Thai Journal of Mathematics
Volume 20 Number 3 (2022)
Pages 1227–1238

http://thaijmath.in.cmu.ac.th

ISSN 1686-0209

Human Age Estimation from Multi-angle Gait Silhouettes with Convolutional Neural Networks

Kotcharat Kitchat, Piya Limcharoen, Nirattaya Khamsemanan* and Cholwich Nattee*

Sirindhorn International Institute of Technology, Thammasat University, Pathum Thani, Thailand
e-mail : nirattaya@siit.tu.ac.th (N. Khamsemanan); cholwich@siit.tu.ac.th (C. Nattee)

Abstract Knowing the age of an unknown person is useful information in many fields such as forensics and commercial fields. Authorities can use a correct age estimation technique to narrow down a search for a suspect. Conventionally, facial features are used to estimate the age of a person. However, obtaining the facial features of a subject needs a high-resolution camera with a close-up image, which is hard to do in a real-world environment. Unlike facial features, gaits, a pattern of locomotion, can be observed from a far and unobtrusively. In this work, we propose a new age estimation technique to predict the age of a person from gaits using multi-angle GEIs and a modified Convolution Neural Network (CNN) with Gaussian Mixture (GM). The proposed system consists of two main parts: the moving direction classification model and the age estimation model. The Gaussian Mixture is used as the loss in our age regression model to estimate the age. Our age estimation model contains a modified CNN and follows with three sub-networks that calculate three parameters of the Gaussian Mixture. The proposed methods perform well in multi-viewpoints. The proposed model achieves a Mean Absolute Error (MAE) of 4.08 years in a multi-viewpoint dataset (SIIT-CN-B), which outperforms the existing techniques.

MSC: 62J12; 68T05; 68T10

Keywords: age estimation; gait recognition; gait enery image

Submission date: 01.11.2021 / Acceptance date: 26.01.2022

1. Introduction

Age estimation is a technique to automatically predict the age of a human subject from his/her biometric characteristics, e.g. face images. Accurately predicting the age of an unknown subject is useful in various application domains, e.g. forensic and commercial domains. The estimated age of an unknown suspect helps to narrow down a search by law enforcement agencies in the forensic domain. An age estimation technique helps department store executives or shop owners know their customers. They can better allocate their budgets to directly respond to their customers. A number of techniques, [1–4], have been proposed to conduct the age estimation from face images. To accurately estimate ages from face images, high-resolution images are required. However, collecting these images are not practical in a real-world situation.

*Corresponding author.

Published by The Mathematical Association of Thailand.

Instead of using face images, gait is a type of biometrics that can be used in the age estimation process. Gait is the locomotion information of a person. Gait features show the characteristics of a person when he/she walks, e.g., stride length, arm swing angle, etc. These features can be collected from afar and in any direction. A subject does not need to be near to or facing a camera. Gait data can also be obtained from a low resolution camera, e.g., security cameras, without sophisticated hardware. They are much more accessible in real-world situations than other types of biometrics. Moreover, it is hard for a person to modify his/her gait for a long period of time. Gait features have been used to identify a human subject. This is performed by constructing predictive models from the collected gait data of known subjects. Various techniques have been proposed for human identification, and they can achieve high prediction accuracy.

A number of works have studied the age estimation from gait data. We review them in detail in the next paragraphs. In summary, most of them use silhouette images of persons as input. For feature extraction, most of the existing techniques combine the input images into a Gait Energy Image (GEI) since a varying number of images can be collected from one walk. GEI, first proposed by Han and Bhanu [5], is a technique to extract gait features by averaging silhouette images in one walk. This is one of the most common techniques for silhouette-based gait recognition. For the output, the age estimation in the existing techniques is performed in two different manners, i.e. (1) predicting an age range as a categorical value, and (2) predicting a person's age as a continuous value.

Lu et al. [6] introduce an ordinary preserving manifold analysis approach to seek a low-dimensional discriminative subspace for age estimation. The GEI technique is applied for the feature extraction. The technique is evaluated using the USF database which contains only fixed-direction walk data captured from two viewpoints. The experimental results show a mean absolute error (MAE) of 4.28 years.

