International Journal of Engineering Research and General Science Volume 4, Issue 2, March- April, 2016
ISSN 2091-2730

Fingerprint Recognition Improvement Using Histogram Equalization and Compression Methods

Nawaf Hazim Barnouti

Baghdad, Iraq

E-mail-nawafhazim1987@gmail.com, nawafhazim1987@yahoo.com

Abstract— Biometrics refers to metrics related to human characteristic. Biometric systems work on physiological and behavioral biometric data to recognize people. Fingerprints are common biometric that have been used to identifying people for many years. Fingerprint identification also known as hand print is the process of comparing two instances of friction ridge skin impressions from human fingers or toes to determine whether these impressions have come from the same person. FVC2000 fingerprint database is used in this work. Pre-processing methods are applied on the images database to increase the recognition rate. Histogram equalization is used and shows improvement in result. Compression methods are used to decrease the redundancy of the image and also to store or transfer date in an efficient structure. PCA is used for feature extraction and dimension reduction. PCA is popular and will generate training set of images database. City block distance classifier is the final step of fingerprint recognition.

Keywords— Fingerprints, FVC2000, TIFF, HE, LZW, PCA, City block, Euclidean distance, Recognition rate

INTRODUCTION

Fingerprints is an impression in its narrow sense left by the friction ridges and actually widely examined from several points of view. Scientific studies on their embryogenesis can be found. Several papers have been written on the inheritance of specific fingerprint features and they include statistically linked to all kinds of common human features (gender) and some more obscure ones (sexual orientation, high blood pressure) [1]. Finger scan technology is deployed biometric with a number of different vendors offering a wide range of solutions. Fingerprints are deposited on a good suitable surface such as metal or glass. Biometric systems work on physiological and behavioral biometric data to recognize a person. The behavioral biometric parameters (signature, speech, gait and keystroke) change with environment and age. Although physiological features such as fingerprint, palm print, face and iris will still unchanged during the life time of a person. The biometric system work as identification mode or verification mode based on the need of an application. The verification mode validates a person identity by comparing taken biometric data with pre made template databases. The identification mode also known as hand print identification recognize a person identity by performing matches against several fingerprints biometric templates.

Fingerprints are commonly applied in daily life for more than 100 years because of its feasibility, reliability, distinctiveness, permanence, accuracy, and acceptability. Fingerprint is a pattern of ridges, furrows and minutiae that are taken using inked impression on a paper or sensor. Fingerprint sensor is an electronic device used to capture a digital image of a fingerprint pattern. High quality fingerprint included 25 to 80 minutiae based on sensor resolution and finger position on the sensor. The false minutiae are the false ridges breaks because of insufficient amount of ink and cross connections as a result of over inking. It is not easy to extract reliably minutia from poor quality fingerprint impressions arising from very dry fingers and fingers mutilated by scars, scratches due to accidents, injuries. Minutia based fingerprint recognition includes Thinning, Minutiae matching, Minutiae extraction and Computing matching score [2].

Fingerprint image obtaining is without question the most important and critical step in an automated fingerprint authentication system, because it determines the final fingerprint image quality which has a drastic effect on overall system performance. Captures the digital image of a fingerprint pattern known as live scan and it can be done by placing the finger on the surface of a fingerprint reader as shown in figure 1. This live scan is digitally done to generate a biometric template (a collection of extracted features) that is collected and used for matching process. The commonly used fingerprint sensor technologies are optical, ultrasonic and capacitive. Siemens ID Mouse comes under the capacitive type of fingerprint sensors. Capacitance sensors use the principles related to capacitance in order to form fingerprint images [3].

Figure 1: Fingerprint Readers.

FVC2000 FINGERPRINT DATABASE

The Fingerprint Verification Competition (FVC) is an international competition involved in fingerprint verification software examination. A subset of fingerprint impressions obtained with different sensors was presented by registered people to help them adjust the parameters of their methods. People were requested to provide registration and match executable files of their methods. The analysis was executed at the organizers facilities applying the submitted executable files on database, obtained with the same sensors as the training set. Table 1 shows the database most important information.

Table 1: FVC2000 Information.

