The Effectiveness of Using Depth Images for Gait-based Gender Classification

Azim Zaliha Abd Aziz, *Member, IAENG*, Muhammad Syazmil Mohd Sabilan, Nur Farraliza Mansor

Abstract— **Generally, gender can be identified based on someone's appearance or voice. However, the walking pattern of a person has become another measurable characteristic in determining gender. This paper presents a new approach for gait-based gender classification using depth images. The main contribution of this study is a new fusion formula named as LETH with higher accuracy rates. Depth images from OU-ISIR Gait database was used. Gait Energy Image (GEI) was then used to normalize the depth image. After that, the images were projected into their PCA plane, gaining a very strong cyclical picture of a change of perspective. Features were then extracted by measuring three distances from three parts of the GEI image (feet, toe, and hand swing). Fine Gaussian Support Vector Machine (Fine Gaussian SVM), Decision Tree, and Weight K-Nearest Neighbor Classifier (KNN) were separately used as the classification method. Experimental results showed that the proposed fusion formula, LETH produced significant results, obtaining 96% accuracy using Fine Gaussian SVM, 98% accuracy using Weight KNN, and 78% accuracy using Fine Decision Tree classifiers, compared to previous studies.**

Index Terms— **gait-based gender recognition, gender classification, depth image, gait energy image.**

I. Introduction

GAIT can be defined as a way to identify and authenticate a person based on her or his way of walking specifically on the physical and behavioral characteristics of that person. Numerous studies have recognized the ability of gait characteristics to be utilized in many areas. For instance, gait can be applied in three major fields which are gait analysis, forensic and biometrics [1]. Moreover, gait also can be applied in visual identification, visual surveillance and monitoring [2] - [4]. Each human is believed to have a unique walking pattern which could be used as one of the biometric classification characteristics. For instance, gait-based analysis is already applied in identifying humans, human recognition or gender classification despite the face-based identification.

Gait-based recognition offers many advantages against other biometric systems especially in passive monitoring which the identification can be done without the knowledge of the person. Additionally, gait can even be implemented using low-resolution devices and the characteristics of the person related to the gait features are difficult to be mimicked [5]. Generally, gait-based gender identification systems can be categorized into three main phases namely data capture and pre-processing phase, feature extraction phase and lastly is the classifiers phase [6]. Various methods were used for data collection. For example, data was captured using a camera by separating at moving person from the static background to produce a silhouette. Some authors used devices such as radar technologies [7] and Microsoft Kinect sensors [8] to acquire human gait characteristics because they claimed that this equipment could control the body motion and could infiltrate any material to get precise data.

Therefore, in this study, depth images obtained from OU-ISIR Gait Database Large Population dataset are used as trials to classify between genders. Currently, the dataset includes over 4,000 subjects with a wide range of ages. However, only images with subjects aged between 20 to 30 years old were selected in this research. This is because a person's walking style is immensely varies between age ranges [9].

The paper is organized as follows. Section 2 includes related work on gait-based recognition. In Section 3, the method for gait-based gender recognition is presented. In Section 4, experiments and results are discussed. Finally, conclusion and future work are provided in Section 5.

II. Related work

Generally, there are two main approaches for gait-based analysis named as model-free and model-based approaches [10]. Model-based methods encode gait information using body and motion models, and extract features from the parameters of the model [11]. While model-free techniques make use of the whole silhouette to provide a concise representation of walking motion [12]. Among the model-free methods, one of the most effective representations for gait is the Gait Energy Image (GEI) proposed by Han and Bhanu [13].

The GEI is a 2D single image that represents both spatial and temporal information of gait silhouettes [14]. This energy image has dramatically reduced the storage space of gait comparing with the original binary a sequence of silhouettes [15]. In addition, it will take much less time to process gait than to analyze the original sequence of

Manuscript received August 15, 2022; revised June 23, 2023.

This research was supported by Ministry of Higher Education of Malaysia (MOHE) through Fundamental Research Grant Scheme (FRGS) (FRGS/1/2020/TK03/UNISZA/02/1) and Universiti Sultan Zainal Abidin (UniSZA).

