Multi-Object Face Recognition Using Local Binary Pattern Histogram and Haar Cascade Classifier on Low-Resolution Images

R. Rizal Isnanto[*], Adian Fatchur Rochim, Dania Eridani, Guntur Dwi Cahyono

Department of Computer Engineering, Diponegoro University, Semarang, Indonesia

Abstract

This study aims to build a face recognition prototype that can recognize multiple face objects within one frame. The proposed method uses a local binary pattern histogram and Haar cascade classifier on low-resolution images. The lowest data resolution used in this study was 76×76 pixels and the highest was 156×156 pixels. The face images were preprocessed using the histogram equalization and median filtering. The face recognition prototype proposed successfully recognized four face objects in one frame. The results obtained were comparable for local and real-time stream video data for testing. The RR obtained with the local data test was 99.67%, which indicates better performance in recognizing 75 frames for each object, compared to the 92.67% RR for the real-time data stream. In comparison to the results obtained in previous works, it can be concluded that the proposed method yields the highest RR of 99.67%.

1. Introduction

Object identification by artificial intelligence is a research topic that is attracting growing interest. Various methods are used for object identification in the field of artificial intelligence. One of the implementation concepts under development is the identification of human faces. The development and implementation of face recognition technology is an interesting field. In this modern era, emerging technologies for smartphones utilize face recognition technology in their security systems.

Face recognition technology uses a camera to capture image data, which can have various resolutions. Low-resolution images are those that lack sharp focus or fine detail. Compared to high-resolution images, low-resolution images have fewer pixels and greater compression. This compression compromises the quality of smaller image files [1]. Furthermore, the acquisition of low-resolution images can produce noise that looks like black or white dots on the original image known as salt-and-pepper noise. This noise is generally caused by software failure, hardware failure in the image capturing, transmission, or the defect of a camera sensor. A standard salt-and-pepper noise value may be either minimum (pixel value 0) or maximum (pixel value 255) [2].

This paper presents a prototype face recognition system based on the test data with four face objects within the same frame. This prototype system implements an LBPH algorithm for generating a model for face recognition purposes. The training data were generated by a low-resolution webcam implemented by the Haar cascade method, which is used to detect specific objects based on a model loaded in the program. Additional methods are used to improve the quality of the image data, including histogram equalization and median filtering. Unlike previous research, in this study, the images with multiple

[*] Corresponding author. E-mail address: rizal@ce.undip.ac.id

Tel.: +62-811-266-1508; Fax: +62 24 76480609

face objects were used and focused on the implementation of the LBPH algorithm and Haar cascade classifier on low-resolution images. The limitations of previous research were that they used single face objects and they did not focus on a comparison of the local data and real-time streaming video data for RR testing.

The LBPHs were utilized to extract features from the training and testing datasets and the Euclidean distance (ED) between image histograms to indicate accuracy. This paper is organized as follows: Section 2 presents the literature review. Section 3 describes the methodology and the implementation process. Section 4 presents the study results and a discussion and section 5 concludes.

2. Literature Review

Taskiran, Kahraman, and Erdem recently reported that facial biometric data is one of the most preferred types of biometric data as it generally does not require the cooperation of the user and can be obtained without violating private space [3]. Wei, Chen, Li, Jiang, Zhao, and Li proposed an adaptive dictionary-learning method based on local configuration patterns for face recognition. Extensive experiments on representative face databases with variations in lighting, expression, pose, and occlusion demonstrated that the performance of the proposed approach was superior for both recognition time and accuracy [4].

In 2018, Ahmed, Guo, Ali, Deeba, and Ahmed implemented a local binary pattern histogram (LBPH) using 500 data sets of each object and an HD 1080p camera to obtain 90% recognition of single-face data in a controlled environment without any additional process to improve the image data [5]. Younis and Muhammad presented the indexed face recognition system (IFRS) based on face recognition and radio-frequency identification technologies. IFRS uses LBPH as a feature vector and a Haar cascade classifier for face detection. The experimental results indicated that 4×4 image divisions provide better results than 8×8 image divisions for database access time, recognition time, and storage capacity. The results also indicated that IFRS can reach 100% accuracy with negligible time delay [6].

Niaraki and Shahbahrami improved the accuracy of face recognition systems using a co-occurrence matrix of local median binary pattern (LMBP) features. The authors combined two algorithms, i.e., the LMBP operator and the calculation of a co-occurrence matrix of the LMBP image, and their experimental results on the ORL and Yale databases showed that their proposed algorithm obtained accuracies of 96.25% and 100% on these respective image databases [7].

Angulu, Tapamo, and Adewumi reported that in addition to the use of LBP to describe texture for face recognition, gender classification, face detection, and facial component tracking, LBP has also been used to characterize texture features for age estimation [8]. Meng, Gao, Wang, Lin, and Zhang investigated a novel and efficient face recognition technique based on a local binary pattern (LBP) with a threshold for resolving the traditional LBP weakness in extracting global features. In their research, the authors combined local and global features using LBP and used principal component analysis (PCA) and linear discriminant analysis (LDA) to reduce dimensionality and optimize discriminative recognition, respectively. Their experimental results revealed this method to be valid and feasible [9]. Anbarjafari presented a face recognition system using color LBPs based on a mutually independent color channels system. The author compared the performance of that system with those of conventional and the state-of-the-art techniques, and the recognition rates (RRs) obtained using the feature vector fusion approach on the FERET database was 99.78%, compared with 79.60% and 68.80% for the conventional gray-scale LBP and PCA-based face recognition techniques, respectively [10]. Dimensionality reduction (DR) is an important and necessary technique for treating high-dimensional data. Currently, PCA and LDA remain the two most popular DR approaches. Yang and Li proposed a novel algorithm using LBP-based discriminant graph construction for reducing dimensionality, which is applicable to face recognition [11].

The fusion of two useful descriptors, i.e., Zernike moments (ZMs) and the LBP/local ternary pattern (LTP) was proposed by Singh, Mittal, and Walia. The recognition results achieved by their method were approximately 10% to 30% higher than those obtained by these descriptors separately. The RRs of the proposed method was also found to be significantly better (i.e., from 8% to 24%) in the case of a single example image per person in the training [12]. The method proposed by Holder and Tapamo for facial expression recognition, which involves improved gradient local ternary patterns (GLTPs), was extensively tested on the CK + and JAFFE datasets using a support vector machine. Their method was shown to further improve the accuracy and efficiency of the GLTPs obtained by other common and state-of-the-art methods [13].