Call for participation	November, 1999
Submission Deadline	June 1st, 2000
Registration Deadline	March 1st, 2000
Evaluation Period	July–August, 2000
Notes	Anonymous part. not allowed
Registered Participants	25 (15 withdrew)
Results Presentation	Barcelona, September 2000
Website	http://bias.csr.unibo.it/fvc2000
HW/SW Implementation	Pentium III (450 MHz) FVC Test suite v1.0 Windows NT

FVC2000 competition attempted to set up the first common benchmark, helping academic institutions and companies to unambiguously evaluate overall performance and monitoring improvements in their fingerprint recognition algorithms. In FVC2000, the "universal" sensor is in fact a collection of four various sensors to more beneficial handle the latest advances in fingerprint sensing techniques. Databases 1 and 2 was collected by applying two small size and low cost sensors. Database 3 was collected by applying a high quality optical sensor. While, images in Database 4 were synthetically generated. Each database was divided into a sequestered "test" set of 800 images (set A) and an open "training" set of 80 images provided to people for method tuning (set B). The four different databases (DB1, DB2, DB3 and DB4) was collected using different sensors technology [4]. Table 2 shows the features of these databases.

Table 2: FVC2000 (DB1, DB2, DB3 and DB4) Features.

#	Image	Set A	Set B	Resolution

	Size	(w*h)	(w*h)	
DB1	300x300	100*8	10*8	500 dpi
DB2	256x364	100*8	10*8	500 dpi
DB3	448x478	100*8	10*8	500 dpi
DB4	240x320	100x8	10x8	About 500 dpi

IMAGE FORMAT

TIFF (Tagged Image File Format) is a file format for mainly strong images, as well as photographs and line art. It is probably the most common and flexible of the present public domain raster file formats. Originally created by the company Aldus, jointly with Microsoft, for use with PostScript printing. TIFF is a common format for high color level images, along with JPEG and PNG. TIFF format is generally supported by image manipulation applications, and by faxing, scanning, optical character recognition, and many other applications. TIFF is a complex format that includes a very wide range of choices. Even if this makes it useful as a general format for interchange between professional image editing applications, it makes helping it in additional application include web browsers difficult [10]. Figure 2 shows samples taken from DB1 and DB3.

