WINDOWING SYSTEM FACIAL DETECTION BASED ON GABOR KERNEL FILTER, FAST FOURIER TRANSFORM, AND PROBABILISTIC LEARNING VECTOR QUANTIZATION

Arif Muntasa[1*]

[1]*Informatics Engineering Department, Faculty of Engineering, University of Trunojoyo Madura, Bangkalan, Madura 69162 Indonesia*

ABSTRACT

Facial detection is a crucial stage in the facial recognition process. Misclassification during the facial detection process will impact recognition results. In this research, windowing system facial detection using the Gabor kernel filter and the fast Fourier transform was proposed. The training set images, for both facial and non-facial images, were processed to obtain the local features by using the Gabor kernel filter and the fast Fourier transform. The local features were measured using probabilistic learning vector quantization. In this process, facial and non-facial features were classified using label 1 and -1. The proposed method was evaluated using facial and non-facial image testing sets, which were taken from the MIT+CMU image database. The testing images were enhanced first before the detection process using four different enhancement methods: histogram equalization, adaptive histogram equalization, contrast limited adaptive histogram equalization, and the single-scale retinex method. The detection results demonstrated that the highest average accuracy was 83.44%.

1. INTRODUCTION

Over the last decade, biometrics has been an interesting issue in computer science and computer vision. The results of biometrics research have been implemented in many areas, including the corporate sector, defense and security, governance, and even in small gadgets. Biometrics have also been developed to obtain new breakthroughs, such as fingerprint recognition (Ito et al., 2006; Tan & Bhanu, 2006), facial detection (Yang et al., 2002; Rowley et al., 1998; Hjelmas & Low, 2001; Viola & Jones, 2004), automatic face-image data acquisition system (Wahab et al., 2015), and facial recognition (Muntasa, 2014; Muntasa, 2015; Shehzad et al., 2014). Facial detection plays an important role in facial recognition, as the results of facial detection will affect facial recognition such that an incorrect detection of a facial image will result in an incorrect facial recognition.

Facial detection based on color distribution has been conducted by various researchers (Zhang et al., 2014; Ghimire & Lee, 2013; Singh et al., 2003; Hsu et al., 2002; Tieu & Viola, 2000; Kienzle et al., 2004; Zou, 2010; Vijay & Patil, 2010; Ali & Sedigheh, 2011; Yang et al., 2002). This method has several limitations. First, non-facial image grayscale similarities with facial image grayscale have caused incorrect detection. Second, as people have different facial colors,

*Corresponding author's email: arifmuntasa@if.trunojoyo.ac.id, Tel. +62- 031-3011146, Fax. +62- 031-3011506

if the facial color found was significantly different than the threshold value given, the system would produce a false detection. Third, the detection results are significantly affected by lighting. Fourth, the method does not work correctly for multiple faces in an image. Last, the use of the threshold value cannot guarantee the accuracy of detection results, because of lighting effects.

Color-based face detection uses the color distribution and statistical theories, such as averages and standard deviations, as the bases for identifying whether an image is facial or non-facial has been performed by many researchers (Zou, 2010; Ghimire & Lee, 2013; Ali & Sedigheh, 2011). This method does not consider features that can represent more detailed characteristics, whereas the characteristics of the object viewed from various angles will produce more dominant features than the statistical method, such as Gabor kernel features. The weakness of this method is that areas similar to the face will be deemed facial because there are no parameters other than color distribution to classify the area.

In addition to color distribution, edge detection, such as the Canny method, is also used in facial detection (Vijay & Patil, 2010). This method works optimally for a single face, because the method uses object indentation as a reference to identify whether an image is facial or non-facial. However, the results of the edge detection often produce edges of objects that are not connected in an object. Therefore, it is necessary to connect the edges of disconnected objects. This method cannot be applied to detect many faces in one image, however, because many edges of objects are disconnected. Accordingly, many objects are falsely identified as faces. In addition to not being suitable for multi-face images, the method is also not suitable for face detection in real time, because the computation time is quite long.

In this research, the Gabor kernel filter and the fast Fourier transform were proposed to detect the facial location in images. The Gabor kernel filter was utilized to gain the dominant facial features; therefore, facial and non-facial dominant features were separated by projection of the Gabor kernel filter and the fast Fourier transform. During the testing process, the test images were searched to find the centers of facial location candidates. The centers of the facial location candidates that were as large as the size of the training set were captured based on the windowing system. The dominant features were then established. The dominant features results from the testing sets were compared to the dominant features results from the training sets.