Azim Zaliha Abd Aziz is a senior lecturer of Faculty of Informatics & Computing, Universiti Sultan Zainal Abidin (UniSZA), Terengganu, Malaysia (phone: +609-6993200; fax: +609-6993299; e-mail: azimzaliha@unisza.edu.my).

Muhammad Syazmil Mohd Sabilan is a postgraduate student of Faculty of Informatics & Computing, Universiti Sultan Zainal Abidin (UniSZA), Terengganu, Malaysia. (e-mail: ashazmil@gmail.com).

Nur Farraliza Mansor is a senior lecturer of Faculty of Informatics & Computing, Universiti Sultan Zainal Abidin (UniSZA), Terengganu, Malaysia (e-mail: farralizamansor@unisza.edu.my).

silhouettes. Fig. 1 shows an example of the silhouette images of a cycle, while in Fig. 2 the corresponding GEI is plotted.

Fig. 1. Binary silhouettes of a cycle.

As mentioned earlier, experiments performed by Han and Bhanu showed GEI to be an effective and efficient gait representation. This has been used as a baseline for recent gait-based gender recognition methods. For example, a five-part partition of the GEI, achieved discriminative results than other components [16]. On the other hand, Li et al. [17] proposed a partition of the GEI into seven components. Each of the seven components is analyzed for the discriminability power of each part for both gender and subject classifications. A novel temporal template, called Chrono-Gait Image (CGI) is proposed for subject recognition [18].

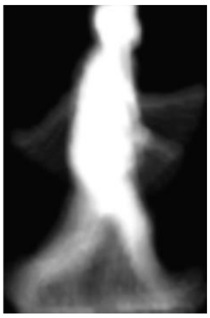


Fig. 2. Example of GEI.

There are some previous works that use depth information on gait analysis. For instance, depth-related data are assigned to the binary image silhouette sequences using the 3-D radial silhouette distribution transform and the 3-D geodesic silhouette distribution transform [19]. Furthermore, a new approach was presented in which the concept of the GEI was extended to 3D, and the gait energy volume (GEV) was created [20]. To further analyze gait recognition using depth images, we present a new fusion formula which combines three distances measurements to differentiate between genders.

III. Gait-Based Gender Recognition

As mentioned earlier, this study used depth images from a publicly available dataset, OU-ISIR. The dataset consists of depth images of walking people in one direction with a specific speed. Each silhouette image is evaluated more than two times where any required changes are made. Therefore, the quality of each silhouette image is reasonably good. Fig. 3 depicts a block diagram of the processes throughout this analysis.

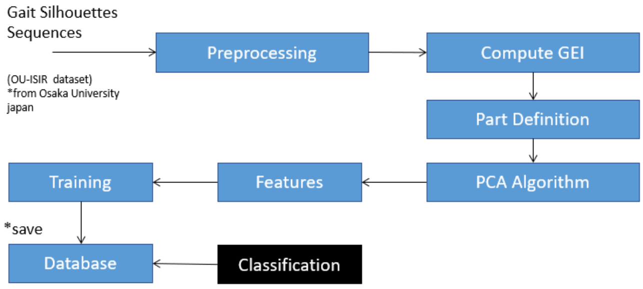


Fig. 3. Flow diagram of the classification processes.

A. Data Preprocessing

Preprocessing was done for all images in one cycle. GEI images obtained from the OU-SIR dataset used in this study had already been preprocessed. The GEI were resized to the same height of 180 pixels. Fig. 4 shows a sample of a normalized silhouette. This image was then used to compute GEI.

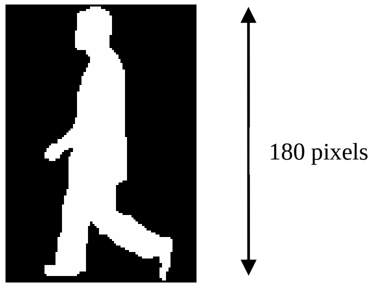


Fig. 4. Example of a normalized silhouettes.

