Sensitivity Analysis of Culture-specific and Cross-cultures Facial Expression Resemblance

Umar Bashir, Ghulam Ali*, Ramzan Talib, Ibrahim Abdul Rehman

Department of Computer Science, Government College University Faisalabad
Contact author*

Abstract: Cross-cultural judgment studies of facial expressions of emotion constitute one of the empirical pillars of the universality of emotional expression. In this paper, cross-culture sensitivity analysis approach is applied to determine the effect of cultural variations on the performance of automated facial expression recognition system. It evaluates how cultural variations of individuals from different cultural groups including Japanese, Taiwanese, Moroccans, Caucasians, Swedish, Europeans, Euro Americans, and Afro Americans affect the facial displays of emotions. The sensitivity analysis approach determines the cross-cultural resemblance in facial displays of emotions while considering different combinations of facial expressions. Different experiments are performed to identify which expression combinations are most influential for the performance of automated facial expression recognition system. Experimental results demonstrate that expression recognition accuracy of facial expression recognition system is higher than there is a subtle resemblance in expressions combination. The expression of sadness has more resemblance to anger as compared to surprise and happiness similarly expression of fear has more resemblance to surprise as compared to anger and happiness. It indicates that in case of anger and happiness the culture variations do not affect the facial displays of emotions.

Keywords: Sensitivity Analysis, Facial Expression, Principal Component Analysis, Support Vector Machines, Histogram of Oriented Gradients.

1. Introduction

In the recent digitally globalized world, social and business communications involving people from different geography, ethnicity, and cultural backgrounds is not unusual. Even though the fact that requires rules and practices for interacting globally vary in marked ways across cultures. One aspect of non-verbal communication that is fundamental for interpersonal affinity is the ability to correctly interpret the facial displays of emotion. This study represents one of the first attempts to investigate whether culture variations affect the performance of automated facial expression recognition system. In this regard, a multi-culture dataset is constructed which contains the facial images of people that belong to Japanese, Taiwanese, Caucasians, Moroccans, Europeans, Euro Americans, Afro Americans, Swedish, and Asians cultures. This dataset will provide a vast variety of cultural variations in facial displays of emotions that are an integral part of cross-culture communication.

The sensitivity analysis approach is applied to determine the resemblance between participants of cross-culture expressions. The classifiers accuracy decreases where different expression combinations have more resemblance. It pointed out that classifiers are sensitive with respect to different expression combinations in cross-culture expression representation. Sensitivity analysis is performed using stacking ensemble technique where Support Vector Machines (SVM) used as base level classifier while Naive Bayes (NB), K-Nearest Neighbor (KNN), and Artificial Neural Network (ANN) as a meta-level classifier. We also introduced a hybrid feature extraction technique HOG-PCA for facial features representation with reduced dimension. The effectiveness of HOG-PCA is also compared with other state-of-the-art techniques like principal component analysis (PCA) and local binary patterns (LBP) by varying combinations of facial expressions to determine the effect of cultural diversity on facial displays of emotions.

Universality hypothesis claims that six basic facial expressions are common across all cultures in the world [1]. Recently, Da Silva and Pedrini [2] suggested that facial expression is innate and universal across all cultures with few minor differences. They also pointed that there is confusion between different expressions due to minor variations among facial displays of emotions in different cultures. Yan et al. [3] applied photo sorting task in which Chinese and Caucasian participants were asked to sort photos by identity or by expression. Caucasian participants faced difficulty in identifying the expressions of Chinese. Similarly, Chinese participants also faced difficulty in recognizing Caucasian faces identities and expressions. They conclude that members of other cultures are unable to learn the cultural differences. Zia and Jaffar [4] proposed facial expression recognition system that can offer incremental learning. The region of interest was extracted to minimize the cultural variations and classified using template matching algorithm. Five different datasets are used for experimental purpose. Although, Ali et al. [5] proposed a technique for multi-cultural facial expression classification that is based on neural network ensemble. PCA, LBP, and HOG feature extraction techniques are employed. Similarly, experiments are performed on JAFFE, TFEID, and RaFD datasets that achieved better recognition accuracy as compared with others.

For machine learning methods, the accurate recognition of facial expression is a major issue due to a different representation of emotions across all cultures. Similarly, images of the same person in same facial expression can vary in terms of background, pose, and brightness. Therefore, different machine learning and pattern recognition algorithms have been introduced to recognize the facial expression efficiently [6-9].

To enhance facial expression recognition accuracy, ensemble techniques are being used. Yu and Zhang [10] introduced a novel approach that based on multiple deep convolutional neural network for better facial expression recognition that achieved higher recognition accuracy using image sequences and static facial expressions in the wild.

2. Literature review of culture-dependent and cross-cultures systems

The cultural-dependent and cross-cultures expression recognition accuracies on JAFFE, CK+, RaFD, TFEID, KDEF, and cross-cultural datasets presented in the literature are illustrated in tables 1,2,3,4,5, and 6 respectively.

