

Multiple Feature Extraction and Multiple Classifier Systems in Face Recognition

Azamossadat Nourbakhsh¹(✉)
and Mohaddeseh Mohammad Hoseinpour²

¹ Department of Computer Engineering, Technical and Engineering College,
Islamic Azad University-Lahijan Branch, Lahijan, Iran

Nourbakhsh@liau.ac.ir

² Department of Computer Engineering,
Islamic Azad University-Rasht Branch, Rasht, Iran

Abstract. Nowadays, face recognition is one of researchable issues in machine vision. In this research a new method using Multiple Feature extraction and Multiple Classifier system (MFMC) has been presented for face recognition. At First, images are gathered from Cohn–Kanade database and segmented via masks. Then image features are extracted via Local Binary Pattern (LBP), gradient histogram and masks. For classifying extracted features, Naïve Bayes, K Nearest Neighbors and Support Vector Machine classifiers are employed by using MFMC Classifier system. The results obtained show that this research significantly boosts the accuracy of face recognition in contrast with previous methods. The proposed method achieved 99.63% recognition accuracy on Cohn–Kanade database.

Keywords: Face recognition · Multiple classifier systems · Local binary patterns · Naïve Bayes · KNN and SVM classifiers

1 Introduction

1.1 A Subsection Sample

One of the appropriate methods for improving the accuracy of sample classifiers, is the use of different multiple classifiers and then combining their output results, which are often called as “Multiple Classifier Systems” or “ensembles systems” which in machine learning resources with different names such as “Classifier ensembles”, “Committees of learners”, “Mixture of experts” and “consensus theory” are called. Researchers of various areas including pattern recognition, machine learning and statistics, have investigated the use of ensembles systems. To create a more accurate classifier, combination of Multiple Classifiers is considered as one of the most popular possible approaches. Multiple Classifier Systems focus on a combination of classifiers that provide a background of mirror modeling systems for final decision making. The old classifiers, such as decision tree, Naïve Bayes (NB), Artificial Neural Network (ANN) and support vector machine (SVM), work with this assumption that the data sample is intended a reliable set of population, which means the need for balance sampling distribution.

When it is faced with the batch distribution, the traditional classifier, have often unsatisfactory performance. The application of compound classifier in some issues is so difficult that due to the lack of proper training or parameters optimization, one classifier cannot classify the data in a correct way, alone. The combination of taken decisions by multiple classifiers can statistically reduce the risk of a wrong decision. In a situation that multiple classifier, due to time or place restrictions, should be taught by different training data sets, a hybrid system can offer reasonable solution. In addition, a hybrid system can also solve uncertainty issues that regularly occurs in a single-classifier such as neural networks.

This study focuses on Multiple Feature extractors and Multiple Classifiers (MFMC), because it can show superior results compared to the multiple classifier and single classifiers that have a Feature Extractor, and against of several researches which have tried to find the manual optimization fixed combination of Feature Extractor and classifiers, so that it is attempted to optimize Feature Extractor and classifiers and their compounds. A new cumulative framework is proposed that by using hierarchical method manages the complexity of MFMC and for improving the accuracy, combines reinforcement machine learning and Bayesian network modeling.

2 Review of the Research Literature

In 1990, during a research in the field of multiple classifier systems, it was showed that the neural networks performance can be improved by using a group of neural networks with similar configuration [2]. In the same year, Schapire proved that a strong classification can be obtained by weak classifiers combination [3].

The best researches in the field of independent classifiers combination were for optimization and have allocated to introduce of bagging and boosting. These two cases are capable of producing strong classifiers in terms of approximate probability theory. According to the conducted researches, a hybrid system compared to conventional systems with only single-classifiers can be achieve to more accurate results [4].

In some researches tissue features have been used to extract properties. The local binary pattern is one of these methods that largely has been paid in the studies related to face and face recognition. For example, Franco and Maio, have offered a method for search of face images using LBP operator. In this study, face features using LBP was extracted and retrieving the face images have conducted using these features [5]. In 2009, Riaz and his colleagues could identify the human face using an approach based on the highlighted poses in face. For using the extracted features from face image sequences, the parameters of an Active Appearance Model (AAM) for the face image can be modeled and the face can be identified by using the optical flow based on the time features for estimating the facial expressions changes and classifier actions [6].