There are three important processes in detecting a face: enhancing the image quality, extracting the features, and measuring the resemblance of the object. To enhance the image quality, the single-scale retinex is proposed to remove the lighting effects. Although some other methods were used to enhance the image, including histogram equalization, adaptive histogram equalization, and contrast limited adaptive histogram equalization. However, the single-scale retinex method produced the best image result.

To extract the features, a new approach is proposed using projection of the Gabor kernel filter and the fast Fourier transform to separate facial and non-facial features. Projection of the Gabor kernel filter and the fast Fourier transform applies eight orientations and five scales to gain 40 different viewpoints. These viewpoints represent the object, for both facial and non-facial features.

A new method to measure the resemblance of the object is also proposed: probabilistic learning vector quantization (Prob-LVQ). Prob-LVQ is a solution for the learning vector quantization (LVQ) weakness in measuring the resemblance of two objects, where changes in weight are based on Bayesian characteristic values by taking the greatest probability value. LVQ is a supervised learning method that uses the shortest distance as the basis for updating weights. LVQ cannot separate two objects that are geometrically similar, but are far apart in distance. Prob-LVQ overcomes this limitation.

The remainder of the paper includes sections on related work, the proposed method, the experimental results and discussion, and conclusions.

2. RELATED WORK

The Gabor kernel filter and the fast Fourier transform are widely used in signal analysis. The Gabor kernel filter is a sinusoidal complex that is combined with the Gaussian envelope, whereas the fast Fourier transform is the efficient form of the discrete Fourier transform in the signal domain.

2.1. Gabor Kernel Filter

The training sets can be classified into two categories: facial and non-facial images. A facial or non-facial image is represented by using $I(x, y)$. In this case, image input $I(x, y)$ is true color or grayscale image. The use of the Gabor kernel filter in facial representation has been widely developed by many researchers (Avinash & Raina, 2010). The Gabor kernel filter can characterize the dominant facial features in the space domain. In this research, to distinguish the facial and non-facial images, the proposed method uses the Gabor kernel filter, which has five scales and eight orientations. The Gabor kernel filter $G(x, y)$ is the result of multiplication of the Gaussian envelope and the Gabor sinusoidal complex as follows:

$$G(\sigma, \omega, \theta, x, y) = s(\sigma, x, y) \times w(\omega, \theta, x, y) \tag{1}$$

In this case, the Gaussian envelope and the Gabor sinusoidal complex can be defined as:

$$s(\sigma, x, y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{2}$$

and

$$w(\omega, \theta, x, y) = \exp(j(2\pi\omega(x\cos\theta + y\sin\theta))) \tag{3}$$

The parameter of σ depicts the standard deviation of the Gaussian envelope. The parameter of θ represents the orientation of an object. In this case, the orientation values are 0°, 22.5^o, 45°, 67.5^o, 90°, 112.5^o, 135°, 147.5^o, and 180°.

2.2. Discrete Fourier Transform

The two-dimensional fast Fourier transform was used to obtain the signal domain of the Gabor kernel filter. The results of the Gabor kernel filter were further processed using the fast Fourier transform. The fast Fourier transform is a method to calculate the discrete Fourier transform. The two-dimensional discrete Fourier transform can be defined as follows:

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-i2\pi(\frac{ux}{M} + \frac{vy}{N})} \tag{4}$$

To reverse the original object, the discrete Fourier transform has inverse expressed as follows:

$$f(x, y) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} F(u, v) e^{i2\pi(\frac{ux}{M} + \frac{vy}{N})} \tag{5}$$

The discrete Fourier transform has $O(N^2)$ time complexity. The fast Fourier transform has a smaller time complexity than the DFT, which is $O(N \log N)$. It can be explained from the one-dimensional discrete Fourier transform as follows:

$$F(u) = \frac{1}{2}(\eta + \rho) \tag{6}$$

where

$$\rho = \frac{1}{N}\sum_{x=0}^{N-1} f(2x)k_{2N}^{ux} \text{ and } \eta = \frac{1}{N}\sum_{x=0}^{N-1} f(2x+1)k_N^{ux}k_{2N}^{u}$$

Furthermore, Equation 6 is separated into two parts: F_{even} and F_{Odd}.

$$F(u) = \frac{1}{2}(F_{even}(u) + k_{2N}^{u} F_{odd}(u)) \tag{7}$$

Equation 7 can be decomposed hierarchically. The decomposition result is called the fast Fourier transform.