Louis and Plataniotis proposed the use of multiple instances of a rotational LBP of pixels as features instead of the histogram bins of the LBP of pixels. The multiple features are selected using the sequential forward selection algorithm known as the co-occurrence of the LBP (CoLBP). The experimental results demonstrated that the CoLBP face features performed better than state-of-the-art Haar-like features and various other LBP feature extensions. In addition, CoLBP features were found to tolerate a wide range of blurring and changes in illumination [14]. Apart from LBP-based face recognition, an LBP-based on the periocular recognition algorithm has also been developed. Mahalingam and Ricanek developed a novel face-based matcher composed of a multi-resolution hierarchy of patch-based feature descriptors for periocular recognition, i.e., recognition of the soft tissue surrounding the eye orbit. Their work showed that the performance of the proposed periocular matcher (using only a small section of the face about the eyes) compared favorably to that of a commercial full-face matcher [15]. Although LBP features have been shown to be discriminative in face recognition, pixel-level descriptions of LBP features are sensitive to change in the scale of the images. Sapkota and Boult extended the utility of a generalized variant of the LBP feature descriptor, which they called the generalized region assigned to binary (GRAB), and revealed that it handles challenges associated with a scale. On face recognition at multiple scales, the GRAB operator has significantly outperformed the basic LBP and is more stable than multi-scale block LBP in cases of reduced scale on subsets of a well-known published database of labeled faces in the wild (LFW) [16].

A parallel ensemble LBP and convolutional neural network for face recognition were proposed by Tang, Su, Su, Fong, Cao, and Gong. In their experiments, their approach was found to enhance tolerance to posture, expression, and illumination, and improve the accuracy of face recognition. However, their approach was found to have poor model generalization performance, which is normally caused by the learning algorithm being trapped in a local minimum [17]. Zhou, Constantinides, Huang, and Zhang proposed an improved center symmetric LBP for face recognition, with the experimental results on some face datasets indicating that a higher RR could be obtained by employing the proposed method with nearest neighbor classification [18].

Zhou, Wang, Lin, Hao, and Lu proposed a face recognition method according to LBP and improved pairwise-constrained multiple metric learning. Their experimental results showed that their proposed method achieved high RRs and was robust to variations in illumination and facial expression, especially for misaligned facial images [19]. A new method for describing texture images using a hybrid combination of local and global texture descriptors was proposed by García-Olalla, Alegre, Fernández-Robles, García-Ordás, and García-Ordás. In this regard, the authors presented a new adaptive LBP (ALBP) descriptor for performing local descriptions, which they built by adding oriented standard deviation information to the ALBP descriptor to achieve a more complete image representation. Hence, it has been referred to as an adaptive LBP with oriented standard deviation [20].

Muqeet and Holambe presented an LBP according to the directional wavelet transform for expression and pose-invariant face recognition. The authors proposed a novel facial feature extraction method that utilizes interpolation-based directional wavelet transform (DIWT) and LBP, in which LBP histogram features are extracted from selected top-level DIWT sub-bands to obtain a local descriptive feature set [21]. An experiment on high-dimensional feature extraction was

conducted by Yoo, Kim, Jung, and Ko, who used the bit-plane decomposition of LBPs for robust face recognition. The authors observed that the LBP can be decomposed into bit planes, each of which has scale-specific directional information about the face image. Each bit plane was reported to contain the inherent local structure of the face image and a robust illumination characteristic [22].

For facial expression recognition, Zhao and Zhang proposed a novel kernel-based supervised manifold learning algorithm based on locally linear embedding (LLE), which they called discriminant kernel locally linear embedding (DKLLE). The experimental results on two benchmark facial expression databases demonstrated the effectiveness and promising performance of DKLLE [23]. Chai, Mendez-Vazquez, He, Sun, and Tan explored semantic-pixel-set-based local patterns with information entropy for face recognition and evaluated their method on three face recognition databases: AR, FERET, and LFW, using the popular LBP operator and some of its extensions. The experimental results showed that their proposed method outperformed methods that use the same local patterns with traditional divisions, as well as some of the best-performing state-of-the-art methods [24]. Massoli, Amato, and Falchi reported that although specific datasets for low-resolution facial images have been proposed and less attention has been paid to the task of cross-resolution matching. Despite their focus on cross-resolution face recognition, using their training algorithm enabled an improvement in the performance of state-of-the-art models that considered low-resolution matches [25].

Dahmouni, Moutaouakil, and Satori proposed the application of the electric virtual binary pattern (EVBP) to face recognition. In their tests, the EVBP descriptor reduced the feature vector size by 93.75%, which enabled them to go from 255 bin-histograms for the LBP to 16 bin-histograms for the EVBP descriptor [26]. Zhou, Liu, and Lu proposed a novel face feature extraction approach according to LBP and two-dimensional locality-preserving projections (2DLPP) to enhance texture features and preserve the spatial structure properties of a face image. Their proposed face feature extraction and recognition method achieved better performance than similar techniques, and the proposed dimension selection rule was reported to provide an appropriate feature dimension in 2DLPP [27]. Liang, Tu, Liu, Zhao, and Jain reported 3D face reconstruction from mugshots as an application for arbitrary-view face recognition. The authors stated that the mugshots routinely acquired by law enforcement agencies are under-utilized by automated face recognition systems, and they proposed a regression-based approach for reconstructing textured full 3D face models from multi-view mugshot images [28].

Chude-Olisah, Sulong, Chude-Okonkwo, and Hashim reported that their proposed method for the edge-based Gabor feature representation of plastic-surgery-altered images was robust against plastic-surgery-induced face variations, despite expression and illumination problems, and outperformed plastic-surgery face recognition methods reported in the literature [29]. Bolotnikova, Demirel, and Anbarjafari proposed a real-time face recognition system based on block processing of the LBPs of face images captured by NAO humanoid robots, which are being used in many human–robot interaction applications. The experimental results showed that the proposed face recognition algorithm outperformed conventional and state-of-the-art techniques [30].

Liu, Song, and Tang developed a framework that fuses virtual mirror synthesized training samples as bases and hierarchical multi-scale LBP features for classification. More specifically, in the first stage of their proposed method, the sampling uncertainty of the linear approximation model is effectively alleviated by generating extra synthesized mirror training samples from the original images [31]. In another study, the use of local polynomial contrast binary patterns for face recognition was proposed by Xu, Jiang, Wang, Zhou, Li, and Liao. From their experimental results, they inferred that their method could efficiently estimate the underlying local surface information, which is approximately represented as linear projection coefficients of the pixels on a local patch. Decomposition using polynomial filters could also capture valuable image information at multiple orientations and frequency bands [32].