In 2015, Lu and his colleagues have proposed a face detection algorithm which it is operated based on GABOR feature combination and Sparse Representation based Classification (SRC). At first, the improved GABOR filter was applied to extract these features. Then using of Principal Component Analysis (PCA) could reduce the size of GABOR features in order to reducing the replication. Finally, also they were applied SRC for face recognition [7]. In 2016, Patil and his colleagues, unlike local binary

patterns or conductive pyramids -which a feature vector precisely is created from the location area and transformation-, proposed a method which is utilized from the spatial features as well as the Contourlet transfer domain. In particular, contourlet transfer has properties such as direction and anisotropy and thus can extract the important features. In addition, they have proposed an improved algorithm that is able to robust the system more, by improving skin area, under the contourlet sub-bands [8].

3 The Proposed Method

Using multiple extractor (Feature) and multiple classifiers of MFMC, a hierarchical framework is provided. The previous researches had focused on finding the best combination of a small number of Feature Extractors and classifiers category. In this study, an automatic optimization of the optional number Feature Extractors and classifiers is offered that can combine newly developed Feature Extractors and classifiers together. The complexity of the optimization problem grows exponentially by joining of the classifiers and new of Feature Extractors in the form of combination. This research suggests that, this problem be developed according to hierarchical method (Which is proposed in the [9]) and leads to improve the Face recognition. In Fig. 1, at first the input image is received and segmented. Then, its features are extracted and is classified by using the image structure and the use of hierarchical method reduces the complexity of MFMC and finally the face recognition is conducted (Fig. 2).

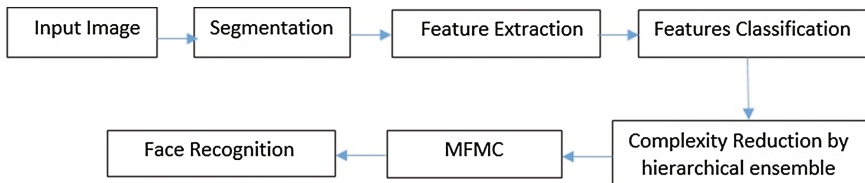


Fig. 1. The face recognition framework using ensembles systems of multiple feature extractor and multiple classifier.

3.1 Pre-processing

These images are gray scale and should be improved before the feature extraction phase implementation in terms of the image quality. In image processing, some commands are used in order to improve image quality and retrieval. An image may have noises. These noises may be created due to weakness of receiver sensor in an image. So there should be a preprocessing stage before beginning of each research.

3.2 Segmentation

The segmentation is to identify the major areas of the image and their classification to one of the background, face, hair and clothes category. The segmentation approach is operated based on a multiple classifier system which uses the color information and

