

RESEARCH ARTICLE

A Comparison of Four Classification Algorithms for Facial Expression Recognition

Hivi I. Dino*, Maiwan B. Abdulrazzaq

Department of Computer Science, Faculty of Science, University of Zakho, Zakho, Kurdistan Region, Iraq

***Corresponding author:**

Hivi I. Dino, Department of Computer science, Faculty of Science, University of Zakho, Zakho, Kurdistan Region, Iraq.

E-mail: hivi.dino@uoz.edu.krd

Received: 12 November 2019

Accepted: 16 June 2020

Published: 30 June 2020

DOI

10.25156/ptj.v10n1y2020.pp74-80

ABSTRACT

Facial expression recognition (FER) has achieved an extreme role in research area since the 1990s. This paper provides a comparison approach for FER based on three feature selection methods which are correlation, gain ratio, and information gain for determining the most distinguished features of face images using multi-classification algorithms which are multilayer perceptron, Naïve Bayes, decision tree, and K-nearest neighbor (KNN). These classifiers are used for the mission of expression recognition and for comparing their proportional performance. The main aim of the provided approach is to determine the most effective classifier based on minimum acceptable number of features by analyzing and comparing their performance. The provided approach has been applied on CK+ dataset. The experimental results show that KNN is proven to be better classifier with 91% accuracy using only 30 features.

Keywords: Correlation; Facial expression recognition; Feature selection; Gain ratio; Information gain

INTRODUCTION

Facial expression recognition (FER) in recent years has a considerable interest in research area due to its significance in human concentrated interfaces. FER applications include machine vision, robotics, security system, computer games, forensics, video conferencing, web services, and broadcasting (Wang et al., 2018). Various techniques have been proposed to analyze face expressions from both static and sequence images (Anil and Suresh, 2016). This paper provides a computationally effective approach to feature selection and classification for FER from sequence facial images.

FER generally goes through three main stages which are face detection, feature extraction, and classification (Hemalatha and Sumathi, 2014). Feature selection phase is considered as selecting a convenient subset of features for representing the input signal which the classification quality and computational complexity are effected by. Basically, the information implicated with selected features is adequate for an accurately determining of the input class. The complexity of classification task and training is caused by a large number of unnecessary features. While, indigent determination of features may detrimentally affect the classification results. The techniques of feature selection that are sufficient to modest division of patterns having a place with various classes fail in classification missions with overlapping boundaries and increasingly

complex distribution. Techniques, such as correlation expect linear conditions among data, can't deal with self-assertive relations between the coordinates of pattern and the distinctive classes. The most common data reduction techniques are not invariant when goes through linear transformations such as scaling of data utilized in preprocessing stage (Li et al., 2010). In the presented method, three different techniques of features selection are used as an objective criterion to produce the most characteristic features of facial images. The combination of first 10 ranked features of each feature selection techniques are used to classify the facial expressions using multilayer perceptron (MLP) (Boughrara et al., 2016), Naïve Bayes (NB) (Lajevardi and Hussain, 2010), decision tree J48 (Yan Nie et al., 2015), and K-nearest neighbor (KNN) (Wang et al., 2015) classifiers.

Many research papers were presented with the ability to recognize only few facial expressions, whereas this paper is provided to recognize all eight basic expressions of human faces which are normal, angry, contempt, disgust, fear, happy, sad, and surprise. Furthermore, many research papers were torturing utilizing a large number of features which was leading to complexity in time consuming. The main goal of the presented method is to provide an accurate recognition rate of facial expressions with less number of features which leads to less time consuming. This paper presents a comparison of four classification algorithms, namely, NB, MLP, decision tree J48, and KNN algorithm based on three feature selection methods, namely,

correlation feature selection (CFS), gain ratio (GR), and information gain (IG) using the extended Cohn-Kanade (CK+) dataset. The aim of this comparison is to find out the best classifier among the utilized ones that can give an acceptable accuracy rate using a minimum number of selected features.

The rest of paper is described as follows: Section II of the paper gives an overview of the related work. Section III describes the dataset been used for validating the provided approach. Section IV presents the methodology of the paper. Section V describes the performance evaluation and results of the provided work. Finally, Section VI shows the conclusion of the paper.

RELATED WORK

Li et al. (2010) proposed an approach that examines the method of feature selection which depends on the decrease of mutual data among the selected features. This paper implements and compares two techniques of feature selection which are symmetric Kullback–Leibler divergence (IG) and probability density function (pdf) to build FER system with reduced number of features. The proposed approach takes 373 feature as input that was selected using probability density function and achieved higher recognition rate than symmetric divergence method.