3. Proposed Method

3.1. *LBP histogram*

LBPH is a new version of the LBP method that improves face recognition results by the use of a histogram to represent the number of occurrences of each binary code in patches of the image data. The recognition rates (RRs) of LBPH algorithms can be reduced by the lighting conditions and changes in expression and attitude deflection. The LBP operator was originally designed for texture descriptions. Formally, given a pixel at (x_c, y_c), the resulting LBP can be expressed in decimal form as follows:

$$LBP_{P,R}(x_C, y_C) = \sum_{P=0}^{R-1} s(i_P - i_C)2^P \tag{1}$$

where i_C and i_P are the gray-level values of the central pixel and the P surrounding pixels in the circle neighborhood with radius R, respectively. The function $s(x)$ is defined as:

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{2}$$

where x is the difference between i_P and i_C. By the above definition, the basic LBP operator is invariant to monotonic grayscale transformations that preserve the pixel intensity order in the local neighborhoods. The histogram of the LBP labels calculated over a region can be used as a texture descriptor. The operator $LBP(P, R)$ produces 2^P different output values, which correspond to the 2^P different binary patterns formed by P pixels in the neighborhood [33].

3.2. *Haar cascade classifier*

The Haar cascade algorithm is a machine-learning model often used to perform object detection in images. In obtaining object detection values, feature values are calculated based on image integrals. These integrals can calculate values accurately and relatively quickly by presenting new images according to previously scanned area values:

$$s(x, y) = i(x, y) + s(x, y-1) + s(x-1, y) + s(x-1, y-1) \tag{3}$$

where $s(x, y)$ is the value of the number of areas in (x, y), $i(x, y)$ is the intensity of the original image matrix, $s(x, y\text{-}1)$ is the value of neighboring pixels of y, $s(x\text{-}1, y)$ is the value of neighboring pixels of x, and $s(x\text{-}1, y\text{-}1)$ is the sum of the values of neighboring diagonal pixels.

The Haar cascade algorithm implements a cascade function to perform image training in four main stages: determination of Haar features, obtaining an integral image, AdaBoost training, and classification using a cascading classifier [34].

3.3. *Histogram equalization*

Histogram equalization is an equalization process for distributing gray degree values in image data in which the cumulative distribution function of the histogram is implemented. Image pixels with values ranging from 0 to 255 are read based on the sum of each value, which is then flattened to form a straight diagonal line on the graph [35].

3.4. *Median filtering*

Median filtering is a method that is used to improve image quality by refining the original image and eliminating disturbances from salt-and-pepper noise in the image data [36]. This method replaces the pixel value in an area of operation

by applying a formula to utilize the value of a neighboring pixel. The median filter provides a high quality of salt-and-pepper noise reduction because of the low values of RMSE and high values of PSNR [2].

3.5. Euclidean distance

The Euclidean Distance (ED) is a method that is used to measure the straight-line distance between two points, which involves measurement of two histograms here, as defined in the paper "Centralized binary patterns embedded with image ED for facial expression recognition" [37]. The output of the algorithm is the calculated distance from the histogram of one image to the nearest histogram of another. Using the ED approach, the distance value can be used as a measure of the confidence level regarding the accuracy of facial recognition values:

$$D = \sqrt{\sum_{i=1}^{n} (hist1_i - hist2_i)^2} \qquad (4)$$

where D is the ED value, $hist1$ is the histogram value of the first image, $hist2$ is the histogram value of the second image, and n is the histogram length (255 for gray-level images).

3.6. Research procedure

Fig. 1 shows the research procedure adopted in this study. First, a dataset of training data is created in which the image data are then augmented using histogram equalization. The images are then processed using median filtering to obtain the final training data and the training process is initiated using the LBPH algorithm. The output of the training process is a model file with the extension yml (YAML). After creating the model, the face recognition prototype is implemented and its performance is tested using the test data.
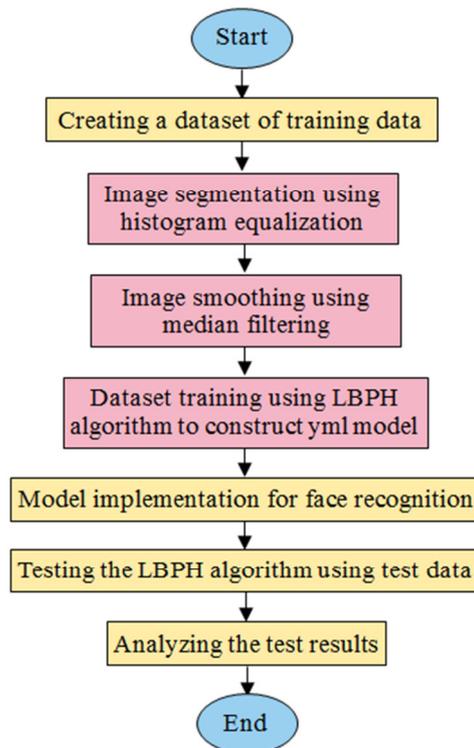

Fig. 1 Block diagram of research procedure

The lowest data resolution captured is 76×76 pixels and the highest is 156×156 pixels. The data in this study consisted of 480 facial images with four identifications for testing local data and four for testing real-time data. The labels

connecting the training data are stored as identification labels, which are then processed as a model of the face recognition system. Fig. 2 shows an example of the collecting process of facial images into the dataset.

For this study, 60 face images from volunteers were collected. Image data were acquired using a VGA webcam to detect parts of the image frame as a specific facial object using the Haar cascade classifier as a feature selection technique, which is a function of the Haar-like feature. A region or some of the pixels around the detected facial object were cropped. The coordinate region of the face in the upper left and right corners and the bottom left and right corners were drawn using a bounding box as shown in Fig. 3.
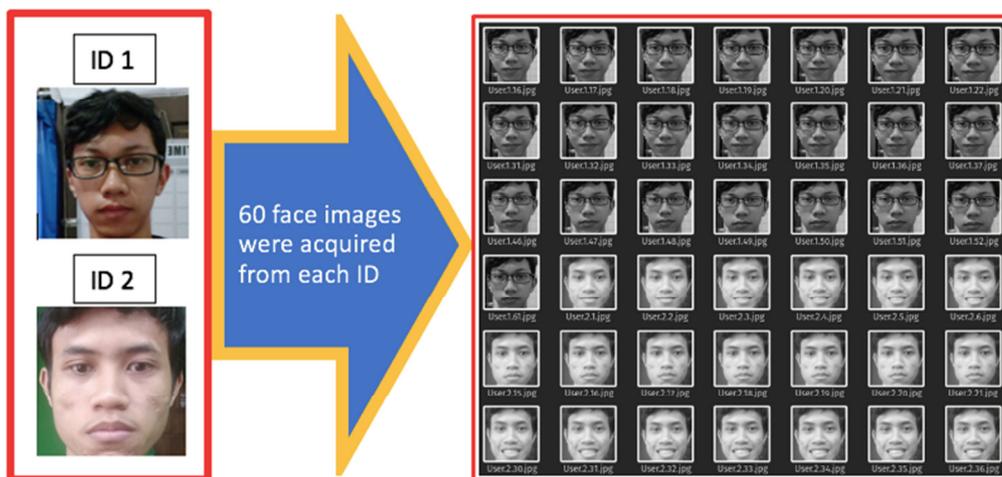

Fig. 2 Example of collecting process of dataset

Fig. 3 Implementation of Haar cascade classifier

Next, the facial image data were subjected to a histogram equalization process, in which the gray-level range is increased to increase the contrast in the image data. Then, the image data were subjected to an image smoothing process by median filtering to reduce noise. The data pre-training process used to create the LBPH model is described in Algorithm 1. Fig. 4 shows the results of the pre-training process from an original image to a median-filtered image.