B. GEI Computation

GEI is defined as the average of silhouettes in a gait circle as follows:

$$Gt(m,n) = \sum_{t=1}^{Ns} St(m,n) \tag{1}$$

where N is the total number of frames in the silhouette sequence cycle (s), t is the frame number in the series, and m and n are the 2D image coordinates.

The GEI was divided into three main parts corresponding to head and shoulders, chest and buttocks, and legs as shown in Fig. 5. However, only two parts were used in this study: the chest and buttocks (B); and legs (C). For each part, Principal Component Analysis (PCA) was computed by maintaining principal components that represent 98% of the variance in the data. PCA is often known as an un-supervised linear dimensionality reduction technique used to minimize the dimension of a dataset containing large amount of information [21]. In addition, PCA uses a collection of uncorrelated variables to represent a set of observations [22]. Features were retrieved from parts B and C of reconstructed GEI images after applying PCA as displayed in Fig. 6.

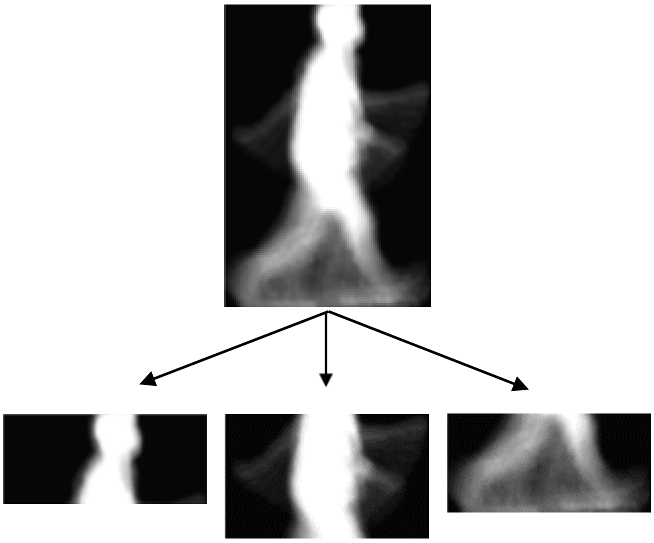


Fig. 5. The GEI with three body parts delimited, only parts B and C were used to compute discriminative features.

C. Newly Proposed Formula (LETH)

This newly gait-based gender recognition formula is called the Leg-Toe-Hand (LETH). In the LETH formula, three measures were calculated and fused as one final classification measure. The first measure is the distance between two feet. The second measure is toe height from the ground and the last one is hand swing distance. Features retrieved using PCA algorithm were used to calculate each measure in the LETH formula.

Distance Between Two Feet

Measurement of the distance between two feet is defined as distance from one foot to the other foot when both feet are on the ground while walking. The distance was measured using the following formula:

$$dl = \sqrt{((y2 - y1)^2 + (x2 - x1)^2)} \tag{2}$$

where dl is the distance between two feet, $x1$ and $y1$ are defined as the toe's coordinates of the first foot, while $x2$ and $y2$ are the heel's coordinates of the second foot. All coordinates were taken when both feet are on the ground. In this study, the first foot is referred to the front foot, while the second foot is at the back side. Fig. 6 illustrates three measures that were taken into the calculation of the LETH formula. As can be seen in Fig. 6, distance between two feet is labelled as *i*.