Table 1. Facial expression recognition accuracy (%) for JAFFE dataset

Method	Facial Features	Classifier	Anger	Happy	Surprise	Sad	Fear	Average
2016 [11]	LBP, 3DH-LLBP, NBP	SVM	98.7[3]	98.1[5]	99.2[1]	98.4[4]	98.9[2]	98.51
2016 [12]	Gabor filter	SVM	93.33[4]	93.33[4]	100[1.5]	100[1.5]	93.33[4]	97.14
	Average Rank		3.5	4.5	1.25	2.75	3	

Table 2. Facial expression recognition accuracy (%) for CK+ dataset

Method	Facial Features	Classifier	Anger	Happy	Surprise	Sad	Fear	Average
2015 [13]	LBP	SVM	87.8[5]	94.2[3]	98.46[1]	96.42[2]	93.33[4]	94.10
2017 [14]	Spatio-Temporal	SVM	91.1[3]	98.6[1]	97.6[2]	89.3[4]	80[5]	89.35
	Average Rank		4	2	1.5	3	4.5	

Table 3. Facial expression recognition accuracy (%) for RaFD dataset

Method	Facial Features	Classifier	Anger	Happy	Surprise	Sad	Fear	Average
2015 [15]	Gabor filter	SVM	96.7[5]	98.6[2.5]	100[1]	98.1[4]	98.6[2.5]	98.11
2015 [16]	HOG	SVM	86.8[5]	100[1]	97.4[3]	98.2[2]	93.9[4]	94.1

Average Rank			5	1.75	2	3	3.25		

Table 4. Facial expression recognition accuracy (%) for TFEID dataset

Method	Facial Features	Classifier	Angry	Happy	Surprise	Sad	Fear	Average
2014 [17]	LBP, Gabor	MPC	94.11[3]	100[1.5]	100[1.5]	82.05[5]	85[4]	92.8
2014 [18]	Face++	SVM	78.43[5]	94.15[1]	91.26[2]	80.81[4]	81.45[3]	85.22
Average Rank			4	1.25	1.75	4.5	3.5	

Table 5. Facial expression recognition accuracy (%) for KDEF dataset

Method	Facial Features	Classifier	Angry	Happy	Surprise	Sad	Fear	Average
2005 [19]	AAM	ANN	80[5]	97[1]	85[3.5]	85[3.5]	93[2]	89.00
2012 [20]	Constrained Local Model	SVM	71.43[4]	92.14[1]	85.71[2]	80[3]	70.71[5]	87.00
Average Rank			4.5	1	2.75	3.25	3.5	

Table 6. Facial expression accuracy (%) for cross-cultural datasets

Method	Facial Features	Classifiers	Angry	Happy	Surprise	Sad	Fear	Average
2009 [21]			41.7[4]	91.7[1]	74.2[2]	52.5[3]	30.8[5]	52.93
2015 [22]	LBP, HOG, Gabor filter	SVM, ANN, KNN	36.7[3]	58.1[2]	83.3[1]	16.1[4.5]	16.1[4.5]	41.2
Average Rank			3.5	1.5	1.5	3.75	4.75	
Average Rank of all Methods			4.08	2	1.79	3.37	3.75	

We used rank analysis to calculate the expression recognition accuracies that shown in tables, which is computed as:

$$R_b = \frac{1}{N}\sum_a r_a^b \qquad (1)$$

In equation 1, r_a^b belongs to expression recognition accuracy rank of an a-th expression and R_b is the average rank of expression recognition accuracy on a particular dataset. The average rank of all methods is the average of all datasets average ranks. The lower values represent the higher rank and the higher values represent the lower rank. The average rank of all methods indicates that the expressions of happiness and surprise are easy to recognize as compared with other expressions similarly the expressions of anger, sadness, and fear are more confusing than happiness and surprise expressions. It also demonstrates that happy and surprise expressions are least confusing expressions as compared to anger, sadness, and fear expressions.

3. Materials and methods

The proposed technique that is applied in this research, consists of following phases.

3.1. Pre-processing

In the first phase, image pre-processing techniques are applied to minimize computational cost and faster facial expression classification. Although, there is need to obtain such images that contain uniformed size and shape, noise-free as well as a facial region having normalized intensity. Therefore, in image pre-processing localization and alignment of facial images is focused. Different approaches for face detection and alignment have been introduced and applied for providing an appropriate solution but computational cost is considerable for evaluation of each new facial image.

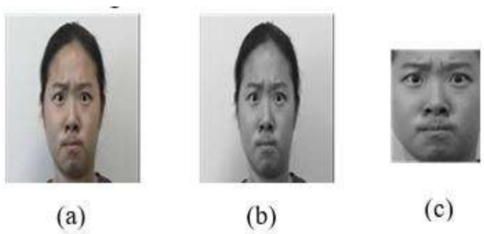


Fig. 1. Cross-culture dataset (a) Original image (b) Gray scaled image (c) Cropped image

Cross-culture dataset original image is shown in figure 1(a) having dimensions 720 pixels width and 480 pixels height and contains 345,600 features. We performed RGB to grayscale conversion on original dataset image that is represented in figure 1(b). To reduce image dimensions and for feature extraction: outliers, edges, hairs, and other extra features are eliminated. The region of interest is gained by performing cropping on the image as shown in 1(c). After performing cropping, original image dimensions are reduced to 269 pixels width and 297 pixels height and now contains 79,893 features. Image having lower dimensions take less computational time in processing but larger dimensions image takes longer time in processing that's why image size is reduced to minimize computational power and time.

Histogram Equalization is implemented to normalize original image to represent uniform intensity throughout the database. Image having histogram equalized is defined as:

$$h_{a,b} = floor((M-1)\sum_{m=0}^{k_{a,b}} q_m) \qquad (2)$$

In equation 2, q_m is normalized histogram that is calculated by the number of pixels with intensity divided by a total number of pixels.