Algorithm 1 Pre-trained image data

Input: Video or stream webcam.

1: Capture each image frame.

2: Convert the RGB image to grayscale.

3: Process the image using histogram equalization.

4: Process the image using median filtering with an addition of 30% noise.

5: Scaling to reduce the size of the original image by 30%.

6: **for** (coordinate) in detected faces

7: Put label or tagging with ID.

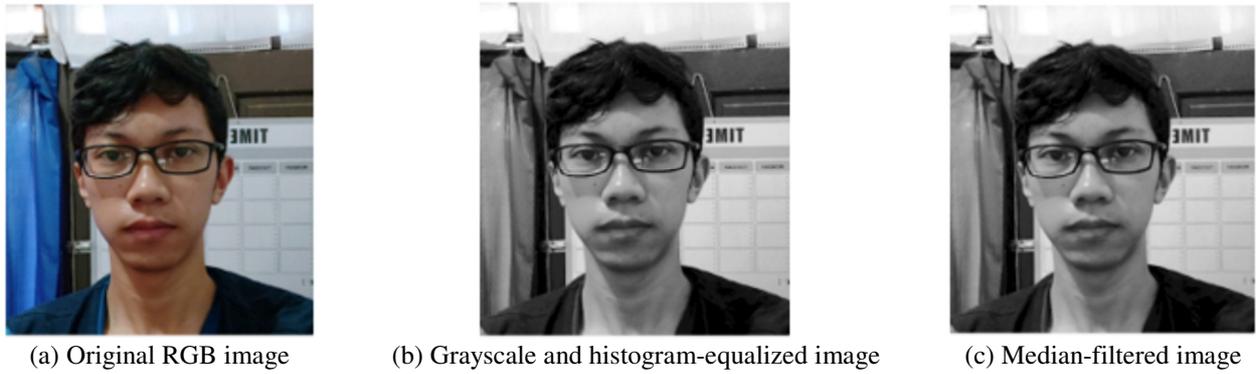| (a) Original RGB image | (b) Grayscale and histogram-equalized image | (c) Median-filtered image |

Fig. 4 Face image processing steps

The images were then processed by the LBPH algorithm, which implements LBP operations. Image data are represented in a matrix in pixel size for each point in the image. Then, the midpoint of the matrix is used as a threshold to determine a new value for the neighboring points around the threshold. Next, for each neighbor of the central value (threshold), a new binary matrix is set. For values equal or higher than the threshold, the values are set 1, otherwise the values are set 0. With the new binary matrix, the binary data for each neighbor are sorted clockwise to obtain a specific binary value, which is converted to decimal numbers as a new value for the initial matrix. Fig. 5 shows the LBPH workflow starting from an original sample image.

(a) LBP descriptor process

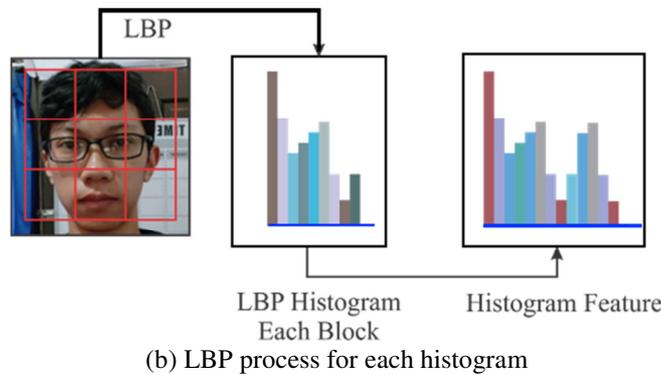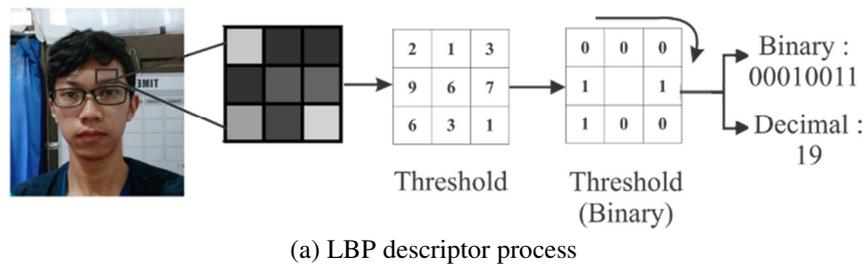

(b) LBP process for each histogram

Fig. 5 Workflow of the LBPH

The LBP operation is used to process a histogram for the data of each image to represent the intensity of each pixel. The average histogram continues to evolve into a new histogram that represents the characteristics of the training image data. The LBPH algorithm used in this study performs calculations to distinguish the characteristics of an image. Face recognition is used to compare the histogram values of two images and find the closest or smallest value. To compare two features, the ED as a similarity measure was used. Fig. 6 shows the detection process using LBPH and ED.