Toe-ground Distance

Toe-ground distance is the distance from the ground to the raised toe of the front foot when the foot is on the ground in the walking cycle. As shown in Fig. 6, *ii* represents toe-ground distance. This distance was measured using the formula as below:

$$dT = \sqrt{(Jy2 - Jy1)^2 + (Jx2 - Jx1)^2} \tag{3}$$

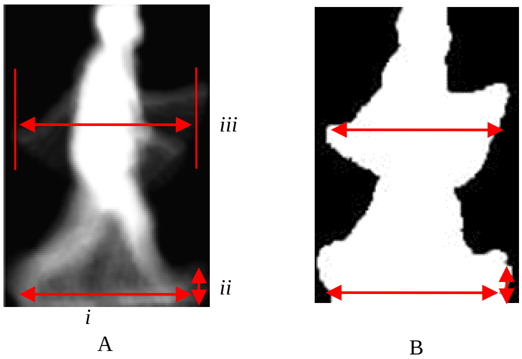


Fig. 6. A: The GEI with three measures: i refers to feet distance, ii is toe-ground distance, and iii is hand swing distance. B: The reconstructed GEI image after applying PCA.

where dT is toe-ground distance, $Jx1$ and $Jy1$ are coordinates of the raised toe, while $Jx2$ and $Jy2$ are coordinates of the ground parallel to the toe's position.

Hands Swing Distance

The distance of the hands swing was measured from the tip of the front hand to the tip of the rear hand while walking. This is illustrated as C in Fig. 6. The formula used is as follows:

$$dH = \sqrt{(Hy2 - Hy1)^2 + (Hx2 - Hx1)^2} \tag{4}$$

where dH is hand swing distance, $Hx1$ and $Hy1$ are coordinates of the tip of front hand, $Hx2$ and $Hy2$ are coordinates of rear hand's tip, in one walking cycle.

Leg-Toe-Hand (LETH) Formula

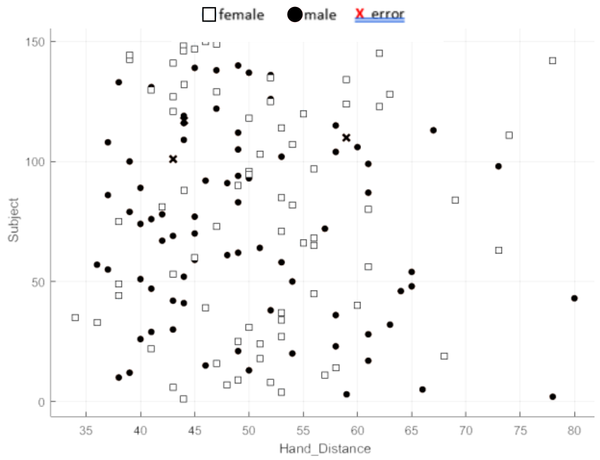
Our newly proposed fusion formula known as LETH is the combination of three measures as shown below:

$$LETH = (2) + (3) + (4) \tag{5}$$

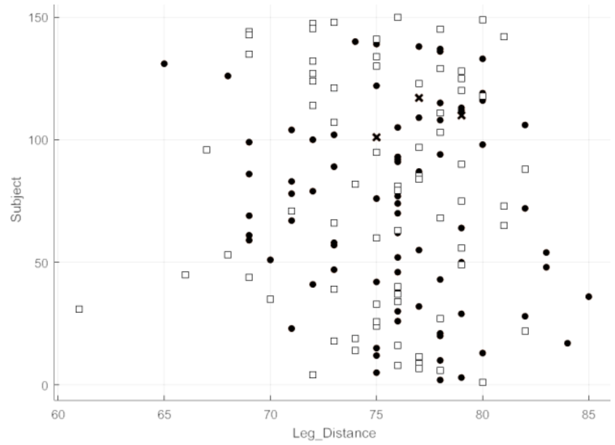
IV. Experiments and Results

As informed earlier in Section I, depth images from OU-ISIR database which age ranging from 20 to 30 years old were used. A total number of 150 images consist of 75 male and 75 female were included as training subjects. Those images were preprocessed to a normalized silhouette GEI. The GEI was divided into three parts in which only two parts were selected. Then, PCA was computed to each selected part. Features obtained were used to measure three distances using equations (2), (3) and (4). These three distances were finally fused to the new LETH formula as in (5) to classify between male and female individuals.