Xu et al. [7] propose an age estimation technique based on the Gaussian Process Regression (GPR). Their technique estimates the posterior probability distribution of the observed age. GEIs are used as the features for the age estimation. In their paper, Xu et al. also present a gait dataset for the age estimation called the OULP-age dataset. The data in this dataset are collected from a large population of 63,846 subjects. However, the subjects are asked to walk in one fixed direction. Therefore, the viewpoint issue is not addressed in their work.

Mansouri et al. [8] propose an age classification technique. The technique classifies silhouette images into two classes, i.e. < 60 years old, and ≥ 60 years old. The Support Vector Machine (SVM) technique is used to create the classification model using three types of features, i.e., the Silhouette Model (SM), Frame to Exemplar Distance (FED), and GEI. The SM features are calculated from the number of white pixels in each column of a silhouette image and the number of white pixels in each row of the silhouette image. The FED feature is the Euclidean distance between a silhouette centroid and points on the image. Their technique achieves a predictive accuracy of 76.76% on the OULP-age dataset.

Hema et al. [9] propose an age classification model called the Gait energy image Projection Model (GPM). This model classifies a subject into two age classes: young and old. The GPM consists of two feature types: the number of white pixels in each row, and the number of white pixels in each column. A Support Vector Machine (SVM) is used

---

http://www.eng.usf.edu/cvprg/Gait_Data.html

FIGURE 1. Example of a GEI generated from a walk with the viewpoint issue.

to construct the age classification model which yields 91.8% accuracy on fixed-direction walks from the OULP-age dataset.

Sakata et al. [10] propose an age estimation model based on the three-stage Convolution Neural Network (CNN) using a GEI as the input of the model. The first stage is designed to identify the gender of a subject. The second stage classifies the subject in each gender into five age groups: 0–5, 6–10, 11–15, 16–60, and >60. Finally, the third stage estimates the age of a subject in each age group. Their technique is evaluated using the OULP-age dataset and achieves a mean absolute error of 5.84 years.

Sakata et al. [11] present an age estimation model using the Densely Connected Convolutional Network (DenseNet) as the estimation model and a GEI as the input. A DenseNet is a CNN with dense blocks. In their work, a number of experiments are conducted to select the most appropriate number of dense blocks and their parameters. The results show that a DenseNet using five dense blocks with five composite layers for each dense block yields the best predictive performance with an MAE of 5.79 years.

Zhang et al. [12] propose a model that performs gender classification and age estimation using the residual neural network (ResNet). The model is trained by using the transfer learning technique. ResNet-18, which was originally created and trained from the image classification problem, is used as the base model. The last two layers of the ResNet-18 are replaced by a dropout layer and a fully connected layer before fine tuning the model. A new loss function is also proposed for training the multi-task learning network. This model achieves an MAE of 5.47 years in the OULP-age dataset.

All the existing techniques for age estimation from gait data are evaluated only on fixed-direction walks where persons do not turn in front of the camera. This is not practical in real-world situations where persons may walk in different directions that pass in front of a security camera. Moreover, they may turn in a different direction while they are in the view angle of the camera. A person sometimes walks so close to the camera that the entire body of the person cannot be captured. Constructing a GEI by averaging the silhouette images from different directions and partial body images together causes the mixing of color values. This results in a GEI that may not clearly show the characteristics of that person. This decreases the predictive performance of an age estimation model. Fig. 1 shows an example of a GEI generated from a walk with the viewpoint issue.