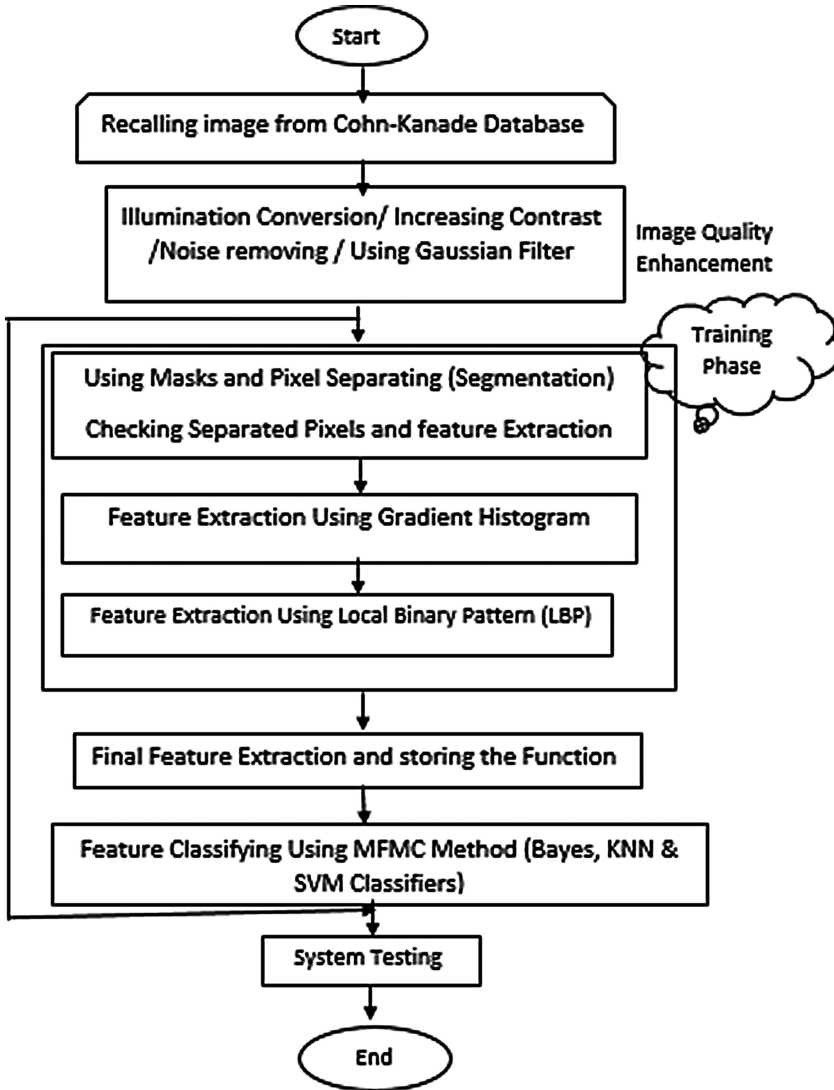


Fig. 2. The research proposed model for feature extraction of face images

texture with basic knowledge about the image structure and classifies the different areas of the image. The image segmentation process is a multiple classification system based on color, texture and spatial data for the classification different areas of the image. Then, the class labels is awarded to each area.

Some areas of the image can be classified with a reasonable amount of certainty based on the subject basic knowledge. A portrait image has a particular arrangement, which can be summarized by introduced masks for each area. This knowledge is coded by series of binary masks.

A class label based on their common degree with binary masks, is allocated to each of the areas. At first, the background is investigated and then, the study of the other three classes are paid. Each mask displays an area of the image that statistically has the most corresponding class pixels (Face, hair, background or clothes). A single area can be allocated to a class, even if its share with the mask assigned to the class, is more than 70% [10].

$$M = \{m_{\text{face}}; m_{\text{hair}}; m_{\text{background}}; m_{\text{clothes}}\} \quad (1)$$

In the basic selection process of background, a binary mask -with name of exclusive background mask- is obtained by background mask subtraction that are related to other three classes. Thus, a part of the image that are likely contain only the background areas, is identified easily. All areas of the image, which are shared with the exclusive background mask, are classified as the background.

After selecting these primary areas, if the image background is almost uniform, the background class label, according to an area growth mechanism and based on the color information and the area texture, spreads into neighboring areas and their detail is determined with zero value (Black). If the image background is not uniform, the area growth process, due to the many changes that occur even between neighboring areas, is not very impressive. In this case, low-level areas using Expectation Maximization (EM) algorithm are clustered. In this algorithm, the location and sharing low-level areas with different binary class mask is used.

After selecting multiple background areas, the areas without labels are subsequently analyzed. This areas by the other three remaining classes (Face, hair, clothes) are marked.

3.3 Classification

The area classification process is performed mainly based on a multiple classifier system. Four different classifier have been trained on the different image features. Two classifiers which use the pixel level information (Location and pixel color) and two other classifier which act based on the more general features of area level (Texture and color histogram).

Since that the classification process is performed in the area level (All pixels of an area have the same class label) a confidence value for each area is considered for classifiers based on pixel. This confidence level is calculated by averaging of each pixel belonging portion to a particular area [10].