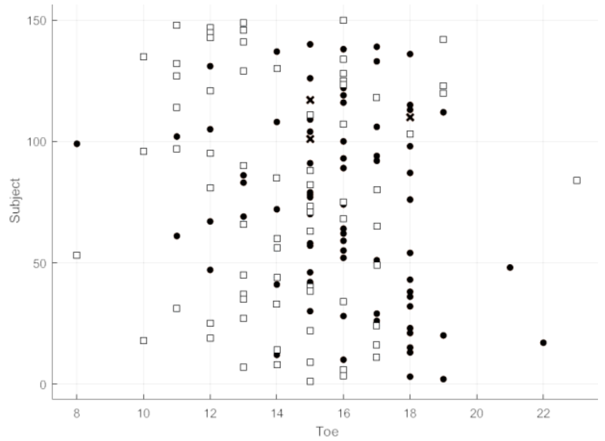
Performance of LETH was evaluated using three classification algorithms: Fine Gaussian Support Vector Machine (Fine Gaussian SVM); Weight K-Nearest Neighbor (Weight KNN); and Fine Decision Tree. Fig. 7 visualizes distance measurements obtained using equations (2), (3), and (4) by using KNN, in which square represents female participants while round bullet refers to male. X is subjects that were erroneously detected as a different gender. The distance is measured in pixels unit.



(a)



(b)



(c)

Fig. 7. KNN distance measurement of each part: (a) Hand, (b) Leg, and (c) Toe for both male and female subjects.

As shown in Fig. 7, no significant difference between male and female for all three distances. A threshold value could not be determined as a classification method. Even though male and female have visibly different walking patterns, to train a machine, system or computer is a challenge. Therefore, these distances were then fused using the LETH formula as in equation (5). The fusion results are plotted and shown in Fig. 8. The data were then trained using KNN classifier and the results were represented in Fig. 9 in the form of Confusion Matrix.

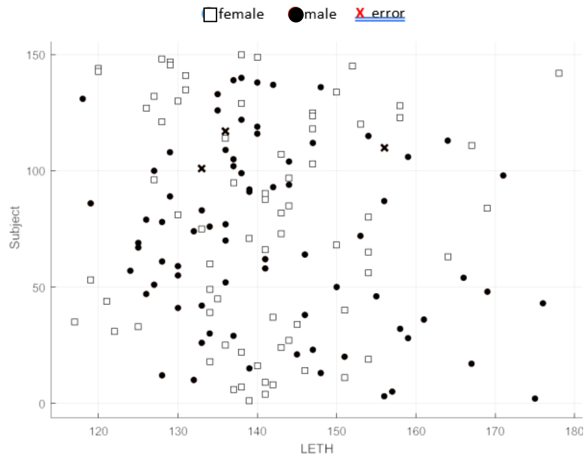


Fig. 8. KNN distance measurement of fusion LETH formula for both male and female subjects.

As shown in Fig. 9, all 75 male subjects were correctly identified as male using Weight KNN classifier. Meanwhile, only three female subjects were wrongly classified as male. Therefore, classification accuracy rate using Weight KNN is high with 98.0%.

True class \ Predicted class	F	M
F	72	3
M		75

Fig. 9. Confusion Matrix of Weight KNN

Performance of LETH was then tested using second classifier, SVM. Results are shown in Fig. 10. The trend of data distribution in Fig. 10 are similar to Fig. 8 in which no significant threshold is shown. True and predicted classes of data in Fig. 10 were presented in the Confusion Matrix table in Fig. 11.

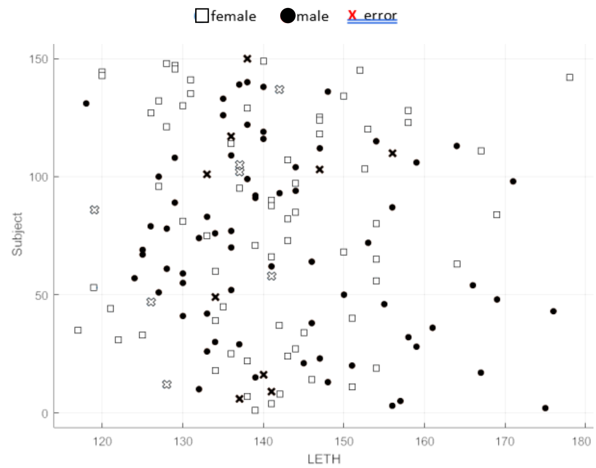


Fig. 10. SVM distance measurement of fusion LETH formula for both male and female subjects.