Properly handling the viewpoint issue in gait data is a key challenge for using gait-based age estimation works in real-world situations. In this paper, we propose a novel technique to estimate the age of a person from gait features using two-stage convolutional neural networks. We deal with the viewpoint issue by proposing a new GEI-based feature representation technique called multi-angle GEIs. For the age estimation model, we apply the skip connection used in ResNet-18 in our proposed model. Moreover, the

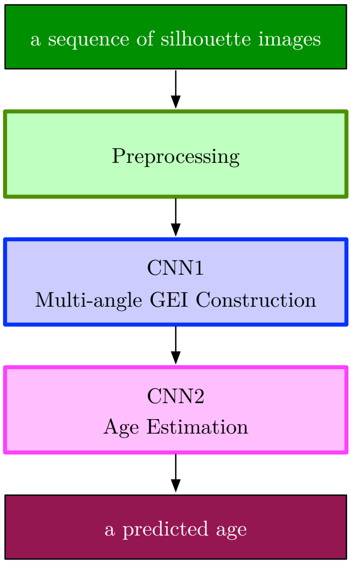


FIGURE 2. Overall process of the proposed technique

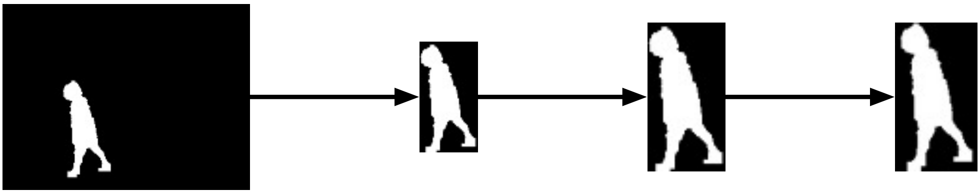
age estimation layers of the proposed model are designed based on the Gaussian Mixture (GM) model. Our idea behind this design is to make the model more general without any assumptions on the age distribution.

2. Materials and Methods

Our proposed gait-based age estimation technique accepts a sequence of silhouette images captured from a walk as an input and returns a predicted age of the person specified by the images. It is composed of three main steps, i.e. (1) *preprocessing* for preparing the input silhouette images, (2) *multi-angle GEI construction* for extracting meaningful features from the input data, and (3) *age estimation* for predicting a person's age from a multi-angle GEI. Both the multi-angle GEI construction and the age estimation are conducted using convolutional neural networks. Fig. 2 shows the overall process of the proposed technique.

2.1. Preprocessing

Since silhouette images captured from a walk with multiple directions and turns may have different sizes of silhouettes, we preprocess them to make them the same size and remove noise. Each silhouette image is cropped in such a way that the boundaries of all four sides of the image are at the farthest white pixel of that side. Since the proposed technique works with a fixed size image, we resize each cropped silhouette to make the height 600 pixels. The resized image is then padded with black pixels or cropped to make the width 280 pixels. The sequence of preprocessed silhouette images is then fed into the multi-angle GEI construction process. Fig. 3 shows an example of the preprocessing steps.



a silhouette image   cropping   resizing based on height   padding/clipping

Figure 3. Example of the preprocessing steps

2.2. Multi-angle GEI Construction

As stated previously, a GEI is not suitable to represent a walk with the viewpoint issue. We propose a multi-angle GEI which is an extension of GEI to cope with the viewpoint issue. Based on observations, we found that a GEI constructed from silhouette images where a subject moves in one direction without turning provides meaningful characteristics for an age estimation.

We construct a convolutional neural network (CNN) that classifies a silhouette image according to the moving direction of the subject. To reduce the complexity of the model, we separate the moving direction into 10 classes according to the range of the angles of the moving directions, i.e. $[0, 36)$, $[36, 72)$, $[72, 108)$, $[108, 144)$, $[144, 180)$, $[180, 216)$, $[216, 252)$, $[252, 288)$, $[288, 324)$, $[324, 360)$. An additional class for rejected silhouette images is added to make the CNN capable to reject unusable images, such as an image with non-human-like shapes. Therefore, the proposed CNN for silhouette classification based on the moving direction is designed to take a silhouette image as an input and return one of the moving direction classes or the rejected class. This construction is adapted from our previous work, [13], on the gender recognition.