In this research, 100 as the threshold value for the ED measurement was used. If the proximity value of the test data input is lower than this threshold, then it can be assumed that the LBPH algorithm successfully performed face recognition. The recognition accuracy percentage was calculated for each face in this process by subtracting the value of the computational ED variable from 100%. Algorithm 2 shows the face recognition prototype process used in this research.
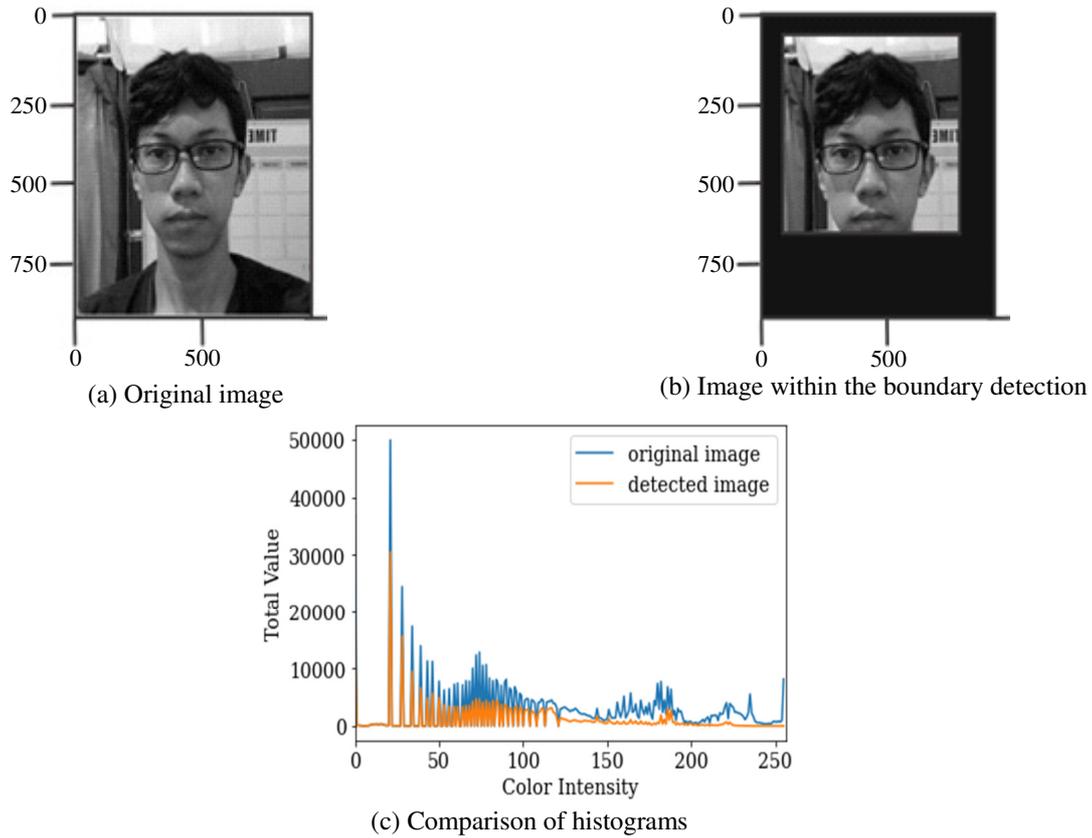
(a) Original image

(b) Image within the boundary detection

(c) Comparison of histograms

Fig. 6 Detection process using the LBPH algorithm and ED similarity measure

---

Algorithm 2 Face Recognition Prototype

Input: Video or stream webcam, number of time intervals (default set to 0), and number of frames to process (default set to process all frames).

1: Load the trained LBPH model for face recognition and Haar cascade classifier for face detection.

2: Start the system timer.

3: Capture every frame from the input.

4: Convert the detected frame to grayscale and reduce the size to 30% of the original.

5: **if** (interval time is set).

6:      **for** (coordinate) in detected faces.

7:          Calculate the LBPH of the detected face and predict the ED with the loaded LBPH model.

8:          **if** (proximity LBPH value <= 100).

9:              Set the ID of matched recognition from the database.

10:         **if** (number of frame count to max).

11:             Restart the system timer.

Next, the performance of the prototype was tested using the time interval recognition process and the RR. The time intervals between float points ranging from 0.1 to 1.0 seconds were used to analyze the accuracy rate for each condition. The RR test determines the consistency of success in the recognition of each object using 300 frames. The face recognition test used four scenarios, i.e., a single face, two faces, three faces, and four faces, as presented in Fig. 7. The RR is then calculated using the following formula:

$$RR = \frac{\text{Number of correct recognition}}{\text{Number of test data}} \times 100\% \qquad (5)$$
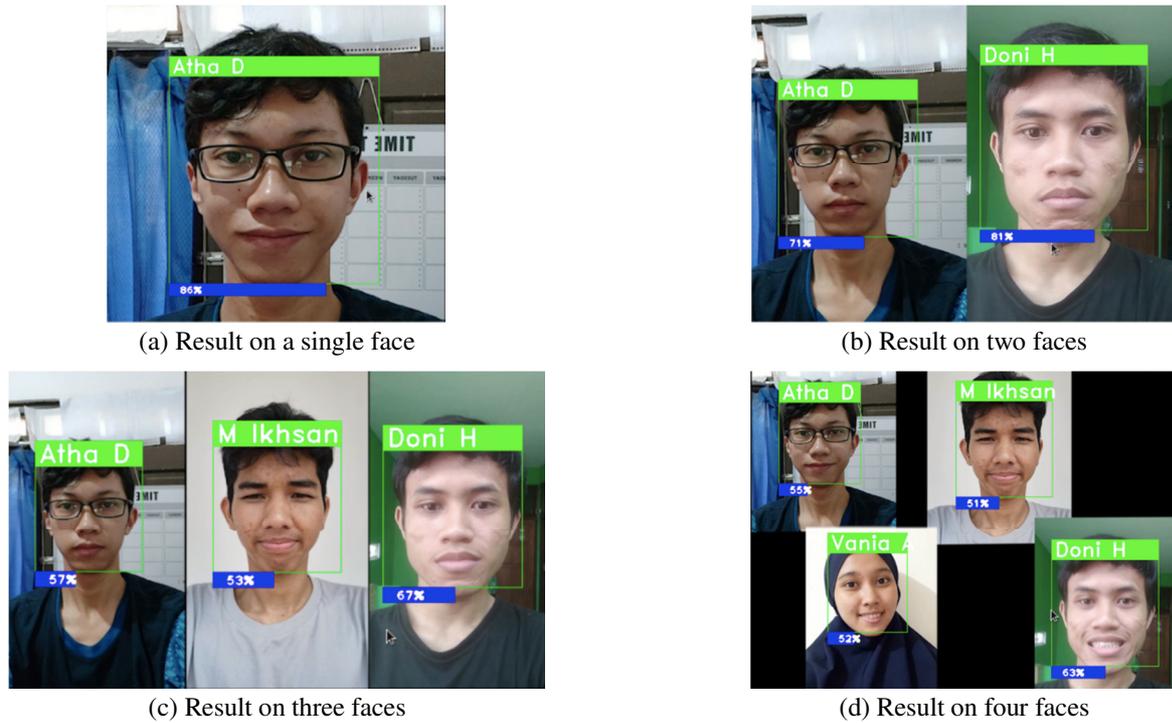
(a) Result on a single face

