International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 3 Issue 5, August 2019 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Face Recognition for Human Identification using BRISK Feature and Normal Distribution Model

Khin Mar Thi

University of Computer Studies, Shan, Lashio, Myanmar

***How to cite this paper**:* Khin Mar Thi "Face Recognition for Human Identification using BRISK Feature and Normal Distribution Model" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-3 | Issue-5, August 2019, pp.1139-1143, **https://doi.org/10.31142/ijtsrd26589**

IJTSRD26589

ABSTRACT
Face recognition is a kind of automatic human identification from face images has been performed widely research in image processing and machine learning. Face image, facial information of the person is presented and unique information for each person even two-person possessed the same face. We propose a methodology for automatic human classification based on Binary Robust Invariant Scalable Keypoints (BRISK) feature of face images and the normal distribution model. In our proposed methodology, the normal distribution model is used to represent the statistical information of face image as a global feature. The human name is the output of the system according to the input face image. Our proposed feature is applied with Artificial Neural Networks to recognize face for human identification. The proposed feature is extracted from the face image of "the Extended Yale Face Database B" to perform human identification and highlight the properties of the proposed feature.

***KEYWORDS:** Face recognition; human identification; face images; Binary Robust Invariant Scalable Key points (BRISK); Normal distribution model; Artificial Neural Networks (ANN); The Extended Yale Face Database B*

1. INTRODUCTION
Face recognition is one of the important approaches for human identification because it is one of the most successful applications of image analysis and understanding.

Face recognition technology (FRT) has a range of prospective applications in information security, law compliance and monitoring, smart cards, authentication, and others, as one of the few biometric approaches that hold the advantages of both elevated accuracy and small intrusion. For this purpose, both scholarly and manufacturing groups have gained significantly enhanced exposure from FRT over the previous 20 years. Recently, several writers studied and assessed the present FRTs from various aspects. Binary Robust Invariant Scalable Keypoints (BRISK) feature is the keypoints based approach for object matching and scene matching. The BRISK feature contains a feature vector for each keypoint in an image. Due to the descriptor's binary existence, the BRISK keypoints can be combined very efficiently. BRISK also takes advantage of the velocity benefits provided in the SSE instruction set, which is commonly endorsed on today's architectures, with a powerful concentrate on computation efficiency. In a small number of variables, a strong distribution model must be accurate but also capable of describing the characteristics of the typical image. In this face recognition approach, it is proposed to model images for BRISK descriptors using the Normal distribution to extract the feature of the facial image. And this proposed feature is applied in Artificial Neural Network (ANN) and validation is performed on Yale face database B.

Most of the face recognition systems have been developed and propose many feature and methodology. Since Convolutional Neural Networks (CNNs) had taken the computer vision community by storm, deep face recognition was proposed to tradeoff between data purity and time. LFW and YTF face benchmark databases were used to show the proposed methodology can handle a large amount of image data with a high recognition rate [1]. FaceNet scheme was proposed to learn a mapping from facial images to a compact Euclidean distance where dimensions straight relate to a metric of facial resemblance [2]. A deep convolutional network was trained to straight optimize the embedding itself rather than an intermediate bottleneck layer as in past deep learning approaches. For training, the triplets of approximately aligned matching was used that non-matching face patches produced using a novel online triplet mining technique to train.

A new supervision signal was proposed and it is called center loss, for the face recognition task. The cluster loss concurrently learns a center for each class ' profound characteristics and penalizes the gaps between the deep characteristics and their respective class centers. In the train of CNNs, softmax loss and center loss were used to train robust CNNs to obtain the deep features with the two key learning objectives, inter-class dispersion and intra-class compactness as much as possible, which are very essential to face recognition [3]. A fresh video-based classification technique intended to reduce the necessary storage space of information samples and speed up the sampling method in large-scale face recognition systems. The image sets gathered from recordings were approximated with kernelized convex hulls in their suggested technique and it has been shown that it is adequate to use only the samples involved in modeling the image establish boundaries in this setting. The kernelized Support Vector Data Description (SVDD) is used to obtain significant samples that shape the limits of the picture set. A binary hierarchical decision tree method was suggested to boot the classification accuracy level [4].

