Received: July 5, 2021. Revised: August 24, 2021.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Feature Representation Using Multi-scale Block Adaptive Local Ternary Pattern for Face Recognition

Syavira Tiara Zulkarnain[1] **Nanik Suciati[1]***

Department of Informatics, Institut Teknologi Sepuluh Nopember, Indonesia
* Corresponding author's Email: nanik@if.its.ac.id

Abstract: Face recognition is a personal identification system based on facial biometric data. By looking at the characteristics of the image data that has lighting variations and contains Gaussian/Poisson/Quantization noise when taken using a camera, facial recognition systems working in the real world with uncontrolled environmental conditions is a challenge. Local Ternary Pattern (LTP) is a representation of image features that are invariant to lighting, resistant to noise, but require manual threshold determination. In this study, we propose an automatic calculation of adaptive LTP's threshold using the statistical characteristic of image histograms. We also propose a Multi-scale Block Adaptive Local Ternary Pattern (MBALTP) which combines local features extracted using adaptive LTP and global features captured using multi-scale block for face recognition. We experimented on the Extended Yale-B dataset, which were collected under different lighting variations and contains noise. The experimental result demonstrated that the proposed method can improve the recognition performance with an accuracy of 98.87%.

Keywords: Face recognition, Multiscale block, Local ternary pattern, Adaptive threshold.

1. Introduction

A process that has goal to identify the object's characteristics and assigns label (expressed in words, images, memories of associated events, or in other form) is called recognition [1, 2]. In the last two decades, facial recognition became an active topic in image processing, computer vision, and pattern recognition [3]. This is caused by demand and wide exploration of commercial face recognition systems for application in law enforcement, human-computer interaction, and multimedia [4], forensics, e-learning, biometric authentication, health monitoring, and surveillance [5, 6]. Face recognition system can be either a verification or identification system. Though the advancement and the spread of those system reached out numerous fields, its robustness is still considered a challenge in the real world, especially to work in an uncontrolled environment [7].

The uncontrolled environment such as differences in lighting, noise, expression, and scaling, led to high variation in facial images. Lighting variations are caused by changes in illumination day and night or differences in lighting intensity. This variation is unavoidable and can hinder face recognition [8]. Other variation in an uncontrolled environment are noise [9]. Noise is caused by some congenital of the electronic devices, coloring, and sharpness of the image, some have the form of invisible fine grains or visible noise to obscure image information [10]. The types of noise found in an electronic device such as television are Gaussian, Poisson, and Quantization; noise found on coherent lighting are speckle and quantization [10]. Even though there is already a deep learning method to decrease those variations, but a discriminative facial descriptor that is well-defined still plays a dominant role in facial recognition application [4].

Facial recognition has two critical sub-problem: description feature and classifier [7], both of which have been significantly used as a subject of research and development. Generally, face extraction performs an important role. For example, if the generated facial features are not representative, then it is possible that even best classifier would fail to recognize faces. Finding and designing an effective

descriptor is still a big challenge considering the three objectives: computational efficiency, effective feature discriminant choice, and robustness against intra-person variation (including different lighting variations, head poses, expression, age, blurry images, noise, and occlusion) [7]. There are two feature extraction approaches, namely holistic feature and local feature [11, 12]. The holistic approach uses the whole face region to create a sub-region as a representation of the facial image. Local features involve local statistical values from the image sub-region and can recognize texture.

Local feature extraction that focuses on pattern became an active research field in the last few years because of its effectiveness and ease of extraction [13]. The handcrafted descriptor method such as Local Binary Pattern (LBP) is known to be more effective to retain statistical characteristics [14]. The problem of LBP is that it can't capture global features at uniform areas such as cheek and forehead [7, 8], [15] that may become a dominant representation. In addition, it is more sensitive to noise compared with other LBP variations such as Local Ternary Pattern (LTP), which was introduced by Tan and Triggs [8] for face recognition.

LTP comes with the advantage of being less sensitive to noise, therefore it is doing well when handling images with variety of noise. However, the problem of LTP is that we must manually assign the t threshold, which are usually done by fine-tuning process that takes a lot of time and resource. In addition, the t threshold obtained from fine-tuning doesn't always capture the uniform area that are usually obtained through global feature representation such as multi-scale block. Furthermore, capturing the global feature representation give a better result [16]. Consequently, there is a need to automatically assign the t threshold that are adaptive according to each image characteristics and able to capture global feature representation.

In this research, we present a new representation for face recognition that consists of two methods. First, we propose an adaptive LTP threshold that is calculated automatically based on statistical characteristics of the image histogram. The adaptive LTP representation is robust against noise and lighting variation. Second, we propose Multi-scale Block Adaptive Local Ternary Pattern (MBALTP) which enriches the feature with a global feature representation using combination of the adaptive LTP and multiscale block. We use the Extended Yale B dataset for experiment.

This paper is arranged into several sections: Section 2 is a related work that refers to normalization, state-of-the-art local binary pattern, and its variations. Section 3 contains a brief explanation of the process that we used and the proposed methods. The own experimental setting, results, and discussion are presented in Section 4. And the rest section containing the conclusion.

2. Related works

Preprocessing as an initial state, useful for preparing and produce stable images. Preprocessing consists of various methods, such as normalization. That method can be done by reducing contrast or enhance the contrast. Based on the research of Lee, Zhang, Li and He (2020) different lighting variation will bring discarded or unknown number of gamma distortion and yet estimating the number are very important to normalization process [17]. Nevertheless, automatically knowing the estimation of gamma distortion is a challenge. From research [17] the estimation of gamma distortion can be solved with a theoretical basis by maximizing differential entropy in blind inverse gamma correction (acronymized as AGT-ME, or Adaptive Gamma Transformation Method).