In additional, Boughara, Chtourou, Amar, and Chen (2016) proposed a deductive training algorithm based on MLP classifier which require to find composition parameters as the quantity of patterns relating for each class subsets to be exhibited at first in the training phase, hidden neurons initial quantity, iteration quantity during the training phase, and predefined value of mean square error. Perceived Facial Images descriptor is used as a feature extraction method which is a based on a biological vision.

Moreover, Kumbhar et al. (2012) suggested that using the specific facial parts perform better accuracy than the holistic methods. Along these lines, authors selected a

couple of dynamic patches from face dependent on the area of the landmark points and utilize the appearance patches as our feature vector. Authors used histogram of orientated gradients to extract features from the facial landmark points instead of whole face for building a robust system against pose and scale variations.

Peng and Yin (2018) suggested blending the photorealistic expression manifolds to extend the exhibition set. By means of blending expression pictures from neutral faces, more inside subject variability can be acquired. Eigen transformation is used for producing the details of shape and expression for novel subjects. Face recognition and expression classification then performed on the expanded training set with blended expressions. The proposed method performed robustness to recognize facial expressions with different degree of emotions.

Furthermore, Zhong et al. (2018) declared that subspace learning methods are critical to compound changes such as face expression, pose, and illumination but have low proportion features while local texture methods are powerful to compound changes but have high proportion features. Authors made full use of both method advantages by building a characteristic analysis from two various patterns of local texture (LTP). The provided framework ensured the effectiveness of both methods with lower proportion features.

Researchers in the related papers used various methods of feature extraction and classification with different numbers of facial expressions. The provided approach obtains a good recognition rate with less number of features and more recognized facial expressions. While Li et al., 2010, and Zhong et al., 2018, obtained a good recognition rate but used a larger number of features with the ability to recognize fewer facial expressions. Boughrara et al., 2016, Kumbhar et al., 2012, and Peng and Yin, 2018, could gain a high accuracy using less number of features but with the ability of recognizing fewer expressions. Table 1 shows the comparison summery of related works.

Table 1: Comparison summery of related works

Literature	Feature extraction	Classifier	Emotion No.	Feature No.	Dataset	Advantages
(Li et al., 2010)	pdf and info gain	SVM	7	373	JAFFE	Better performance
(Boughrara et al., 2016)	PFI	MLP	6	12	GEMEP FERA 2011, CK and FER 2013	More reliable and efficient
(Kumbhar et al., 2012)	Gabor filter and PCA	NN	6	20	JAFFE	Feasibility in computer vision applications
(Peng and Yin, 2018)	PCA	KDA and LCA	7	29	CK+, AR, Bosphorus, JAFFE, MUG, and MMI	Improved performance
(Zhong et al., 2018)	LTP	NN	4	65–120	Ar and CAS-PEAL-R1	Better accuracy

DATASET

To validate the provided approach, CK+ dataset has been used which consists of 593 image sequences of 123 adults (Lucey et al., 2010). Images consist of eight basic facial expressions which are normal, angry, contempt, disgust, fear, happy, sad, and surprise, the image sequence of each expression starts with natural face until peek expression. About 31% of participants are male and 69% are female, they are 18–50 years old. About 13% are Afro-American, 81% are Euro-American, and 6% different groups. From the whole dataset, 17 persons have all eight facial expressions, 22 have six facial expressions, 28 have five facial expressions, 26 have four facial expressions, 8 have three facial expressions, and 20 of them have two facial expressions. The obtained images are taken under different illuminations with the resolution of 640×480 or 640×490 . About 37% of images are taken with bright illumination, 18% with dark illumination, and 45% with normal illumination. In this work, 3203 instances from CK+ dataset are randomly selected for validating this approach. Samples from CK+ dataset are shown in Figure 1.

METHODOLOGY

This paper provides an approach to identify the expression of human faces with eight expressions. The process of this system goes through four basic stages which are face detection, feature extraction, feature selection, and classification, as shown in Figure 2.

