Facial Expression Recognition Through Machine Learning

Nazia Perveen, Nazir Ahmad, M. Abdul Qadoos Bilal Khan, Rizwan Khalid, Salman Qadri

Abstract: Facial expressions communicate non-verbal cues, which play an important role in interpersonal relations. Automatic recognition of facial expressions can be an important element of normal human-machine interfaces; it might likewise be utilized as a part of behavioral science and in clinical practice. In spite of the fact that people perceive facial expressions for all intents and purposes immediately, solid expression recognition by machine is still a challenge. From the point of view of automatic recognition, a facial expression can be considered to comprise of disfigurements of the facial parts and their spatial relations, or changes in the face's pigmentation. Research into automatic recognition of the facial expressions addresses the issues encompassing the representation and arrangement of static or dynamic qualities of these distortions or face pigmentation. We get results by utilizing the CVIPtools. We have taken train data set of six facial expressions of three persons and for train data set purpose we have total border mask sample 90 and 30% border mask sample for test data set purpose and we use RST- Invariant features and texture features for feature analysis and then classified them by using k- Nearest Neighbor classification algorithm. The maximum accuracy is 90%.

Keywords CVIPtools, RST- Invariant, KNN, Human-Machine Interfaces

1. INTRODUCTION

Facial expression is the most effective form of non verbal communication and it provides intimation about emotional state, mindset and intention. Facial expressions not only can change the flow of conversation but also provides the listeners a way to communicate a wealth of information to the speaker without even uttering a single word. When the facial expression does not match with the spoken words, then the information pass on by the face gets more power in interpreting the information.

- *Nazia Perveen, GCUF Layyah Campus, Pakistan, PH-03451652331. E-mail: Naziazulfaqar@gmail.com*
- *Nazir Ahmad, The Islamia University Bahawalpur, Pakistan, PH-03457937789. E-mail: nazeerrana@gmail.com*
- *M. Abdul Qadoos Bilal Khan, GCUF Layyah Campus, Pakistan, PH - 03217033258. E-mail: aqkhan_iub@yahoo.com*
- *Rizwan Khalid, GCUF Layyah Campus, Pakistan, PH-03336201747. E-mail: rizwankhalid201@gmail.com*
- *Salman Qadri, The Islamia University of Bahawalpur, Pakistan*

From the perspective of automatic recognition, a facial expression can be considered to consist of deformations of facial components and their spatial relations, or changes in the pigmentation of the face. Facial expressions represent the changes of facial appearance in reaction to a person's inside emotional states, social communications or intentions. To communicate the emotions and express the intentions the Facial expression is the most powerful, natural, non verbal and instant way for humans. It is faster to communicate the emotions through facial expressions than through verbalization. The requirement for proficient communication channels between machines and humans becomes progressively imperative in light of the fact that machines and individuals start to share a variety of tasks. Systems to form these communication channels are known as human machine interaction (HMI) systems. Progresses in technology make it possible the development of more useful HMI systems which no more depend on regular devices for example keyboard, mouse and displays but take commands directly from user's voice and mimics. Such systems intend to simulate human-human interaction by only utilizing communication channels utilized between humans and not requiring artificial equipment. Human-machine interaction should be enhanced to more nearly simulate human to human interaction before machines take more places in our lives. As change of expressions on human face is an intense method for passing emotions, facial expression recognition (FER) will be one of the best steps for improving HMI systems. An automatic facial expression recognition system generally comprises of three main parts: face detection, facial feature points extraction and facial expression classification. In the first step, system obtains input image and performs some image processing techniques on it in order to locate the face region. In static images it is called face localization whereas in videos it is called face tracking. After the face has been located in the image or video frame, it can be evaluated in terms of facial action happening. A feature is a point of interest or a piece of information. There are two types of features that are generally used to represent facial expression: geometric features and appearance features. Geometric features evaluate the displacements of certain parts of the face for instance brows or mouth corners, whereas appearance features represent the change in texture of face when particular action is performed. The task of geometric feature measurement is typically associated with face region analysis, particularly finding and tracking key points in the face region. Potential problems that take place in face decomposition task could be occlusions and occurrences of facial hair or glasses. Moreover, defining the feature set is difficult, because features should be descriptive and possibly not correlated. The last part of the Facial Expressions Recognition system is based on machine learning theory; specifically it is the classification task. A set of features which were retrieved from face region in the previous stage is the input to the classifier. The set of features is created to explain the facial expression. Classification needs supervised training, so the training set should consist of labeled data. Once the classifier is trained, it can recognize input images by assigning them a specific class label. The most frequently used facial expressions classification is done both in terms of Action Units, proposed in Facial Action Coding System and in terms of universal emotions: happiness, sadness, anger, surprise, disgust and fear. There are several different machine learning techniques for classification task for instance K-Nearest Neighbors, Artificial Neural Networks, Support Vector Machines, Hidden Markov Models, Expert Systems with rule based classifier, Bayesian Networks or Boosting Techniques. Three main issues in classification task are: choosing good feature set, competent machine learning technique and different database for training. Feature set should be composed of features that are discriminative and characteristic for specific expression. Machine learning technique is selected usually by the type of a feature set. The database used as a training set should be big enough and include a variety of data. In facial expression recognition Region of Interests represent the eye pair, nostrils, and the mouth area. Region of Interest is related to define a large region which contains the point that we want to detect. In Facial Expression Recognition Systems, only particular regions of the face are used for discrimination. The areas of the eyes, eyebrows, mouth, and nose are the main features in any Facial Expression Recognition System. Some facial recognition algorithms perceive facial components by extracting landmarks, or elements, from a picture of the subject's face. For instance, an algorithm may assess the relative position, size, and state of the eyes, nose, cheekbones, and jaw. These features are then used to search for other images with identical features. The majority of the facial expression recognition methods reported yet are focused on recognition of six primary expression categories such as: happiness, sadness, fear, anger, disgust and grief.

