.54	84.88	90.03
KNN			
$SVME_{50}$	86.21	84.05	89.87
$SVME_{100}$	86.54	84.22	**90.20**
$SVME_{150}$	86.38	**85.05**	90.37
$SVME_{200}$	86.71	84.05	89.53
$SVME_{250}$	86.54	84.39	89.70
$SVME_{300}$	86.88	83.72	89.70
$SVME_{350}$	86.54	84.39	89.70
$SVME_{400}$	**87.04**	83.72	89.70
$SVME_{450}$	86.38	84.22	89.37
$SVME_{500}$	86.54	83.72	90.03

Table 2: Confusion matrix for the universal emotion classifier with NB predictor and HOG filter

	AN	**HA**	**SU**	**SA**	**FE**	**Total**
AN	92.5	1	1	7	0	120
HA	3	96.7	0	1	0	124
SU	1	1	91.4	2	8	141
SA	12	1	3	85.9	1	121
FE	6	1	1	6	85.4	96
Total	133	124	134	120	91	602

The detailed description of each emotion with a different combination of SVMEs, final predictor, and feature representation techniques are presented in following confusion matrixes. In the confusion matrixes, the column headings specify the detected emotion and the row headings representing the true emotion. The labels (AN stands for anger, HA for happiness, SU for a surprise, SA for sadness, and FE for fear) are used for row and column heading. Whereas, the total is the sum of rows and columns. The diagonal values represent the correct classification and off-diagonal values represent the misclassified emotions. Each row of these matrices provides a clear indication about the strength of confusion on each expression pair.

Table 3: Confusion matrix for the universal emotion classifier with KNN predictor and HOG filter

	AN	**HA**	**SU**	**SA**	**FE**
AN	91.6	1	1	8	0
HA	4	95.9	0	1	0
SU	1	1	92.1	1	8
SA	13	1	3	85.1	1
FE	7	1	2	5	84.3

Table 4: Confusion matrix for the universal emotion classifier with NB predictor and PCA filter

	AN	**HA**	**SU**	**SA**	**FE**	**Total**
AN	83.3	2	3	13	2	120
HA	5	91.9	1	1	3	124
SU	4	0	90.0	4	5	141
SA	11	0	3	84.2	5	121
FE	3	4	3	3	86.4	96
Total	123	120	137	123	98	602

Table 5: Confusion matrix for the universal emotion classifier with KNN predictor and PCA filter

	AN	HA	SU	SA	FE	Total
AN	81.6	2	9	11	0	120
HA	3	91.1	2	3	3	124
SU	1	1	91.4	4	6	141
SA	7	0	6	85.9	4	121
FE	4	1	5	6	83.3	96
Total	123	117	151	128	93	602

Table 6: Confusion matrix for the universal emotion classifier with NB predictor and LBP filter

	AN	HA	SU	SA	FE	Total
AN	88.3	1	5	7	1	120
HA	2	94.3	3	1	1	124
SU	6	1	90.7	0	6	141
SA	17	0	2	78.5	7	121
FE	10	6	8	5	69.7	96
Total	141	125	146	108	82	602

The confusion matrixes presented in Table 2, Table 3, Table 4, Table 5, Table 6 and Table 7 illustrates that emotions of fear and sadness are incorrectly recognized as most similar emotions surprise and anger. This resembles the expression recognition accuracies presented in [26], where it is mentioned that the confusion is owing to the emotions pair fear-sadness have similar facial appearance around eyes similarly to emotions pair surprise-anger.

The maximum confusion occurs between emotions sadness and anger, fear and surprise. Moreover, from these results, it can be observed that among the five emotions, the emotions happiness and surprise are easy to recognize as compare to sadness, fear, and anger. It demonstrates that the use of SSVMEs trained with HOG features outperforms for all emotions. Moreover, the highest recognition accuracy (96.7%) achieved on emotion happiness and the lowest (69.7%) is on emotion fear.

Table 7: Confusion matrix for the universal emotion classifier with KNN predictor and LBP filter

	AN	HA	SU	SA	FE	Total
AN	79.2	1	9	14	1	120
HA	1	94.3	2	1	3	124
SU	1	0	93.6	0	8	141
SA	11	0	4	80.1	9	121
FE	5	5	9	7	72.9	96
Total	113	123	156	119	91	602

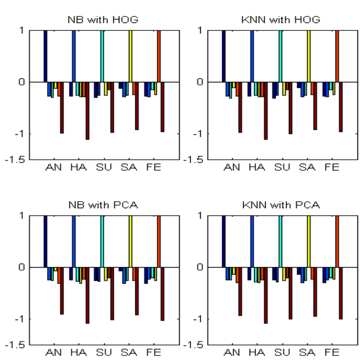

Fig. 1 Correlation coefficients based on comparing recognized and actual emotions for each of the five emotions.