The proposed CNN consists of multiple convolutional and max pooling layers. It accepts a silhouette image a the width and height of 280 and 600 pixels, respectively. The input is fed through three groups of convolutional layers with max pooling and batch normalization. Then, the output is flattened before being fed to two fully connected layers. The rectified linear unit (ReLU) is used as the activation function for these layers. Finally, the output layer of the network is a fully connected layer with 11 units that represent the 10 moving direction classes and one rejected class. The activation function of the output layer is set to the softmax function. Fig. 4 shows the architecture of the CNN for silhouette image classification.

After classifying all preprocessed silhouette images according to the moving direction, the silhouette images of the same class are collected together and fed into the GEI construction process. In the GEI construction, all silhouette images in the same class are averaged as shown in the following equation.

$$GEI(x, y) = \frac{1}{n} \sum_{t=1}^{n} I_t(x, y), \tag{2.1}$$

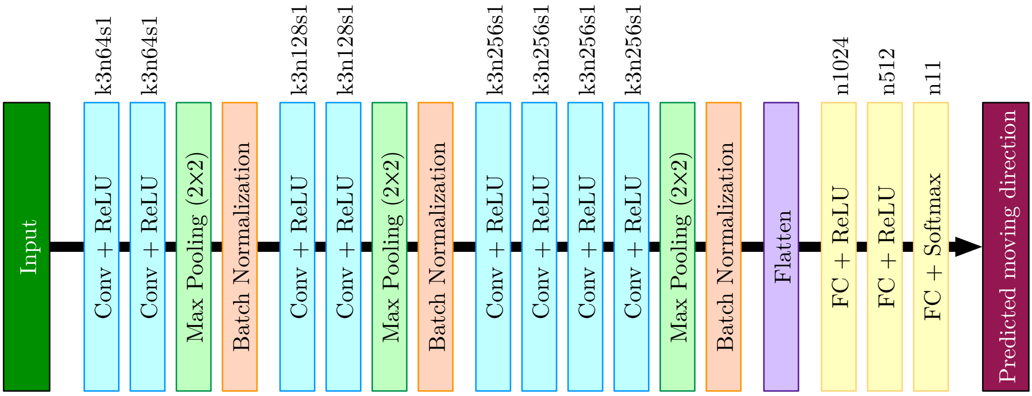


FIGURE 4. Architecture of the convolutional neural network for moving direction classification of a silhouette image; k denotes the size of a kernel, n denotes the number of kernels or units, and s shows the stride

where $GEI(x, y)$ returns a GEI value at the location (x, y), n is the number of silhouette images used in the GEI construction, and $I_t(x, y)$ is the color value at location (x, y) of the t-th image.

After the GEI construction, all 10 GEIs are stacked to create a three-dimensional tensor. We call this tensor a multi-angle GEI. Fig. 5 shows the multi-angle GEI construction steps.

Training the moving direction classification model To train the moving direction classification model from examples, we need a collection of silhouette images labeled with moving directions of the persons shown in the images. To avoid huge efforts on labeling the images, we use Microsoft Kinects to capture silhouette images for training the model. With the Kinect SDK, skeletal data can be extracted from each captured silhouette image. These skeletal data composed of 20 joint coordinates can be used to calculate the moving direction of a person. Moreover, the Kinect SDK also determines if each joint coordinate can be tracked with high confidence. If a joint type is set to *tracked*, this means that the coordinate of the joint can be extracted from the image with high confidence. We use this information to identify if a silhouette image is an acceptable one. Here are steps to prepare training examples for the moving direction classification model:

(1) Retrieve a silhouette image with skeletal data from a Kinect and check the number of *tracked* joints.
(2) If the number of *tracked* joints is less than or equal to 15, the silhouette image is labeled as *rejected*.
(3) Otherwise, calculate the moving direction of the person in the image using the coordinates of the head joint of two consecutive frames:

$$\theta = \arccos \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \tag{2.2}$$