The extraction of geometric and appearance feature was proposed to identifier age and gender. In their feature extraction approach, cumulative benchmark approach was used. For gender clustering, both of supervise and unsupervised approaches were used. While supervised machine learning approach was used for gender classification. To compare the performance of the classifier with the proposed feature, SVM, neural network, and adobos were used [5]. An extended kernel discriminant analysis framework for Face Recognition is proposed based on Image Set (FRIS) to overcome the problem of FRIS. To handle the underlying non-linearity in data storage, an image set from the original input space is mapped into model space and described with Support Vector Domain Description (SVDD). In model space, most of the mapped data is contained in a hyper-sphere and the outliers are outside the hyper-sphere. By researching an efficient information metric in model space [6], a kernel function moves information from model space to a high-dimensional feature speed.

According to these related work, several features, models, machine learning algorithms and deep learning approaches are proposed for face recognition. In our proposed methodology, the key points of the face image are extracted from the face image by BRISK keypoints generation algorithm and then we derive the statistical values from these extracted keypoints by normal distribution model. Our proposed feature is derived from the combination of key points based feature and probability distribution model called BRISK and normal distribution. The extracted features Yale face database B are trained by Artificial Neural Network (ANN). The 10 Fold-cross validations are used for classifier performance to show the advantages of proposed features.

There are four main sections of our paper. Introduction and related works are presented in section 1. The proposed methodology is presented in section 2. Experimental results and dataset are described in section 3 and the conclusion is presented in the final section.

2. Proposed Methodology

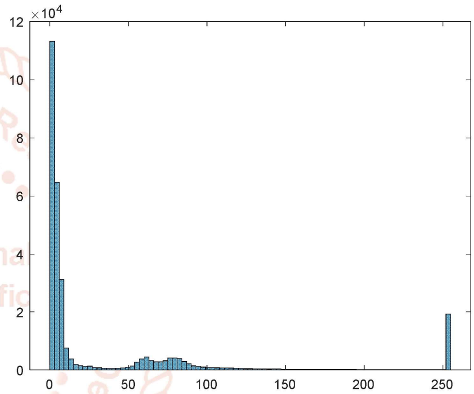
In the proposed methodology, there are three main steps: pre-processing, feature extraction and recognition. Among these steps, feature extraction is the main contribution of this paper. In preprocessing, the output is the pre-processed image for the input face image. The proposed features are extracted from the pre-processed image and the Artificial Neural Network is trained by using extracted proposed features.

This section presents statistical information of BRISK feature and how normal distribution fits with BRISK keypoints by measuring Goodness of Fitting (GOF) test. The overview of the BRISK feature and Normal distribution are also presented in this section. The input to our proposed feature extraction is the face image and the output is the name of the person.

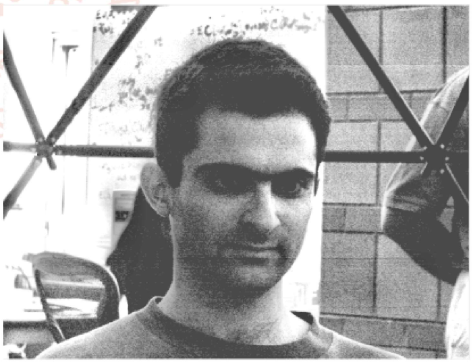
2.1. Image Pre-processing

Image enhancement is performed to highlight the different parts of the face in an image. Histogram equalization is performed to enhance the contrast of images by transforming the values in an intensity image so that the histogram of the output image approximately matches a specified histogram. The histogram equalization result of the face image is shown in Figure 1.