3.2. Feature extraction

After performing pre-processing on images, the second phase is feature extraction from facial images. The main objective of feature extraction is to reduce dimensions for more accurately facial expression classification while considering less computational cost and time. In this phase, different feature extraction approaches are applied to represent interesting parts of an image as a compact feature vector. Feature extraction plays a vital role when large image size and a reduced feature representation is needed to quickly complete tasks such as image retrieval and matching. Feature extraction is widely used for finding the solution of computer vision, object retrieval, face detection, and recognition problems. For representation of feature vectors PCA, HOG, and LBP are employed. Features from images are extracted using HOG while PCA is further used for dimension reduction. After original image cropping, we have 79,893 features and applied HOG-PCA to reduce these features. Finally, we selected fifty prominent features for further processing.

3.2.1 HOG-PCA features

We applied a novel approach for feature extraction and reduction that is named as HOG-PCA features for improvement of facial expression recognition performance, accuracy, faster matching, and reduction of computation cost. Therefore, we performed two level feature extraction. In our proposed work, we can also call them as HOG-PCA features. At first level, HOG features are extracted from images belonging to a vast variety of cultures. After feature extraction, PCA is implemented on these HOG features vectors to require HOG-PCA features. In this method, dimensions of features are minimized because PCA is a well-known technique for dimension reduction. It has various applications in computer vision problems. From an input image, the histogram is calculated by involving orientation of the gradients in HOG because edges and the shapes of objects can be well-characterized by the intensity of individual local gradients using HOG. First, the input image is divided into a grid of cells to obtain each histogram of the gradient, minimize computational power, and time.

Let I be an intensity function representing an input image to be analyzed. The image is divided into cells of size $M \times M$ pixels as well as the orientation $\theta_{x,y}$ of gradient in every pixel is calculated as follows:

$$\theta_{x,y} = \tan^{-1}\frac{I(x,y+1)-I(x,y-1)}{I(x+1,y)-I(x-1,y)} \tag{3}$$

After performing the calculation, cells are categorized into regions known as blocks. Then normalization is performed on each block to reduce the response with respect to brightness. For performing detection, a descriptor is computed for every block. To analyze an image at multiple scales, these operations are performed on windows having a finite size. By obtaining a descriptor belonging each window is further given to SVM for providing a prediction about present or absent of a pedestrian.

Block normalization is computed on the base of following equations.

L2-norm:

$$w \longleftarrow w/\sqrt{\|w\|_2^2+\varepsilon} \tag{4}$$

L1-norm:

$$w \longleftarrow \sqrt{w/(\|w\|_1+\varepsilon)} \tag{5}$$

At last, all histograms are ordered and concatenated into a single histogram for further processing. PCA is also termed as orthogonal projection of data onto a lower dimensional linear subspace that is termed as principal subspace so that variance of the projected samples is maximized. It can also be shortened as a linear projection that reduces the mean squared distance between the data points and their projection. Let $\{a_j \mid j=1,...,N\}$ be a set of features. These training samples are given as input for PCA. Projection matrix P is defined using following equation:

$$b = P^T(a_j-\bar{a}) \tag{6}$$

Mean vectors are calculated as $\bar{a}=\frac{1}{N}\sum_{j=1}^{N}a_j$.

Then projection matrix P is gained as:

$$\sum P = P\wedge,\ P\,P^T = I \tag{7}$$

The confusion matrix is obtained using following equation:

$$\sum = \frac{1}{N}\sum_{n=1}^{N}(a_j-\bar{a})(a_j-\bar{a})^2 \tag{8}$$

3.3. Cross-culture facial expression dataset

We developed the cross-culture facial expression dataset by combining datasets including JAFFE, TFEID, KDEF, RaFD, and CK+ shown in figure 2. The participants of JAFFE dataset belong to Japanese culture. The subjects of TFEID belong to Taiwanese culture. RaFD contains the facial images of Moroccans and Caucasians cultures while KDEF presents the facial images of Swedish objects. CK+ contains the facial images of a vast variety of cultures including Sothern Europeans, Euro Americans, Asians, and Afro Americans.

Fig. 2. Cross-culture dataset original images

3.4. Stacked generalization

In this work, multiple models are combined together for improving classification accuracy by applying Stacked Generalization (or stacking) that consist of two levels. The first level is base level (or level-0) and the second level is meta-level (or level-1). SVM is applied at level-0 similarly to minimize base level errors at level-1 NB, KNN, and ANN are employed as meta-level classifiers. The most prominent fifty features are used to train the base level classifier.

Algorithm of stacked generalization:

Input:

Training Data $C = \{a_j, b_j\}_{j=1}^{n} (a_j \in \mathbb{Z}^r, b_j \in B)$

Output:

G is an ensemble classifier

Step 1: Apply Cross-Validation technique in order to prepare a meta-level classifier

Randomly divide training data C into M equal subsets: $C = \{C_1, C_2, C_3, ..., C_M\}$

for $m \longleftarrow 1$ to M **do**

Step 1.1: Initially, base level classifiers learning take place

for $p \longleftarrow 1$ to P **do**

Perform classifier learning G_{mp} from $C \setminus C_m$

end for

Step 1.2: Build a training set for meta-level classifier

for $a_j \in C_m$ **do**

Obtain a record $\{a'_j, b_j\}$, where $a'_j = \{g_{m1}(a_j), g_{m2}(a_j), g_{m3}(a_j), ..., g_{mp}(a_j)\}$

end for

end for

Step 2: Perform meta-level classifier learning

Also perform new classifier learning g' by using the collection of $\{a'_j, b_j\}$

Step 3: Repeatedly perform base level classifiers learning

for $p \longleftarrow 1$ to P **do**

Train a classifier g_p that consist on C

end for

Return $G(a) = g'(g_1(a), g_2(a), g_3(a), ..., g_P(a))$

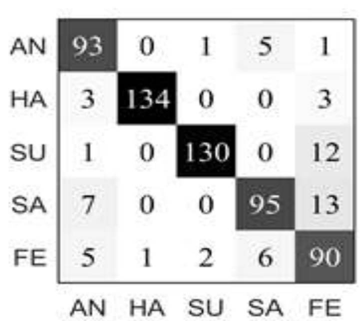


Fig. 3. Sample confusion matrix of universal dataset

Figure 3 represents the expression recognition accuracy on the universal dataset using proposed classification technique. It represents average expression recognition accuracy of 90.03% with NB as a meta-level classifier and 1404 SVM at base level classifier while divided by a total number of pixels.