4.1 Inter-Expression Resemblance Analysis

Based on confusion matrices represented in Table 2, Table 3, Table 4, and Table 5, we performed the inter-expression resemblance analysis between different expression pairs. The inter-expression resemblance was computed based on correlation coefficients as shown in Fig. 1. The correlation coefficients were computed using eq. (3),

$$R(i,j) = \frac{C(i,j)}{\sqrt{C(i,i)C(i,j)}} \tag{3}$$

Fig. 1 shows that there is no positive correlation between two emotions. Though, little negative correlations found for some expressions, for example, surprise-fear, anger-sadness, and sadness-anger. These values represent the intensity of resemblance between two expressions which have been observed independently of each other. When emotions of sadness-anger are considered, it is observed that these two expressions have low resemblance than other expression pairs. It indicates that the expressions of sadness-anger are antagonistic. In contrast, when we compare the other expression pairs such as anger-happiness, anger-surprise, and happiness-surprise a different correlation could be inferred. Furthermore, while considering the expression of fear-anger, that have less resemblance as compare to fear-surprise, we can conclude that, in the occurrence of fear emotion, the emotion pair anger-fear have lower resemblance as compare to expression fear-surprise. In the same way, the combination of anger-sadness suggests that these two expressions have more resemblance as compare to other emotion pairs. Thus, the significance of the correlation coefficients comes from their relative values instead of absolute ones. Whereas, "SUM" presents the aggregate value of negative correlation coefficients of all emotions against a particular emotion. The higher negative sum coefficient represents the lower resemblance between the respective expression and others. Therefore, it could be concluded that expressions of happiness, anger, and surprise are innate and universal across all cultures.

Table 8: Comparisons with state-of-the arts methods

Reference	Feature	Classifier	Dataset	Accuracy
Zavaschi et al. [20]	LBP	SVM	JAFFE	96.2
Farajzadeh et al. [21]	Gabor, LBP	SVM	JAFFE	96.2
			CK	88.9
Meguid et al. [7]	LBP	SVM	KDEF	46.17
Ali et al. [22]	MPC	SVM	CK	93.1
Zavaschi [19]	Semantic	SVM	CK+	94.7
Murray et al. [11]	Gabor	SVM	RaFD	98.1
Zia et al. [14]	HOG	SVM	Multi-culture	79.8
Matsumoto et al. [15]	LBP	ANN	Multi-culture	86.3
Zheng [23]	HOG	NNE	Multi-culture	93.05
Proposed method	HOG	SVME	Multi-culture	90.70

4.2 Comparison with existing approaches

Expression recognition rate in terms of accuracies obtained from the test set is higher than the results represented in [14, 15, 23], as shown in Table 8. The proposed method achieved 90.7% expression recognition accuracy using the combination of SSVM ensembles and HOG features. Whereas, expression recognition approach suggested by [14] using the combination of SVM and HOG features achieved accuracy of 79.8%. Similarly, achieved accuracy for similar aims represented in [15] was 86.3%, which used an incremental learning approach with ANN classifier and LBP features. Moreover, the results presented in [23] are better than the expression recognition accuracy achieved in this research. The variations between the results presented in this research and [23] are due to the difference in a number of facial expression databases. In this research, we used the five facial expression databases instead of three databases, which represent more cultural diversity as compare to [23]. The variation in expressions recognition accuracy obtained in this research is because of cultural differences in facial expression databases.

5. Conclusion

The construction of Stacked Ensembles focuses on the investigation that how the classifier will be trained to accurately learn the cross-cultural expression variations. Specifically, the classifier should be able to learn the expression similarities primarily from Moroccan, Caucasian, Swedish, American, and European subjects to the Japanese, Afro-Americans, and Taiwanese subjects. In this regard, we introduced Stacked SVM ensembles. The introduction of Stacked SVM ensembles is entirely a novel structure that gains the advantage of individual SVME. The Stacking ensemble technique combines multiple SVMEs which are a complement to one another. Boosting encourages new SVMEs to more accurately classify the instances misclassified by earlier ones. However, SVM ensembles provide high generality by using a large number of SVMs for ensemble learning. The binary output of an SVME tells about the possibility of the presence of respective expression. The boosting technique gives the chance to optimize the SVME construction process, which ultimately leads to the development of more structured ensemble classifier. Therefore, it is a challenging task for the SVM ensemble to recognize the facial expressions accurately, in the presence of variations in expression representation and facial appearance. The results obtained using the proposed ensemble approach are significantly comparable to the performance of multi-culture facial expression recognition techniques presented in the literature [14, 15, 23]. Consequently, we observed that the combination of Stacked SVM ensembles with NB predictor and HOG features were particularly important as it provides the best combination of ensemble classifier and feature representation technique.