The area labelling based on classifier is done as follows: The confidence value of compound relative class is calculated for each of the image areas and then, the class labels are assigned gradually to them. The label assign when that the highest class confidence value is more than a predefined threshold is possible. Thus, an alternative process has applied. At each stage, the class threshold reduces by a fixed decline factor. After the labelling stage, an area growth stage is implemented that allows that more areas be classified with regard to the similarity in color and texture with the labeled areas. (These areas are not selected based on the class reliability).

After classification stage, label of many low-level areas is identified. However, some areas still may be remaining without labels, due to the limit confidence threshold that have been applied. In this case, a preprocessing stage is performed that its aim is completing the labelling process. Since that the probability of error and mistake in this stage is very low, a heuristic approach is used. In particular, post-processing includes labelling operations of large areas, labelling based on vicinity and incompatibilities correction [10].

3.4 Feature Extraction

Feature Extraction is a process in which by performing an operation on the data, their clear and determinant characteristics is determined. The purpose of feature extraction is that raw data is formed with a more usable state for the next statistical processing. The feature extraction process, converts the pixel data into a higher level view of the shape, form, color, texture and space configuration from the face or its components. This extracted view is used for future classification. Generally, feature extraction reduces the input space dimensions. This dimension reduction process should maintain the essential information which contain high power of the separation capability and high stability. To do so in this paper, a local binary pattern has been used [11].

For features classification in this paper, the three classifiers has been used as follow: Naïve Bayes (NB) Classifier, Nearest Neighbor and Support Vector Machine.

- Naive Bayes Classifier is an estimation for Bayesian classifier that classifies a sample based on any given class of the sample feature variables probability. The main assumption is that the features are independent of any given class of sample [12].
- K Nearest Neighbors Classifier stores the complete training data. The new samples are classified by the main class selection among the K nearest sample in the training data [13].
- Support Vector Machine which is a binary separator, also has been used as the third classifier. In the Support Vector Machine classifier for problem complexity matching, the effective number of parameters are automatically adjusted and the solution is presented as linear combination of support patterns. These patterns, are a subset of the training patterns that are closer to the borders of the decision [14].

3.5 Hierarchical Hybrid System

As it is observed in Fig. 3, the proposed hybrid system consists of three steps:

- Making all possible couples of Feature Extractor classifier;
- Making a set of local combinations according to these couples and using reinforcement machine learning;
- Performing a final decision-making by creating a global combination based on Bayes network.

In the first stage, each Feature Extractor according to the input image, creates a feature set in the form of a vector. Feature vectors made by a Feature Extractor, has been used for training and testing by a classifier which with that certain Feature

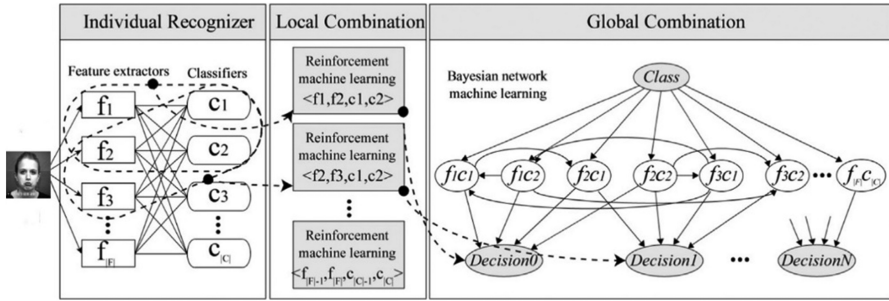


Fig. 3. The design framework for hierarchical hybrid MFMC [9].

Extractor has been paired. This stage is similar to the conventional approach which is used to create unique detectors. In the next stage, to reduce complexity, due to the exponential number of possible combinations, the hierarchical method is used. Thus, in the second stage, only a limited couple number of classifier-Feature Extractor are combined together and form a group (There may be many different combinations and each combination, forms its own group). The weight of each pair in the group is determined, according to their effectiveness and by reinforcement learning. Then, in the final stage, for final decision-making, the groups are combined together in a unit decision-making structure with name of Bayes network.

