Two Feature Selection Methods Comparison Chi-square and Relief-F for Facial Expression Recognition

Mayyadah R. Mahmood
Department of Computer Science, University of Zakho, Duhok, Kurdistan Region – Iraq.
mayyadah.mahmood@uoz.edu.krd

Abstract. Feature selection method represents one of the main keys that has direct influence on classification accuracy. During the last two decades, researchers have given a lot of attention in feature selection approaches due to their importance. This paper provides a comparative approach between the two feature selection methods: Chi-Square and Relief-F. The two methods rank the features according to their score. The first highest six emotion features from the both methods are selected. The six features are used to compare the accuracy ratio among the four classifiers: Support Vector Machine, K-Nearest, Decision Tree, and Radial Base Function. These classifiers are used for the mission of expression recognition and to compare their proportional performance. The ultimate aim of the provided approach is to use minimum number of features from the both methods in order to distinguish the performance accuracy of the four classifiers. The provided approach has been applied on CK+ facial expression recognition dataset. The result of the experiment illustrates that K-Nearest Neighbor is the most accurate classifier on the both feature selection methods according to the employed dataset. The K-Nearest Neighbor accuracy average rate for Chi-square is 94.18% and for Relief-F is 94.93%.

1. Introduction
As a biometric identification, Face Expression Recognition (FER) is worldwide applied to know someone among others [1]. Face Recognition techniques depend on the difference in specific physical or behavioural characteristics among people [2]. This technology is recently being implemented for various forensic purposes, emergency and security [3], machine learning [4], computer vision [5], computer games [6], Real-Time video [7] and Web Services [8]. Various researches are being carried out in terms of face detection and tracking, methods of expression classification, and the mechanisms of feature expression [9].

Peng and Yin [10] offered a system of FER via combining the photorealistic expression manifolds in order to expand the exhibition set. More inside subject variability can be acquired via blending expression images from neutral faces. Eigen, as a transformation approach, is applied to produce the details of expression and facial form for new objects. Expression classification and face-recognition are mainly applied on the extended training dataset with blended facial features to test the accuracy of the suggested approaches such as MMI, MUJ, CK+, JAFFE, AR, Bosporus datasets. The experiment illustrates that the suggested method represents a favourable face expression recognition rate.

Fekri [11] suggested a novel approach in which consists of two phases. In the first phase, applying ILBP, a feature vector is extracted for human face. In the concluding phase, gender-classification is carried out applying T-KL on two data bases: ICPR and self-collected. The suggested method offers 94.76 percent of accuracy on ICPR dataset. Munir et al. [12], in order to compensate the weak
illumination, suggests application of Fast Fourier Transform (FFT) and Contrast Limited Adaptive Histogram Equalization (CLAHE) approach. In the next step, in order to generate each pixel merged binary pattern code is used. To form a 16-bit code for each pixel, two bits per neighbour are generated. This approach elicits the best quality to find the edge and the permanent attributes around mouth, eyebrows, eyes, and facial wrinkles. The suggested method results are compared with other variants of LGC and LBP methods for zoned and holistic images. The method applies SFEW dataset. The approach achieved the total accuracy rate of 96.5% for holistic and 67.2% for division-based approach.

Zhong et al [13] suggested a novel face recognition sub-space learning based approach on the local texture attributes. Both of the methods, CAS-PEARL-R1 and AR database demonstrate that either of the approaches are more accurate than the traditional methods. The suggested approaches deserve higher consideration due to their capacity and potential of achieving more accuracy. Bilkhu et al. [14] presented another model for Facial Expression Recognition I order to recognize the basic six emotional expressions. In order to extract features, this method applies cascade regression. To implement this task, the method utilizes three machine learning algorithms for the classification of the features. The technique applied Logistic Regression, Support Vector Machine, and NN. The dataset applied in this system was CK+ and the achieved result for each algorithm was compared. The result illustrates 89% of accuracy for SVM, 80% for NN, and 77.06% accuracy for Logistic Classifier.

The current article aims to evaluate the performance of the selected classifiers through utilization of two feature selection method for facial expression recognition via a set of facial images. This paper provides a computationally effective approach to feature selection and classification for FERT from sequence facial images.