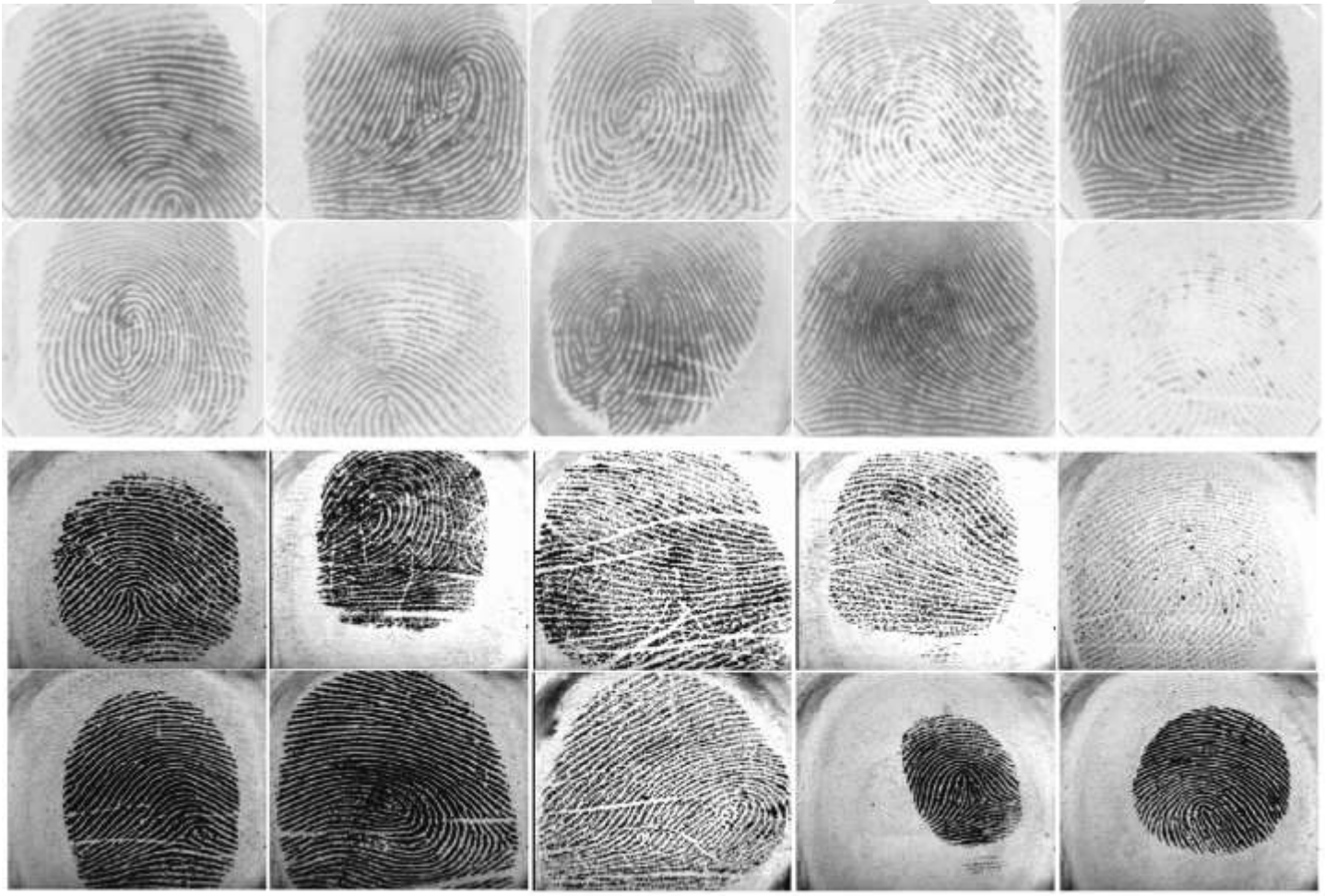

Figure 2: Sample Images Taken from DB1 and DB3.

IMAGE PRE-PROCESSING

Image enhancement improvement is a basic problem in computer vision and image processing. It's the process of modifying the pixel's intensity of the input image by using mathematical operations to generate better output images. Many methods have been developed in image processing during the last decades. It's nearly become a pre-processing part of many areas like medical image processing and as a preprocessing part of texture synthesis, speech recognition, and a lot of additional image and video processing applications [6]. Contrast is created by the difference in luminance reflectance from two adjacent areas. Enhancement methods mainly fall under two basic categories spatial and frequency domain methods. Special domain methods are usually more popular than the frequency based methods, as they are developed on direct manipulation of pixels in image. Special domain methods have been developed for visualizing the impact effect. Some of these methods makes use of simple linear or nonlinear intensity level transformation functions, while others use complex analysis of different image features that include edge and connected component information [7].

1. HISTOGRAM EQUALIZATION

Contrast enhancement improvement issue in digital images was resolved by using different methods, but Histogram Equalization (HE) method is the most popular. Histogram Equalization method flattens the histogram and extends the dynamic selection of intensity values by applying the cumulative density operation. Even so, you will discover major draw backs in Histogram Equalization especially when executed to process digital images. First of all, HE transforms the histogram of the input image into a uniform histogram by importing the whole selection of gray levels uniformly over the histogram of an image, with a mean value that could be in the center of gray levels range. Consequently, the mean brightness of the output image is usually at the center or close to it when it comes to discrete implementation. In terms of images with high and low mean brightness value, you can find an important change in the view of the enhanced image. Secondly, histogram equalization performs the enhancement depending on the global content of the image [7]. Basically, histogram equalization highlights the edges and borders between different objects, but may decreases the local details of these objects and not adequate for local enhancement. Another result to do this mergence is the production of over saturation artifacts and enhancement [8]. For image *I (x, y)* with K discrete level, the gray values histogram is determined using the occurrence probability of the gray level *i*, as shown in the equation below:

$$Prob\ (i) = \frac{n_i}{N} \qquad Where\ 0 \leq i\ \leq K-1 \qquad (1)$$

Where N is the total number of pixels in the image, n is the total number of pixels with the same intensity level, and K is the total number of gray level in the image. Figure 3 shows fingerprint image before and after applying the histogram equalization.

Figure 3: Fingerprint Image Before and After Applying HE.