4 Simulation Results

In this study, regard to a specific set of images, the implementation of the mentioned ideas is discussed by using MATLAB software. The used input images, from Cohn-Kanade database are related to the University of Pittsburgh in 2013. This database consists of 100 people, in 18–30 years old ages, which it is included about 85% of men and 15% of women. The images are sized in 640×640 pixel. Each person has 6 poses [15].

This database includes the people face images with a specified ID. Each of the images was gray scale and after operations for improving image quality and removing noises, features are extracted. These images have a gray color intensity matrix which should be conduct the feature extraction over it. In this research, among the 600 present images, the number of 70% (420 images) are selected as training category and 30% remaining (180 images) as test category.

4.1 The First Phase - Image Recalling

In this study, the following flow process is used to improve image quality and noise elimination. The output images will be with very low noise that causes more appropriate feature extraction.

- One of the functions of intensity is the `imadjust`. By this action, the present image intensity can be changed. In this case, by finding the lowest and highest number in the images, it is going to change its range into zero to 255.

- If the image has low Contrast, for achieving a histogram with the appropriate distribution, contrast correction is needed. For this purpose, stretching contrast method for image data has been used in this research.
- For eliminating salt and pepper noise, the medfilt2 algorithm has been used. In this case, first windows with sizes of 3×3 or more are determined and by specifying the median value of each of them, the image noise value is reduced. On the other hand, the noise elimination by Gaussian method is one of the image pre-processing and its noise elimination. The use of median filter causes to remove the present stroke in the images.

This research likes to find and isolate the image components, so that the image background and the face features and so one are extracted. In this case, the masks based on research [10] has been created. These masks have been prepared in size of 300×200 which is consist of 0 and 1 numbers. The number (0) means the absence of color intensity and the number (1) means the presence of color intensity. These masks are multiplied in the original image and act as a filter which can be used for the image segmentation. Each of the images are multiplied in the relevance mask and an average of the color intensity is determined in it. These numbers indicate the index of the face image segments (Background, face, hair and clothes).

One of the database images has been selected for the output display of this part. As it is observed, the designed algorithm can separate the face details from each other well and show in the output. Each of these segments are separated and stored in the database. Then each of them is used for feature extraction. For better implementation, the images are converted into squares and then the areas masks are applied on the images. Therefore, the output image is displayed in the squared form and its size is slightly different with the original image. pixels are separated in terms of existing in different areas and after the study, required characteristics are extracted of them. In this stage, the face components have been isolated so well from the image (see Fig. 4).

For feature extraction, in addition to masks, gradient histogram and Local Binary Pattern (LBP) are also used. In each image pixel, a $W \times W$ neighbor is considered. This neighbor is called cell. In each cell a gradient histogram with n bin is calculated.

The Local Binary Pattern algorithm output is a matrix same size with the original image that it has been provided its binary properties. This matrix is extracted from the histogram which the numbers levels of this matrix is among 0 to 255. Due to the high volume of information, these numbers have converted into 10 category of numbers. Thus, category 1 means the total intensity of numbers between 0 and 25. Each bar shows the number of matrix in the presented category. So, each image has a vector of the image LBP intensity that causes the differences between the images.

4.2 The Second Phase - Classifiers Implementation

According to the proposed model (Fig. 2), the improved images after the feature extraction should be placed in three classifiers of simple Bayesian, nearest neighbors and support vector machine and be tested. In this study, features are selected with better perspective by taking the variance of each characteristic in each class. If the intended feature variance is zero, this feature is removed from the database.