2. RELATED WORK:

Giorgana and Ploeger (2012) exhibited a completely programmed FERS to perceive the feelings of sadness, joy and surprise. For feature extraction they used Gabor filters and utilized AdaBoost for feature selection. A recognition rate of 87.14% has been accounted for utilizing the one-versus. - One support vector machine (SVM) and Error-Correcting Output Codes (ECOC) with standardized face pictures that are 96 x 96 in size. Wang and Xiao (2013) proposed a new technique of facial expression classification based on neural network ensembles. To extract features from face images they used Principal Component Analysis and Gabor filters. MD-Adaboost is used to combine the results of neural network classifiers. Classifiers are trained using different feature sets to improve the classification rate and stability of the classifier. The approach was evaluated in JAFFE image database. The experiments demonstrate the positive effect of the neural network ensemble based classifier, and show that the MD-Adaboost based classifier was more efficient and firm than other algorithms. Ghimire and Lee (2013) introduced two methods for recognizing dynamic facial expressions, either directly by using multiclass AdaBoost, or by using Support Vector Machine on the boosted geometric features. The geometric features are extracted from the sequences of facial expression images. The landmark initialization and tracking is based on the Elastic Bunch Graph Multi-resolution method. Multi class AdaBoost with dynamic time warping similarity distance between the feature vector of input facial expression and prototypical facial expression, is used as a weak classifier to select the subset of discriminative feature vectors. The recognition rate of 95.17% using feature selective multi class AdaBoost, and 97.35% using SVM on boosted features, is achieved on the Extended Cohn-Kanade (CK+) facial expression database. Abidin and Harjoko (2012) focused on static images of single person for facial expression recognition. They used integral projection method for localization and segmentation of face portion from the images. The features have been extracted by using fisherface method. To recognize the facial expression the neural network was used as a classifier. The Japanese Female Facial Expression (JAFFE) database has been used for experiments. The recognition rate of proposed system was 86.85%. Zavaschi et al., (2013) introduced SVM ensemble classifier based on the combination of Gabor filters and LBP feature vector. Multi Objective Genetic Algorithm is used for searching the best combination of SVM classifiers for facial expression recognition. Experimental results on JAFFE and Cohn-Kanade databases have shown the effectiveness of the proposed approach in finding powerful ensembles, which improves the recognition rates between 5% and 10% over usual approaches that utilize single feature sets and single classifiers. Sarawagi and Arya (2013) designed a competent system for recognition of facial expression in an image sequences. They emphasized on the color normalization and facial feature extraction. They used Local Binary Pattern for feature detection. To discover facial landmarks automatic landmark detection technique is introduced which used Local Binary Pattern. On the Indian database the accuracy of 94.7% is observed. Hablani et al., (2013) extracted features based on Local Binary Pattern. They have chosen some important facial parts like sub parts of eyes, nose and mouth as facial features. They used the templates of extracted facial features as template matching to classify the expression. Their proposed approach is better than approaches that use the whole face image. The proposed method integrates person identity to perform better than usual expression systems. The proposed person dependent approach attains higher recognition rates. For classification Chi square distance is used as measure of similarity. They used Support Vector Machines (SVM) as classification algorithm. Lingareddy and Haritha (2013) proposed a semi-supervised version of the classical Principal Component Analysis-based face recognition algorithm, based on the self-training method, to utilize unlabelled data for off-line updating of the Eigen space and the templates. Results show that the exploitation of unlabelled data by self-training can significantly improve the performances