3.   PROPOSED METHOD

In this research, windowing model facial detection based on the Gabor kernel filter and the fast Fourier transform is proposed. The proposed method was initiated by generating the Gabor filter bank using five scales and eight orientations, followed by the fast Fourier transform process. Second, the training sets were enhanced using the single scale retinex method. Third, the dominant features were trained using probabilistic learning vector quantization. Last, the facial area candidates were tracked and followed by detection of the facial locations. See the following training algorithm:

Input: Facial and non-facial images
Output: Facial region of interest
Process:
  1.   Enhance the training sets;
  2.   Create Gabor filter bank and calculate fast Fourier transform; and
  3.   Create the features and training process using probabilistic learning vector quantization

3.1.   Image Enhancement of the Testing Sets
Four image enhancement methods were used to improve the testing set: histogram equalization, adaptive histogram equalization (AHE), contrast limited adaptive histogram equalization (CLAHE), and the single-scale retinex method. CLAHE reduces the image noise by the clipping limit. The clipping limit and block size will affect the image enhancement results. CLAHE is an improvement of AHE, and AHE is an improvement of the histogram equalization method. CLAHE can reduce the noise of an image so that the image can be displayed clearly, but CLAHE can overcome images with lighting effects. The single-scale retinex method is the best image enhancement method for improving image quality due to lighting effects.

3.1.1.   Histogram equalization
The image quality can impact the recognition process. In the real word, the captured results often produce low-quality images. The lower the image quality is, the less recognition or detection can be performed. Histogram equalization is an image enhancement method in which

the global contrast is increased by adjustment of the image intensity so that the enhanced image has better gray distribution than the original image.

3.1.2. Adaptive histogram equalization

Adaptive histogram equalization (AHE) is an improvement of the histogram equalization method, as AHE produces better contrast enhancement. AHE computes some image histograms and corresponds them to distinct image areas. Each pixel is updated based on the surrounding pixel area. AHE is calculated through an interpolation process as follows:

$$m(i) = a\big[bm_{-,-}(f) + (1-b)m_{+,-}(f)\big] + \big[1-a\big]\big[bm_{-,+}(f) + (1-b)m_{+,+}(f)\big] \tag{8}$$

In this case, $a=(y-y_-)/(y_+-y_-)$ and $b=(x-x_-)/(x_+-x_-)$. If f states the image intensity on the (x, y) position, then the right upper mapping of $x_{+,-}$ can be stated by using $m_{+,-}$. The right lower mapping of $x_{+,+}$ can be represented by $m_{+,-}$. The right lower mapping of $x_{-,+}$ is stated by $m_{-,+}$, whereas the left lower mapping of $x_{-,-}$ is stated by $m_{-,-}$.

3.1.3. Contrast limited adaptive histogram equalization

Adaptive histogram equalization has a limitation. It also enhances image noise, so the image enhancement result is not optimal. The CLAHE method is an improvement to the AHE method, as it can suppress the noise during the enhancement process. The CLAHE method enhances the image quality based on a small region, and it is adaptively conducted by an interpolation process. The CLAHE results are better than the AHE and histogram equalization results.

3.1.4. Single-scale retinex

Sometimes, an image with a lighting effect cannot be enhanced using CLAHE. The single-scale retinex method can be used to handle such images. As mentioned above, the single-scale retinex method was utilized here to enhance image quality and normalize lighting effects. The output of the process depends on the input and the nearest neighbor values. The single-scale retinex method can be modeled by the following equation:

$$R(x, y) = \alpha\big(\log(I(x, y)) - \log(I(x, y) \times F(x, y))\big) - \beta \tag{9}$$

The result of an image enhancement was further processed using the fast Fourier transform method and followed Gabor filter bank by using the dot product.

3.2. Creating Features of the Training Sets

Facial and non-facial features must be generated by multiplication between the Gabor filter bank and the fast Fourier transform of the training sets as follows:

$$F = G_{if}(\sigma, \omega, \theta, x, y) \times I_f(x, y) \tag{10}$$

In this case, $I_f(x, y)$ depicted the fast Fourier transform of the training set images, for both facial and non-facial images.