Face Detection and Feature Extraction

The first step in FER process is to standardize images which includes noise removal, resizing, and transformation of the input image. The images taken from CK+ dataset are transformed to black and white images and fed to Viola-Jones algorithm for face detection. Viola-Jones algorithm is the most successful face detector algorithm due to its robustness in face detection rate and extremely

less false-positive rate, also it can be used for real time detection (Paul et al., 2018). The algorithm goes through four stages including selection of Haar features, integral image creation, AdaBoost training, and cascading classifiers. After face is detected, face images are cropped and resized to 80×80 , as shown in Figure 3a. By dividing cropped face horizontally into two halves, we take only left half with size 80×40 . From the left half, two box regions with size 25×25 are determined which are left eye regions and half mouth region, as shown in Figure 3b. Fifty features are extracted from both determined box regions. Features are then fed to three feature selection methods to determine the most distinct features for classification.

Feature Selection

After the extraction of 50 features from both box regions of the left eye and half mouth region, three different methods of feature selection are applied including CFS, IG, and GR to get most distinct features of each method. After testing our approach on these features based on trail and test method, it is found that selecting the first 30 ranked features from 50 features is the minimum numbers of features that will not effect on the quantity of accuracy and performance, as illustrated in Table 2. The accuracy of provided approach decreases when using <30 features and remains the same with very low and unnoticeable ratio when using more than first 30 ranked features. The expressed faces are then classified to eight facial expressions using only 30 features.

Classification

To identify the expression of face images, four techniques of supervised data mining are used and compared using 30 features gained from all three feature selection techniques in addition with class label. The technique of 10-fold cross-validation is used to train and test data.

MLP

MLP is a special model of artificial neural network with feed forward model which maps set of input data into

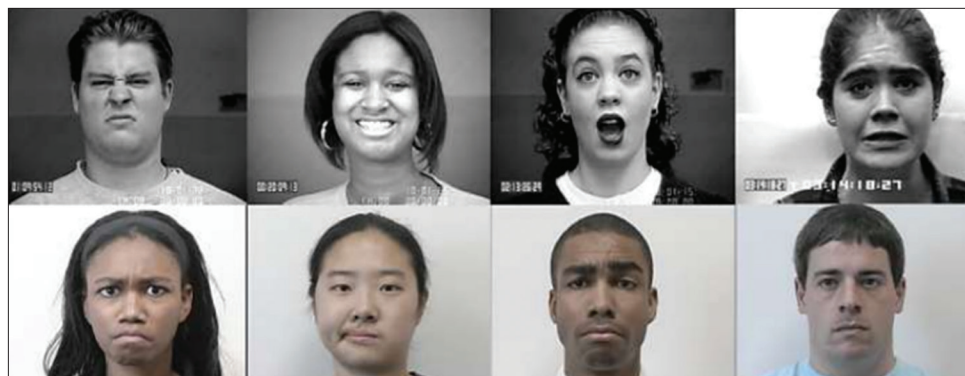


Figure 1: Sample expressions from CK+ dataset (Sadeghi et al., 2014)

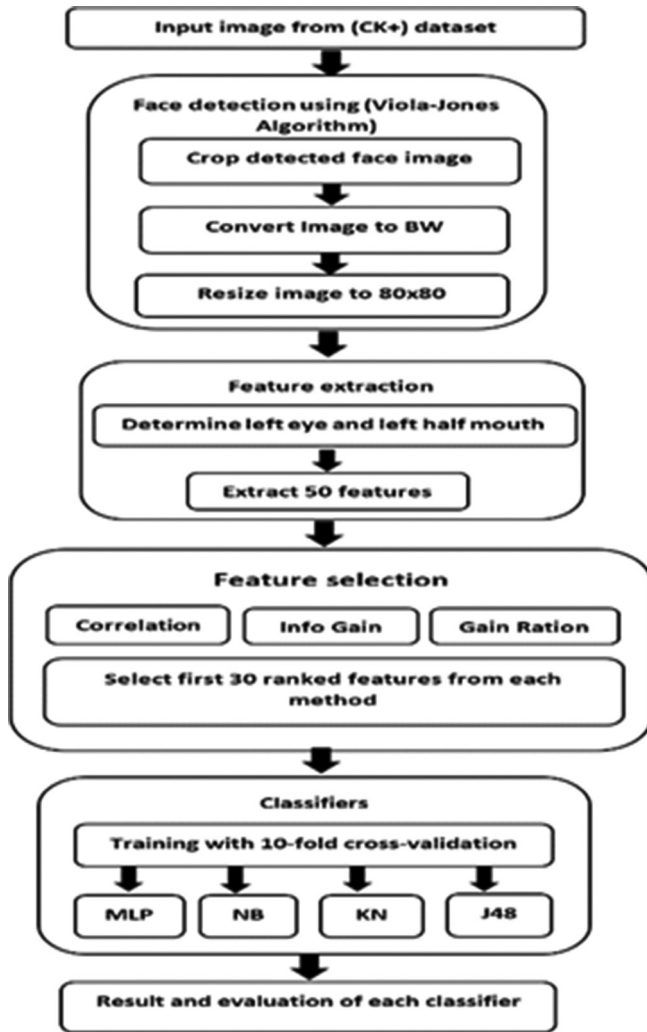
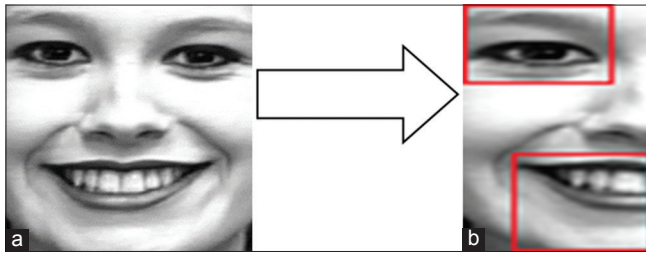