achieved with a small set of labeled training examples. Self-training performances are achieved with a small set of labeled training examples. Thus the self-trained is used to keep the system up-to-date and made the better Facial Recognition System. Zhang et al., (2013) proposed an approach for facial expression by discovering associations between visual feature and Local Binary Pattern (LBP). The proposed approach automatically tracks the facial area and segments face into significant areas based on description of Local Binary Pattern. Then it accumulates the probabilities throughout the frames from video data to capture the temporal characteristics of facial expressions by analyzing facial expressions. Through the proposed approach, the temporal variation of facial expression can be quantified in individual areas. So, the recognition process of facial expression tends to be more comprehensible without sacrificing results of recognition. Das (2014) declared that to recognize facial expressions from real time is a challenge. They proposed a real time automatic facial expression recognition system which used a geometrical feature- based approach. To detect a face from an image and to segment it into different regions modified Active Shape Model (ASM) has been used. The Census Transformation (CT) based feature histogram has been used to represent facial region. The proposed system works in real time on video data and has better performance for the JAFFE dataset. Hai et al., (2015) proposed a model for facial expression classification by using ANN and KNN. They used Independent Component Analysis to extract facial features. The input of Artificial Neural Network classifier is all feature vectors processed by ICA. The input of K-NN classifier is the distance ratios of the local region of face. To combine the output of ANN and K-NN classifier the minimum function is used. They applied ANN_KNN model for seven basic facial expressions classification on JAFEE database. The classifying accuracy was 92.38%. De et al., (2015) modeled eigenface approach to recognize the human facial expressions. To encode variation in the eigenfaces the best feature space is Eigen space that characterized the overall variations among face images. To detect the face in an image they proposed the Hue Saturation Value (HSV) color model. For reducing the high dimensionality of the eigenspace Principal Component Analysis (PCA) has been used. By calculating the Euclidean distance between the input test image and the mean of the eigenfaces of the training dataset the various emotions are recognized. Chao et al., (2015) enhanced the performance of the popular LBP feature. They applied the techniques of expression specific extension and symmetric extension and proposed the es-LBP feature. That captures the better local information of faces on important fiducial points. Furthermore, the LPQ feature, which was initially used for texture classification, is also integrated in the proposed facial expression recognition system. They proposed the class-regularized locality preserving projection (cr-LPP) method for dimensionality reduction to better separate different expressions and match the usage of classifiers. The recognition rate of the proposed method is 96.19%. Yu and Liu (2015) combined the appearance descriptors and geometric features of the image for facial expression recognition. To represent facial appearance the covariance descriptors containing different textural features are computed. Then to present a general facial movement description of the facial expression the geometric features are detected. To form a vector representation of the facial expression the appearance and geometric features are combined. The proposed method shows encouraging performance on the CK+ database. Zhang and Tjondronegoro (2011) proposed a facial expression recognition system using features extracted from facial movements. The features are obtained by extracting patch-based 3D Gabor features, selecting the salient patches, and performing patch matching operations. To recognize the six basic emotions the resultant distance features are fed into a multiclass support vector machine (SVM). The experimental results also display significant performance improvements because of the consideration of facial movement features and promising performance under face registration errors.