FERT fundamentally undergoes through three main phases. The first phase implements the process of face detection. Feature extraction is done within the second phase. Then, the training and testing via classification method is applied for the third phase [15]. The first phase phase is considered as selecting a convenient subset of features for representing the input signal [16] that is influenced by the instant impact of the computational complexity and classification quality [17]. Naturally, the implicated rundown with the elicited attributes is quite enough for the accurate determination of the input class. The complexity of classification task and training is caused by a big quantity of redundant attributes. The methods of feature selection which are competent in division of patterns having a position with different classes fail classification tasks with overlapping edges and progressive complex distribution Techniques, such as correlation expect linear conditions among data, and 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 pre-processing stage [18]. In the suggested approach the features selection method is applied as a Chi-Squared and Relief-F that select the highest 6 attributes. The previously elicited attributes will be used through the process of training and testing the dataset using the appointed classifiers which include KNN [19], J48 [20], RBF [21] and SVM [22].

2. Dataset

CK+ dataset is one of the most well-known and highly applied set of data which is gathered by the participation of 210 adult individuals of both genders [23]. Images consist of 8 fundamental facial-expressions that are (surprised, sad, happy, scared, disgusted, contempt, angry, and neutral) [24]. The dataset has the total percentage of 31% males and the rest females. [25].
Figure 1. CK+ Facial Expressions Dataset Sample [27]

The dataset consists of individuals of different ethnicities including Euro-American, Afro-American, and other races which respectively are 81%, 13%, and 6%. The number of facial expression image sequence is 139 from 123 subjects, which among them 327 people show all the eight facial expressions. Images’ resolution is originally 640 x 480 which are captured under different illumination. This article applies 4090 random selected samples from the appointed dataset. Samples from CK+ dataset are shown in Figure 1 [26]. Table 1 demonstrates all the emotions detected and face labelled for Chi-Square and Relief-F feature selection methods.

| No. | Expressions | No. of Instances |
|-----|-------------|------------------|
| 1   | Angry       | 527              |
| 2   | Contempt    | 47               |
| 3   | Disgust     | 389              |
| 4   | Fear        | 458              |
| 5   | Happy       | 614              |
| 6   | Normal      | 913              |
| 7   | Sad         | 540              |
| 8   | Surprised   | 602              |

3. Methodology

The main approach of this paper goes through four steps to identify human facial expression recognition. The four steps are data processing are Data Processing, Face Detection, Feature Selection and Classification, as illustrated in Figure 2.
3.1. Data preprocessing
The process initiates with standardization of the images which means removal of the noise, resize, and transformation. The black and white pictures from the CK+ dataset is processed by Viola-Jones.

3.2. Face Detection and Feature Selection
Viola-Jones could be considered as one of the most applicable and competent approaches that is used particularly for the purpose of face-detection [28], add to that, it is used in Realtime detection [29]. Viola-Jones [23] is widely used for face detection due to its robustness in face detection rate and being outstandingly accurate among other techniques [24]. Add to this, it is a real time detection tool [25]. This method consists of the following steps: Haar feature selection, integral image creation, and Adaboost training and cascading classifiers. The detected faces from Viola-Jones are cropped and resized to 28x28. The 784 attributes are fed in Relief-F feature selection method to rank the features based on their positional importance. Then, the most distinctively high ranked features will be elicited from the rest in order to be used in six classifiers to distinguish the most accurate one.

3.3. Feature selection
Feature selection is used to order the features according to their ranks [30]. This paper uses two types of feature selection methods that are Chi-Square and Relief-F.

3.3.1. Feature selection via Chi-square
Chi-Square method is one of the most useful machines learning tools. Chi-Square equation is:

$$\chi^2(t, c) = \frac{N(AD-CB)^2}{(A+C)(B+D)(A+B)(C+D)}$$
When A is the variant frequency containing t belonging to class c, B contains t and is not a subordinate part of c, C illustrates the frequency of the document which does not contain t and does not belong to the class c, and N is the valiant of the document in the quantity [31]. The method has applied CK+ dataset and the optimum 6 features are being chosen as illustrated in Table 2.