Figure 2: Block diagram of the provided approach

Figure 3: (a) Sample of cropped detected face with size 80×80 . (b) Sample of determined regions of the left half face with size 80×40

set of convenient output data (Nasr et al., 2011). MLP includes three layers which are input layer, hidden layer, and output layer (Kinnikar et al., 2016). In this work, MLP is constructed using 30 nodes as input layer and 20 hidden layers, in addition with eight output nodes which represent eight facial expression labels.

KNN

KNN is a simplest algorithm among all machine learning algorithms. KNN is a classification algorithm of non-

Table 2: Feature selection based on correlation, gain ratio, and info gain

No.	Correlation		Gain ratio		Info gain	
	Attribute	Ranking	Attribute	Ranking	Attribute	Ranking
1	2	0.1722	11	0.17068	11	0.26281
2	4	0.1413	50	0.12579	16	0.2186
3	6	0.1408	6	0.12103	6	0.14976
4	7	0.1392	48	0.11411	48	0.12579
5	10	0.1386	16	0.11314	44	0.11785
6	11	0.1366	44	0.11045	43	0.11376
7	12	0.1355	45	0.09907	37	0.10141
8	13	0.1272	49	0.0981	49	0.10063
9	16	0.1201	37	0.08295	32	0.09618
10	17	0.1187	40	0.08158	50	0.09161
11	19	0.1167	43	0.07845	45	0.08639
12	20	0.1166	35	0.07154	31	0.08587
13	21	0.1158	32	0.0607	26	0.0814
14	22	0.1146	12	0.06042	40	0.0765
15	23	0.1096	17	0.05738	21	0.07632
16	26	0.098	26	0.05505	19	0.0759
17	29	0.0927	19	0.04927	35	0.06261
18	31	0.0924	31	0.04835	17	0.0626
19	32	0.0867	29	0.03854	12	0.05074
20	34	0.0819	4	0.03683	23	0.04657
21	35	0.0818	7	0.03591	4	0.04455
22	37	0.0808	23	0.03497	14	0.04194
23	40	0.0781	34	0.03411	29	0.03731
24	42	0.0764	10	0.0333	30	0.03588
25	43	0.076	22	0.0313	20	0.03543
26	44	0.0755	2	0.03099	2	0.03503
27	45	0.071	20	0.02965	22	0.03481
28	48	0.0698	42	0.02796	9	0.03406
29	49	0.0697	21	0.02675	10	0.03372
30	50	0.0688	13	0.02647	34	0.03339

parametric type. It is a type of lazy learning because the function is locally approximated and the computations are postponed until classification. K is a number of input training instances in feature space (Wang et al., 2015). K is typically a small, positive, and integer number. The output is class labels. An item can be classified by the majority vote of an item neighbors with the item being appointed to a class which is closest to its KNNs. Assume there is a set of training data, KNN receives an unknown record X' and searches for the KNN of X record among the training set by predicting its class value and assigning to X' the most widely recognized class among its k closest neighbors.

$$d(x, y) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2} \quad (1)$$

NB

NB algorithm is a type of simple probabilistic classifier of machine learning based on setting the Naïve (strong) independent hypothesis with Bayes theorem between the features. NB is extremely scalable, requires a various parameters linear in the quantity of features in a learning problem (Lajevardi and Hussain, 2010). Most extreme training are prepared by assessing a closed structure expression which require linear time as opposed to by costly iterative estimate as utilized for some different sorts classifiers.

$$P(M / X) = P(X / M) \cdot (M) / (X) \quad (2)$$

X is the input record. M represents that X belongs to C class. $P(M)$ is the prior probability, $P(M/X)$ is the posterior probability of M conditioned on X , and $P(X/M)$ is the posterior probability of X conditioned on M .