3. METHODOLOGY:

We use these approaches: Support Vector Machine, Artificial Neural Network and K Nearest Neighbours (KNN). In pattern recognition, the k-Nearest Neighbors algorithm (or k-NN for short) is a non-parametric method used for classification and regression. In both cases, the input consists of the k closest training examples in the feature space. In Preprocessing we take the image and preprocessed with using these techniques then after this features extraction. In features extraction technique some predefined positions as facial features. In next step feature selection and then classified them.

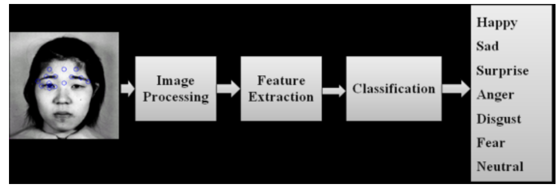


Figure 1. *Overview of Facial Expression Recognition System*

3.1. Image Acquisition

Images used for facial expression recognition are static images. To take the images of expressions of people we use a Panasonic camera (Model DMC- LS5) with focal length of 5mm is used. The format of images is 24 bit color JPEG with resolution of 4320x 3240 pixels. The distance between the camera and person was four feet and images of six basic expressions of each person were taken.

3.2. Image Preprocessing

The image preprocessing procedure comes as a very important step in the facial expression recognition task. The objective of the preprocessing phase is to take images which have normalized intensity, uniform size and shape, and represent only a face expressing certain emotion. The preprocessing procedure should also reduce the effects of illumination and lighting. Expression representation can be delicate to translation, scaling, and rotation of the head in a picture. To battle the effect of these pointless changes, the facial image may be geometrically institutionalized before classification.

3.3. Feature Extraction

In developing accurate facial expression recognition system feature extraction is the most important stage. Unprocessed facial images hold vast amounts of data and feature extraction is required to decrease it to smaller sets of data called features. Feature extraction change pixel information into a more elevated amount representation of color shape, motion, texture, and spatial configuration of the face or its features. The separated representation is utilized for further expression categorization. Feature extraction ordinarily decreases the information's dimensionality space. The reduction procedure ought to keep up essential data having high segregation force and high security.

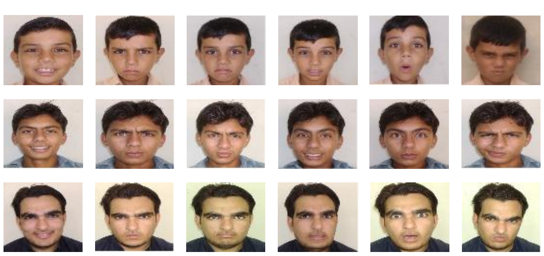
3.4. Feature Selection

Feature selection is concerned with choosing of a subset of features perfectly necessary to perform the classification task from a larger set of candidate features. The feature selection step has an effect on both the computational complexity and the quality of the classification results. It is essential that the information contained in the selected features is adequate to correctly verify the input class. Too many features may unnecessarily raise the complexity of the training and classification tasks, while a poor, inadequate selection of features may have a detrimental effect on the classification results. The process of selecting a sub set of features improves the efficiency of classifier and reduces execution time.

3.5. Classification

The last step of Facial Expressions Recognition systems is to recognize facial expression based on the extracted features. Classification refers to an algorithmic approach for recognizing a given expression as one of a given number of expressions. We use K- Nearest Neighbor classifier for classification. The K-Nearest Neighbor algorithm is a non parametric method used for classification and regression. The input comprises of K closest training examples in the feature space. The output is class participation. By a majority vote of its neighbors an object is classified, with the object being allotted to the class most common among its k nearest neighbors.