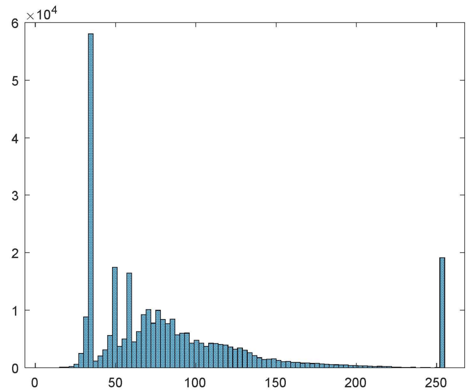
(i). Input image (yaleB11_P08_Ambient.pgm)



(ii) Histogram of input image



(iii) Enhanced Image



(iv) Histogram of enhanced Image

Fig.1. Histogram equalization of the face image

In the histogram equalization process, the value of contrast enhancement limit is 0.05 and creates a bell-shaped histogram of an input image for a more enhanced image.

2.2. Binary Robust Invariant Scalable Keypoints (BRISK)Feature

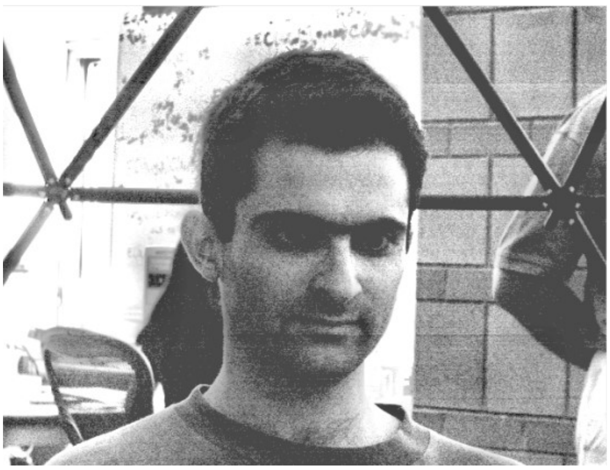
The intrinsic difficulty in extracting appropriate features from an image resides in balancing two conflicting objectives: high-quality description and low computing demands. In 2011, Leutenegger, Stefan, Margarita Chli, and Roland Siegwart proposed BRISK methodology. To achieve robustness and low computational cost for image feature extraction. Among keypoints generation methods, BRISK achieves the comparable quality of matching at much less computation time. There are two main steps in BRISK methodology: keypoints detection and keypoints description. The keypoints detection step consists of [7]:

- Generate scale space
- Calculate FAST score using scale space.
- Pixel level non-maximal suppression.
- Calculate sub-pixel maximum across patch.
- Calculate continuous maximum across scales.
- Re-interpolate image coordinates from scale space feature point detection.

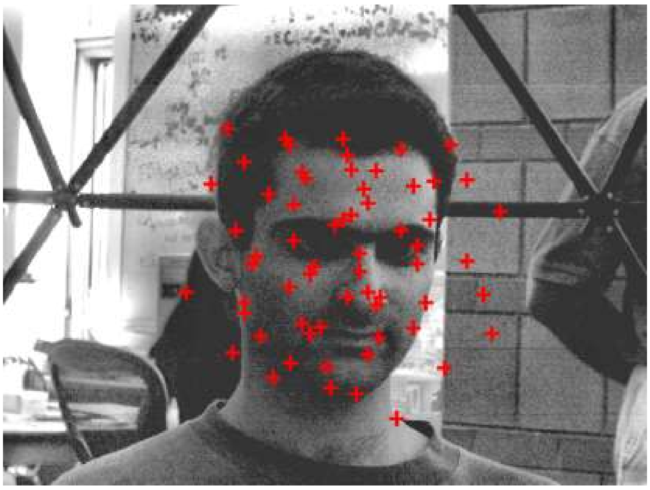
Given a set of keypoints (consisting of sub-pixel refined image locations and associated floating-point scale values), the BRISK descriptor is composed as a binary string by concatenating the results of simple brightness comparison tests. The keypoints description step consists of [9]:

- Sample pattern of smoothed pixels around the feature.
- Generate short-distance pairs and long-distance pairs for pairs of pixels
- Calculate the local gradient between long-distance pairs.
- Calculate total gradients to determine feature orientation.
- Rotate short-distance pairs using orientation.
- Generate binary descriptor from rotated short-distance pairs.