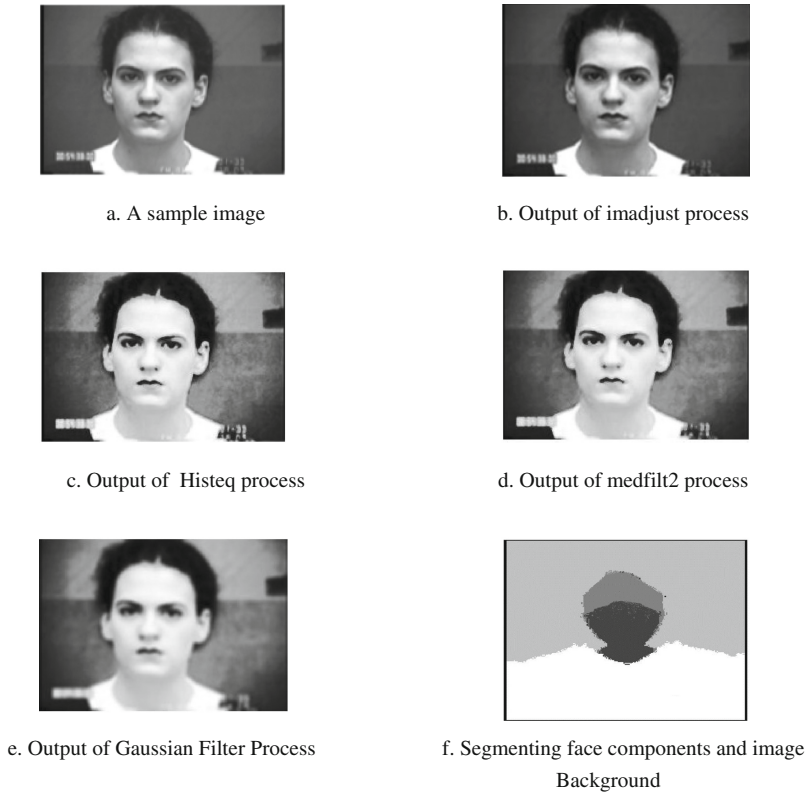


Fig. 4. Each of the implemented steps output during the first phase, respectively (a to f steps).

4.3 The Third Phase - Features Investigation

After passing stages of image quality improve and noise elimination, each of the face images have been classified. An ID is allocated to each of these images. For every person who has several images, one ID has been considered. The considered database is created as input variables X and Class Y . In fact, the aim is finding features that can identify the relevant class with higher accuracy.

The research data are divided randomly into two parts of training and testing, that shows the ratio of correct response of the designed system to be compared against the inputs. About 70% of the images as training data and 30% of the images as the test data have been selected. After training the data by amount of 70%, the error value of this system is calculated by 30% of the remaining data.

For evaluating the obtained results of the present research and its comparison with the proposed tests in the research [10], the obtained error value with respect to the two parts images of men and women -according to Table 1- shows almost the same results (Fig. 5).

Table 1. Accuracy rate determination about implementation of face detection components.

Research	Face		Hair		Clothes		Backgrounds		Final error
	FP	FN	FP	FN	FP	FN	FP	FN	
Ferrara et al. [10]									
Men	2.44	2.13	8.15	10.39	4.17	6.43	1.49	0.86	2.41
Women	5.52	4.28	7.19	19.34	13.37	4.03	1.49	1.90	4.38
Current research									
Men	3.14	1.82	10.17	9.10	3.56	5.62	2.27	2.39	1.48
Women	8.63	3.20	10.30	12.39	3.89	5.67	2.10	2.40	3.34

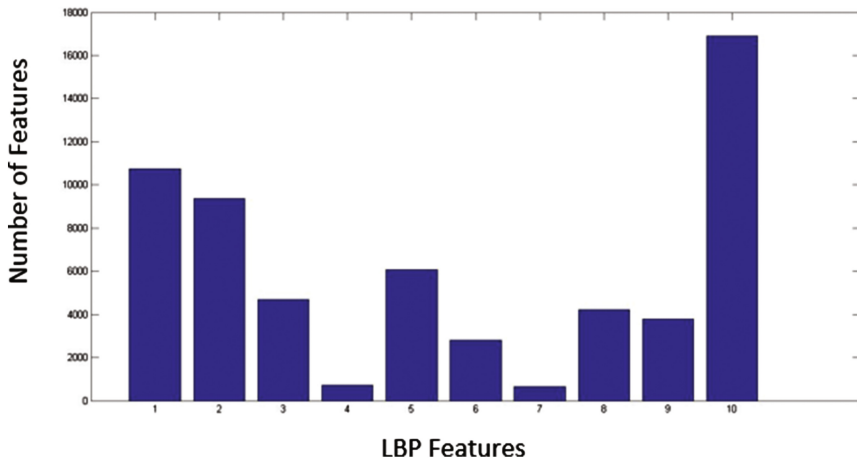


Fig. 5. Chart of LBP feature extraction process.