4. Cross-culture sensitivity analysis

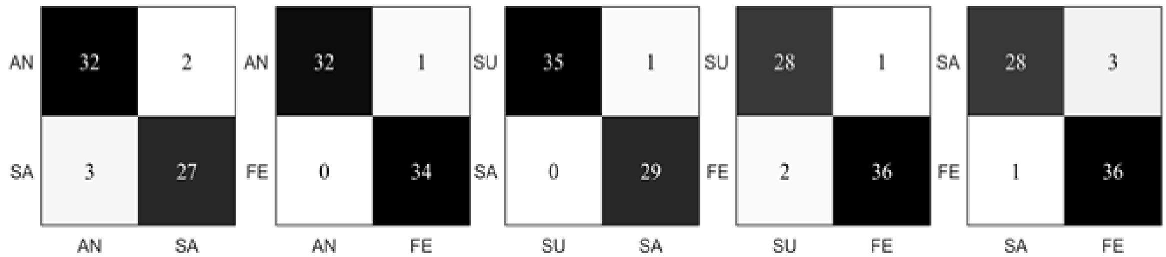
The cross-culture facial expression sensitivity analysis is performed by combining datasets that contain similarity in culture, geography, and ethnicity. These experiments are categorized into three different groups. The first group is Universal dataset which is a combination of five facial expressions databases that contain facial images of a vast variety of cultures including Americans, Europeans, Afro Americans, Euro Americans, Caucasians, and Moroccans. JAFFE and TFEID datasets are used in the second group. These datasets are used for cross-cultural sensitivity analysis because these two datasets having a same facial structure as compared with other cultures. Japanese and Taiwanese facial structure is different from Americans, Europeans, Afro Americans, Euro Americans, Caucasians, and Moroccans due to a geographic region that is why these datasets are combined together for performing cross-culture sensitivity analysis. In the third group, CK+, KDEF, and RaFD datasets are used for experimental purpose. These groups of images are the combination of all cultures including Americans, Europeans, Afro Americans, Euro Americans, Caucasians, and Moroccans having a same geographic region that is why these datasets are combined together for performing sensitivity analysis.

Table 7. Proposed approach achieved accuracy in percentage (%)

Angry	Happy	Surprise	Sad	Fear	Universal Dataset			JAFFE & TFEID			CK+ & KDEF & RaFD		
					HOG	PCA	LBP	HOG	PCA	LBP	HOG	PCA	LBP
✔	✔				100	96	98	100	98	97	100	99	99
✔		✔			99	95	98	97	97	94	100	99	98
✔			✔		92	88	87	92	86	75	97	92	96
✔				✔	98	94	97	99	88	94	99	93	97
	✔	✔			99	98	99	98	92	94	100	99	100
	✔		✔		100	98	98	97	100	97	100	97	98
	✔			✔	99	97	98	97	99	94	99	95	97
		✔	✔		98	96	98	98	95	94	100	92	95
		✔		✔	95	91	92	96	90	88	95	97	93
			✔	✔	94	87	93	94	96	90	97	91	90
✔	✔	✔			99	96	97	92	95	93	99	97	96
✔	✔		✔		96	93	90	93	92	82	99	93	93
✔	✔			✔	96	95	95	98	90	90	96	93	98
✔		✔	✔		94	89	91	89	89	79	95	94	95
✔		✔		✔	94	92	88	87	83	93	96	92	92
✔			✔	✔	90	87	84	90	89	93	97	94	90
	✔	✔	✔		97	93	96	99	92	91	99	95	97
	✔	✔		✔	96	91	92	99	90	87	98	91	94
	✔		✔	✔	95	91	89	95	92	86	97	93	95
		✔	✔	✔	95	86	84	86	85	78	94	81	87
✔	✔	✔	✔		95	89	92	86	89	82	98	92	95
✔	✔	✔		✔	94	89	91	95	87	77	96	92	94
✔	✔		✔	✔	92	89	87	87	80	87	96	94	92
✔		✔	✔	✔	89	85	81	89	80	74	93	88	86
	✔	✔	✔	✔	95	89	85	89	93	83	95	89	88
✔	✔	✔	✔	✔	90	88	84	88	83	83	93	89	88