**Table 2.** The highest ranked six feature from Chi-square feature selection for each expression Chi-square feature selection

| Feature Numbers | Anger | Contempt | Disgust | Fear | Happy | Normal | Sad | Surprise |
|-----------------|-------|----------|---------|------|-------|--------|-----|----------|
| 1               | 95    | 378      | 150     | 595  | 499   | 545    | 37  | 605      |
| 2               | 96    | 431      | 151     | 596  | 510   | 550    | 38  | 627      |
| 3               | 101   | 436      | 160     | 609  | 511   | 551    | 39  | 628      |
| 4               | 102   | 564      | 177     | 610  | 512   | 570    | 433 | 633      |
| 5               | 103   | 565      | 178     | 623  | 526   | 571    | 437 | 634      |
| 6               | 130   | 592      | 179     | 637  | 527   | 572    | 601 | 655      |

### 3.3.2. Feature Selection via Relief-F
Relief-F is applied to process binary classification. After matching the similarly between the observations and categories to finds the nearest miss. In addition, this method analyses the relevancy index that imposes a negative weight to the attributes that are similar in different classes. Finally, the positive weight of the index is given to similar features in the same class. Hence, the operation result assigned with maximum relevancy index [32]. Relief-F feature selection method is applied on CK+ dataset and the highest rank six features are being elicited as shown in Table 3.

**Table 3.** The highest ranked six features from Relief-F feature selection for each expression

| Feature Numbers | Anger | Contempt | Disgust | Fear | Happy | Normal | Sad | Surprise |
|-----------------|-------|----------|---------|------|-------|--------|-----|----------|
| 1               | 93    | 64       | 121     | 120  | 526   | 543    | 93  | 571      |
| 2               | 94    | 65       | 149     | 568  | 539   | 570    | 104 | 572      |
| 3               | 103   | 431      | 150     | 569  | 540   | 571    | 121 | 577      |
| 4               | 104   | 436      | 151     | 580  | 541   | 572    | 132 | 579      |
| 5               | 121   | 564      | 159     | 581  | 553   | 573    | 133 | 599      |
| 6               | 131   | 784      | 160     | 757  | 554   | 574    | 404 | 600      |

### 3.4. Classification
The applies KNN, J48, SVM, and RBF for classification purposes depending on the resulted six features acquired from Relief-F and Chi-Square. The training and testing process used ten-fold-cross validation technique.

#### 3.4.1. K-Nearest Neighbour (KNN)
KNN is one of the broadly applied classifiers universally depending on its simplicity. The learning process, as far as concerned with KNN, is slow due to the local computational traits and also because of the postponed during the classification stage. K represents the number of categories in the domain. The output is class that has a label. In this approach the distance of a stance and its neighbouring categories are determined according to the most specified features in each group. This classifier tests the unlabelled X and attempts to figure out the nearest category belongs to. The equation below illustrates KNN’s calculational operation [33].

\[
    d(x, y) = \sqrt{\sum_{i=1}^{m} (x_i - y_i)^2}
\]
3.4.2. Decision Tree (J48)
J48 is an advanced model of C4.5 algorithm and is considered as widely known as tree algorithm. It is constructed upon the information entropy bedrock [34]. The tree consists of three main nodes. During the process of the attributes, the highest entropy is considered the root of the tree. Similar entropy concept is applied on the rest subsets which form internal nodes and the leaf nodes. During the process, it is the leaf node that determines the label of a class. A set of conditions determines the that a tested record belongs to which class.

3.4.3. Radial Basis Function (RBF)
As an offshoot of ANN, RBF is a non-linear and consists of 3 layers: input layer, hidden layer, and the output layer. Input layer relays on the number of selected features in which are feed into the algorithm. The fed features are processed within the hidden layer. The output layer provides us with the ultimate classes. This technique is considered as a feedforward neural network [35].

3.4.4. Support Vector Machine (SVM)
This classifier is designed upon the concept of two classes and a margin. This technique is mostly applicable for biometric information, data mining, computer vision, intrusion detection in which most often are involved with the processing of a huge number of data [36]. As a Kernel-based approach, the method is widely used for the classification process. It attempts to construct the optimum plane margin where the distance from the nearest of the training set of instances is maximized. The optimal hyperplane, when involved with linearly separable cases, leads to eye-catching accurate results on different kinds of datasets [37].

4. Performance evaluation and results
The confusion matrix is used for analysing and evaluating the performance of each classifier. The average weights values of TP rate, FP rate, Precision, Recall, F-measure and the processing time in seconds (sc.) for each instances of each classifiers based on each utilized feature selection methods are shown in. Table 4, Table 5, Table 6 and Table 7 represent the Chi-square feature selection. Table 8, Table 9, Table 10 and Table 11 represent Relief-F feature selection.