Decision Tree (J48)

J48 decision tree is the enhanced version of C4.5 algorithm. It is the most common classifier of tree algorithms. It build up on the concept of information entropy (Bhargava et al., 2013). It utilizes the way that each trait of the data can be utilized by dividing data into small subsets which include of nodes that structure the rooted tree. Each tree has three sorts of nodes including root node, internal node, and leave node, first one has no incoming edges and the last one determines the class labels. Inward nodes contain feature test conditions to divide records that have various attributes. The pseudo-code of constructing the decision tree is written below:

Determine the class labels

For each feature f , find the info gain from separating f .

Let F be the best feature with the highest info gain.

Make a decision node that separates F .

Repeats on the sub-records acquired by separate f best and include those hubs as offspring of nodes.

PERFORMANCE EVALUATION

For analyzing and evaluating the performance of each classifier which are utilized in this work, the following measurements are used (Behaine and Scharcanski, 2012):

$$\text{True positive (TP)} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{False positive (FP)} = \frac{FP}{FP + TN} \quad (4)$$

$$\text{True negative (TN)} = \frac{TN}{TN + FP} \quad (5)$$

$$\text{False negative (FN)} = \frac{FN}{FN + TP} \quad (6)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (8)$$

$$\text{Recall} = \frac{TP + TN}{TP + FN} \quad (9)$$

$$F - \text{Measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (10)$$

When:

- Accuracy: Correctly classified instances divided by the total number of instances.
- Precision: The quantity of the instances which are true to a specific class divided by overall instances classified as that class.
- Recall: It refers to the quantity of correctly classified instances divided by the total number of instances present in the class.
- F-measure: The combination measure of precision and recall.
- True positive (TP): The correctly classified instances.
- False positive (FP): The incorrectly classified instances.
- True negative (TN): The correctly rejected instances.
- False negative (FN): The incorrectly rejected instances.

The average weight values of TP rate, FP rate, precision, recall, F-measure, and the processing time (P. time) in seconds (sc.) for each instance of each classifier based on each utilized feature selection method are shown in Tables 3-5, respectively.

The distinctions and similitudes of correctly and incorrectly classified expressions in each classifier for each feature selection method are outlined and summarized in Figures 4-6.

The objective of this work is to provide the comparison between the performances of used classifiers. The

experimental results show that when using 30 features from CFS method, KNN gets the highest recognition rate with 90.17%, MLP gets 73.56% of recognition rate, J48 gets 81.11% of recognition rate, and NB gets the lowest recognition rate with 43.87%. When using 30 features from IG method, KNN gets the highest recognition rate with 89.45%, MLP gets 73.56% of recognition rate, J48 gets the recognition rate of 40.64%, and NB gets the recognition rate with 40.72%. When using 30 features from GR method, KNN gets the highest recognition rate with 91.07%, MLP gets 73.68% of recognition rate, NB gets 48.33%, and J48 gets the recognition rate of 81.14%. The accuracy performance of each classifier based on all three feature selection methods is illustrated in Table 6.

Table 3: Performance detail for classifiers with 10-fold cross-validation based on correlation

Classifier	TP rate	FP rate	Precision	Recall	F-measure	P. time (sc.)
KNN	0.902	0.021	0.903	0.902	0.902	0.02
MLP	0.743	0.052	0.747	0.743	0.745	16.77
NB	0.439	0.101	0.476	0.439	0.45	0.03
DT	0.81	0.04	0.81	0.81	0.81	0.3

Table 4: Performance detail for classifiers with 10-fold cross-validation based on info gain

Classifier	TP rate	FP rate	Precision	Recall	F-measure	P. time (sc.)
KNN	0.894	0.023	0.897	0.894	0.895	0.02
MLP	0.736	0.054	0.741	0.736	0.738	16.77
NB	0.451	0.098	0.484	0.451	0.462	0.02
DT	0.784	0.039	0.787	0.784	0.784	0.03