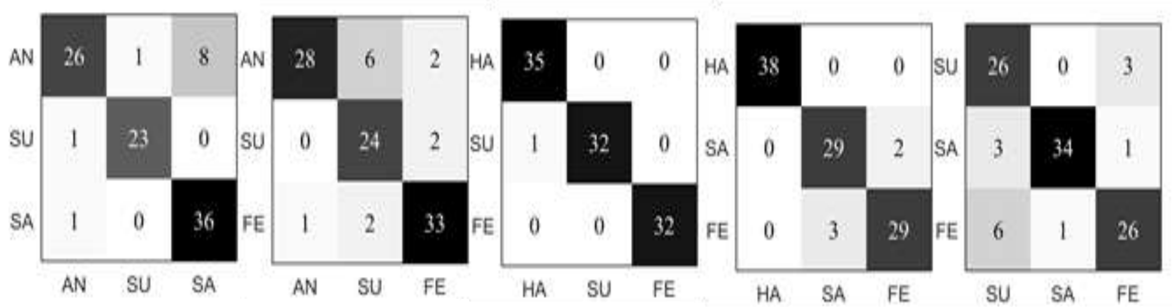
HOG, PCA, and LBP approaches are used for facial feature representation. Similarly, we performed experiments with different combinations of facial expressions as represented in Table 7. Labels used in confusion matrixes: (AN for angry, HA for happy, SU for the surprise, SA for sad, and FE for fear). Confusion matrixes of two combinations: Angry-Sad, Angry-Fear, Surprise-Sad, Surprise-Fear, and Sad-Fear are shown in figures (4, 8, and 12). Figures representing confusion matrixes of three combinations: Angry-Surprise-Sad, Angry-Surprise-Fear, Happy-Surprise-Fear, Happy-Sad-Fear, and Surprise-Sad-Fear are (5, 9, and 13). Four combinations confusion matrixes: Angry-Happy-Surprise-Sad, Angry-Happy-Sad-Fear, and Angry-Surprise-Sad-Fear are represented in figures (6, 10, and 14). Angry-Happy-Surprise-Sad-Fear that are five combinations confusion matrixes represented in figures (7, 11, and 15). Confusion matrixes of all above three categories are shown further.

4.1. Japanese and Taiwanese cultures



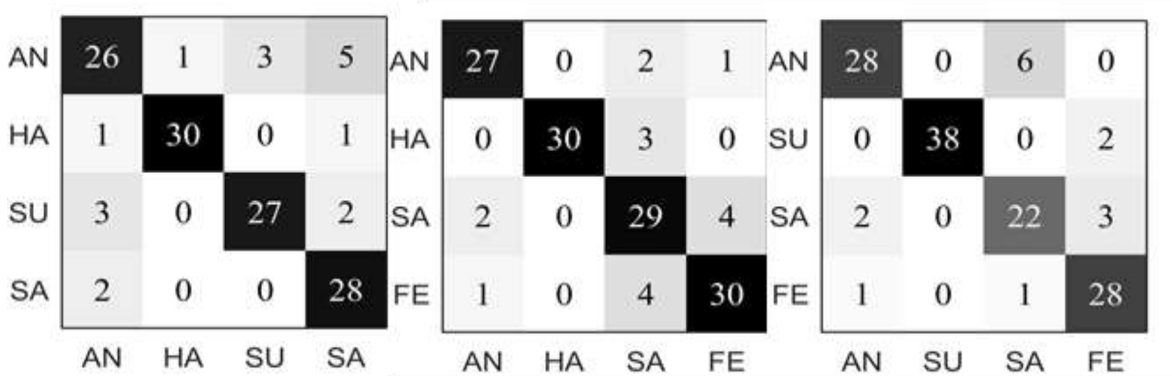
(a) (b) (c) (d) (e)

Fig. 4. Two combinations confusion matrixes (a) Angry-Sad (b) Angry-Fear (c) Surprise-Sad (d) Surprise-Fear (e) Sad-Fear



(a) (b) (c) (d) (e)

Fig. 5. Three combinations confusion matrixes (a) Angry-Surprise-Sad (b) Angry-Surprise-Fear (c) Happy-Surprise-Fear (d) Happy-Sad-Fear (e) Surprise-Sad-Fear



(a) (b) (c)

Fig. 6. Four combinations confusion matrixes (a) Angry-Happy-Surprise-Sad (b) Angry-Happy-Sad-Fear (c) Angry-Surprise-Sad-Fear

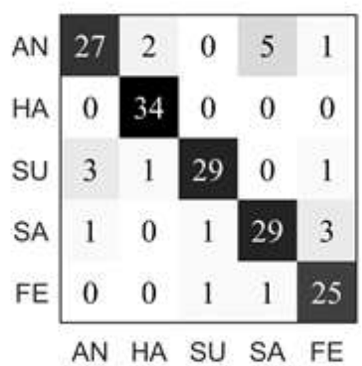


Fig. 7. Five combinations confusion matrix: Angry-Happy-Surprise-Sad-Fear

From these confusion matrixes, we can say that expressions: angry-sad, angry-surprise, sad-fear, and fear-surprise have difficulty in expression recognition as compared with other expression combinations. Similarly, in most experiments angry is difficult to recognize in presence of sad, sad is difficult to recognize in presence of fear while fear is also difficult to recognize in presence of surprise. On the other hand, the expression happy is easier to recognize as compared with angry, surprise, sad, and fear due to facial structure. In JAFFE and TFEID datasets confusion matrixes, happy is easier to recognize while angry, surprise, sad, and fear expressions are most difficult to correctly recognize due to similar muscular movement.