3.3. Training Process using Probabilistic Learning Vector Quantization (Prob-LVQ)

The similarity measurements of feature extraction have been conducted by many researchers (Ai et al., 2001; Li et al., 2001). Learning vector quantization (LVQ) is a similarity measurement based on supervised learning. It classifies objects by a learning process. The classification results usually depend on the minimum distance. However, the minimum distance cannot correctly classify when the class members are not near the center of the class. The probabilistic method is proposed to obtain the maximum probability value and to enhance the limitation of the class members. The maximum probability used is influenced by all similarity measurement parameters. In this case, the naïve Bayesian method is proposed to obtain the

highest probability of the class members. The proposed method can resolve those classification cases where the class members are not close to the class center. Suppose an $x(m,n)$ input with a $T(1, n)$ target. Furthermore, the J value can be determined by the maximum probability of the naïve Bayesian equation as follows:

$$p_j(x = v \,|\, c) = \frac{1}{\sqrt{2\pi\sigma_c^2}} e^{\frac{(v-\mu_c)^2}{2\sigma_c^2}} \tag{11}$$

Modification of the minimum distance to the maximum probability can be used to handle cases where the member locations move away from the class center, as shown in Figure 1. It shows that the training sets are the facial images, and the training results are labeled with "*Face=1*" and "*Non-Face=-1*." If the training sets are non-facial images, then the value of "*Face=-1*" and "*Non-Face=1*." The maximum probability value correlates to the c_j value. Equation 11 represents the probability for the jth object, where $\forall j, j \in 1. . .n$. In this case, n represents the number of training sets. If it is implemented as referenced to update the weight, then the weight can be updated by using the equation:

$$w = \begin{cases} w + \alpha(x - w) & if \quad T = p_{max}(x = v \,|\, c) = \dfrac{1}{\sqrt{2\pi\sigma_c^2}} e^{\frac{(v-\mu_c)^2}{2\sigma_c^2}} \\[4mm] w - \alpha(x - w) & if \quad T \ne p_{max}(x = v \,|\, c) = \dfrac{1}{\sqrt{2\pi\sigma_c^2}} e^{\frac{(v-\mu_c)^2}{2\sigma_c^2}} \end{cases} \tag{12}$$

The difference of the proposed method compared to the LVQ is that the LVQ uses the minimum distance, while the proposed method uses the maximum probability; therefore, the value of $p_{max}(x = v \,|\, c)$ can be found as the maximum value of the probabilistic learning for the training sets.

$$\begin{aligned} p_{max}(x = v \,|\, c) &= \max \arg\big(p_j(x = v \,|\, c)\big) \quad \text{and } \forall j, j \in 1, 2, \cdots, n \\ &= \max \arg\big(p_1(x = v \,|\, c) \quad p_2(x = v \,|\, c) \quad \cdots \quad p_n(x = v \,|\, c)\big) \end{aligned} \tag{13}$$

Based on Equations 12 and 13, the new architecture for the training process using Prob-LVQ can be described as in Figure 1.
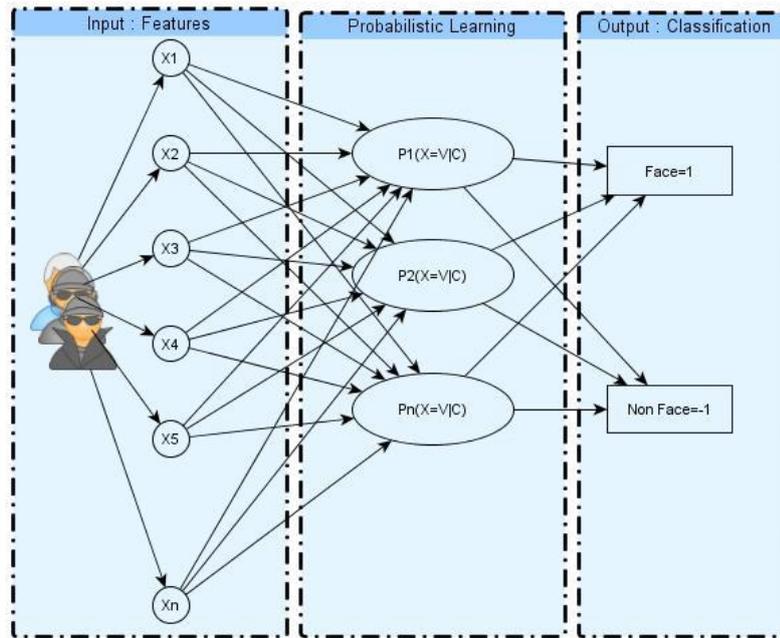
Figure 1 The proposed method, new architecture of the probability learning vector quantization