While the Speed up Robust (SURF) descriptor is also assembled via brightness comparisons, BRISK has some fundamental differences apart from the obvious pre-scaling and pre-rotation of the sampling pattern. The BRISK descriptor has Rotation invariant and scale-invariant. The BRISK feature extraction is shown in Figure 2.



(i) Enhanced Image



(ii) BRISK key points of an image

Fig.2. BRISK key points of the face image

In Figure 2, there are 650 key points for the enhance Image and the BRISK feature descriptor consists of 650x128 structure of a matrix for image information.

2.3. Normal Distribution Model

In the domain of statistics, the normal probability distribution is really prevalent. When the height, weight, wage, views or votes of people are measured, the resulting graph is almost always a normal curve. The normal distribution applies to a broad spectrum of events and is the most commonly used distribution in statistics. It was initially created as an estimate of the binomial distribution once the amount of tests is big and the Bernoulli probability p is not near to 0 or 1. It is also the exponential type of the total value of random variables under a large variety of circumstances. The normal distribution was first defined in 1733 by the French mathematician De Moivre. The growth of the allocation is most often attributed to Gauss, who introduced the concept to the motions of celestial bodies [8]. The probability density function of Normal distribution is:

$$f(x) = \frac{\sigma(2\pi)^{1/2}}{e^{\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)}} \qquad (1)$$

where f(x) is the distribution of x value, σ is the standard deviation and μ is the mean. Thus for the normal distribution the mean, μ, is a location parameter (the locating point is the midpoint of the range) and the standard deviation, σ, is a scale parameter. The normal distribution does shape parameter. In a normal distribution model, the method of moments is used to estimate the two parameters of its distribution.

$$(\hat{\mu}) = \frac{1}{n}\sum_{k=1}^{n} E(X_k) \qquad (2)$$

$$E(\widehat{\sigma^2}) = \frac{1}{n}\left(1-\frac{1}{n}\right)\sum_{k=1}^{n} E(X_k^2) - \frac{2}{n^2}\sum_{k=2}^{n}\sum_{m=1}^{k} E(X_k X_m) \qquad (3)$$

where $E(\hat{\mu})$ is the estimated mean value, $E(\widehat{\sigma^2})$ is the estimated variance value, X is the extracted BRISK feature values and n is the total number of values in feature BRISK. After getting estimated mean and variance values, standard deviation, kurtosis, skewness and root mean square are derived using these estimated values.

2.4. Proposed Feature Extraction
The BRISK feature is extracted from the preprocessed face image. The numerical representation of the BRISK feature is matrix structure and difficult to handle in classification. To directly represent the information of face keypoints, the extracted BRISK feature is model by Normal distribution model. The flow of the proposed feature extraction is shown in Figure 3.

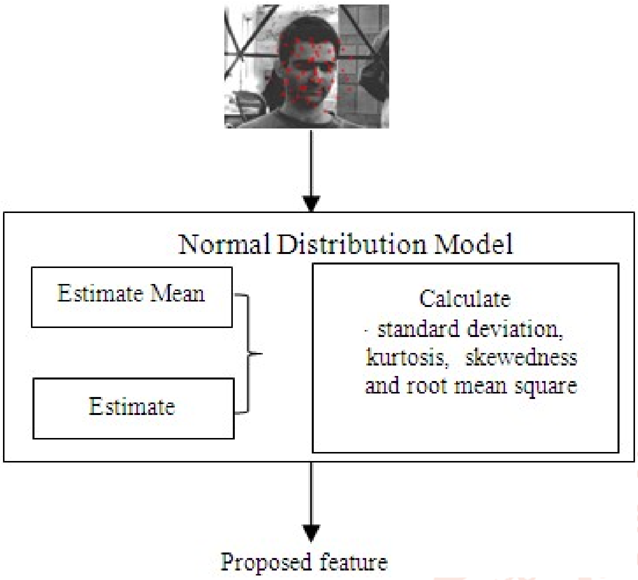


Fig.3. The flow of the proposed feature extraction

2.5. Artificial Neural Network (ANN)
Artificial neural networks are the modeling of the human brain with the simplest definition and building blocks are neurons. In multi-layer artificial neural networks, there are also neurons placed in a similar manner to the human brain. Each neuron is connected to other neurons with certain coefficients. During training, information is distributed to these connection points so that the network is learned. A neural network consists of three layers: an input layer, an intermediate layer and an output layer as shown in Figure 4.