After reviewing and implementing several tests, it was shown that by selecting the face mask features, less than 8% error can be achieved. In fact, this error is because of low number of each person images and the more number of classes. Figure 6 has been plotted according to information of Table 2. This bar chart shows that the proposed

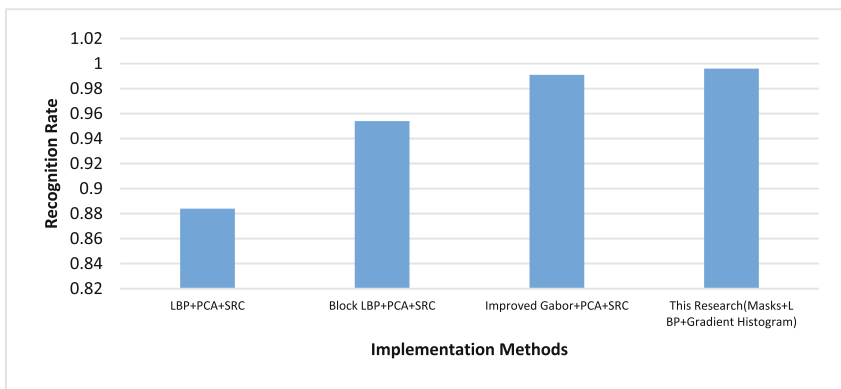


Fig. 6. Comparison chart of the face recognition algorithms

Table 2. Recognition rate of different features.

Row	Algorithm	Recognition rate
1	LBP+PCA+SRC [7]	88.40
2	Block LBP+PCA+SRC [7]	95.48
3	Gabor+PCA+SRC [7]	99.12
4	This research (Mask+LBP+Gradient Histogram)	99.63

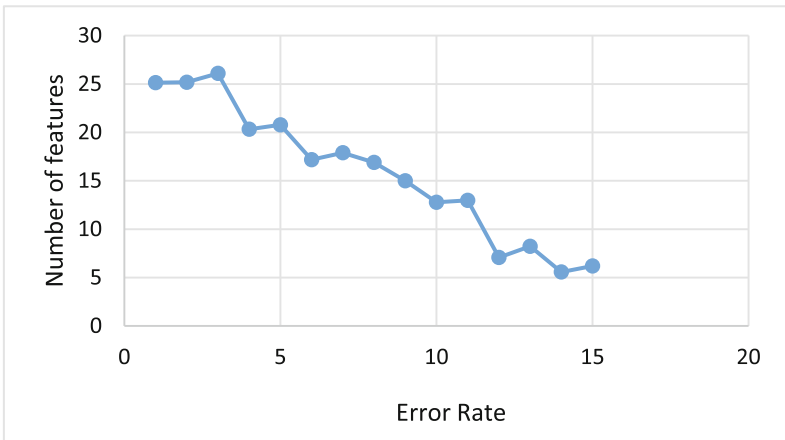
method compared to “LBP+PCA+SRC method”, causes about 12% improvement in the images recognition. Moreover, this proposed method compared to the second and third method shows 4% and 1% better performance, respectively. Based on the performed observations, it can be said that the proposed method of this study show better performance. In addition to, the proposed approach compared to the specified algorithm in Table 3, reached to better recognition rate in Cohn–Kanade database.

Table 3. Comparison of recognition rate on Cohn–canade database.