There is a tremendous scope of research in the area of cross-cultural emotions recognition. The significant focus of future research will be on the development of ensemble techniques, facial feature representation techniques and pre-processing techniques to overcome the limitations of this research. Thus, the focus of future research should be the recognition of micro-expressions such as tension, like, dislike, agree, disagree, inconvenience, stress, fake smile, pleasant, unpleasant, interest, satisfaction, disappointment, convenience, impressed, appreciation and exciting. These micro-expressions will be useful in many applications like measuring consumers' emotions for product packaging design, the emotional response to products price tagging, offering discounts to customers, advertisement, health care, and social collaborations.

References

[1] D. Lundqvist, A. Flykt, and A. Öhman, "The Karolinska directed emotional faces (KDEF)," CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, vol. 91, p. 630, 1998.

[2] J. Fitzgerald, "Editor's Introduction," New Zealand Journal of Psychology, vol. 46, p. 5, 2017.

[3] M. J. Lyons, J. Budynek, and S. Akamatsu, "Automatic classification of single facial images," IEEE transactions on pattern analysis and machine intelligence, vol. 21, pp. 1357-1362, 1999.

[4] X.-F. Bai and W.-J. Wang, "An approach for facial expression recognition based on neural network ensemble," in Machine Learning and Cybernetics, 2009 International Conference on, 2009, pp. 19-23.

[5] L.-F. Chen and Y.-S. Yen, "Taiwanese facial expression image database," Brain Mapping Laboratory, Institute of Brain Science, National Yang-Ming University, Taipei, Taiwan, 2007.

[6] H.-C. Lee, C.-Y. Wu, and T.-M. Lin, "Facial Expression Recognition Using Image Processing Techniques and Neural Networks," in Advances in Intelligent Systems and Applications-Volume 2, ed: Springer, 2013, pp. 259-267.

[7] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression," in Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on, 2010, pp. 94-101.

[8] M. K. A. El Meguid and M. D. Levine, "Fully automated recognition of spontaneous facial expressions in videos using random forest classifiers," IEEE Transactions on Affective Computing, vol. 5, pp. 141-154, 2014.

[9] D. Ghimire and J. Lee, "Extreme Learning Machine Ensemble Using Bagging for Facial Expression Recognition," JIPS, vol. 10, pp. 443-458, 2014.

[10] O. Langner, R. Dotsch, G. Bijlstra, D. H. Wigboldus, S. T. Hawk, and A. Van Knippenberg, "Presentation and validation of the Radboud Faces Database," Cognition and emotion, vol. 24, pp. 1377-1388, 2010.