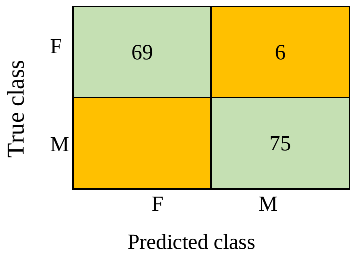


Fig. 11. Confusion Matrix of Fine Gaussian SVM.

In Fig. 11, SVM classifier also correctly predicted all 75 male subjects. However, six out of 75 female subjects were falsely identified as male hence the accuracy rate of using SVM is 96.0%. Next, the LETH data were trained using Fine Decision Tree classifier. Fig. 12 presents the results distribution which were trained using Fine Decision Tree.

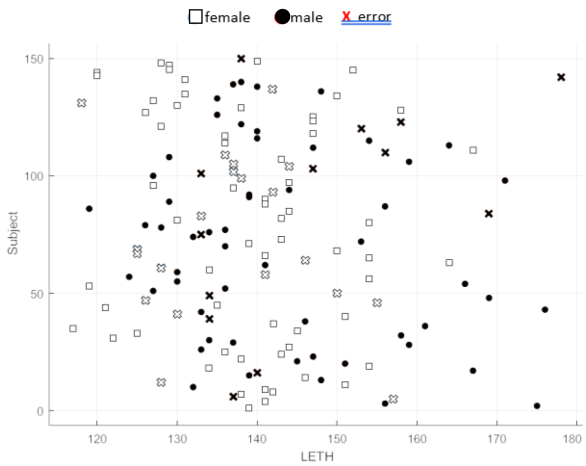


Fig. 12. Fine Decision tree distance measurement of fusion LETH formula for both male and female subjects.

Finally, data in Fig. 12 were trained with Fine Decision Tree classifier. Confusion Matrix of Fine Decision Tree classifier is illustrated in Fig. 13. As shown in Fig. 13, 20 male subjects out of 75 males in total were falsely identified as female. In addition, only 62 female subjects were correctly classified as female while the other 13 genuine females have been identified as male, erroneously. With this, the accuracy rate for Fine Decision Tree classifier is the lowest with 78.0% compared to the other two.

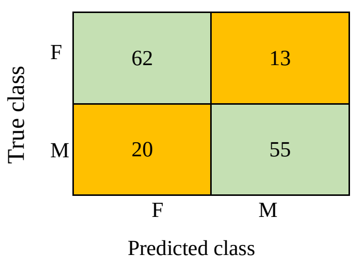


Fig. 13. Confusion Matrix of Fine Decision Tree.

In this study, depth images were used. To prove that depth image has better accuracy rate in classifying genders, comparison was carried out with skeleton images used in [2]. In [2], the authors developed their own database consists of 18 participants (9 male and 9 female) recorded using Microsoft Kinect sensor. Skeleton images were used in their study to differentiate between genders. The third classifier used was Linear Discriminant (LDC) while Fine Decision Tree was used in this study.

To further validate the effectiveness of using depth images in gait-based gender classification, accuracy rates achieved in this study were then compared to previous study which used skeleton images. Table I compares gender classification accuracy rates using skeleton images in [2] with the accuracy results achieved by our newly proposed formula, LETH which used depth images.