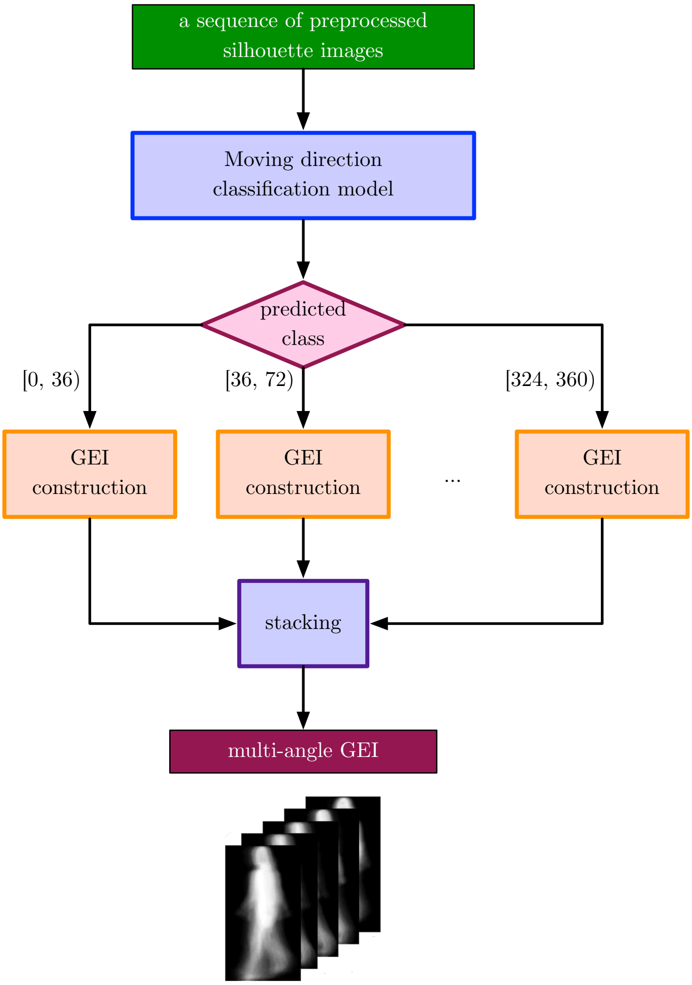


FIGURE 5. Multi-angle GEI construction

where θ represents the moving direction in the form of an angle, $\mathbf{u}$ is the vector representing the coordinates of the head joint in the current frame, and $\mathbf{v}$ is the vector representing the coordinates of the head joint of the previous frame.

(4) The moving direction θ is converted into degrees before dividing by 36 since the proposed network identifies a range of angles.

2.3. Age Estimation

We propose a new age estimation model based on a CNN. The proposed model accepts a 3D tensor representing a multi-angle GEI as explained in the previous section. The shape of the input tensor is therefore $280 \times 600 \times 10$. We treat each GEI in the input tensor as one channel in the convolution operation. Therefore, the input tensor can be viewed as an image with 10 different primary values.