(b) Result on two faces

(c) Result on three faces

(d) Result on four faces

Fig. 7 Face recognition results on single and multiple objects

4. Results and Discussion

First, the time interval used in the recognition process of the test data sourced from local video files and real-time capture was tested. The time interval is the time lag in seconds of the program's face recognition function. The local video undergoes an editing process to adjust the quality of the video acquired by the VGA webcam (640×480 pixels) and the real-time video stream uses a webcam of the same quality. To determine the average accuracy level, the test process was performed on the four research objects containing from one to four faces using a minimum interval value.
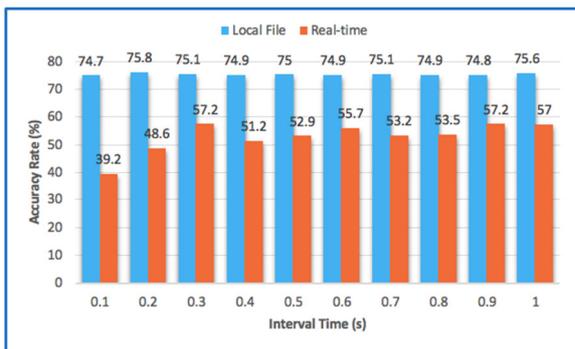

Fig. 8 Accuracy rate for a single face in one frame
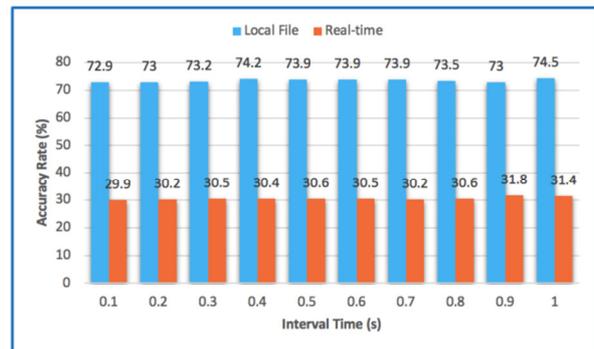

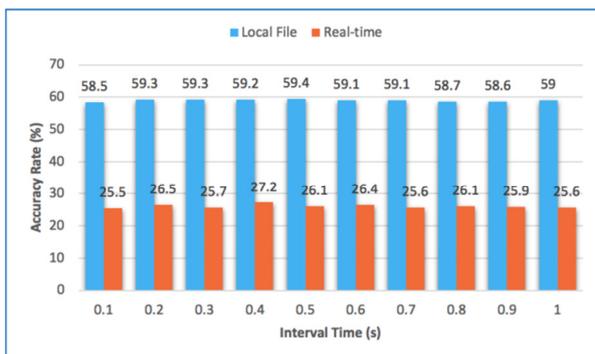Fig. 9 Accuracy rate for two faces in one frame

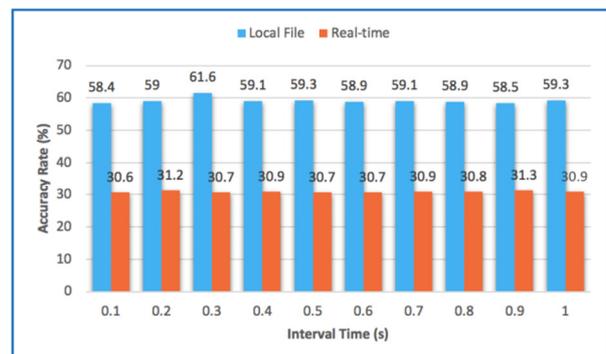Fig. 10 Accuracy rate for three faces in one frame

Fig. 11 Accuracy rate for four faces in one frame

The identification accuracy values for the local and real-time data were significantly different. These data sources were compared in terms of their resolution with the local data having been edited to match the real-time data acquired using webcams that produce VGA quality. Figs. 8-11 indicate that although they were converted to lower resolution, the test data acquired using a high-resolution camera still realized good accuracy with an average RR of four face objects higher than 50%. A comparison of the data also displays similarities in the local and real-time data files, whereby the smallest recognition time interval of 0.1 seconds produced the lowest accuracy value.

The difference in the results is also offset by differences in the test data source. To determine the recognition performance regarding the right face, the RRs were determined by testing 75 frames on each object. Each face identified in the recognition process and still counted in the sequence was adjusted according to its call identification in the database. An appropriate and orderly correct answer (if the data are regular, such as local data) indicates the success of the recognition process in recognizing the owner's face. Table 1 shows that the average RR of each test is obtained using the same average test data in 300 frames for each source.

Table 1 RR results from each test

Data Source	Object Identification	Number of Correct Recognitions	RR (%)
Local File	ID 1	75	100
	ID 2	75	100
	ID 3	74	98.67
	ID 4	75	100
Real-time	ID 1	75	100
	ID 2	69	92
	ID 3	75	100
	ID 4	59	78.67

Figs. 12 and 13 respectively show the difference in RRs for the local file data and real-time video stream data. In Fig. 12, the local test data percentages illustrates the success in recognizing the owner of the face in the test with only one false recognition count each time. In Fig. 13, the real-time test data displays differences in the RRs and the success in recognizing the owner of each face. There are also recognition errors for the object being tested.
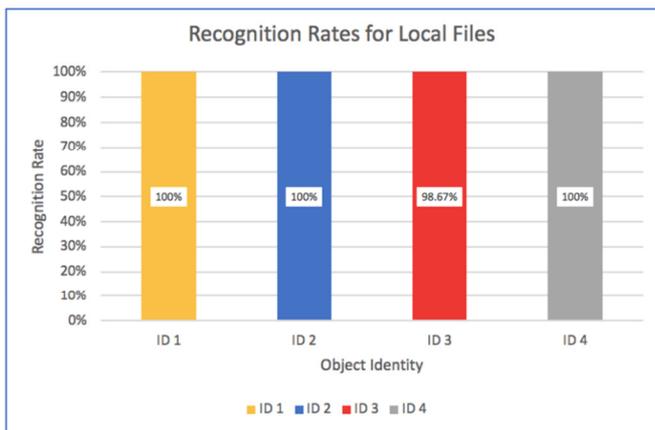

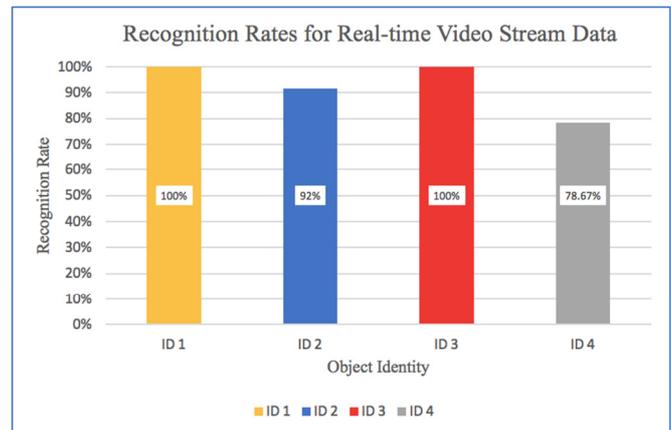Fig. 12 Recognition rates using local file data

Fig. 13 Number of successful recognitions using real-time video stream data

The standard error values in Table 2 demonstrate the standard deviations of the face recognition test results. The success rates for both data sources in the recognition of each face are high, but a mismatch is evident in the recognition process, whereby the real-time data source incurred more error values.