For each iteration process, the value of α will be decreased. The process is stopped when the value of $\alpha > \xi$, where ξ represents the minimum error rate desired. Figure 1 displays the new architecture of the probability learning vector quantization, where the weight update is based on the maximum probabilistic values of the training sets. The proposed method can overcome the limitation of the LVQ. In the proposed method, though two objects have longer distances than the other, but they have the greatest probabilistic, then the two objects are considered similar.

3.4. Find the Facial Location Candidate and Detect Facial Location

Three facial images of the training sets were normalized into a particular range: -1 until 1. A similar process was also conducted for the testing sets. Next, the testing set was convoluted using three masking matrixes, followed by the selection of maximum value. The selection result was filtered to select the values. The pixel $P(x,y)$ was then checked, followed by feature extraction and mapping using probabilistic learning vector quantization. If the value was equal to 1, then the value was considered as the center of a facial area candidate. Otherwise, the value was assumed as a non-facial image. The center of a facial region candidate was used to capture the original image based on a windowing system as significant as the size of the training sets. The facial area candidate's dominant features were computed using Equation 7. The dominant features were compared to the training sets using probabilistic learning vector quantization. If the output of the probabilistic learning vector quantization was smaller than a threshold value, then the center of the facial region candidate was assigned to the center of the facial area.

Input: The testing set images
Output: Facial region of interest
Process:
1. Image enhancement using the single-scale retinex method;
2. Output of the 1st stage is normalized (F) into -1 until 1 range;
3. For each k, $k \in$ 1, 2, 3, and 4 do the following
   a. Normalize the training set image (I_k); and
   b. Perform the testing set convolution using I_k as the masking matrix;
4. Select the maximum values for each pixel position $P(x,y)$ of the testing convolution results;
5. Scan the results of the 3rd stage;

6. Initialize c=1;
7. Check the pixel value $P_c(x,y)$;
   a. If the pixel position $P_c(x,y)$ is equal to 1, then $P_c(x,y)$ is considered as the center of a facial location candidate, and the value of $c=c+1$;
   b. Otherwise, the value is assumed as a non-facial image candidate;
8. For each m, $m \in 1, 2, \ldots, c$ do the following:
   a. Capture the original image based on a windowing system as big as the size of the training sets using pixel position $P_c(x,y)$;
   b. Calculate the dominant features using Equation 7;
   c. Test the dominant features using probabilistic learning vector quantization as in Equations (11), (12), and (13);
   d. Check the value of the testing result (V_c);
      i. If $V_c \leq$ threshold value, then the center of the facial location candidate is assigned as the center of a facial location;
      ii. Otherwise, the center of the facial location candidate is assumed as non-facial.

3.5. Image Threshold

The threshold of the testing image was required to separate the facial image from the background. The Otsu method was proposed to obtain the best threshold. The Otsu method can be assumed in stationary statistics and modified into locally adaptive values. The Otsu method can be started by calculation of within class variance as follows:

$$\sigma_W^2 = W_b^2 \sigma_b^2 + W_f^2 \sigma_f^2 \tag{14}$$

Between class variance can be computed by the following equation:

$$\begin{aligned}
\sigma_B^2 &= \sigma^2 - \sigma_W^2 \\
&= W_b \left(\mu_b - \mu\right)^2 + W_f \left(\mu_f - \mu\right)^2 \\
&= W_b W_f \left(\mu_b - \mu_f\right)^2
\end{aligned} \tag{15}$$

The best threshold takes the largest value of the "between class variance" or the smallest "within class variance."