Table 4. Performance Result for J48 Algorithm Based on Chi-Square

| Expressions | TP Rate | FP Rate | Precision | Recall | F-Measure | Accuracy |
|-------------|---------|---------|-----------|--------|-----------|----------|
| Anger       | 0.97    | 0.42    | 0.94      | 0.97   | 0.95      | 91.93    |
| Contempt    | 1.00    | 0.64    | 0.99      | 1.00   | 1.00      | 99.12    |
| Disgust     | 0.99    | 0.51    | 0.95      | 0.99   | 0.97      | 93.94    |
| Fear        | 0.99    | 0.62    | 0.93      | 0.99   | 0.96      | 92.47    |
| Happy       | 0.99    | 0.17    | 0.97      | 0.99   | 0.98      | 96.26    |
| Normal      | 0.95    | 0.58    | 0.85      | 0.95   | 0.90      | 83.01    |
| Sad         | 0.98    | 0.62    | 0.91      | 0.98   | 0.95      | 90.02    |
| Surprise    | 0.99    | 0.19    | 0.97      | 0.99   | 0.98      | 96.19    |
| Avg. Rate   | 0.98    | 0.47    | 0.94      | 0.98   | 0.96      | 92.87    |
The objective of this work is to provide the comparison between the performances of used classifiers. The experimental results show that when using six features from Chi-square method, gets the highest recognition rate with 94.18% by KNN classifier, J48 gets 92.87% of recognition rate, SVM gets 90.48% of recognition rate and RBF gets the lowest recognition rate with 90.38% that shown in Tables 4, 5, 6 and 7 respectively.

The result for Chi-square feature selection applied to recognize the eight types of face detection. The highest accuracy rate for Anger gets 94.47%, Contempt gets 99.14%, Disgust gets 96.14%, Fear gets 92.54%, Normal gets 87.95%, Sad gets 90.69% and Surprise gets 96.94% by using KNN classifier while Happy gets 96.26% by using J48. The best classifier for Chi-square is KNN because the highest result detected from it, just happy detected from J48.

| Expressions | TP Rate | FP Rate | Precision | Recall | F-Measure | Accuracy |
|-------------|---------|---------|-----------|--------|-----------|----------|
| Anger       | 0.99    | 0.35    | 0.95      | 0.99   | 0.97      | 94.48    |
| Contempt    | 1.00    | 0.75    | 0.99      | 1.00   | 1.00      | 99.14    |
| Disgust     | 1.00    | 0.37    | 0.96      | 1.00   | 0.98      | 96.14    |
| Fear        | 1.00    | 0.66    | 0.92      | 1.00   | 0.96      | 92.54    |
| Happy       | 0.99    | 0.26    | 0.96      | 0.99   | 0.97      | 95.58    |
| Normal      | 0.96    | 0.41    | 0.89      | 0.96   | 0.93      | 87.95    |
| Sad         | 0.99    | 0.66    | 0.91      | 0.99   | 0.95      | 90.69    |
| Surprise    | 1.00    | 0.19    | 0.97      | 1.00   | 0.98      | 96.94    |
| Avg. Rate   | 0.99    | 0.46    | 0.94      | 0.99   | 0.97      | 94.18    |

| Expressions | TP Rate | FP Rate | Precision | Recall | F-Measure | Accuracy |
|-------------|---------|---------|-----------|--------|-----------|----------|
| Anger       | 0.99    | 0.83    | 0.89      | 0.99   | 0.94      | 88.17    |
| Contempt    | 1.00    | 0.77    | 0.99      | 1.00   | 1.00      | 99.12    |
| Disgust     | 0.99    | 0.67    | 0.93      | 0.99   | 0.96      | 92.91    |
| Fear        | 0.99    | 0.77    | 0.91      | 0.99   | 0.95      | 90.81    |
| Happy       | 0.96    | 0.29    | 0.95      | 0.96   | 0.95      | 92.20    |
| Normal      | 0.95    | 0.80    | 0.81      | 0.95   | 0.87      | 78.41    |
| Sad         | 0.99    | 0.87    | 0.88      | 0.99   | 0.93      | 87.75    |
| Surprise    | 1.00    | 0.41    | 0.93      | 1.00   | 0.96      | 93.67    |
| Avg. Rate   | 0.98    | 0.68    | 0.91      | 0.98   | 0.95      | 90.38    |