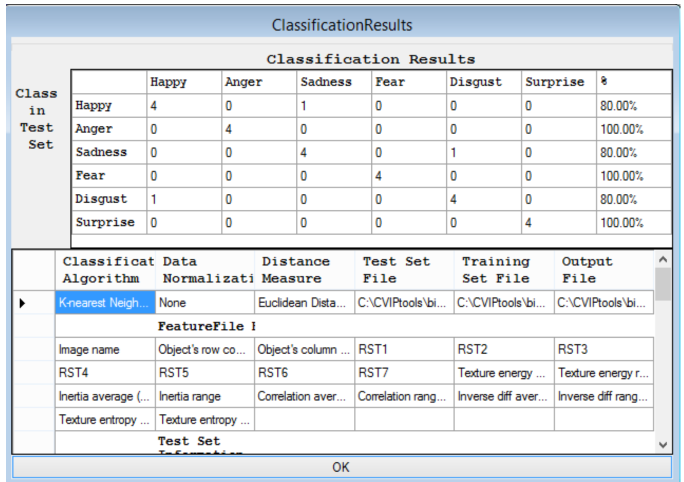
3.6. Physically Data Set



4. RESULTS

We get results by using the CVIPtools.

Table 1. *Confusion Table for Person*



ClassificationResults

Classification Results

Class in Test Set

	Happy	Anger	Sadness	Fear	Disgust	Surprise	%
Happy	4	0	1	0	0	0	80.00%
Anger	0	4	0	0	0	0	100.00%
Sadness	0	0	4	0	1	0	80.00%
Fear	0	0	0	4	0	0	100.00%
Disgust	1	0	0	0	4	0	80.00%
Surprise	0	0	0	0	0	4	100.00%

Classificat Algorithm	Data Normalizati	Distance Measure	Test Set File	Training Set File	Output File
K-nearest Neigh...	None	Euclidean Dista...	C:\CVIPtools\bi...	C:\CVIPtools\bi...	C:\CVIPtools\bi...
	FeatureFile I				
Image name	Object's row co...	Object's column ...	RST1	RST2	RST3
RST4	RST5	RST6	RST7	Texture energy ...	Texture energy r...
Inertia average (...	Inertia range	Correlation aver...	Correlation rang...	Inverse diff aver...	Inverse diff rang...
Texture entropy ...	Texture entropy ...				
	Test Set				

OK

The above table is a confusion table of six facial expressions of person A. We take train data set. We have ratio of 30, 70 test data set for classification. The feature analysis is performed by using RST- Invariant features and texture features. The texture difference is 2. After that pattern classification is performed. The data normalization is no normalization. The distance measure is Euclidean Distance. The classification algorithm is k- Nearest Neighbor and k= 2. The algorithm testing is Training/Test. The accuracy of Happy is 100%, Anger is 80%, Sadness is 80%, Fear is 100%, disgust is 80% and surprise is 100%. The total accuracy rate is 90%.

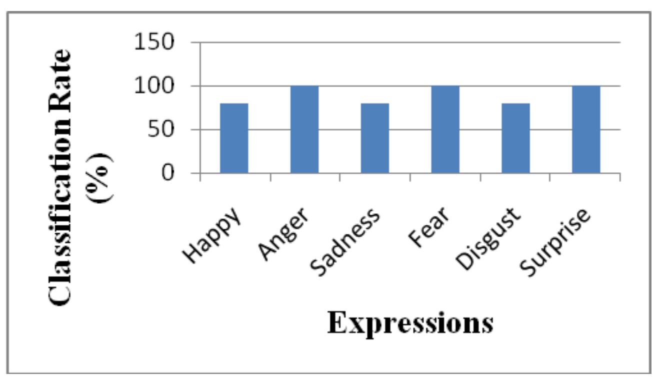


Figure 2*. Classification Result for Facial Expressions of Person A*

The above graph shows the classification results of six facial expressions of person A. The accuracy of Happy is 100%, Anger is 80%, Sadness is 80%, Fear is 100%, disgust is 80% and surprise is 100%.

Table 2. *Confusion Table for Person B*

ClassificationResults

Classification Results

Class in Test Set	Happy	Anger	Sadness	Fear	Disgust	Surprise	%
Happy	4	0	0	0	1	0	80.00%
Anger	1	4	0	0	0	0	80.00%
Sadness	0	0	4	0	1	0	80.00%
Fear	0	0	0	4	0	0	100.00%
Disgust	1	0	0	0	4	1	66.67%
Surprise	0	0	0	0	0	4	100.00%