4. RESULTS AND DISCUSSION

In this research, two image categories were used as the training sets: facial and non-facial images. Both facial and non-facial images were taken from the MIT+CMU image database, including 4,916 facial images and 7,960 non-facial images. Each image has a dimension, i.e. 24 pixels for height and 24 pixels for width. Figures 2 and 3 demonstrate the facial and non-facial image training sets. In the training process, the features searching process took 417.21 seconds for the 4,916 facial images and 665.37 seconds for the 7,960 non-facial images. The average training process thus required 0.08 seconds for each facial image and 0.083 seconds for each non-facial image. The training process using probabilistic learning vector quantization took 712.4 seconds. Thirty images from the MIT+CMU test set image database were randomly selected and used to evaluate the performance of the proposed method. The performance evaluation was conducted for each image. The values for sensitivity, specificity, and accuracy were calculated.

Figure 2 Facial image sample from the MIT+CMU training set

Figure 3 Non-facial image sample from the MIT+CMU training set

The experiment was conducted using four different scenarios. The testing set images were enhanced using histogram equalization, adaptive histogram equalization, contrast limited adaptive histogram equalization, and the single-scale retinex method. During the scanning process, if the detection results produced false detections, the facial candidate locations could not be detected correctly because of the many pixel similarities around the facial candidate locations. It can be proved that the detection results have found a non-facial object, but it is detected as a facial object. The experimental results have demonstrated that the single-scale retinex method produced the highest accuracy, as compared to the other methods (see Table 1). The proposed method demonstrated the facial detection results on the testing set images. The detection results proved that the detection accuracy was affected by the image enhancement results.

Table 1 The facial detection results of the proposed method (Prob-LVQ)

Experimental Scenario	Percentage		
	Sensitivity	Specificity	Accuracy
Histogram Equalization	81.20	72.31	76.76
AHE	83.26	72.88	78.07
CLAHE	84.44	74.95	79.70
Single-Scale Retinex	87.13	79.74	83.44

The experimental results using histogram equalization produced the lowest accuracy at 76.76%, whereas the facial detection results with adaptive histogram equalization, contrast limited adaptive histogram equalization, and single-scale retinex enhancements were 78.07%, 79.70%, and 83.44%, respectively. The proposed method has shown that an image enhancement process influences facial detection results. In Figure 4, five facial images could not be detected, and three non-facial images were identified as facial images. The similarity of the grayscale of a non-facial and a facial image can cause such false detection. Similar results also occurred as

seen in Figure 5. Feet and hands have similar grayscales as facial images, thus they have also caused false detections.
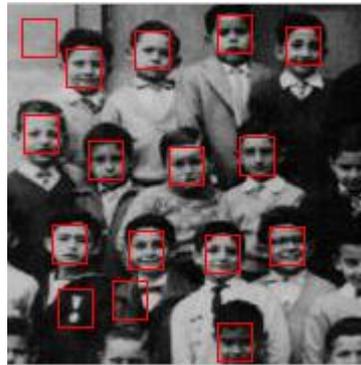
Figure 4 Facial detection results using the proposed method (True Positive=13, False Positive=3, and False Negative=5)
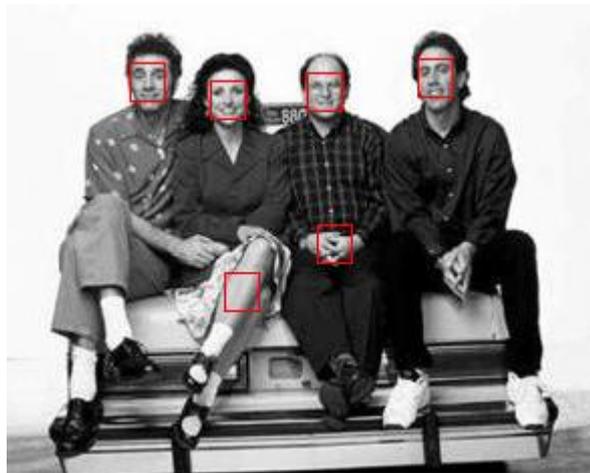
Figure 5 Facial detection results using the proposed method (True Positive=4, False Positive=2)