Ojala, Pietikäinen, and Harwood (1996) [18] introduced LBP, known to be the most successful statistical approach for texture representation and invariance against monotonous illumination [7]. The calculation of the original LBP value is based on a sliding window that is 3 × 3 sub-regions of images to get the neighborhood. The next step is comparing the center pixel p_c of sub-region with other neighbor's p_i (8 neighbors) using Eq. (1) and illustrated in Fig. 1 to get the LBP value. If the result of comparing p_c and p_i higher or equal to 0, the LBP code becomes 1. Then the LBP code becomes 0 with a vice versa condition.

However, the LBP method has an open challenge: it cannot capture a bigger scale structure and might

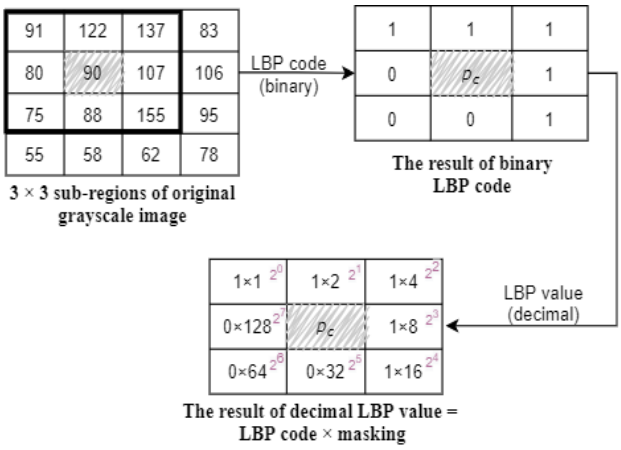

Figure. 1 LBP process flow

be the dominant facial feature, it is also unable to handle variations in appearance that are generally happen because of a pose, partial occlusion, and expression [19], LBP is also sensitive towards random noise at uniform areas such as cheek and forehead [7, 8, 15].

$$\sum_{j=0}^{8} 2^j \times f(p_i - p_c), f(\Delta p) = \begin{Bmatrix} 1, \Delta p \geq 0 \\ 0, otherwise \end{Bmatrix} \quad (1)$$

A new variation of LBP called Local Ternary Pattern (LTP) was introduced by Tan and Triggs [8] and is used for facial recognition, by using t threshold parameter [13,8]. LTP are proven less sensitive against noise compared to LBP [13], [14], [16]. LBP gives label to neighboring pixel as 0 or 1 using the middle pixel as a reference, and therefore sensitive towards the noise that appears in the image, such as Gaussian, quantization, or Poisson noise. On the other hand, LTP decreases the bad effect from noise because it quantizes in ternary code (1, 0, -1) that are produced by comparing value from center pixel p_c and pixel neighbor's p_i with threshold t into 2 patterns (high and low). A high pattern of LTP contains a ternary code with a constraint other than quantizing 1 replaced by 0. A low pattern of LTP contains ternary code with a constrain change quantize code 1 with 0 and change quantize code -1 with 1. In Farooque and Rohankar's research (2013) about testing on the proposed method and LTP, both of them successfully withstood against Gaussian noise [9]. LTP methods are illustrated in Fig. 2.

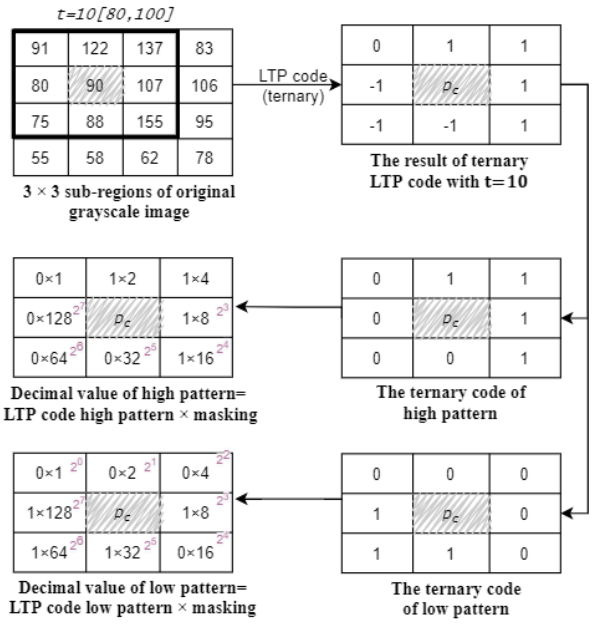

Figure. 2 LTP process flow

$$\sum_{j=0}^{8} \times f(p_i, p_c, t); f(p_i, p_c, t) \begin{Bmatrix} 1, p_i \geq p_c + t \\ 0, |p_i - p_c| < t \\ -1, p_i \leq p_c - t \end{Bmatrix} \quad (2)$$

The multi-scale block method [16] has been introduced to solve the LBP's limitations in capturing global features to enrich the features with macrostructure characteristics [20]. Those concepts are introduced to multi-scale block LBP [16] that are used to enlarge LBP's sub-region. The main difference between the multi-scale block with the original LBP is comparing the average intensity of the sub-region neighboring with the selected pixel and require scale determination constraining multiple of 3. If the scale of multi-block is 3 the sub-region is 1, those called the original LBP.