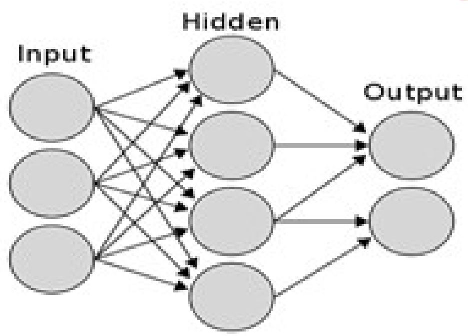


Fig.4. Three Layers of Artificial Neural Network

In our proposed methodology, Artificial Neural Network is used for training because it adapts to unknown situations, it can model complex functions and ease of use, learns by example, and very little user domain-specific expertise needed.

3. Experimental Results
In this section, we perform an experiment to show the advantages of the proposed feature. "The Extended Yale Face Database B" is used and measure classifier performance by True Positive Rate (TPR) and False Negative Rate (FNR).

3.1. The Extended Yale Face Database B
The Extended Yale Face Database B contains 5760 single light source images of 10 subjects each seen under 576 viewing conditions (9 poses x 64 illumination conditions). For every subject in a particular pose, an image with ambient (background) illumination was also captured [10].

3.2. Experiment
In our experiment, classifier performance is measured for each pose of 10 subjects in the structure of 10-fold cross-validation. The setting of a neural network used in this experiment is shown in Figure 5.

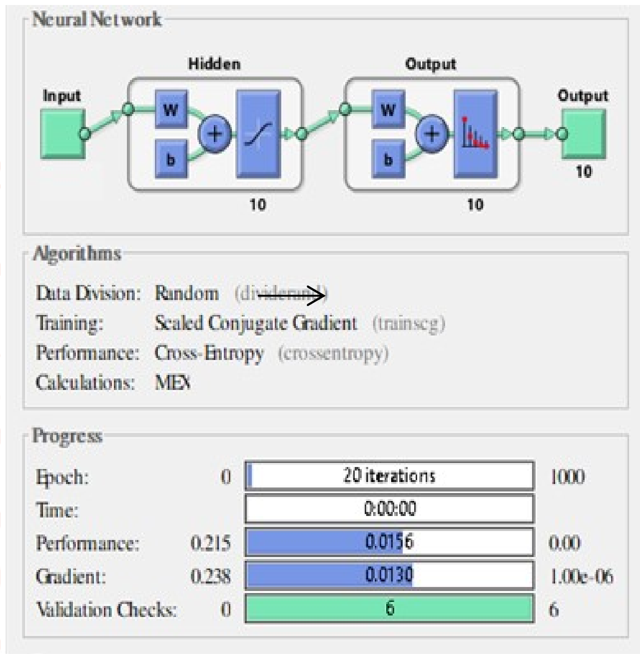


Fig.5. Training Artificial Neural Network

In this figure, the number of hidden layers is 10, training type is Scaled Conjugate Gradient and maximum epoch is 1000. According to the 10-fold cross-validation structure, the dataset is divided into 10 groups, nine groups are used as training and the remaining one is used as testing for each validation time. The validations are performed for 10 times and calculate average classification accuracy, average true positive rate and average false-negative rate as shown in table1.

3.3. Results and Discussion
Our experiment is performed in the structure of 10-fold cross-validation to show much more sincere information about our proposed feature and Artificial Neural Network. Although our average classification accuracy reaches 81.6%, the classification accuracy is 62.4% because of weak training data. The true positive rate reaches 80.7% but the true positive rate of validation6 is 63.0%. The average classification accuracy and average true positive rate are acceptable and reasonable to apply our proposed feature in face recognition with an Artificial Neural Network.