TABLE I
GENDER CLASSIFICATION ACCURACY RATES BETWEEN DEPTH AND SKELETON IMAGES

Classifiers	Depth Image	Skeleton Image
Fine Gaussian SVM	96.0%	90.0%
Weight KNN	98.0%	96.7%
Fine Decision Tree / LDC	78.0%	91.1%

Similarly, both studies applied three classifiers for the measurements, except for the third one. From the results shown in Table I, the accuracy rates of using depth image are higher than skeleton image obtained by SVM and KNN. As for depth image, KNN recorded 98.0% compared to skeleton image with 96.7%. While SVM also gave higher scores for depth image with 96.0% compared to only 90.0% for skeleton image. However, the accuracy rate of skeleton image is higher than depth image as measured by the third classifier. This could be excluded since different classifiers were used in the third classification method.

V. CONCLUSION

This paper argues that depth image can provide promising gender classification results. Depth images from OU-ISIR Database Large Population Dataset were used for the gait-based gender classification analysis. Depth images of subjects in the age range between 20 to 30 years old were selected. The depth images were then gone through a preprocessing process before being computed into GEI images. Each GEI image was processed and divided into three parts: head and shoulders; chest and buttocks; and legs. Only two parts were selected for further analysis in this study which were chest and buttocks and legs. PCA was applied to reduce image dimension, resulting in a rebuilt image.

Three distances from the two selected parts (chest and buttocks; and legs) were measured. These three distances are distance between two feet, toe-ground, and hand-swing. Measurements from these three distances were then fused and named as the LETH formula. Results from LETH were tested using three different classifiers: SVM; KNN; and Fine Decision Tree. Comparing the classification results between the LETH formula based on depth image and results in [2], depth images provide significant distinguishable values between genders. The proposed LETH formula is robust for depth image. These results can be explained by the nature of features set that locates diverse aspects of human body that are essential for classification between male and female based on their walking style.

In our future investigations, we will conduct further studies in distinguishing between genders using our own captured dataset. This dataset will record depth image of selected subjects using Microsoft Kinect sensor. In addition, our proposed formula will be tested against any genders spoofing trials.

REFERENCES

[1] Connor, Patrick, and A.A. Ross. "Biometric Recognition by Gait: A Survey of Modalities and Features." Computer Vision and Image Understanding 167: 2018, pp. 1–27.

[2] Ahmed, M. H., & Sabir, A. T. "Human gender classification based on gait features using kinect sensor" in 3rd International Conference on Cybernetics (Cybconf), IEEE, 2017, pp. 1-5.

[3] Huang, J., Di Troia, F., & Stamp, M. "Acoustic Gait Analysis using Support Vector Machines" in ICISSP, 2018, pp. 545-552.

[4] Shelke, P. B., & Deshmukh, P. R. "Gait based gender identification approach" in International Conference on Advanced Computing & Communication Technologies, IEEE, 2015, pp. 121-124.

[5] Binsaadoon, A. G., & El-Alfy, E. S. M. "Enhanced method for recognizing gender in smart environments from gait biometric" in Smart Cities Symposium, IET, 2018, pp. 1-6.

[6] Liu, T., Ye, X., & Sun, B. "Combining convolutional neural network and support vector machine for gait-based gender recognition" in Chinese Automation Congress (CAC), IEEE, 2018 pp. 3477-3481.

[7] Teklehaymanot, F. K., Seifert, A. K., Muma, M., Amin, M. G., & Zoubir, A. M. "Bayesian target enumeration and labeling using radar data of human gait" in 26th European Signal Processing Conference (EUSIPCO), IEEE, 2018, pp. 1342-1346).

[8] Abd Aziz, A. Z., Wei, H., & Ferryman, J. "On evaluation of depth accuracy in consumer depth sensors" in Eight international conference on machine vision (ICMV 2015), SPIE, vol. 9875, pp. 47-52.

[9] Reimann, H., Ramadan, R., Fettrow, T., Hafer, J. F., Geyer, H., & Jeka, J. J. "Interactions between different age-related factors affecting balance control in walking" in Frontiers in Sports and Active Living, 2020, vol. 2, pp. 94.