The histogram is a visualization of the distribution graph to determine the gray level variation of the image. Analyzing based on the histogram can be used for thresholding process to split up object and background pixel. The well-known thresholding approach is an Otsu [21]. That method applied the idea of reducing intra-class variance based on the bimodal histogram. Because the Otsu method assumes can work well in bimodal distribution (i.e., two populations), so there is a problem if bimodal has different sizes of classes. Also, the method doesn't work well with lighting variation [22]. Nevertheless, the Otsu thresholding is a powerful combination cause used gray-level characteristic based on statistic approach using variance as homogeneity of the region. Consider it, we decide to adopt other statistical calculations to define threshold improvement. Statistical measurement such as local variance known as standard deviation is a measure of the intensity of contrast with its neighbors [1]. That calculation quite favor cause based on [1] the standard deviation is more informative because it captures much less variability in intensity.

3. Method

Section 3, present the description of our proposed method, which is those flow illustrated in Fig. 3. Overall, our proposed method involves of 4 major steps as presented below:

1. Apply preprocessing method on the dataset to subtract and suppress noise and different lighting variation.
2. Initialize global sub-region scale.
3. Combining the global representation with local representation using the proposed method.
4. Use Support Vector Machine to classify and get evaluation result of face recognition.

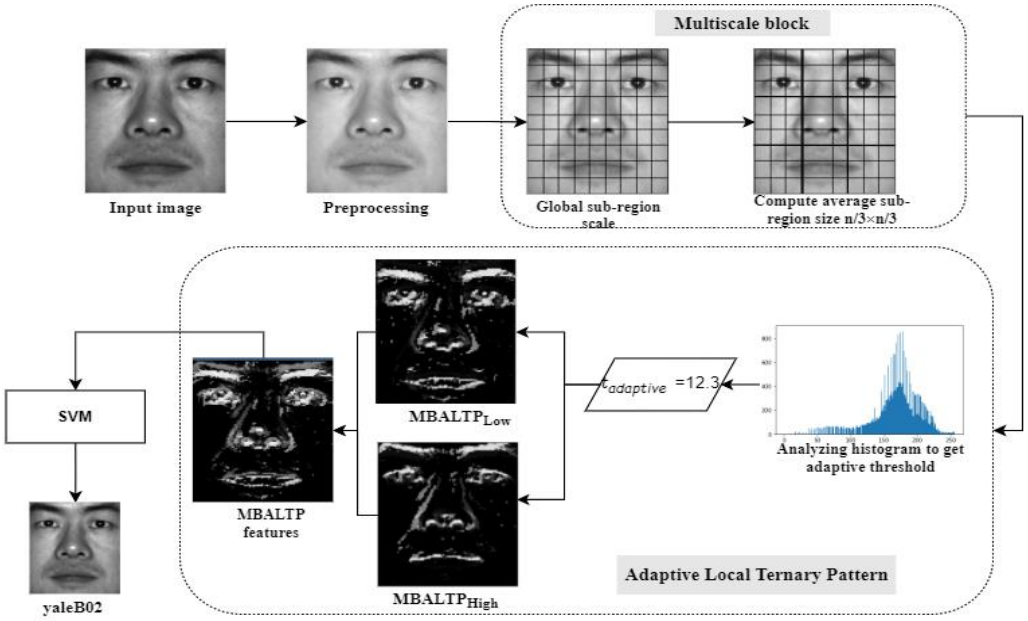

Figure. 3 Our proposed method's flow

3.1 Dataset

For experimental purposes and testing the robustness of the MBALTP method against different lighting conditions and noise, we use the Extended Yale B database [23] is an update from the previous database namely the Yale B dataset. Extended Yale B popular with face dataset that has challenging lighting variation. This dataset involves 38 participants, with a size image of 168 × 192 for each participant and divided into 5 subsets based on different point of view angle between light source and camera. The description of point-of-view's angle are located on the filename of each image on the dataset, formatted as Number of Subject_Number of Pose_Point of view angle with respect to the camera axis at azimuth degrees_Point
of view angle of elevation degrees. For example, filename yaleB01_P01A+005E+10, belongs to subject 01, Pose 01, the angle between light source and camera axis is at 5 degrees azimuth ('A+005') and 10 degrees elevation ('E+10'). The splitting of subset is shown in Table 1. If the angle formed between light source and camera (azimuth and elevation) is less than 12 degrees, than it is categorized as Subset 1, Subset 2 are from 12 degrees through 25 degrees, Subset 3 are form 26 degrees until 50 degrees, Subset 4 are form 51 degrees to 77 degrees and Subset 5 are above 78 degrees [24]. Therefore, the Subset 5 contains face image with the most extreme variation in lighting conditions, while the Subset 1 contains the less lighting variation images. The sample image of