| Expressions | TP Rate | FP Rate | Precision | Recall | F-Measure | Accuracy |
|-------------|---------|---------|-----------|--------|-----------|----------|
| Anger       | 1.00    | 0.92    | 0.88      | 1.00   | 0.93      | 87.78    |
| Contempt    | 1.00    | 0.75    | 0.99      | 1.00   | 1.00      | 99.10    |
| Disgust     | 1.00    | 0.79    | 0.92      | 1.00   | 0.96      | 92.35    |
| Fear        | 0.99    | 0.76    | 0.91      | 0.99   | 0.95      | 90.95    |
| Happy       | 0.98    | 0.32    | 0.95      | 0.98   | 0.96      | 93.77    |
| Normal      | 1.00    | 1.00    | 0.78      | 1.00   | 0.87      | 77.68    |
| Sad         | 1.00    | 1.00    | 0.87      | 1.00   | 0.93      | 86.80    |
| Surprise    | 0.99    | 0.25    | 0.96      | 0.99   | 0.97      | 95.43    |
| Avg. Rate   | 1.00    | 0.72    | 0.91      | 1.00   | 0.95      | 90.48    |
Table 8. Performance Result for J48 Algorithm Based on Relief-F

| Expressions | TP Rate | FP Rate | Precision | Recall | F-Measure | Accuracy |
|-------------|---------|---------|-----------|--------|-----------|----------|
| Anger       | 0.97    | 0.33    | 1.00      | 0.97   | 0.96      | 93.35    |
| Contempt    | 1.00    | 0.75    | 1.00      | 1.00   | 1.00      | 99.14    |
| Disgust     | 0.98    | 0.47    | 1.00      | 0.98   | 0.97      | 93.72    |
| Fear        | 0.98    | 0.53    | 1.00      | 0.98   | 0.96      | 92.18    |
| Happy       | 0.97    | 0.19    | 1.00      | 0.97   | 0.97      | 94.67    |
| Normal      | 0.94    | 0.68    | 1.00      | 0.94   | 0.88      | 80.42    |
| Sad         | 0.96    | 0.55    | 1.00      | 0.96   | 0.94      | 89.61    |
| Surprise    | 0.98    | 0.23    | 1.00      | 0.98   | 0.97      | 95.06    |
| Avg. Rate   | 0.97    | 0.47    | 1.00      | 0.97   | 0.96      | 92.27    |

Table 9. Performance Result for KNN Algorithm Based on Relief-F

| Expressions | TP Rate | FP Rate | Precision | Recall | F-Measure | Accuracy |
|-------------|---------|---------|-----------|--------|-----------|----------|
| Anger       | 0.99    | 0.12    | 0.98      | 0.99   | 0.99      | 97.58    |
| Contempt    | 1.00    | 0.75    | 0.99      | 1.00   | 1.00      | 99.14    |
| Disgust     | 1.00    | 0.31    | 0.97      | 1.00   | 0.98      | 96.77    |
| Fear        | 0.99    | 0.46    | 0.95      | 0.99   | 0.97      | 94.25    |
| Happy       | 0.77    | 0.01    | 0.96      | 0.77   | 0.85      | 96.04    |
| Normal      | 0.96    | 0.47    | 0.88      | 0.96   | 0.91      | 85.92    |
| Sad         | 0.98    | 0.26    | 0.96      | 0.98   | 0.97      | 94.96    |
| Surprise    | 1.00    | 0.33    | 0.95      | 1.00   | 0.97      | 94.74    |
| Avg. Rate   | 0.96    | 0.34    | 0.95      | 0.96   | 0.96      | 94.93    |

Table 10. Performance Result for RBF Algorithm Based on Relief-F

| Expressions | TP Rate | FP Rate | Precision | Recall | F-Measure | Accuracy |
|-------------|---------|---------|-----------|--------|-----------|----------|
| Anger       | 0.98    | 0.83    | 0.89      | 0.98   | 0.93      | 87.75    |
| Contempt    | 1.00    | 0.75    | 0.99      | 1.00   | 1.00      | 99.14    |
| Disgust     | 0.98    | 0.77    | 0.92      | 0.98   | 0.95      | 91.10    |
| Fear        | 1.00    | 1.00    | 0.89      | 1.00   | 0.94      | 88.80    |
| Happy       | 0.96    | 0.38    | 0.93      | 0.96   | 0.95      | 91.20    |
| Normal      | 0.95    | 0.74    | 0.82      | 0.95   | 0.88      | 79.24    |
| Sad         | 1.00    | 0.97    | 0.87      | 1.00   | 0.93      | 87.04    |
| Surprise    | 0.99    | 0.40    | 0.94      | 0.99   | 0.96      | 92.91    |
| Avg. Rate   | 0.98    | 0.73    | 0.91      | 0.98   | 0.94      | 89.65    |
Table 11. Performance Result for SVM Algorithm Based on Relief-F