2. IMAGE COMPRESSION

Images become very important and used in many application areas. Because of the large amount of data images need to be compressed. Image compression is an application of data compression that encodes the original image with number of bits. Data compression has an important part in the different areas of storage. The purpose of image compression would be to decrease the redundancy of the image and also to store or transfer date in an efficient structure [11]. The primary objective of such system is to decrease the storage amount as much as possible, and the decoded image shown in the monitor can be similar to the original image as much as can be [9]. Image Compression can be either lossy or lossless. Lossless compression methods is suitable for archival purposes and for medical imaging. Lossy compression methods are suitable for natural images such as photographs in applications. The high quality at a given compression rate is the main purpose of images compression. Now consider an encoder and a decoder. When the encoder will receive the original image file, the image file is going to be converted into a series of binary data, frequently known as

the bit stream. The decoder after that receives the encoded bit stream and decodes it to form the decoded image. If the total data amount of the bit stream is lower than the total data amount of the original image, then this is called image compression [13] [14]. Figure 4 shows the full compression process. TIFF (Tagged Image File Format) is the image format used in FVC2000 database.

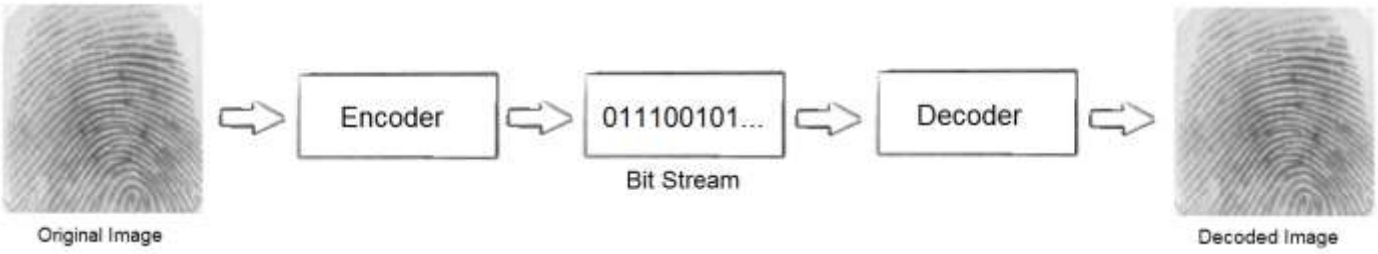

Figure 4: Image Compression Process.

PRINCIPAL COMPONENT ANALYSIS (PCA)

Building the image space by using the PCA method. The PCA is feature extraction and dimension reduction method. PCA is a popular linear projection technique, and also known as eigenspace projection, Karhunen and Loeve (KL) transformation, or Hotelling. It reduces dimensionality by extracting the principal components of multidimensional data. PCA can extract the important features, capture the nearly variable data components of samples, and then select numerous significant individual from all the feature components [5]. Through PCA, an efficient and simple recognition process can be obtained compared with other methods. Recognitions that use PCA features also perform better in singular variation cases for each individual. Raw intensity data are used for recognition and learning without mid-level or low-level processing [12].

PCA method Steps are as follows:

1. Collect images from fingerprint database.
2. Represent images as a single vector.
3. Compute average mean.

$$AverageMean = \frac{1}{M}\sum_{n=1}^{M} TrainImage(n) \quad (2)$$

Where M is the total number of the training images.

4. Subtract original image from average mean.

$$S = TrainImage - AverageMean \quad (3)$$

5. Compute covariance matrix.

$$Covarance = \frac{1}{M}\sum_{n=1}^{M} S(n)\, S^{T}(n) \quad (4)$$

6. Calculate eigenvalues and eigenvectors.
7. Sort and eliminate eigenvalues.
8. Build training matrix.

DISTANCE MEASUREMENT

The distance between two images is a major concern in image recognition and computer vision. The final step of fingerprint recognition is measuring the distance between two images. Image similarity is the distance between the vectors of two images. The distances among feature space representations are used as the basis for recognition decisions. Many existing image distance methods suffer from complicated measure computations, leading to a difficulty in combining the metric with some fingerprint recognition methods.