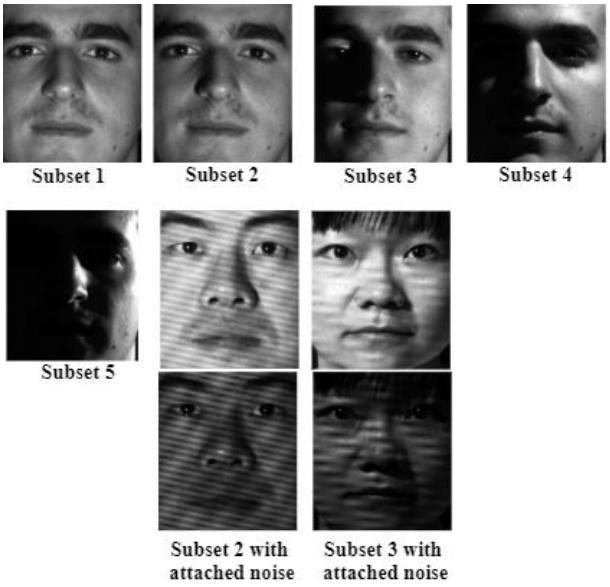

Figure. 4 The sample image of every subset in extended yale B dataset

Input: grayscale image I with pixel $\{(px_i, py_i)\}_{i=0}^{N}$

1 Check dimension of the image I
2 Apply preprocessing with normalize = min $\left(255 \times (I - \min I)/(\max I - \min I + 0.1)\right)$
3 Normalization pixel $\{(px_i, py_i)\}_{i=0}^{N}$ range 0-1 with formula Eq. (3)
4 Calculate optimal gamma using binary mask full image I
5 Apply gamma correction to image I with pixel $\{(px_i, py_i)\}_{i=0}^{N}$ based on previous gamma

Output: gamma estimation for contrast enhancement

Figure. 5 Normalization AGT-Me pseudocode [17]

Table 1 Splitting subset on Ext yale B dataset

Number of subsets	Point of view (angle of camera and light)	Total images
1	$\measuredangle < 12°$	263
2	$12° < \measuredangle < 25°$	456
3	$25° < \measuredangle < 50°$	455
4	$50° < \measuredangle < 77°$	526
5	$\measuredangle > 77°$	714

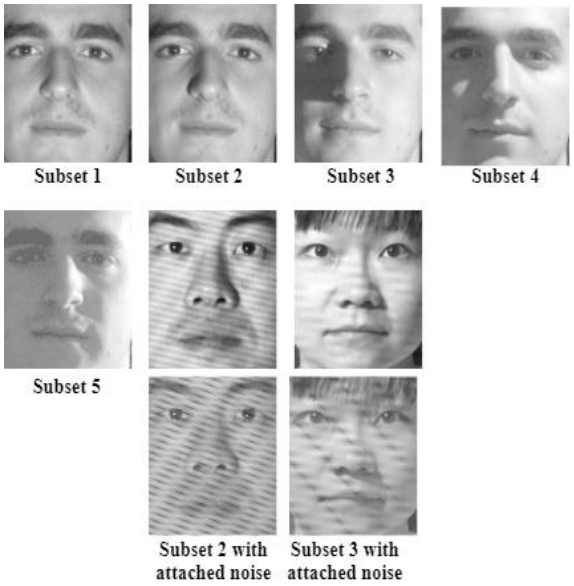

Figure. 6 The result of AGT-Me applied in sample image of every subset

each Subset presents in Fig. 4. The presented image shows that the dataset not only has extreme lighting variation but also has noise attached to images.

3.2 Preprocessing

The first step is preprocessing, which aims to stabilize the large variation of lighting on the dataset based on entropy calculation to get gamma distortion using the AGT-Me method. The preprocessing input is an image with size of 168 × 192. Pseudocode of normalization present in Fig. 5. Output image that applied using AGT-Me shown in Fig. 6. From the results, it can be seen that AGT-Me increases the contrast level of the images. But the noise attached in Subset 2 and Subset 3 still appeared.

Fig. 7 shows pseudocode of the proposed method, and we split the explanation into 2 parts as presented below:

3.3 Global sub-region scale

The next step is getting macrostructure pixels for enriching feature representation using multiscale blocks. It is called macrostructure because coverage of sub-region captured by multiscale blocks is wide.

To get the global feature representation, we used input from the preprocessed image and the processed features will be fed to the next learning step. The output features is an 8-bit string of binary or if converted to histogram 256 bin (2^8). In this research, we didn't convert feature multiblock into the histogram. The pseudocode presents in Fig. 7, steps 1-3. The step begins with initializing size n for next step, those n divide image I into sub-region multiple of 3. After created the sub-region, next to calculate mean $\bar{x}$ for the output.

3.4 Combining the global and local representation using MBALTP

The next step for getting a feature that combining the global and local representation requires the adaptive LTP threshold determination. Fig. 7 step 4-21 illustrates the pseudocode for obtaining the threshold. The input of adaptive LTP is the processed feature from the multi-scale block, then we calculate the normalized frequency and cumulative distribution from the image's histogram. After that, we calculate the mean for probabilities of normalize histogram, cumulative sum in class, and weight followed by variance. Then we proceed to calculate the standard deviation as the last calculation. That output is the adaptive LTP threshold.

$$\frac{\{(px_i, py_i)\}_{i=0}^{N} + 0.5}{\sum bins}, bins = 0,1,2, \dots ,255 \quad (3)$$

$$CDF(px) = \sum_{x_j \le px} f(x_j) = \sum_{x_j \le px} px_j \quad (4)$$

The next stage is to get the local representation using LTP. That pseudocode is presented in Fig. 7 step 22-25. The output for step 24 $MBALTP_{High}, MBALTP_{Low}$ shown in Fig. 8.

4. Result and discussion

4.1 Experimental setup

For the experiment's requirement, we split every subset into 2 sets there are training and testing. In order to make sure the fairness of evaluation, all experiments carried out based on each subset division, used Support Vector Machine (SVM) classifier with Radial Basis Function kernel, and splitting data used stratified K-fold with k=7.

We conduct a performance evaluation using 2 scenarios. The first scenario is internal comparison.