| Expressions | TP Rate | FP Rate | Precision | Recall | F-Measure | Accuracy |
|-------------|---------|---------|-----------|--------|-----------|----------|
| Anger       | 1.00    | 1.00    | 0.871     | 1.00   | 0.93      | 87.12    |
| Contempt    | 1.00    | 0.79    | 0.991     | 1.00   | 1.00      | 99.71    |
| Disgust     | 1.00    | 1.00    | 0.905     | 1.00   | 0.95      | 90.49    |
| Fear        | 1.00    | 1.00    | 0.888     | 1.00   | 0.94      | 88.80    |
| Happy       | 0.97    | 0.38    | 0.935     | 1.00   | 0.95      | 91.96    |
| Normal      | 1.00    | 1.00    | 0.777     | 1.00   | 0.87      | 77.68    |
| Sad         | 1.00    | 1.00    | 0.868     | 1.00   | 0.93      | 86.80    |
| Surprise    | 0.99    | 0.40    | 0.935     | 0.99   | 0.96      | 92.91    |
| Avg. Rate   | 0.99    | 0.82    | 0.90      | 1.00   | 0.94      | 89.43    |

Relief-F method using six features, KNN gets the highest recognition rate with 94.93%, J48 gets 92.27% of recognition rate, RBF gets the recognition rate of 89.65%, and SVM gets the lowest recognition rate with 89.43%. The accuracy performance of each classifier is shown in Table VIII, IX, X, XI respectively.

The result for Relief-F feature selection applied on eight types of face detection. The highest result for Anger gets 97.58%, Disgust gets 96.77%, Fear gets 94.25%, Happy gets 96.04%, Normal gets 85.92% and Sad gets 94.96% by using KNN classifier while Contempt gets 99.71% by using SVM classifier. Surprise emotion gets 95.06% by using J48 classifier. The best classifier for Relief-F feature selection is KNN because the highest result detected from it, just contempt and surprise detected from SVM and J48 respectively.

Based on the performance evaluation results illustrated for each classifier, the optimum classified facial emotion is Contempt with the total ratio of 99.71 percent. Simultaneously, the minimum recognition rate is normal with the total ratio of 77.68 percent.

Table 12 shows the comparison summary of the related works. From this table, it is clear that other related experiments have utilized different approaches of classification and feature selection on various datasets according to different number of facial expressions. Compared to the related works, the provided approach obtains a good recognition rate with fewer features and more recognized facial expressions. However, researchers in [10] and [14] obtained a good recognition rate ranged (77.06%-90%) but using different number of features (29, 68) respectively with different classifiers. Also, researchers in [12], and [13] could gain a high accuracy using large number of features (<120 or 120) but with the ability to recognize fewer expressions from different dataset. In the other hand research [11] depends on local binary patterns to extract the features then classify the face with accessory 94.15%. This work uses fewest features applied on Chi-square and Relief-F with four classifiers to reach excellent accuracy.
Table 12. Comparison with other Works

| Reference | Dataset | Emotion No. | Feature No. | Feature Selection | Classifier | Result |
|-----------|---------|-------------|-------------|------------------|------------|--------|
| [10]      | CK+     | 7           | AUs 29      | Correlation Coefficient and Normalized Distance | EMS        | MMD & MDA, PCA 90% |
| [11]      | self-collected, ICPR | 2           | local binary patterns | LBP, ILBP | Kullback-Leibler | 94.15% |
| [12]      | SFEW    | 7           | 40, 60, 80, 100, 120 | MSBC | Holistic, division | 96.5% 67.2% |
| [13]      | Ar and CAS PEAL-R1 | 4           | Ar<120 CAS<65 | LTP | NN | Ad<99.0% CAS 91.4% |
| [14]      | CK+     | 5           | 68          | FER | SVM, Logistic Regression, NN | SVM 88% Logistic Regression 80% NN 77.06% |
| This work | CK+     | 8           | 6           | Chi-Square Relief-F | J48, KNN, RBF, SVM | Chi-square KNN 94.18% Relief-F KNN 94.93% |