Table 2 Comparison of RR values, standard deviations, and standard errors

Data Source	RR (%)	Standard Deviation	Standard Error
Local File	99.67	0.8164965809	0.047140
Real time	92.67	13.73559852	0.793025

Compared with the work of Ahmed, Guo, Ali, Deeba, and Ahmed, who studied low-resolution images with 45×45 and 35×35 pixels, which resulted in 94% and 90% RRs, respectively [3], the study yielded a higher RR of 99.67%. The proposed method which used the combination of the median filter and Haar cascade classifier also obtained a higher RR than the co-occurrence matrix of LMBP features (COM + LMBP) with 96.25% RR [14], the LBP + PCA + LDA with 98% RR [21] and CoLBP features with 97.28% RR [7].

In the comprehensive survey by Huang, Shan, Ardabilian, Wang, and Chen of the LBP methodology, including recently proposed variations. The best RR was obtained by the Gabor + LBP method with the FERET database, i.e., 99% [33], which is still lower than the RR value obtained in this study. This means that this proposed method using the median filter and Haar cascade classifier has enhanced the RRs of previous studies.

The proposed method may be improved with the latest methods, such as machine learning (ML) or deep learning (DL) algorithms. However, several papers explain the disadvantages of ML and DL face recognition. The problems or limitations for an ML face recognition system are: illumination variation, facial expression change, aging, pose change, scaling factor (i.e. size of the image), presence and absence of spectacles, mustache or beard, and occlusion due to mask, scarf or obstacles in front [38]. With all its popularity, ML is not perfect. It requires massive data sets to train on and these should be inclusive/unbiased and of good quality. Also, there will be times where it must wait for new data to be generated. ML also needs enough time for the algorithm to learn to fulfill its purpose with a considerable amount of accuracy. It also needs massive resources to function. This can mean additional requirements of computer power [39]. While DL face recognition has been successfully applied on many user-cooperated applications. However, the ubiquitous recognition applications are still an ambitious goal. In practice, it is difficult to collect and label sufficient samples for innumerable scenes in the real world [40].

Some limitations on the research results using this proposed method are that the maximum number of objects studied in one frame is 4 (four) faces. RR test results using objects from the video stream are also not as high as using the local data. For further development, it is necessary to test objects with more than 4 (four) faces per frame. Further research is needed using other algorithms, for example, algorithms that are included in both machine learning and deep learning methods, to then compare it with the proposed method so that the most suitable method for multi-face recognition is obtained.

5. Conclusions

The face recognition prototype proposed in this paper successfully was recognized four face objects simultaneously. The results obtained by the LBPH algorithm were comparable for local and real-time stream video data for testing. The RR obtained with the local data test was 99.67%, which indicates better performance in recognizing 75 frames for each object, compared to the 92.67% RR for the real-time data stream. However, this prototype does not consider conditions such as illumination, pose variant, and more faces than four within the same frame. The academic contribution of this research is the proposed process and face recognition method by inputting still images or video streams using LBPH and Haar cascade classifiers on low-resolution images. Meanwhile, the practical contribution is the use of the proposed method to be used in the real-time application of multi-face recognition from still images or video streams so that it can reduce the computation time compared to sequential recognition of several single faces.

In the future, researches on the effects of illumination, pose variant, and more than four faces within the same frame on recognition rates will be done using the proposed method. Meanwhile, for advanced researches, it is recommended to use other algorithms, for example, those included in the machine learning, deep learning methods, or other methods based on template matching, statistics, or neural networks to be compared with the proposed method. These comparisons should be performed to determine the best method for multi-object face recognition.

Acknowledgment

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] H. Chen, "A Kind of Effective Method of Removing Compound Noise in Image," 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), October 2016, pp. 157-161.

[2] J. Azzeh, B. Zahran, and Z. Alqadi, "Salt and Pepper Noise: Effects and Removal," JOIV: International Journal on Informatics Visualization, vol. 2, no. 4, pp. 252-256, 2018.

[3] M. Taskiran, N. Kahraman, and C. E. Erdem, "Face Recognition: Past, Present and Future (A Review)," Digital Signal Processing, vol. 106, November 2020.

[4] D. Wei, T. Chen, S. Li, D. Jiang, Y. Zhao, and T. Li, "Adaptive Dictionary Learning Based on Local Configuration Pattern for Face Recognition," EURASIP Journal on Advances in Signal Processing, pp. 1-12, May 2020.

[5] A. Ahmed, J. Guo, F. Ali, F. Deeba, and A. Ahmed, "LBPH Based Improved Face Recognition at Low Resolution," International Conference on Artificial Intelligence and Big Data (ICAIBD), May 2018, pp. 144-147.

[6] M. I. Younis and R. S. Muhammad, "IFRS: An Indexed Face Recognition System Based on Face Recognition and RFID Technologies," Wireless Personal Communications, vol. 101, no. 4, pp. 1939-1966, May 2018.

[7] R. J. Niaraki and A. Shahbahrami, "Accuracy Improvement of Face Recognition System Based on Co-occurrence Matrix of Local Median Binary Pattern," 4th International Conference on Pattern Recognition and Image Analysis (IPRIA), March 2019, pp. 141-144.

[8] R. Angulu, J. R. Tapamo, and A. O. Adewumi, "Age Estimation via Face Images: A Survey," EURASIP Journal on Image and Video Processing, June 2018.

[9] J. Meng, Y. Gao, X. Wang, T. Lin, and J. Zhang, "Face Recognition Based on Local Binary Patterns with Threshold," International Conference on Granular Computing, IEEE Press, August 2010, pp. 352-356.

[10] G. Anbarjafari, "Face Recognition Using Color Local Binary Pattern from Mutually Independent Color Channels," EURASIP Journal on Image and Video Processing, January 2013.

[11] B. Yang and Q. Z. Li, "Local Binary Pattern-Based Discriminant Graph Construction for Dimensionality Reduction with Application to Face Recognition," Multimedia Tools and Applications, vol. 78, no. 16, pp. 22445-22462, April 2019.

[12] C. Singh, N. Mittal, and E. Walia, "Complementary Feature Sets for Optimal Face Recognition," EURASIP Journal on Image and Video Processing, July 2014.

[13] R. P. Holder and J. R. Tapamo, "Improved Gradient Local Ternary Patterns for Facial Expression Recognition," EURASIP Journal on Image and Video Processing, June 2017.