Input: normalized image I

1 Initialize image I with size $n \times n$
2 Divide image I into sub-region with size $\frac{n}{3} \times \frac{n}{3}$
3 Compute $\bar{x}$ of sub-region

Output: average value each sub-region

Input: average value each sub-region
4 Analyze histogram for every image input $H(I) \leftarrow 1[px_i, py_i]_{i=0}^{N}, i = 1,2,3, \ldots N-1$
5 Compute $H(I)_{normalized} \leftarrow H(I)max$
6 Find cumulative distribution $Q \leftarrow CDF(H(I)_{normalized})$ using Eq. (4)
7 Set $w \leftarrow \sum bins\ H(I) + 1$ to store weight
8 Initialize $temp_{min}$ peaks ← −0.0001
9 **for** $i \leftarrow$ each w:
10 Get probability $P_{1,2}(H(I)_{normalized}, i)$
11 Get $CDF_{1,2}(Q, Q - w(H(I)))$
12 Get $temp_{weight\ 1,2} \leftarrow$ split (w, i)
13 Find $\mu_{1,2} \leftarrow \sum P_{1,2} \times temp_{weight\ 1,2} \div CDF_{1,2}$
14 Find $\sigma^2{}_{1,2} \leftarrow \sqrt{\sum (temp_{weight\ 1,2} - \mu_{1,2})^2 \times P_{1,2} \div CDF_{1,2}}$
15 Calculate the min peaks of variance ← $\mu_{1,2} \times CDF_{1,2}$
16 **if** min peaks of variance < $temp_{min}$ peaks:
17 $temp_{min}$ peaks ← min peaks of variance
18 $t \leftarrow i$
19 **endif**
20 **endfor**
21 $f(p_i, p_c, t) \leftarrow$ Compare center pixel p_c ,pixel neighbor's p_i with t
22 Quantizes $f(p_i, p_c, t)$into 2 patterns: $f(p_i, p_c, t)_{High}$ and $f(p_i, p_c, t)_{Low}$
23 Convert $f(p_i, p_c, t)High$ and $f(p_i, p_c, t)Low$ into LTP_{High}, LTP_{Low}
24 $MBALTP_{features} \leftarrow$ concatenate $MBALTP_{High}, MBALTP_{Low}$

Output: features of MBALTP

Figure. 7 The proposed method pseudocode

The second scenario is to compare several extractors of texture features including LBP, LTP, COUDSHE with Fuzzy [25], IALTP+ [26], the Proposed Method (MBALTP). For LBP, we implement the steps on a previous study [18] because the original LBP study uses different dataset, therefore the result cannot be compared directly. We implement LBP with

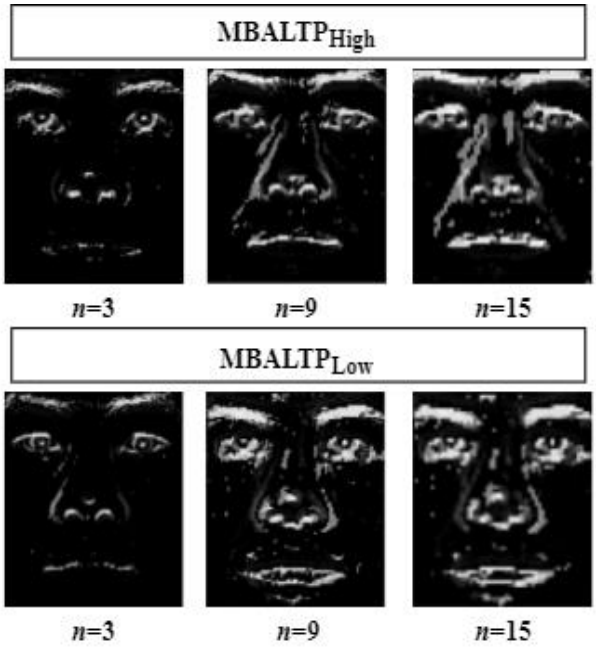

Figure. 8 The result of $MBALTP_{High}$ and $MBALTP_{Low}$ with scale $n = 3$, $n = 9$, and $n = 15$

neighbor = 8 and radius = 1. For LTP, we implement the steps described in previous study [8], with threshold t = 5.

The performance indicator to measure success of proposed method uses accuracy as the main metric. Precision and recall are also used as secondary metric to choose the optimal configuration for the proposed method.

Mathematical equation of performance metrics is denoted in Eq. (5), Eq. (6), and Eq. (7). Accuracy is the ratio of correct predictions (true positives (TP) + true negatives (TN)) to the total number of predictions. Precision is the fraction of positive cases correctly identified (the number of true positives divided by the number of true positives plus false positives (FP)). Recall is the fraction of the cases classified as positive that are actually positive (the number of true positives divided by the number of true positives plus false negatives (FN))

$$accuracy = \frac{TP + TN}{total\ number\ of\ prediction} \quad (5)$$

$$precision = \frac{TP}{FP + TP} \quad (6)$$

$$recall = \frac{TP}{FN + TP} \quad (7)$$

4.2 Comparison internal validation

Internal validation aims to finds the optimum configuration of size n on the multiscale blocks based on the result from the performance of proposed