5. Conclusion
Chi-square and Relief-F are considered as two accurate feature selection approaches that are recently tackled and utilized by numerous scholars. In this study, the highest six ranked features from 784 features are selected through application of Chi-Square and Relief-F. The selected features from the both methods are applied in the four classifiers. The outcome of the experiment demonstrates KNN as the most accurate classifier amongst the four. Utilization of the four classifiers demonstrate various outcomes. KNN shows the optimum average ratio of accuracy with the total percentage of 94.18% through the application of Chi-square and 94.93% for Relief-F. RBF demonstrates the minimum accuracy ratio of 90.38% as concerned with Chi-square and support Vector Machine 89.43% for Relief-F.

References
[1] Dino, H., Abdulrazzaq, M.B., Zeebaree, S.R., Sallow, A.B., Zebari, R.R., Shukur, H.M. and Haji, L.M., Facial Expression Recognition based on Hybrid Feature Extraction Techniques with Different Classifiers. TEST Engineering & Management, 2020, 83, pp.22319-22329.
[2] Xu, Z., et al., Building the multi-modal storytelling of urban emergency events based on crowdsensing of social media analytics. Mobile Networks and Applications, 2017. 22(2): p. 218-227.
[3] Eman Zakaria, W.A.R., Abeer Twakol, Ashraf Shawky, Face Recognition using Deep Neural Network Technique. SL International Conference Giza, 2019(ISBN: 9788192958046).
[4] Abbasnejad, I. and D. Teney. A hierarchical bayesian network for face recognition using 2d and 3d facial data. in 2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP). 2015. IEEE.
[5] Li, S. and W. Deng, Deep facial expression recognition: A survey. IEEE Transactions on Affective Computing, 2020.
[6] Rajawat, A., M.K. Pandey, and S.S. Rajput. Low resolution face recognition techniques: A
survey. in 2017 3rd International Conference on Computational Intelligence & Communication Technology (CICT). 2017. IEEE.

[7] Sajjad, M., et al., Raspberry Pi assisted face recognition framework for enhanced law-enforcement services in smart cities. Future Generation Computer Systems, 2020. 108: p. 995-1007.

[8] Yang, T., et al., Evaluating facial recognition web services with adversarial and synthetic samples. Neurocomputing, 2020.

[9] Dino, H.I. and M.B. Abdulrazzaq. Facial Expression Classification Based on SVM, KNN and MLP Classifiers. in 2019 International Conference on Advanced Science and Engineering (ICOASE). 2019. IEEE.

[10] Peng, Y. and H. Yin, Facial expression analysis and expression-invariant face recognition by manifold-based synthesis. Machine Vision and Applications, 2018. 29(2): p. 263-284.

[11] Fekri-Ershad, S., Gender Classification in Human Face Images for Smart Phone Applications Based on Local Texture Information and Evaluated Kullback-Leibler Divergence Gender Classification in Human Face Images for Smart Phone Applications Based on Local Texture Information and Evaluated Kullback-Leibler Divergence. Traitement du Signal, 2019. 36(6): p. 507-514.

[12] Munir, A., et al., Illumination invariant facial expression recognition using selected merged binary patterns for real world images. Optik, 2018. 158: p. 1016-1025.

[13] Zhong, F., et al. An Effective Face Recognition Framework With Subspace Learning Based on Local Texture Patterns. in 2018 5th International Conference on Systems and Informatics (ICSAI). 2018. IEEE.

[14] Bilkhu, M.S., S. Gupta, and V.K. Srivastava, Emotion Classification from Facial Expressions Using Cascaded Regression Trees and SVM, in Computational Intelligence: Theories, Applications and Future Directions-Volume II. 2019, Springer. p. 585-594.

[15] Wei, P., et al., Research on face feature extraction based on K-mean algorithm. EURASIP Journal on Image and Video Processing, 2018. 2018(1): p. 83.

[16] Zhou, Z., et al., Earthquake Detection in 1D Time-Series Data with Feature Selection and Dictionary Learning. Seismological Research Letters, 2019. 90(2A): p. 563-572.