Row	Algorithm	Recognition rate
1	AAM-BDT [8]	91.17
2	AAM-BN [8]	98.69
3	LBP+Contourlet Transform [8]	99.56
4	This Research (Mask+LBP+Gradient Histogram)	99.63

Number of the recognized correct images has been calculated compared to the total images. So these methods can be compared together. As it is obvious, the proposed algorithm -due to the existence of a stage selection of most optimize indicators- achieved a better output recognition.

In this study, error rate of the research has been paid by different results with different numbers. In this case, if 14 number of indexes have been selected, the obtained error value occurs in the lowest amount. In this case, the obtained error is about 5.58% (see Fig. 7).

**Fig. 7.** Number of features and their error rate

5 Conclusion

According to the proposed method, first the input image was received from Cohn–Kanade database. Images are involved of noises or are in trouble in terms of brightness intensity, contrast, etc.; so the stages of image quality improvement were applied on the image. Then, the images were segmented by face, hair, clothes and background masks. Features of image were extracted by Local Binary Pattern algorithm, gradient histogram and masks. Then, the extracted features were classified with three Bayesian, KNN and SVM classifiers. The proposed method of this research has reached to 99.63% of recognition rate. Also, based on MFMC method, using 14 features had the best result.

References

1. Díez-Pastor, J.F., Rodríguez, J.J., García-Osorio, C.: Random balance: ensembles of variable priors classifiers for imbalanced data. *Knowl. Based Syst.* **85**, 96–111 (2015)
2. Hansen, L.K., Salamon, P.: Neural network ensembles. *IEEE Trans. Pattern Anal. Mach. Intell.* **12**, 993–1001 (1990)
3. Schapire, R.E.: The strength of weak learnability. *Mach. Learn.* **5**, 197–227 (1990)
4. Wozniak, M., Grana, M., Corchado, E.: A survey of multiple classifier systems as hybrid system, pp. 9–11 (2014)
5. Franco, A., Maio, D.: Similarity searches in face databases. In: *International Conference on Image Analysis and Processing, Vietri sul Mare*, pp. 443–450 (2009)
6. Riaz, Z., Mayer, C.H., Wimmer, M., Beetz, M., Radig, B.: A model based approach for expressions invariant face recognition. In: *Lecture Notes in Computer Science*, vol. 5558, pp. 289–298. Springer (2009)
7. Lu, X., Kong, L., Liu, M., Zhang, X.: Facial expression recognition based on gabor feature and SRC. In: *CCBR 2015. LNCS*, vol. 9428, pp. 416–422. Springer, Heidelberg (2015)
8. Patil, H.Y., Kothari, A.G., Bhurchandi, K.M.: Expression invariant face recognition using local binary patterns and contourlet transform, 2670–2678 (2016)
9. Kim, K., Lin, H., Choi, J.Y., Choi, K.: A design framework for hierarchical ensemble of multiple feature extractors and multiple classifiers, 1–16 (2016)
10. Ferrara, M., Franco, A., Maio, D.: A multi-classifier approach to face image segmentation for travel documents, 8452–8466 (2012)
11. Abdur, R., Najmul, H., Tanzillah, W., Shafiul, A.: Face recognition using local binary patterns (LBP). *Glob. J. Comput. Sci. Technol.* **13**(4) (2013). Version 1.0
12. Langley, P., Iba, W., Thompson, K.: An analysis of bayesian classifiers. In: *Proceedings of the Tenth National Conference on Artificial Intelligence*, pp. 223–228. AAAI Press and MIT Press (1992)
13. Duda, R.O., Hart, P.E.: *Pattern Classification and Scene Analysis*. Wiley, Hoboken (1973)
14. Boser, B.E., Guyon, I.M., Vapnik, V.N.: A training algorithm for optimal margin classifier. In: *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, Pittsburgh, pp. 144–152 (1992)
15. Lucey, P., Cohn, J.F., Kanade, T., Saragih, J., Ambadar, Z., Matthews, I.: The extended Cohn–Kande dataset (CK+): a complete facial expression dataset for action unit and emotion-specified expression. In: *The Third IEEE Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB)* (2010)