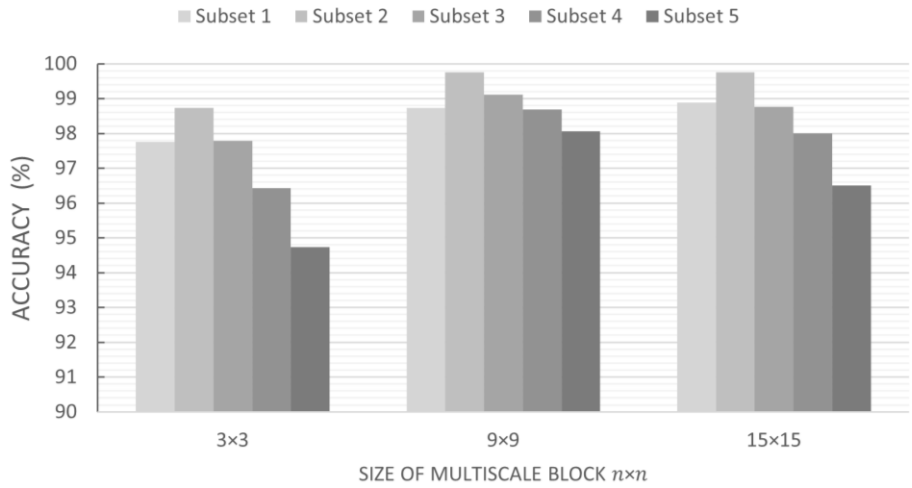

Figure. 9 The comparison of accuracy on the proposed method with different multiscale block sizes

Table 2. Performance evaluation of proposed work with different size of multiscale block n

Subset	**Size of multiscale block** $(n \times n)$	**Accuracy (%)**	**Precision (%)**	**Recall (%)**
1	3×3	97.75	97.37	97.44
	9×9	98.74	98.73	98.74
	15×15	98.88	98.31	98.88
2	3×3	98.73	98.74	98.73
	9×9	99.76	100	99.95
	15×15	99.75	99.76	99.76
3	3×3	97.78	96.43	94.64
	9×9	99.11	99.55	99.6
	15×15	98.77	98.77	98.7
4	3×3	96.43	97.78	94.64
	9×9	98.69	99.13	98.69
	15×15	98	99.1	97.98
5	3×3	94.74	93.9	94.3
	9×9	98.06	97.54	98.06
	15×15	96.5	96.5	93.85

method (MBALTP) on each subset image (Subset 1to Subset 5). The reason for choosing configuration based on the performance on each subset, is because it aims to evaluate the detail robustness with different lighting conditions, and to perform fair evaluations according to the characteristic of each subset. Then, the chosen optimal size is used for the external validation comparison. Keep in mind that size n is a multiple of 3, therefore we use the size $n = 3$, $n = 9$, $n = 15$.

For example, if size $n = 3$, we conduct the experiment 5 times (Subset 1 to Subset 5). Then we split the training data and the testing data for each subset using stratified k-fold validation. If Subset 1 becomes the training data, then the testing data are also taken from the Subset 1.

The internal comparison results are shown in Table 2, and Fig. 9. The X-axis in the figure denotes the size of the multiscale block and Y-axis denotes the accuracy in percent. The evaluation performance (precision, recall and accuracy) in percent is shown detail in Table 2. Generally, the trend of accuracy clearly seen in Fig. 9, the result on Subset 2 is increased, and the other subsets have a decreased rate.

In the graphics, the bars charts are divided into 3 groups based on size n. First, the experiment uses size $n = 3$. For block size 3×3, the sub-region formed with size 1×1. The result of Subset 1 achieved an accuracy of 97.75%, precision of 97.37% and recall 97.44%. And those accuracy result increase 0.98% in Subset 2, reaching 98.73%. But, in Subset 3 the rate starts to decrease until Subset 5, the accuracy of Subset 3 achieved 97.78%. A decrease of

1.69% occurs when the bar chart changes from Subset 4 to Subset 5, the accuracy of Subset 4 achieved 96.43%, decreased into 94.74% in Subset 5. The second experiment uses size $n = 9$. For block size 9×9, the sub-region formed with size 3×3. In Subset 1, the accuracy achieved 98.74%. Then those results increase 1.02% in Subset 2. Same with the previous analysis, the accuracy will decrease in Subset 3 up to Subset 5. With the larger decrease of 0.65% in Subset 2 to Subset 3, the accuracy of Subset 3 becomes 99.11%. Subset 4 and Subset 5 achieved 98.69 and 98.06 of accuracy. So far, the number of reductions in the recognition accuracy for every subset is not as extreme if we compare it with size $n = 3$. And the third experiment uses size $n = 15$. Based on an experiment on this size, Subset 1 achieved an accuracy up to 98.88%. The result on Subset 2 same as the result of Subset 2 with size $n = 9$ is 99.76%. The identification rate decreases in Subset 3 until 5, each have an accuracy 98,77%, 98% and 96,5%.

If analyzed based on the group of subsets, Subset 1 is optimal with configuration $n = 15$, this can occur because there is no noise and lack of lighting variations in Subset 1 are not as extreme as Subset 3, Subset 4, or 5. In addition, with a larger scale, the captures features are more global and visible lines on the face as shown in Fig. 8. It also has the tendency to increase the performance on Subset 2 when compared to Subset 1 and then decreasing. This is caused by several reasons, the first reason is that Subset 2 has larger training data compared to Subset 1, and therefore has more variation. Another reason is that the lighting variation is not as complex compared with Subset 3, 4, or 5. Nevertheless, from Subset 2 to Subset 5 where the exposure variation is greater, the optimal configuration for the multiscale block is $n = 9$. Table 3 illustrates how the different sizes of n determines the average accuracy from Subset 1 until Subset 5.

Analyzing from Table 3, the highest average of accuracy metric based on size n, we get the optimum combination with size $n = 9$ is better than size $n = 3$ or $n = 15$. Based on Fig. 8 the result of feature representation MBALTP with size $n = 9$ captures informative features such as the representation of eye line, nose, mouth, eyebrow on the face and redundant features such as dot or point that don't represent informative features are not captured. Such size was able to get more stable. So that size $n = 9$ is better than size $n = 3$ or $n = 15$.