[17] Mahmood, M.R. and A.M. Abdulazeez. Different Model for Hand Gesture Recognition with a Novel Line Feature Extraction. in 2019 International Conference on Advanced Science and Engineering (ICOASE). 2019. IEEE.

[18] Li, P., et al. Feature selection for facial expression recognition. in 2010 2nd European Workshop on Visual Information Processing (EUVIP). 2010. IEEE.

[19] Abdulrazzaq, M.B. and K.I. Khalaf. Handwritten Numerals’ Recognition in Kurdish Language Using Double Feature Selection. in 2019 2nd International Conference on Engineering Technology and its Applications (IICETA). 2019. IEEE.

[20] Aashkaar, M., P. Sharma, and N. Garg. Performance analysis using J48 decision tree for Indian corporate world. in 2016 International Conference on Research Advances in Integrated Navigation Systems (RAINS). 2016. IEEE.

[21] Shetty, R.P., A. Sathyabhama, and A.A. Rai. Optimized radial basis function neural network model for wind power prediction. in 2016 Second International Conference on Cognitive Computing and Information Processing (CCIP). 2016. IEEE.

[22] Mahmood, M.R. and A.M. Abdulazeez. A Comparative Study of a New Hand Recognition Model Based on Line of Features and Other Techniques. in International Conference of Reliable Information and Communication Technology. 2017. Springer.

[23] Ji, C., S. Song, and H. Xia. Research and Analysis of Learning Rate on Face Expression Recognition. in 2019 International Conference on Electronic Engineering and Informatics (EEE). 2019. IEEE.

[24] Smith, F.W. and S. Rossit, Identifying and detecting facial expressions of emotion in peripheral vision. PloS one, 2018. 13(5).
[25] Kanade, T., J.F. Cohn, and Y. Tian. Comprehensive database for facial expression analysis. in Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580). 2000. IEEE.

[26] Lucey, P., et al. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. in 2010 ieee computer society conference on computer vision and pattern recognition-workshops. 2010. IEEE.

[27] Liliana, D. Emotion recognition from facial expression using deep convolutional neural network. in Journal of Physics: Conference Series. 2019. IOP Publishing.

[28] PAUL, T., et al., A Study on Face Detection Using Viola-Jones Algorithm in Various Backgrounds, Angles and Distances. International Journal of Biomedical Soft Computing and Human Sciences: the official journal of the Biomedical Fuzzy Systems Association, 2018. 23(1): p. 27-36.

[29] Mahmood, M.R., A.M. Abdulazeez, and Z. Orman. Dynamic Hand Gesture Recognition System for Kurdish Sign Language Using Two Lines of Features. in 2018 International Conference on Advanced Science and Engineering (ICOASE). 2018. IEEE.

[30] Zebari, R., et al., A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction. Journal of Applied Science and Technology Trends, 2020. 1(2): p. 56-70.

[31] Sun, J., et al. Efficient method for feature selection in text classification. in 2017 International Conference on Engineering and Technology (ICET). 2017. IEEE.

[32] Urbanowicz, R.J., et al., Relief-based feature selection: Introduction and review. Journal of biomedical informatics, 2018. 85: p. 189-203.

[33] Salih, A.A. and M.B. Abdulrazaq. Combining best features selection using three classifiers in intrusion detection system. in 2019 International Conference on Advanced Science and Engineering (ICOASE). 2019. IEEE.

[34] Bhargava, N., et al., Decision tree analysis on j48 algorithm for data mining. Proceedings of International Journal of Advanced Research in Computer Science and Software Engineering, 2013. 3(6).

[35] Rani, H.R. and A.T. Victoire, Training radial basis function networks for wind speed prediction using PSO enhanced differential search optimizer. PloS one, 2018. 13(5): p. e0196871-e0196871.

[36] Yin, Z., et al. Introduction of SVM algorithms and recent applications about fault diagnosis and other aspects. in 2015 IEEE 13th International Conference on Industrial Informatics (INDIN). 2015. IEEE.

[37] Le, T., et al. Robust support vector machine. in 2014 International Joint Conference on Neural Networks (IJCNN). 2014. IEEE.

[38] Behaine, C.A. and J. Scharcanski, Enhancing the performance of active shape models in face recognition applications. IEEE Transactions on Instrumentation and Measurement, 2012. 61(8): p. 2330-2333.

[39]