Table 5: Performance detail for classifiers with 10-fold cross-validation based on gain ratio

Classifier	TP rate	FP rate	Precision	Recall	F-measure	P. time (cs.)
KNN	0.911	0.019	0.912	0.911	0.911	0.0
MLP	0.81	0.04	0.81	0.81	0.81	16.45
NB	0.452	0.1	0.484	0.452	0.462	0.02
DT	0.818	0.038	0.819	0.818	0.818	0.11

Table 6: Accuracy detail for classifiers with 10-fold cross-validation based on three feature selection methods

Classifiers	Accuracy based on three feature selection techniques (%)		
	Correlation	Info gain	Gain ratio
KNN	90.17	89.45	91.07
J48	81.11	40.64	81.14
NB	43.87	40.72	48.33
MLP	73.56	73.56	73.68

According to the illustrated results, when applying the classifiers on the features illustrated from each technique of feature selection, the highest classified emotion is fear which has 617 correctly classified instances that were classified with KNN classifier, the lowest correctly classified instances are with contempt emotion because of having the least number of instances. Fear, happy, and surprise emotions have the most correctly classified instances when applied to KNN and MLP classifiers. Based on the CFS, the highest correctly classified instances is fear emotion with quantity 617, 544, and 494 when applied on KNN, J48, and MLP, respectively, also happy emotion with 338 correctly instances when applying NB classifier. Furthermore, based on GR, the highest correctly classified instances are fear with quantity 615 and 491 when applied on KNN and MLP, also surprised emotion with quantity 468 and 355 when applied on J48 and NB, respectively. Moreover, based on IG, the highest correctly classified instances are fear with

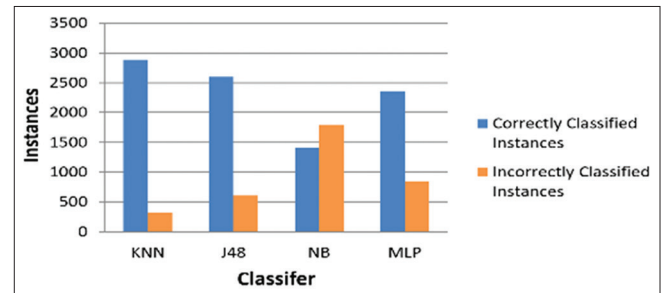


Figure 4: A comparison of correctly classified instances for each classifier based on correlation feature selection

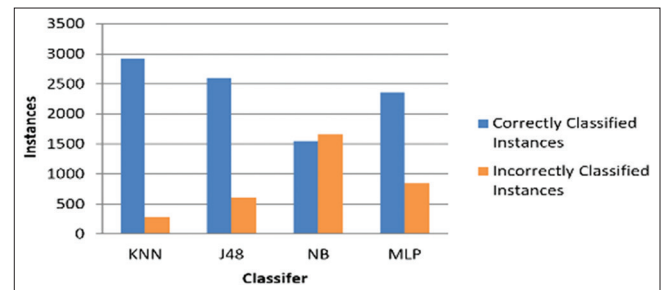


Figure 5: A comparison of correctly classified instances for each classifier based on gain ratio

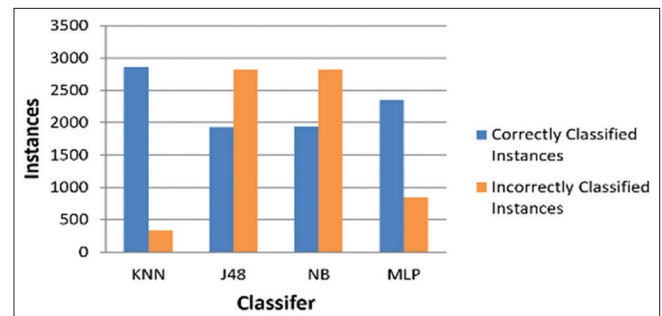


Figure 6: A comparison of correctly classified instances for each classifier based on info gain

quantity 606 and 494 when applied on KNN and MLP, also surprised emotion with quantity 456 and 432 when applied on J48 and NB, respectively.