4.3 Comparison with existing method

Table 4 summarizes the performance evaluation (Accuracy) proposed method with several extractors of texture features. This comparison scenario aims to shows the method's robustness against different lighting variations. In the first method, we implement LBP, which result is clearly seen on the table, achieving an accuracy 89.55%. The result was taken by averaging the accuracy of each subset (Subset 1-5) and are each shown in Fig. 10. The LBP's

Table 3. The average of accuracy on the proposed method with different multiscale block sizes

Multiscale block sizes	The Average Accuracy (%)
3×3	97.08
9×9	98.87
15×15	98.38

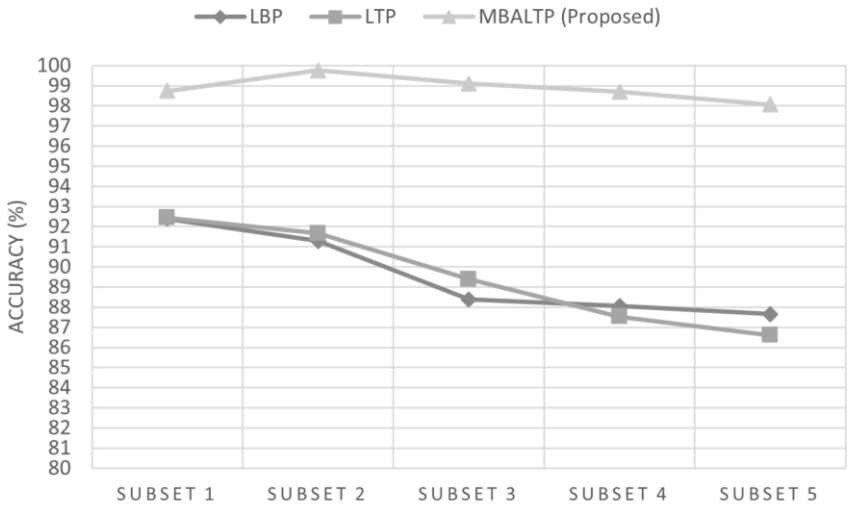

Figure. 10 The comparison of accuracy between the proposed method and the two original method on each Subset data

Table 4. The average accuracy of the proposed method and other related comparison methods

Methods	Accuracy (%)
LBP (Our implementation using step from [18])	89.55
LTP (Our implementation using step from [8])	89.52
COUDSHE with Fuzzy [25]	92.12
IALTP+ [26]	97.39
MBALTP	98.87

line chart decreases as the subset changes to the right (X-axis). Subset 1 almost has the same accuracy with LTP achieving 92.39%. Performance of LBP in Subset 2 reduced by 1.1%, becoming 91.29%, and the performance decrease is quite high about 2.91% from Subset 2 to Subset 3. Subset 4 and 5 achieved an accuracy up to 88.05% and 87.66%.

The second method are through our own implementation of LTP, which are considered less sensitive towards noise. The details of LTP's accuracy for each subset are as follows: The accuracy in Subset 1 achieved 92.45%, which is higher than LBP. The Performance of LTP is better than LBP as seen from the result in Subset 2 and 3 shown by LTP's graph which is above those of LBP. But for the result of Subset 4 and Subset 5, LTP's graph is under LBP's graph. This might happen because the determination of threshold t on the LTP is done manually and fine-tuning is needed to get better performance. In this experiment, we assign $t = 5$. The average of accuracy LTP's achieved 89.52%. Those accuracy smaller than LBP's accuracy.

The third method is COUDSHE (Completely Overlapped Uniformly Decrementing Sub-block Histogram Equalization) with Fuzzy [25]. The method was proposed by S. D. Ganesan and M. A. R. Mohammed. Their experiment on the Extended Yale B Database, could improve in the recognition accuracy of the face recognition using Principal Component Analysis achieved to 92.12%

The four method is IALTP+ [26] which is a fusion of IALTP (Improved Adaptive Local Ternary Pattern) and (2D)2 PCA without preprocessing image such us enhancement. Y. Luo, B. Yu Wang, Y. Zhang, and L. Ming Zhao experimented on the Extended Yale B and AR face databases to solve the different illuminations and random noises problem. Their work achieved accuracy 97.39% on Ext. Yale B databases. In addition, this study [26] already compared the IALTP+ method with their implementation of LBP and finely-tuned LTP, of which IALTP+ achieves the best accuracy.

The last experiment is using our proposed method, MBALTP, showed accuracy above the four existing methods. The final result is that MBALTP has the highest accuracy achieving 98.87%, with the IALTP+ in second place, followed by COUDSHE, LBP and LTP.

5. Conclusion

In this study, we propose a feature representation using Multi-scale Block Adaptive Local Ternary Pattern (MBALTP), that capable of handling lighting variation and noise in images using adaptive LTP and multiscale block. The experiment on the Extended Yale B dataset that has extreme lighting variation and contains noise shows the robustness of the proposed MBALTP feature representation. The experiment result shows that the proposed method achieved an accuracy of 98.87%, which is the highest accuracy compared to the related handcrafted method including LBP, LTP, COUDSHE with Fuzzy, and IALTP+, on the same Extended Yale B dataset.

For future work, an exploration of other statistical approaches that have less time consumption should be considered, as adaptive LTP requires long time consumption.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

Syavira Tiara Zulkarnain: conceptualization, methodology, software, formal analysis, writing – original draft preparation, visualization.

Nanik Suciati: validation, supervision, formal analysis, writing – reviewing and editing, funding acquisition.