4.2. Swedish, Caucassian, Morrocon, European, Euro American, and Afro American cultures

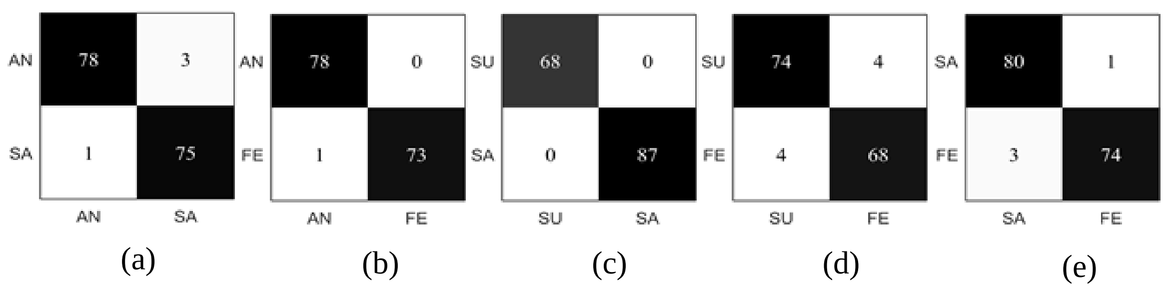


Fig. 8. Two combinations confusion matrixes (a) Angry-Sad (b) Angry-Fear (c) Surprise-Sad (d) Surprise-Fear (e) Sad-Fear

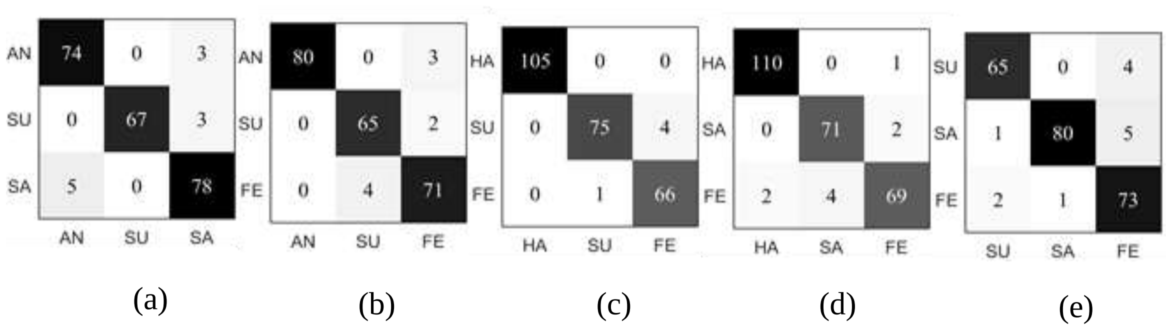
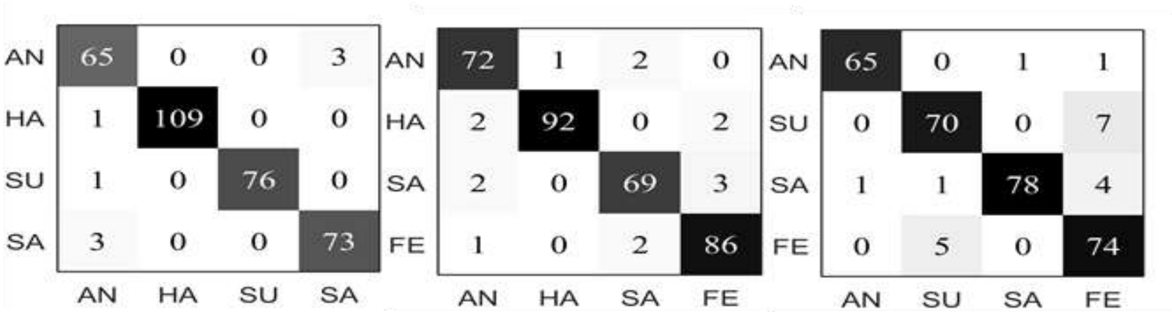


Fig. 9. Three combinations confusion matrixes (a) Angry-Surprise-Sad (b) Angry-Surprise-Fear (c) Happy-Surprise-Fear (d) Happy-Sad-Fear (e) Surprise-Sad-Fear



(a) (b) (c)

Fig. 10. Four combinations confusion matrixes (a) Angry-Happy-Surprise-Sad (b) Angry-Happy-Sad-Fear (c) Angry-Surprise-Sad-Fear

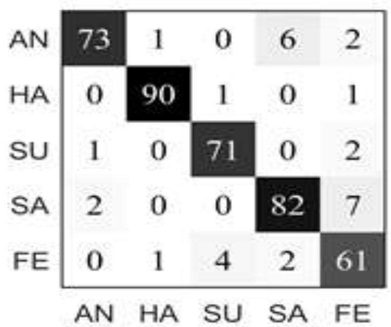
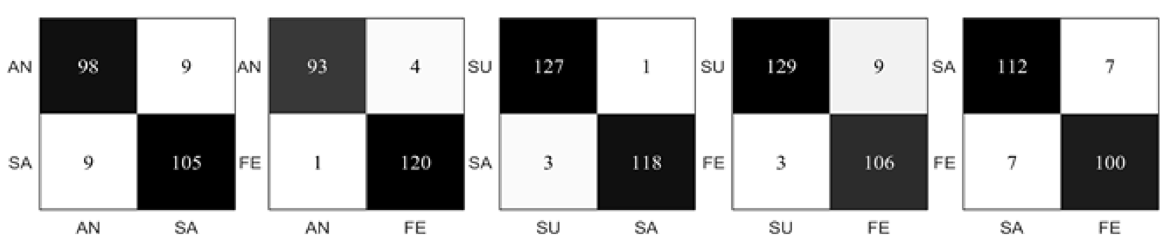