Table 1: Classification Accuracy, True Positive Rate and False Negative Rate over 10-fold cross-validation with "the Extended Yale face Database B"

Validation Times	Training	Testing	Classification Accuracy	True Positive Rate	False Negative Rate
1	5184	576	85.7	84.2	15.8
2	5184	576	79.5	78.4	21.6
3	5184	576	84.2	82.2	17.8
4	5184	576	88.7	88.2	11.8
5	5184	576	82.5	80.2	19.8
6	5184	576	62.4	63.0	37.0
7	5184	576	82.7	83.0	17.0
8	5184	576	88.4	88.2	11.8
9	5184	576	82.9	81.3	18.7
10	5184	576	78.6	77.9	22.1
Average			81.6	80.7	19.3

4. Conclusion

Extracted features are the key to achieving a greater classification efficiency in the automatic face recognition system. And its classification accuracy also relies on generating code books or extracting global features. The BRISK function is modeled on the normal distribution model to resolve global feature generation issues. In our recommended methodology, BRISK feature statistics is used explicitly instead of global feature generation. In addition, the pre-processing of this paper used histogram equalization with a bell-shaped histogram to enhance the input image. After preprocessing model-based statistical values are calculated to represent the facial information of an image. Then, these extracted features are applied in the Artificial Neural Network classifier training and testing. The efficiency of the classifier is evaluated to demonstrate the usefulness of the proposed feature in face recognition. Although our proposed feature has the appropriate classification accuracy, other function and image processing methods need to consider booting the classification accuracy and being implemented in real automatic face recognition mechanism.

Acknowledgments:

The face image dataset used in this paper is supported by "the Extended Yale Face Database B". Specially thank for allowance of free to use the extended Yale Face Database B for research purposes.

References

[1] Parkhi, O. M., Vedaldi, A. and Zisserman, A., 2015, September. Deep face recognition. In bmvc (Vol. 1, No. 3, p. 6).

[2] Schroff, F., Kalenichenko, D. and Philbin, J., 2015. Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 815-823).

[3] Wen, Y., Zhang, K., Li, Z. and Qiao, Y., 2016, October. A discriminative feature learning approach for deep face recognition. In European conference on computer vision (pp. 499-515). Springer, Cham.

[4] Cevikalp, H., Yavuz, H.S. and Triggs, B., 2019. Face Recognition Based on Videos by Using Convex Hulls. IEEE Transactions on Circuits and Systems for Video Technology.

[5] Verma, V. K., Srivastava, S., Jain, T. and Jain, A., 2019. Local Invariant Feature-Based Gender Recognition from Facial Images. In Soft Computing for Problem Solving (pp. 869-878). Springer, Singapore.

[6] Zeng, Q. S., Huang, X. Y., Xiang, X. H. and He, J., 2019. Kernel Analysis based on SVDD for Face Recognition from Image Set. Journal of Intelligent & Fuzzy Systems, 36(6), pp.5499-5511.

[7] Leutenegger, S., Chli, M. and Siegwart, R., 2011. BRISK: Binary robust invariant scalable keypoints. In 2011 IEEE international conference on computer vision (ICCV) (pp. 2548-2555). Ieee.

[8] Forbes, C., Evans, M., Hastings, N. and Peacock, B., 2011. Statistical distributions, Chapter 33, Normal (Gaussian) Distribution, pp.143-148. John Wiley & Sons.

[9] Aragon, M.C., Castillo, R., Agustin, J. and Aguilar, I.B., 2019, March. Utilization of Feature Detector Algorithms in a Mobile Signature Detector Application. In Proceedings of the 2019 2nd International Conference on Information Science and Systems (pp. 49-53). ACM.

[10] Georghiades, A. S., Belhumeur, P. N. and Kriegman, D.J., 2001. From few too many: Illumination cone models for face recognition under variable lighting and pose. IEEE Transactions on Pattern Analysis & Machine Intelligence, (6), pp.643-660.