1. **CITY BLOCK DISTANCE**

City block distance, which also known as L1 distance or Manhattan distance classifier, is the summation of the absolute difference between two vectors. It is especially used for discrete types of descriptors. City block distance is a true distance function because it responds to triangle inequality. It also assumes a triangular distribution [15]. Suppose X is the training images and Y is the test image. City block distance then can be calculated from the equation bellow:

$$d(X,Y) = |X - Y| = \sum_{i=0}^{No.Of\ Images} |X_i - Y_i| \quad (5)$$

2. EUCLIDEAN DISTANCE

Euclidean distance is widely used to classify and compute the similarity between images because it is faster than other classifiers and is simple. Euclidean distance examines the root of square differences between the coordinated of a pair of objects. A minimum Euclidean distance classifier is the most suitable condition for normally distributed classes [15]. Euclidean distance can be calculated from the equation below:

$$d(X,Y) = \sqrt{\sum_{i=1}^{No.Of\ Images} (X_i - Y_i)^2} \quad (6)$$

RESULT AND DISCUSSION

In this analysis, FVC2000 fingerprint database that contain four databases (DB1, DB2, DB3, and DB4) is used with different number of training and testing images to evaluate the performance and recognition rate. TIFF (Tagged Image File Format) is a file format for FVC2000 fingerprint database. Table 3 shows the recognition rate of DB1 and DB3 without applying any pre-processing on images. The recognition rate will increase when the number of training images is increased. The recognition rate was weak and not acceptable. Therefore, images pre-processing methods are applied. Image pre-processing methods are important and applied to perform better results. Histogram equalization and Image compression are applied.

Table 3: Recognition Rates (DB1 and DB3).

No. of Testing Images	No. of Training Images	DB1 Recognition Rate	DB3 Recognition Rate
7	1	22.86 %	18.57 %
6	2	22.67 %	20 %
5	3	24 %	22 %
4	4	27.5 %	24.5 %
3	5	33.33 %	50 %
2	6	45 %	50 %
1	7	45 %	50 %

1. HISTOGRAM EQUALIZATION

The widely accepted histogram equalization cannot correctly enhance all the part of an image. When the original image is irregularly illuminated, certain information will be either too dark or too bright. These methods are generally the most applied histogram adjustment methods. HE is used to create a uniformly distributed image over the whole brightness scale, and HE helps the input image histogram to have a predefined shape. HE improve and increase the recognition rate. Table 4 shows the recognition rate of DB1 and DB3 using HE. Histogram of the image is equalized as shown in figure 5.

Table 4: Recognition Rate (DB1 and DB3) Using HE.

No. of Testing Images	No. of Training Images	DB1 Recognition Rate	DB3 Recognition Rate
7	1	40 %	31.43 %
6	2	43.33 %	32.33 %
5	3	48 %	40 %
4	4	57.5 %	47.50 %
3	5	66.67 %	63.33 %
2	6	75 %	65 %
1	7	75 %	70 %

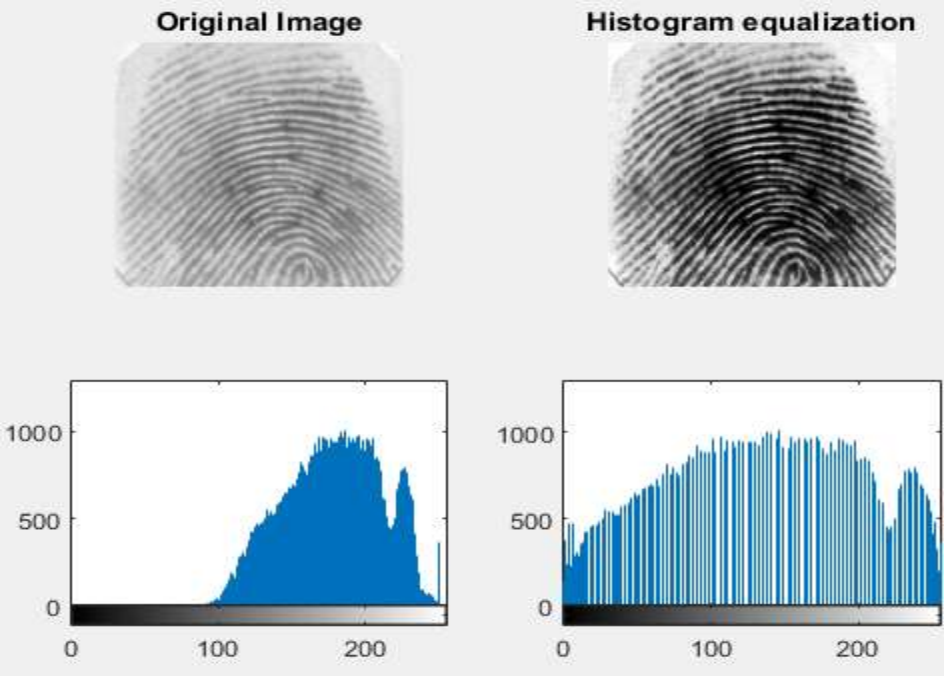

Figure 5: Image After Applying Histogram Equalization.