CONCLUSION

In this research, a comparison approach for FER is provided based on three techniques of feature selection, namely, CFS, GR, and IG using four machine learning algorithm for classification including KNN, MLP, J48, and NB. CK+ dataset is used which is a good data resource to validate the process of this work. The experimental results show that when using only 30 features, KNN is proven to be better classifier with accuracy 91.01%, 90.17%, and 89.45% when applied on GR, CFS, and IG, respectively. NB is the worst classifier according to the lowest accuracy 81.11%, 40.64%, and 81.14% when applied on CFS, IG, and GR, respectively. MLP and J48 classifiers gained the medium range of accuracy 81.11%, 40.64%, and 81.14% for J48, 73.56%, 73.56%, and 73.68% for MLP when applied on CFS, IG, and GR, respectively. In future works, more modern classification methods will be applied on the provided approach such as CNN with comparing their performances.

REFERENCES

- Anil, J. and L. P. Suresh. 2016. Literature Survey on Face and Face Expression Recognition. 2016 International Conference on Circuit, Power and Computing Technologies, IEEE. p1-6.
- Behaine, C. A. and J. Scharcanski. 2012. Enhancing the performance of active shape models in face recognition applications. IEEE Trans. Instrum. Meas. 61: 2330-2333.
- Bhargava, N., G. Sharma, R. Bhargava and M. Mathuria. 2013. Decision tree analysis on j48 algorithm for data mining. Proc. Int. J. Adv. Res. Comput. Sci. Softw. Eng. 3: 77-90.
- Boughrara, H., M. Chtourou, C. B. Amar and L. Chen. 2016. Facial expression recognition based on a mlp neural network using constructive training algorithm. Multimed. Tools Appl. 75: 709-731.
- Hemalatha, G. and C. Sumathi. 2014. A study of techniques for facial detection and expression classification. Int. J. Comput. Sci. Eng. Surv. 5: 27.
- Kinnikar, A., M. Husain and S. Meena. 2016. Face Recognition Using Gabor Filter and Convolutional Neural Network. Proceedings of the International Conference on Informatics and Analytics. Association for Computing Machinery. p113.
- Kumbhar, M., A. Jadhav and M. Patil. 2012. Facial expression recognition based on image feature. Int. J. Comput. Commun. Eng. 1: 117.
- Lajevardi, S. M. and Z. M. Hussain. 2010. Higher order orthogonal moments for invariant facial expression recognition. Digit. Sig. Process. 20: 1771-1779.
- Li, P., S. L. Phung, A. Bouzerdom and F. H. C. Tivive. 2010. Feature Selection for Facial Expression Recognition. 2010 2nd European Workshop on Visual Information Processing, IEEE. p35-40.
- Lucey, P., J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews. 2010. The Extended Cohn-Kanade Dataset (ck+): A Complete Dataset for Action Unit and Emotion-specified Expression. IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, IEEE. p94-101.
- Nasr, G., R. Achkar and F. Fayad. 2011. Stereo-Vision Calibration by Multi-Layer Perceptrons of an Artificial Neural Network. 2011 7th International Symposium on Image and Signal Processing and Analysis, IEEE, p621-626.
- Paul, T., U. A. Shammi, M. U. Ahmed, R. Rahman, S. Kobashi and M. A. R. Ahad. 2018. A study on face detection using viola-jones algorithm in various backgrounds, angles and distances. Int. J. Biomed. Soft Comput. Hum. Sci. 23: 27-36.
- Peng, Y. and H. Yin. 2018. Facial expression analysis and expression-invariant face recognition by manifold-based synthesis. Mach. Vision Appl. 29: 263-284.
- Sadeghi, H., A. A. Raie and M. R. Mohammadi. 2014. Facial expression recognition using texture description of displacement image. J. Inf. Syst. Telecommun. 2: 205-212.
- Wang, N., X. Gao, D. Tao, H. Yang and X. Li. 2018. Facial feature point detection: A comprehensive survey. Neurocomputing. 275: 50-65.
- Wang, X. H., A. Liu and S. Q. Zhang. 2015. New facial expression recognition based on FSVM and KNN. Optik. 126: 3132-3134.
- Yan Nie, C., J. Wang, F. He and R. Sato. 2015. Application of J48 Decision Tree Classifier in Emotion Recognition Based on Chaos Characteristics. 2015 International Conference on Automation, Mechanical Control and Computational Engineering, Atlantis Press.
- Zhong, F., S. Yan, L. Liu and K. Liu. 2018. An Effective Face Recognition Framework With Subspace Learning Based on Local Texture Patterns. 2018 5th International Conference on Systems and Informatics, IEEE. p266-271.