[11] A. Moeini and H. Moeini, "Multimodal facial expression recognition based on 3D face reconstruction from 2D images," in Face and Facial Expression Recognition from Real World Videos, ed: Springer, 2015, pp. 46-57.
[12] I. R. Murray and J. L. Arnott, "Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion," The Journal of the Acoustical Society of America, vol. 93, pp. 1097-1108, 1993.
[13] M. N. Dailey, C. Joyce, M. J. Lyons, M. Kamachi, H. Ishi, J. Gyoba, et al., "Evidence and a computational explanation of cultural differences in facial expression recognition," Emotion, vol. 10, p. 874, 2010.
[14] F. A. M. da Silva and H. Pedrini, "Effects of cultural characteristics on building an emotion classifier through facial expression analysis," Journal of Electronic Imaging, vol. 24, p. 023015, 2015.
[15] M. S. Zia and M. A. Jaffar, "Cross-cultural emotion classification based on incremental learning and LBP-features," Applied Mathematics & Information Sciences, vol. 9, p. 2651, 2015.
[16] D. Matsumoto, A. Olide, J. Schug, B. Willingham, and M. Callan, "Cross-cultural judgments of spontaneous facial expressions of emotion," Journal of Nonverbal Behavior, vol. 33, p. 213, 2009.
[17] E. Lozano-Monasor, M. T. López, A. Fernández-Caballero, and F. Vigo-Bustos, "Facial expression recognition from webcam based on active shape models and support vector machines," in International Workshop on Ambient Assisted Living, 2014, pp. 147-154.
[18] U. Mlakar and B. Potočnik, "Automated facial expression recognition based on histograms of oriented gradient feature vector differences," Signal, Image and Video Processing, vol. 9, pp. 245-253, 2015.
[19] C.-C. Hsieh, M.-H. Hsih, M.-K. Jiang, Y.-M. Cheng, and E.-H. Liang, "Effective semantic features for facial expressions recognition using SVM," Multimedia Tools and Applications, vol. 75, pp. 6663-6682, 2016.
[20] T. H. Zavaschi, A. L. Koerich, and L. Oliveira, "Facial expression recognition using ensemble of classifiers," in Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on, 2011, pp. 1489-1492.
[21] T. H. Zavaschi, A. S. Britto Jr, L. E. Oliveira, and A. L. Koerich, "Fusion of feature sets and classifiers for facial expression recognition," Expert Systems with Applications, vol. 40, pp. 646-655, 2013.
[22] N. Farajzadeh, G. Pan, and Z. Wu, "Facial expression recognition based on meta probability codes," Pattern Analysis and Applications, vol. 17, pp. 763-781, 2014.
[23] G. Ali, M. A. Iqbal, and T.-S. Choi, "Boosted NNE collections for multicultural facial expression recognition," Pattern Recognition, vol. 55, pp. 14-27, 2016.
[24] W. Zheng, "Multi-view facial expression recognition based on group sparse reduced-rank regression," IEEE Transactions on Affective Computing, vol. 5, pp. 71-85, 2014.
[25] A. Renda, M. Barsacchi, A. Bechini, and F. Marcelloni, "Comparing ensemble strategies for deep learning: An application to facial expression recognition," Expert Systems with Applications, vol. 136, pp. 1-11, 2019.
[26] J. Chen, Y. Lv, R. Xu, and C. Xu, "Automatic social signal analysis: Facial expression recognition using difference convolution neural network," Journal of Parallel and Distributed Computing, vol. 131, pp. 97–102, 2019.
[27] E. Hadavandi, J. Shahrabi, and S. Shamshirband, "A novel Boosted-neural network ensemble for modeling multi-target regression problems," Engineering Applications of Artificial Intelligence, vol. 45, pp. 204-219, 2015.
[28] E. Owusu, Y. Zhan, and Q. R. Mao, "A neural-AdaBoost based facial expression recognition system," Expert Systems with Applications, vol. 41, pp. 3383-3390, 2014.

Khurshid Asghar is working as Assistant Professor Computer Science /Head Department of Computer Science University of Okara, Pakistan. Dr. Asghar worked as research associate for one year at Cardiff School of Computer Science and Informatics, Cardiff University, UK. He earned his PhD (Computer Science) from COMSATS University Islamabad, Lahore Campus in the field of artificial intelligence. His current research interest includes image processing, image forensics, video forensics, machine learning, deep learning, network security, biometrics, medical imaging and brain signals.

Mubbashar Saddique is working as Lecturer Computer Science at University of Okara Pakistan. He got merit scholarship from COMSATS University Islamabad, Pakistan where he completed his MS computer science in 2010. Currently, he is PhD Scholar at COMSATS University Islamabad, Pakistan. Currently he is working in video and image forensic domain. Furthermore, his research interest is in the area of image/video processing, computer vision, machine learning, data mining and networks.

Inam ul Haq MS Computer Science is serving as Assistant Professor in Computer Science in University of Okara, Okara with 14+ years of teaching experience at graduate levels. Inam has research interests in the field of Ergonomics for Older People and Related Technologies, Wearable Computing for Health and Safety and Ergonomic IT issues related to Older People. Inam likes to work in close relations with other related researchers. The innovative idea to write such article is originated by Inam and completed and executed along with cooperation of Dr. Jenny Lundberg from Lund University Sweden. University of Okara, Pakistan 56300.

Muhammad Ahmad Nawaz Ul Ghani received the BS degree in software engineering from the University of Management and Technology Lahore in 2015, and the MS in Software Engineering from the University of Management and Technology Lahore in 2017. Muhammad Ahmad Nawaz Ul Ghani is a young researcher, software engineer. Recently working with the “University of Okara” as lecturer in the department of computer sciences. He has been performed duties as lecturer in the department of information & technology at University of Gujrat (Lahore campus). His main areas of research interest are software engineering, agile developments, data mining. He is much too interested to show zeal & zest and imitativeness in Computer science in a higher panoramic environment.

Ghulam Ali is working as Assistant Professor Computer Science at Department of Computer Science University of Okara, Pakistan. He earned his PhD (Computer Science) from University of Central Punjab, Lahore Pakistan in the field of artificial intelligence. His current research interest includes image processing.