Fig. 11. Five combinations confusion matrix: Angry-Happy-Surprise-Sad-Fear

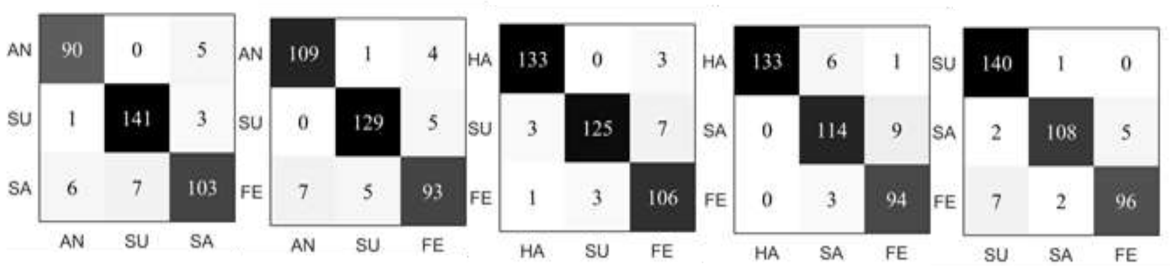
Expressions that showed significant classifier recognition confusion are angry-sad, fear-sad, and surprise-fear as compared with others. Most of the fear samples are incorrectly classified as sad or surprise while angry samples as sad. Experiments indicate that happy is most correctly classified expression while sad and fear are most incorrectly classified expressions due to similar muscular movement. Confusion matrixes of CK+, KDEF, and RaFD show that happy is most accurately classified expression while sad and fear are mostly inaccurately classified expressions among five expressions.

4.3. Japanese, Taiwanese, Swedish, Caucassian, Morrocon, European, Afro American, and Euro American cultures



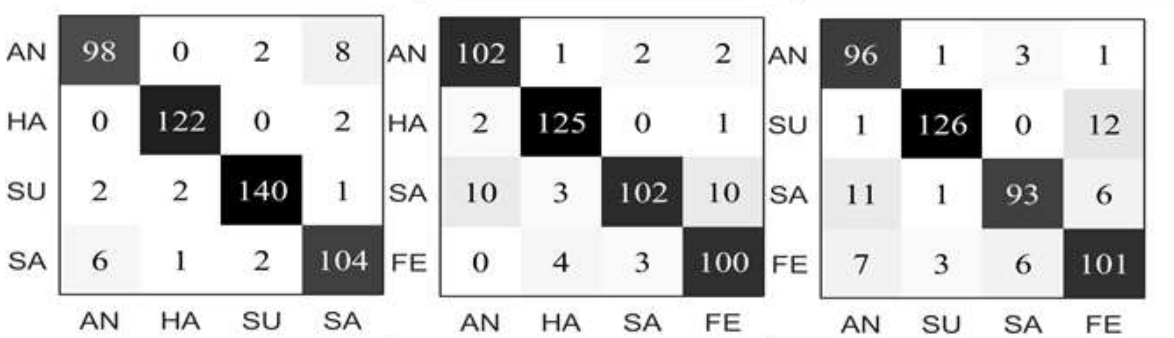
(a) (b) (c) (d) (e)

Fig. 12. Two combinations confusion matrixes (a) Angry-Sad (b) Angry-Fear (c) Surprise-Sad (d) Surprise-Fear (e) Sad-Fear



(a) (b) (c) (d) (e)

Fig. 13. Three combinations confusion matrixes (a) Angry-Surprise-Sad (b) Angry-Surprise-Fear (c) Happy-Surprise-Fear (d) Happy-Sad-Fear (e) Surprise-Sad-Fear



(a) (b) (c)

Fig. 14. Four combinations confusion matrixes (a) Angry-Happy-Surprise-Sad (b) Angry-Happy-Sad-Fear (c) Angry-Surprise-Sad-Fear

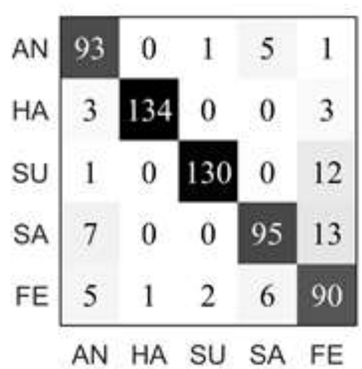


Fig. 15. Five combinations confusion matrix: Angry-Happy-Surprise-Sad-Fear

These combinations: angry-sad, angry-fear, sad-fear, and surprise-fear showed higher confusion. Figures represent that angry is misclassified as sad or fear. Similarly, sad is misclassified as fear while surprise is misclassified as a fear expression. Happy showed relatively high results as compared to other expression results while angry, surprise, sad, and fear showed comparatively low results due to similar facial muscular movement. In the universal dataset, confusion matrixes show that happiness is easier to recognize as compared with angry, surprise, sad, and fear expressions.

5. Results and discussion

From all inter-culture and cross-cultural confusion matrices, it can be concluded that happy achieved best recognition accuracy as compared with other expressions accuracy that means happy showed the best performance among five expressions. Similarly, sad and fear achieved lowest recognition accuracy as compared with recognition accuracy of other expressions that is an indication of the poor performance of these expression combinations. An expression that shows higher recognition accuracy is less sensitive. From this fact, we can say that happy is universal facial expression across all cultures because its presence does not degrade classifiers recognition accuracy of other facial expressions while sad and fear are most sensitive expressions. Their presence degrades recognition accuracy of other classifiers while angry and surprise are little sensitive expressions because their classifiers recognition accuracy is moderate as compared with other classifiers recognition accuracy. The confusion matrices indicate that fear and sadness are most confusing expressions that is the similar case as shown by the average ranks on the individual dataset and the average rank of all methods. The confusion matrices also show that the expression of happiness is the least confusing expression which same as the average ranks of culture- specific methods. From these observations, we can conclude that culture variations degrade the overall classifier performance of the facial expression recognition system. But the confusion between expression combinations remains same as in culture-dependent techniques. We can conclude that cultural variations do not affect the facial appearance of different expressions. In other words, we can say that the expressions of anger, happiness, surprise, sadness, and fear are innate and universal across all cultures.