Classificat Algorithm	Data Normalizati	Distance Measure	Test Set File	Training Set File	Output File
K-nearest Neigh...	None	Euclidean Dista...	C:\CVIPtools\bi...	C:\CVIPtools\bi...	C:\CVIPtools\bi...
	FeatureFile I				
Image name	Object's row co...	Object's column ...	RST1	RST2	RST3
RST4	RST5	RST6	RST7	Texture energy ...	Texture energy r...
Inertia average (...	Inertia range	Correlation aver...	Correlation rang...	Inverse diff aver...	Inverse diff rang...
Texture entropy ...	Texture entropy ...				
	Test Set				

OK

The above table is a confusion table of six facial expressions of person B. We take train data set. We have ratio of 30, 70 test data set for classification. The feature analysis is performed by using RST- Invariant features and texture features. The texture difference is 2. After that pattern classification is performed. The data normalization is no normalization. The distance measure is Euclidean Distance. The classification algorithm is k- Nearest Neighbor and k= 2. The algorithm testing is Training/Test. The accuracy of Happy is 80%, Anger is 80%, Sadness is 80%, Fear is 100%, disgust is 66.67% and surprise is 100%. The total accuracy rate is 84%.

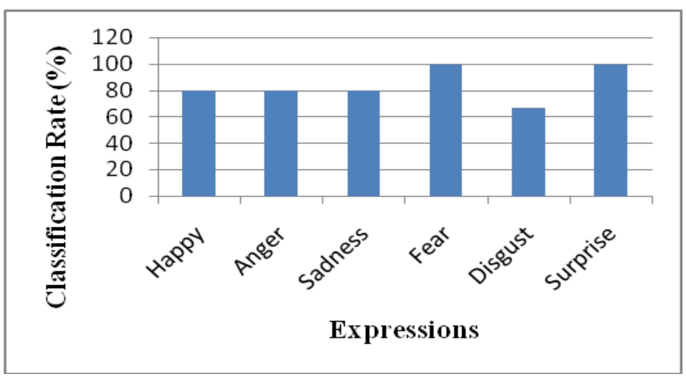


Figure 3. *Classification Result for Facial Expressions of Person B*

The above graph shows the classification results of six facial expressions of person B. The accuracy of Happy is 80%, Anger is 80%, Sadness is 80%, Fear is 100%, disgust is 66.67% and surprise is 100%.

Table 3. *Confusion Table for Person C*

ClassificationResults

Classification Results

Class in Test Set	Happy	Anger	Sadness	Fear	Disgust	Surprise	%
Happy	4	0	0	1	0	0	80.00%
Anger	0	4	1	0	0	0	80.00%
Sadness	0	1	4	0	0	1	66.67%
Fear	1	0	0	4	0	0	80.00%
Disgust	0	0	0	0	4	1	80.00%
Surprise	0	0	0	0	0	4	100.00%

Classificat Algorithm	Data Normalizati	Distance Measure	Test Set File	Training Set File	Output File
K-nearest Neigh...	None	Euclidean Dista...	C:\CVIPtools\bi...	C:\CVIPtools\bi...	C:\CVIPtools\bi...
	FeatureFile I				
Image name	Object's row co...	Object's column ...	RST1	RST2	RST3
RST4	RST5	RST6	RST7	Texture energy ...	Texture energy r...
Inertia average (...	Inertia range	Correlation aver...	Correlation rang...	Inverse diff aver...	Inverse diff rang...
Texture entropy ...	Texture entropy ...				
	Test Set				

OK

The above table is a confusion table of six facial expressions of person C. We take train data set. We have ratio of 30, 70 test data set for classification. The feature analysis is performed by using RST- Invariant features and texture features. The texture difference is 2. After that pattern classification is performed. The data normalization is no normalization. The distance measure is Euclidean Distance. The classification algorithm is k- Nearest Neighbor and k= 2. The algorithm testing is Training/Test. The accuracy of Happy is 80%, Anger is 80%, Sadness is 66.67%, Fear is 80%, disgust is 80% and surprise is 100%. The total accuracy rate is 81%.