Acknowledgments

This research is supported by Institut Teknologi Sepuluh Nopember and is partially funded by the Indonesian Ministry of Research, Technology, and Higher Education under WCU Program managed by Institut Teknologi Bandung.

References

[1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*, 2007.

[2] J. C. Russ, *The Image Processing Handbook, Fifth Edition (Image Processing Handbook*, Fifth Edit. United State: CRC Press, Inc., 2006.

[3] W. Zhao, R. Chellappa, P. Phillips, and A. Rosenfeld, "Face Recognition : A Literature Survey", *ACM Comput. Surv.*, Vol. 35, No. 4, pp. 399–458, 2003.

[4] W. Huang and H. Yin, "Robust face recognition with structural binary gradient patterns", *Pattern Recognit.*, Vol. 68, pp. 126–140, 2017.

[5] L. Chen, Y. H. Wang, Y. D. Wang, and D. Huang, "Face recognition with statistical local binary patterns", *Proc. 2009 Int. Conf. Mach. Learn. Cybern.*, Vol. 4, No. February, pp. 2433–2439, 2009.

[6] L. F. Zhou, Y. W. Du, W. S. Li, J. X. Mi, and X. Luan, "Pose-robust face recognition with Huffman-LBP enhanced by Divide-and-Rule strategy", *Pattern Recognit.*, Vol. 78, pp. 43–55, 2018.

[7] L. Liu, P. Fieguth, G. Zhao, M. Pietikäinen, and D. Hu, "Extended local binary patterns for face recognition", *Inf. Sci. (Ny).*, Vol. 358–359, pp. 56–72, 2016.

[8] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions", *IEEE Trans. Image Process.*, Vol. 19, No. 6, pp. 1635–1650, 2010.

[9] M. A. Farooque and J. S. Rohankar, "Survey on Various Noises and Techniques for Denoising the Color Image", *Int. J. Appl. or Innov. Eng. Manag.* , Vol. 2, No. 11, pp. 217–221, 2013.

[10] A. C. Bovik, J. D. Gibson, S. T. Acton, and G. Arce, *Handbook of Image and Video Processing*. Canada: Academic Press, 2000.

[11] W. Yang, Z. Wang, and B. Zhang, "Face recognition using adaptive local ternary patterns method", *Neurocomputing*, Vol. 213, pp. 183–190, 2016.

[12] Z. Li, N. Yang, B. Xie, and J. Zhang, "A two-phase face recognition method in frequency domain", *Optik (Stuttg).*, Vol. 124, No. 23, pp. 6333–6337, 2013.

[13] I. E. Khadiri, A. Chahi, Y. E. Merabet, Y. Ruichek, and Y. Touahni, "Local directional ternary pattern: A New texture descriptor for texture classification.pdf", *Comput. Vis. Image Underst.*, Vol. 169, pp. 14–27, 2018.

[14] W. Yang, X. Zhang, and J. Li, "A local multiple patterns feature descriptor for face recognition", *Neurocomputing*, Vol. 373, pp. 109–122, 2020.

[15] R. Vedantham and E. S. Reddy, "A robust feature extraction with optimized DBN-SMO for facial expression recognition", *Multimed. Tools Appl.*, Vol. 79, No. 29–30, pp. 21487–21512, 2020.

[16] S. Liao, X. Zhu, Z. Lei, L. Zhang, and S. Z. Li, "Learning multi-scale block local binary patterns for face recognition", *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, Vol. 4642 LNCS, pp. 828–837, 2007.

[17] Y. H. Lee, S. Zhang, M. Li, and X. He, "Blind Inverse Gamma Correction with Maximized Differential Entropy", *ArXiv*, Vol. abs/2007.0, pp. 1–12, 2020.

[18] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions", *Pattern Recognit.*, Vol. 29, No. 1, pp. 51–59, 1996.

[19] S. Murala, R. P. Maheshwari, and R. Balasubramanian, "Local tetra patterns: A new feature descriptor for content-based image retrieval", *IEEE Trans. Image Process.*, Vol. 21, No. 5, pp. 2874–2886, 2012.

[20] D. Huang, C. Shan, M. Ardabilian, Y. Wang, and L. Chen, "Local binary patterns and its application to facial image analysis: A survey", *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, Vol. 41, No. 6, pp. 765–781, 2011.

[21] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms", *IEEE Trans. Syst. Man. Cybern.*, Vol. C, No. 1, pp. 62–66, 1979.

[22] M. Petrou and C. Petrou, *Image Processing: The Fundamentals*, 2nd ed. United Kingdom: J. Wiley, 2010.

[23] K. C. Lee, J. Ho, and D. Kriegman, "Acquiring Linear Subspaces for Face Recognition under Variable Lighting", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 27, No. 5, pp. 684–698, 2005.

[24] G. Y. Chen, T. D. Bui, and A. Krzyzak, "Illumination invariant face recognition using dual-tree complex wavelet transform in logarithm domain", *J. Electr. Eng.*, Vol. 70, No.

2, pp. 113–121, 2019.
[25] S. D. Ganesan and M. A. R. Mohammed, "A hybrid face image contrast enhancement technique for improved face recognition accuracy", *Int. J. Intell. Eng. Syst.*, Vol. 10, No. 6, pp. 106–115, 2017.
[26] Y. Luo, B. Y. Wang, Y. Zhang, and L. M. Zhao, "A novel fusion method of improved adaptive LTP and two-directional two-dimensional PCA for face feature extraction", *Optoelectron. Lett.*, Vol. 14, No. 2, pp. 143–147, 2018.