[10] J. Han and B. Bhanu, "Statistical feature fusion for gait-based human recognition," in Computer Vision and Pattern Recognition (CVPR), in Proceedings of the IEEE Computer Society Conference, 2004, pp. II-II.

[11] Igual, L., Lapedriza, A., & Borras, R. "Robust gait-based gender classification using depth cameras" in EURASIP Journal on Image and Video Processing, 2013, vol. 1, pp. 1-11.

[12] J. Wang, M. She, S. Nahavandi, and A. Kouzani, "A review of vision-based gait recognition methods for human identification" in Digital Image Computing: Techniques and Applications (DICTA) International Conference, 2010, pp. 320-327.

[13] Liu, W., Zhang, C., Ma, H., & Li, S. "Learning efficient spatial-temporal gait features with deep learning for human identification" in Neuroinformatics, 1028, vol. 16(3), pp. 457-471.

[14] Han, J., & Bhanu, B. "Individual Recognition Using Gait Energy Image" in IEEE transactions on pattern analysis and machine intelligence. 2005, vol. 28, pp. 316–322.

[15] Gupta, S. K., Sultaniya, G. M., & Chattopadhyay, P. "An efficient descriptor for gait recognition using spatio-temporal cues" in Emerging Technology in Modelling and Graphics, 2020, pp. 85-97.

[16] Yu, S., Tan, T., Huang, K., Jia, K., & Wu, X. "A study on gait-based gender classification". IEEE Transactions on image processing, 2009, vol. 18(8), pp. 1905-1910.

[17] Li, X., Maybank, S. J., Yan, S., Tao, D., & Xu, D. "Gait components and their application to gender recognition". IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2008, vol. 38(2), pp. 145-155.

[18] Wang, C., Zhang, J., Pu, J., Yuan, X., & Wang, L. "Chrono-gait image: A novel temporal template for gait recognition" in European Conference on Computer Vision, Berlin, Heidelberg, Springer, 2010, pp. 257-270.

[19] Ioannidis, D., Tzovaras, D., Damousis, I. G., Argyropoulos, S., & Moustakas, K. "Gait recognition using compact feature extraction transforms and depth information". IEEE Transactions on Information Forensics and security, 2007, vol. 2(3), pp. 623-630.

[20] Sivapalan, S., Chen, D., Denman, S., Sridharan, S., & Fookes, C. "Gait energy volumes and frontal gait recognition using depth images" in International Joint Conference on Biometrics (IJCB), IEEE, 2011, pp. 1-6.

[21] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "MPCA: Multilinear principal component analysis of tensor objects," IEEE, 2008, Trans. Neural Networks, vol. 19, no. 1, pp. 18–39.

[22] Nanaa, K., Rizon, M., Abd Rahman, M. N., Ibrahim, Y., & Abd Aziz, A. Z., "Detecting mango fruits by using randomized hough transform and backpropagation neural network" in 18th International Conference on Information Visualisation, IEEE, 2014, pp. 388-391.

Azim Zaliha Abd Aziz obtained her Doctor of Philosophy in Computer Science from University of Reading, United Kingdom in 2017. She did her first degree in Computer Science in 2000 at Faculty of Computer Science, Universiti Teknologi Malaysia. This author became a Member (M) of IAENG in 2020. Her research areas are image processing, machine learning, biometrics recognition and object detection.

Muhammad Shazmil is currently pursuing his Master Degree at Faculty of Informatics & Computing, Universiti Sultan Zainal Abidin (UniSZA), Besut Terengganu, Malaysia. He obtained his Degree in Computer Science Software Development from the same university. He also did his first diploma in Information Technology Multimedia in 2017 there. His research areas are software development, image processing, machine learning and object detection.

Nur Farraliza Mansor obtained her Doctor of Philosophy and Master in Computer Science from University Teknikal Malaysia Melaka (UTeM) in 2018 and 2011 respectively. She did her first degree in Electrical Engineering in 2009 at Faculty of Electrical Engineering at Universiti Teknologi MARA. Her research interest is machine learning, internet of things and graph theory.