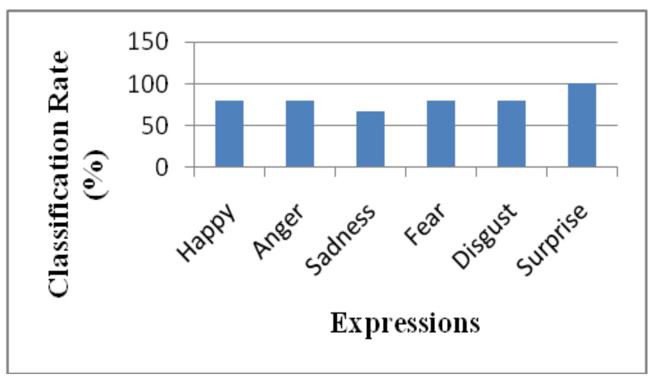


Figure 4. *Classification Result for Facial Expressions of Person C*

The above graph shows the classification results of six facial expressions of person C. The accuracy of Happy is 80%, Anger is 80%, Sadness is 66.67%, Fear is 80%, disgust is 80% and surprise is 100%.

5. CONCLUSION

All Confusion tables have shown different classification results of six facial expressions. Confusion table 4.1 has shown maximum accuracy of 90% at k= 2, where k is k- Nearest Neighbor algorithm. It is shown from all the confusion tables that Surprise expression has better results of 100% accuracy than other expressions. We have taken train data set of six facial expressions of three persons and we have total border mask sample 90 for train data set purpose and 30% border mask sample for test data set purpose and then classified them with better result. For better result we use RST- Invariant features and texture features.

REFERENCES

[1] Giorgana, G., and Ploeger, P. G., 2012, Facial expression recognition for domestic service robots, In Robo Cup 2011: Robot Soccer World Cup XV, pp. 353-364.

[2] Wang, Z. and Xiao, N., 2013, Using MD-Adaboost to Enhance Classifier of Facial Expression Recognition, Journal of Computational Information Systems, 9(3), 923- 932.

[3] Ghimire, D., and Lee, J., 2013, Geometric feature-based facial expression recognition in image sequences using multi-class Adaboost and support vector machines, Sensors, 13(6), 7714-7734.

[4] Abidin, Z., and Harjoko, A., 2012, A neural network based facial expression recognition using fisherface, International Journal of Computer Applications, 59(3), 30-34.

[5] Zavaschi, T. H., Britto, A. S., Oliveira, L. E., and Koerich, A. L., 2013, Fusion of feature sets and classifiers for facial expression recognition, Expert Systems with Applications, 40(2), 646-655.

[6] Sarawagi, V., and Arya, K. V., 2013, Automatic facial expression recognition for image sequences, In Contemporary Computing (IC3), Sixth International Conference on, 278-282.

[7] Hablani, R., Chaudhari, N., and Tanwani, S., 2013, Recognition of Facial Expressions using Local Binary Patterns of Important Facial Parts, International Journal of Image Processing (IJIP), 7(2), 163- 170.

[8] Lingareddy, M. and D, Haritha., 2013, An Optimal Face Recognition Tool, International Journal of Research in Engineering and Technology, 2(3), 351-357.

[9] Zhang, H., Luo, S., and Yoshie, O., 2013, Facial expression recognition by analyzing features of conceptual regions, In Computer and Information Science, 529-534.

[10] Das, D., 2014, Human's Facial Parts Extraction to Recognize Facial Expression, International journal on information theory, 3(3), 65-72.

[11] Hai, T. S., Thai, L. H. and Thuy, N. T., 2015, Facial Expression Classification Using Artificial Neural Network and K-Nearest Neighbor, International Journal of Information Technology and Computer Science (IJITCS), 7(3), 27- 32.

[12] De, A., Saha, A., and Pal, M. C., 2015, A Human Facial Expression Recognition Model Based on Eigen Face Approach, Procedia Computer Science, 45, 282-289.

[13] Chao, W. L., Ding, J. J., and Liu, J. Z., 2015, Facial expression recognition based on improved local binary pattern and class-regularized locality preserving projection, Signal Processing, 4(7), 1-39.

[14] Yu, H., and Liu, H., 2015), Combining appearance and geometric features for facial expression recognition, In Sixth International Conference on Graphic and Image Processing , 944308-944308.

[15] Zhang, L., and Tjondronegoro, D., 2011,. Facial expression recognition using facial movement features, IEEE Transactions on Affective Computing, 2(4), 219-229.