[14] W. Louis and K. N. Plataniotis, "Co-Occurrence of Local Binary Patterns Features for Frontal Face Detection in Surveillance Applications," EURASIP Journal on Image and Video Processing, January 2011.

[15] G. Mahalingam and K. Ricanek, "LBP-Based Periocular Recognition on Challenging Face Datasets," EURASIP Journal on Image and Video Processing, July 2013.

[16] A. Sapkota and T. E. Boult, "GRAB: Generalized Region Assigned to Binary," EURASIP Journal on Image and Video Processing, June 2013

[17] J. Tang. Q. Su, B. Su, S. Fong, W. Cao, and X. Gong, "Parallel Ensemble Learning of Convolutional Neural Networks and Local Binary Patterns for Face Recognition," Computer Methods and Programs in Biomedicine, vol. 197, December 2020.

[18] N. Zhou, A. G. Constantinides, G. Huang, and S. Zhang, "Face Recognition Based on an Improved Center Symmetric Local Binary Pattern," Neural Computing and Applications, vol. 30, no. 12, pp. 3791-3797, 2018.

[19] L. Zhou, H. Wang, S. Lin, S. Hao, and Z. M. Lu, "Face Recognition Based on Local Binary Pattern and Improved Pairwise-Constrained Multiple Metric Learning," Multimedia Tools and Applications, Vol. 79, no. 1-2, pp. 675-691, 2020.

[20] O. García-Olalla, E. Alegre, L. Fernández-Robles, M. T. García-Ordás, and D. García-Ordás, "Adaptive Local Binary Pattern with Oriented Standard Deviation (ALBPS) for Texture Classification," EURASIP Journal on Image and Video Processing, vol. 2013, no. 1, May 2013.

[21] M. A. Muqeet and R. S. Holambe, "Local Binary Patterns Based on Directional Wavelet Transform for Expression and Pose-Invariant Face Recognition," Applied Computing and Informatics, vol. 15, no. 2, pp. 163-171, July 2019.

[22] C. H. Yoo, S. W. Kim, J. Y. Jung, and S. J. Ko, "High-Dimensional Feature Extraction Using Bit-Plane Decomposition of Local Binary Patterns for Robust Face Recognition," Journal of Visual Communication and Image Representation, vol. 45, pp. 11-19, May 2017.

[23] X. Zhao and S. Zhang, "Facial Expression Recognition Using Local Binary Patterns and Discriminant Kernel Locally Linear Embedding," EURASIP Journal on Advances in Signal Processing, vol. 2012, no. 1, January 2012.

[24] Z. Chai, H. Mendez-Vazquez, R. He, Z. Sun, and T. Tan, "Explore Semantic Pixel Sets Based Local Patterns with Information Entropy for Face Recognition," EURASIP Journal on Image and Video Processing, vol. 2014, no. 1, May 2014.

[25] F. V. Massoli, G. Amato, and F. Falchi, "Cross-Resolution Learning for Race Recognition," Image and Vision Computing, vol. 99, July 2020.

[26] A. Dahmouni, K. El Moutaouakil, and K. Satori, "Face Description Using Electric Virtual Binary Pattern (EVBP): Application to Face Recognition," Multimedia Tools and Applications, vol. 77, no. 20, pp. 27471-27489, April 2018.

[27] L. Zhou, H. Wang, W. Liu, and Z. M. Lu, "Face Feature Extraction and Recognition via Local Binary Pattern and Two-Dimensional Locality Preserving Projection," Multimedia Tools and Applications, vol. 78, no. 11, pp. 14971-14987, 2019.

[28] J. Liang, H. Tu, F. Liu, Q. Zhao, and A. K. Jain, "3D Face Reconstruction from Mugshots: Application to Arbitrary View Face Recognition," Neurocomputing, vol. 410, pp. 12-27, October 2020.

[29] C. C. Chude-Olisah, G. Sulong, U. A. Chude-Okonkwo, and S. Z. Hashim, "Face Recognition via Edge-Based Gabor Feature Representation for Plastic Surgery-Altered Images," EURASIP Journal on Advances in Signal Processing, vol. 2014, no. 1, July 2014.

[30] A. Bolotnikova, H. Demirel, and G. Anbarjafari, "Real-Time Ensemble Based Face Recognition System for NAO Humanoids using Local Binary Pattern," Analog Integrated Circuits and Signal Processing, vol. 92, no. 3, pp. 467-475, June 2017.

[31] Z. Liu, X. Song, and Z. Tang, "Fusing Hierarchical Multi-Scale Local Binary Patterns and Virtual Mirror Samples to Perform Face Recognition," Neural Computing and Applications, vol. 26, no. 8, pp. 2013-2026, March 2015.

[32] Z. Xu, Y. Jiang, Y. Wang, Y. Zhou, W. Li, and Q. Liao, "Local Polynomial Contrast Binary Patterns for Face Recognition," Neurocomputing, vol. 355, pp. 1-12, August 2019.

[33] D. Huang, C. Shan, M. Ardabilian, Y. Wang, and L. Chen, "Local Binary Patterns and its Application to Facial Image Analysis: A Survey," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 41, no. 6, pp. 765-781, November 2011.

[34] L. Wisesa, "OpenCV Face Recognition Berbasis Algoritma Haar Cascade," https://medium.com/purwadhikaconnect/opencv-face-recognition-berbasis-algoritma-haar-cascade-1a5e23d9e8cb, May 04, 2020.

[35] A. Mordvintsev and A. Rahman, "Histograms-2: Histogram Equalization," https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_imgproc/py_histograms/py_histogram_equalization/py_histogram_equalization.html, May 04, 2020.

[36] J. J. Priestley, V. Nandhini, and V. Elamaran, "An Improved Trimmed Median Filter for the Restoration of Images Corrupted by High Density Impulse Noise," International Journal of Applied Engineering Research, vol. 10, no. 5, pp. 11587-11598, January 2015.

[37] X. Fu and W. Wei, "Centralized Binary Patterns Embedded with Image Euclidean Distance for Facial Expression Recognition," Fourth International Conference on Natural Computation, IEEE Press, vol. 4, pp. 115-119, October 2008.

[38] K. Kavita and M. Kau, "A Survey Paper for Face Recognition Technologies," International Journal of Scientific and Research Publication, vol. 6, no. 7, pp. 441-445, 2016.

[39] "Advantages and Disadvantages of Machine Learning Language," https://data-flair.training/blogs/advantages-and-disadvantages-of-machine-learning, November 23, 2020.

[40] M. Wang and W. Deng, "Deep Face Recognition: A Survey," https://arxiv.org/abs/1804.06655v9, November 23, 2020.