Figure 6 shows the detection results of the proposed method, which accurately detected eight facial images (100% true positive). The proposed method was also evaluated using other images from the World Wide Web, as seen in Figures 7 and 8. In Figure 7, the proposed method correctly detected seven facial images, two facial images were not detected as facial, and one non-facial image was detected as a facial image. Similar results are shown in Figure 8. The proposed method identified seven facial images as facial images, and one non-facial image was identified as a facial image.
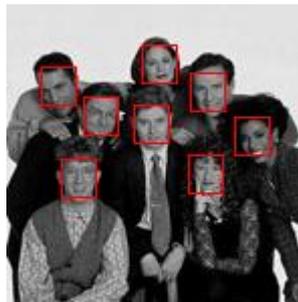
Figure 6 Facial detection results using the proposed method (True Positive=8)
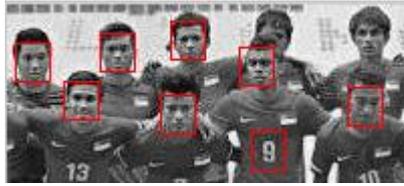
Figure 7 Facial detection results using the proposed method (True Positive=7, False Positive=1, and False Negative=2)

Figure 8 Facial detection results using the proposed method (True Positive=7, False Positive=1)

The results demonstrate that false detections are caused by several factors. First, the window found could have similar characteristics to the facial image training; therefore, the window area would be classified as a facial image, as shown in Figure 8, where the head of a camel is regarded as a face. Second, besides having similar characteristics, the window found may also have a similar grayscale level distribution to the training sets of the facial images. If the window found has similar grayscale levels only, then the window found would be supposed as non-facial, but if it also has similar characteristics, then the window found would be detected as a face, as shown in Figure 7. Last, the Gabor kernel filter and the fast Fourier transform can numerically represent characteristic features, but they cannot geometrically represent the window area found; therefore, the similarity of the grayscale level distribution and characteristics will be more accurate if the features are supported by a geometric base.

To show the performance of the proposed method, the sample sets were also evaluated using an original LVQ, as shown in Table 2. Overall, the Prob-LVQ, as the proposed method, has delivered better results than LVQ, as in Figure 9. The results indicate that the Prob-LVQ method can improve the LVQ method. The same features were used to represent the object and were measured by both methods, Prob-LVQ and LVQ.

Table 2 The facial detection results of the LVQ

Experimental Scenario	Percentage		
	Sensitivity	Specificity	Accuracy
Histogram Equalization	78.40	67.42	72.91
AHE	77.67	73.21	75.44
CLAHE	84.03	72.04	78.03
Single-Scale Retinex	82.65	77.43	80.04
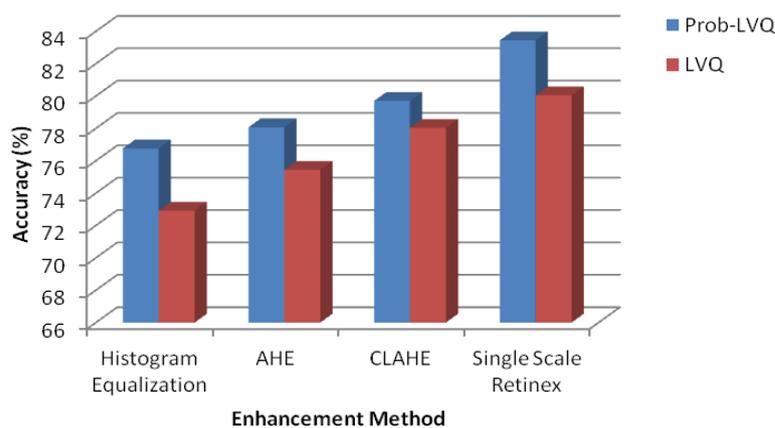
Figure 9 Performance of the Prob-LVQ as the proposed method and the LVQ

5.   CONCLUSION

The proposed method has demonstrated the highest accuracy of facial detection with 83.44%. The detection rate was affected by image quality. Therefore, the image enhancement process played an important role in the facial detection. The better the image enhancement results were, the higher the facial detection results. The results also proved that the LVQ enhancement (the Prob-LVQ) outperformed the original LVQ in all scenarios. The proposed method demonstrated that the highest accuracy was found using the single-scale retinex image enhancement method, whereas the lowest accuracy was produced when using histogram equalization. However, the proposed method also has several limitations. First, the proposed method was only able to detect front-view facial images. Second, the computation time for the training and the testing sets is quite high This result was caused by the large training set, the large number of iterations, and the small error rate value. The larger the training set used, the smaller the error rate value, and the greater the number of iterations, the longer the computing time.