6. Conclusion

The proposed sensitivity analysis approach that applied on cross-culture using stacking ensemble technique gained better recognition accuracy. Our main concern of this work was to identify which facial expression is universal expression among other expressions while which is not universal expression. Experimental results showed that in Ekman six basic expressions happy is less sensitive facial expression while sad and fear are most sensitive expressions among six expressions. It is noticed that some facial expressions that showed confusion may due to difference of culture, geography, and ethnicity in databases samples because these samples belong to Americans, Europeans, Afro-Americans, Euro Americans, Caucasians, Moroccans, Japanese, and Taiwanese including with some other factors involves like facial structure, visual representation, and number of samples used for experiments. In this work, proposed sensitivity analysis is also a novel approach to the literature because it is really challenging task for producing a more accurate decision while facial structure appearance varying.

Acknowledgement

We thank for acknowledging our research work by providing these datasets named as CK+, JAFFE, TFEID, KDEF, and RaFD for experimental purpose.

References

1. Ekman, P. Universal Facial Expressions of Emotion. *California Mental Health Research Digest*, 8(4), 151-158 (1970).
2. Da Silva, F. A. M., Pedrini, H. Effects of cultural characteristics on building an emotion classifier through facial expression analysis. *Journal of Electronic Imaging*, 24(2), 023015-9 (2015).
3. Yan, X., Andrews, T. J., Jenkins, R., Young, A. W. Cross-cultural differences and similarities underlying other-race effects for facial identity and expression. *The Quarterly Journal of Experimental Psychology*, 69(7), 1247-1254 (2016).
4. Zia, M. S., Jaffar, M. A. Cross-Cultural Emotion Classification based on Incremental Learning and LBP-Features. *An International Journal Applied Mathematics & Information Sciences*, 9(5), 2651-2656 (2015).
5. Ali, G., Iqbal, M. A., Choi, T. S. Boosted NNE collections for multicultural facial expression recognition. *Pattern Recognition*, 55, 14-27 (2016).
6. Wingenbach, T. S. H., Ashwin, C., Brosnan, M. Diminished sensitivity and specificity at recognizing facial emotional expressions of varying intensity underlie emotion-specific recognition deficits in autism spectrum

disorders. *Research in Autism Spectrum Disorders*, 34, 52-61 (2017).
7. Fan, X., Tjahjadi, T. A dynamic framework based on local Zernike moment and motion history image for facial expression recognition. *Pattern Recognition*, 64, 399-406 (2017).
8. Sun, Y., Wen, G. Cognitive facial expression recognition with constrained dimensionality reduction. N*eurocomputing*, 230, 397-408 (2017).
9. Yan, H. Transfer subspace learning for cross-dataset facial expression recognition. *Neurocomputing*, 208, 165-173 (2016).
10. Yu, Z., Zhang, C. Image based Static Facial Expression Recognition with Multiple Deep Network Learning. *In Proceedings of the ACM on International Conference on Multimodal Interaction*, 435-442 (2015).
11. Moeini, A., Faez, K., Sadeghi, H., Moeini, H. 2D facial expression recognition via 3D reconstruction and feature fusion. *Journal of Visual Communication and Image Representation*, 35, 1-14 (2016).
12. Hegde, G. P., Seetha, M., Hegde, N. Kernel Locality Preserving Symmetrical Weighted Fisher Discriminant Analysis based subspace approach for expression recognition. *Engineering Science and Technology, an International Journal*, 19(3), 1321-1333 (2016).
13. Happy, S. L., Routray, A. Automatic Facial Expression Recognition Using Features of Salient Facial Patches. *IEEE transactions on Affective Computing*, 6(1), 1-12 (2015).
14. Fan, X., Tjahjadi, T. A dynamic framework based on local Zernike moment and motion history image for facial expression recognition. *Pattern Recognition*, 64, 399-406 (2017).
15. Moeini, A., Moeini, H. Multimodal facial expression recognition based on 3D face reconstruction from 2D images. *Springer International Publishing*, 46-57 (2015).
16. Carcagni, P., Coco, M. D., Leo, M., Distante, C. Facial expression recognition and histograms of oriented gradients: a comprehensive study. *Springer Plus*, 4(1), 645 (2015).
17. Farajzadeh, N., Pan, G., Wu, Z. Facial expression recognition based on meta probability codes. *Pattern Analysis and Applications*, 17(4), 763-781 (2014).
18. Wang, M., Yee, S. Determining mood from facial expressions. (2014).
19. Van Kuilenburg, H., Wiering, M., Den Uyl, M. A model based method for automatic facial expression recognition. *In European Conference on Machine Learning, Springer-Verlag Berlin*, 194-205 (2005).
20. Jeni, L. A., Hashimoto, H., Kubota, T. Robust Facial Expression Recognition using Near Infrared Cameras. *JACIII*, 16(2), 341-348 (2012).
21. Tay, C., Obaid, M., Mukundan, R., Bainbridge-Smith, A. Facial Expressions Using A Quadratic Deformation Model: Analysis and Synthesis. *In Image and Vision Computing New Zealand, IVCNZ'09. 24th International Conference, IEEE*, 50-55 (2009).
22. Da Silva, F. A. M., Pedrini, H. Effects of cultural characteristics on building an emotion classifier through facial expression analysis. *Journal of Electronic Imaging*, 24(2), 023015-9 (2015).