2. IMAGE COMPRESSION

In this analysis TIFF format is used in the FVC2000 database. Many compression methods can be applied. The compression reduces the size of files to save storage space and to save time when transmitting it. Compression play a center role in communication technology and is part of modern life. Lossless compression methods are useful in image archiving. Lossless means there is no quality loss due to compression. LZW data compression is a universal lossless method created by Abraham Lempel, Jacob Ziv, and Terry

Welch. LZW is one of the adaptive dictionary methods. LZW is simple to implement and has the potential for very high throughput in hardware implementation. LZW compression works better with images that have solid colors. LZW is also suitable for compressing text files. Many other lossless methods can be used. Pack Bits is simple and fast method for run length encoding.

CONCLUSION

Fingerprints are important area of biometrics and used in many applications. In this work Databases 1 and 3 of FVC2000 (Fingerprint Verification Competition) are used. PCA used for feature extraction and dimension reduction. City block distance is used for matching process. Different number of training and testing images are used. Increasing the training images will increase the recognition rate. The results without use any Pre-processing methods was weak and not acceptable. Histogram equalization and Image compression methods are used. The recognition rate increased when applying the histogram equalization in each database 1 and 3. Image compression also important and used to reduce the time and the size of files to save storage space.

REFERENCES:

[1] Kücken, Michael, and Alan C. Newell. "Fingerprint formation." *Journal of theoretical biology* 235, no. 1 (2005): 71-83.
[2] Raja, K. B. "Fingerprint recognition using minutia score matching." *arXiv preprint arXiv:1001.4186* (2010).
[3] Mazumdar, Subhra, and Venkata Dhulipala. "Biometric Security Using Finger Print Recognition." *University of California, San Diego* (2008): 3.
[4] Maio, Dario, Davide Maltoni, Raffaele Cappelli, James L. Wayman, and Anil K. Jain. "FVC2000: Fingerprint verification competition." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 24, no. 3 (2002): 402-412.
[5] Kale, Piyush G., and C. S. Khandelwal. "IRIS & Finger Print Recognition Using PCA for Multi Modal Biometric System." (2016).
[6] Rao, Raghuveer M., and Manoj K. Arora. "Overview of image processing." In*Advanced Image Processing Techniques for Remotely Sensed Hyperspectral Data*, pp. 51-85. Springer Berlin Heidelberg, (2004).
[7] Krishna, A. Sri, G. Srinivasa Rao, and M. Sravya. "CONTRAST ENHANCEMENT TECHNIQUES USING HISTOGRAM EQUALIZATION METHODS ON COLOR IMAGES WITH POOR LIGHTNING." *International journal of computer science, engineering and applications* 3, no. 4 (2013): 15.
[8] Raju, A. "A comparative analysis of histogram equalization based techniques for contrast enhancement and brightness preserving." (2013).
[9] Wei, Wei-Yi. "An introduction to image compression." *National Taiwan University, Taipei, Taiwan, ROC* (2008).
[10] Aguilera, Paula. "Comparison of different image compression formats."*Wisconsin College of Engineering, ECE* 533 (2006).
[11] Kaur, Simrandeep, Verma, Sulochana. "Design and Implementation of LZW Data Compression Algorithm." *International Journal of Information Sciences and Techniques (IJIST)*, 3, no. 4 (2012).
[12] Melo, Ernande F., and Hélio M. de Oliveira. "A Fingerprint-based Access Control using Principal Component Analysis and Edge Detection." *arXiv preprint arXiv:1502.01880* (2015).
[13] Vijayvargiya, Gaurav, Sanjay Silakari, and Rajeev Pandey. "A Survey: Various techniques of image compression." *arXiv preprint arXiv:1311.6877*(2013).
[14] Dheemanth, H. N. "LZW Data Compression." Volume 3 Issue, (2014): 22.
[15] Gawande, Mohit P., and Dhiraj G. Agrawal. "Face recognition using PCA and different distance classifiers."