6.   REFERENCES

Ai, H., Liang, L., Xu, G., 2001. Face Detection based on Template Matching and Support Vector Machines. *In:* Proceedings of the International Conference on Image Processing (ICIP'01), October 7-10, 2001, Thessaloniki, Greece, pp. 1006–1009

Ali, A., Sedigheh, G., 2011. Robust Component-based Face Detection using Color Feature. *In:* Proceedings of the World Congress on Engineering 2011, Volume II WCE 2011, 6–8 July, London, U.K.

Avinash, K., Raina, J.P.S., 2010. Face Detection using Neural Network & Gabor Wavelet Transform. *International Journal of Computer Science and Technology*, Volume 1(1), pp. 58–63

Ghimire, D., Lee, J., 2013. A Robust Face Detection Method based on Skin Color and Edges. *Journal of Information Processing System*, Volume 9(1), pp. 141–156

Hjelmas, E., Low, B.K., 2001. Face Detection: A Survey. *Computer Vision and Image Understanding*, Volume 83(3), pp. 236–274

Hsu, R.L., Abdel-Mottaleb, M., Jain, A.K., 2002. Face Detection in Color Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 24(5), pp. 696–706

Ito, K., Morita, A., Aoki, T., Nakajima, H., Kobayashi, K., Higuch, T., 2006. A Fingerprint Recognition Algorithm Combining Phase-based Image Matching and Feature-based Matching. *In:* International Conference on Biometrics, Volume 3832, pp. 316–325

Kienzle, W., Bakir, G., Franz, M., Schölkopf, B., 2004. Face Detection—Efficient and Rank Deficient. *In*: Proceeding of Advances in Neural Information Processing Systems 17, December 2004, Vancouver, Canada. pp. 673–680

Li, S.Z., Fu, Q., Gu, L., Scholkopf, B., Cheng, Y., Zhang, H., 2001. Kernel Machine based Learning for Multi-view Face Detection and Pose Estimation. *In:* Proceedings of the 8th International Conference on Computer Vision and Pattern Recognition (ICCV'01), Volume 2, pp. 674–679

Muntasa, A., 2014. New Modeling of Modified Two-dimensional Fisherface based Feature Extraction. *Telkomnika*, Volume 12(1), pp. 115–122

Muntasa, A., 2015. Facial Recognition using Square Diagonal Matrix based on Two-Dimensional Linear Discriminant Analysis. *International Review on Computer and Software*, Volume 10(7), pp. 718–725

Rowley, H.A., Baluja, S., Kanade, T., 1998. Neural Network based Face Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 20(1), pp. 39–51

Shehzad M.I., Awais, M., Amin, M., Shah, Y.A., 2014. Face Recognition using Average Half Face Template. *International Journal of Technology*, Volume 5(2), pp. 159–168

Singh, S.K., Chauhan, D.S., Vatsa, M., Singh, R., 2003. A Robust Skin Color based Face Detection Algorithm. *Tamkang Journal of Science and Engineering*, Volume 6(4), pp. 227–234

Tan, X., Bhanu, B., 2006. Fingerprint Matching by Genetic Algorithms. *Pattern Recognition*, Volume 39(1), pp. 465–477

Tieu, K., Viola, P., 2000. Boosting Image Retrieval. *In:* Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 15 June 2000, Hilton Head Island, USA

Vijay, L.H.C., Patil, K.S., 2010. Segmentation Algorithm for Multiple Face Detection in Color Images with Skin Tone Regions using Color Spaces and Edge Detection Techniques. *International Journal of Computer Theory and Engineering*, Volume 2(4), pp. 1793–8201

Viola, P., Jones, M.J., 2004. Robust Real-time Face Detection. *International Journal of Computer Vision*, Volume 57(2), pp. 137–154

Wahab, W., Ridwan, M., Kusumoputro, B., 2015. Design and Implementation of an Automatic Face-image Data Acquisition System using IP Based Multi Camera. *International Journal of Technology*, Volume 6(6), pp. 1042–1049

Yang, M.H., Kriegman, D.J., Ahuja, N., 2002. Detecting Faces in Images: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 24(1), pp. 34–58

Zhang, Z.W., Wang, M.H., Lu, Z.M., Zhang, Y., 2014. A Skin Color Model based on Modified GLHS Space for Face Detection. *Journal of Information Hiding and Multimedia Signal Processing*, Volume 5(2), pp. 144–151

Zou, L., 2010. Face Detection in Color Images based on Skin Color Models. *In:* 2010 IEEE Region 10 Conference, pp. 681–686