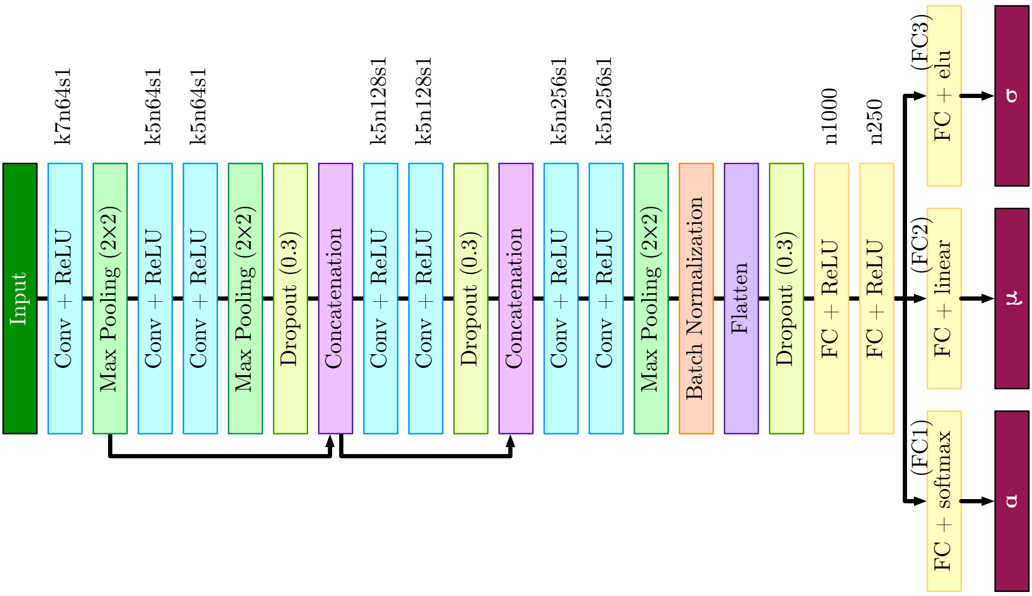


FIGURE 6. Architecture of the convolutional neural network for moving direction classification of a silhouette image; k denotes the size of a kernel, n denotes the number of kernels or units, and s shows the stride

We design the proposed CNN to utilize the idea of skip connections based on the ResNet. The application of skip connections helps to avoid the vanishing gradient problem. Thus, it improves the training process and performance of the proposed model. Fig. 6 shows the architecture of the proposed network for age estimation. The network consists of three groups of convolutional layers with max pooling and batch normalization. Two concatenation operations are added to facilitate skip connections. To avoid overfitting, two dropout layers are introduced into the network. The output of the last convolution group is flattened before being fed to two fully-connected layers. The ReLU activation function is used in all the layers up to this point. The output of the second fully-connected layer, which is a vector with 250 components, is fed into three different fully-connected layers marked as (FC1), (FC2), (FC3). These three layers are added to the network to optimize the parameter values of the Gaussian mixture model for age estimation.

To predict the age of a person, we propose an application of the Gaussian mixture model. This is to avoid parameterizing the model with one Gaussian distribution. The Gaussian mixture model is defined as:

$$p(\hat{y}\,|\,\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\sigma}) = \sum_{i=1}^{K} \alpha_i \mathcal{N}(\hat{y}\,|\,\mu_i,\sigma_i) \tag{2.3}$$

$$\mathcal{N}(\hat{y}\,|\,\mu_i,\sigma_i) = \frac{1}{\sqrt{2\pi\sigma_i}} \exp\left(-\frac{(\hat{y}-\mu_i)^2}{2\sigma_i^2}\right), \tag{2.4}$$

where $\hat{y}$ is a predicted age, and $p(\hat{y} \,|\, \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\sigma})$ denotes the probability of the predicted age given the parameters of Gaussian distributions. Here, $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_K)$ and $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_K)$ are the means and the standard deviations of the distribution, respectively, where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_K)$ is the mixture weights. Therefore, the probability of the Gaussian mixture model is calculated from a weighted sum of K probability values with a constraint:

$$\sum_{i=1}^{K} \alpha_i = 1 \tag{2.5}$$

The network learns the values of the parameters, $\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\sigma}$, by using the last three fully-connected layers. The FC1 layer is designed to find the appropriate values of $\boldsymbol{\alpha}$. Its activation function is therefore set to the softmax function according to Eq. 2.5. The FC2 layer is designed to learn for the value of $\boldsymbol{\mu}$. Since there are no constraints related to the values, we set the activation function to the linear function. The FC3 layer is designed to find the optimal value of $\boldsymbol{\sigma}$. We use a modification of the exponential linear unit (ELU) as the activation function:

$$g(\mathbf{x}) = elu(\mathbf{x}) + 1 \tag{2.6}$$

where $elu(\cdot)$ is the exponential linear unit defined as

$$elu(z) = \begin{cases} z & z > 0, \\ \exp z - 1 & z \leq 0 \end{cases}. \tag{2.7}$$

To train the proposed network, we use the average negative log-likelihood as the loss function. Therefore, the training process tries to minimize

$$\mathcal{L}(y; \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\sigma}) = -\log p(y \,|\, \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\sigma}) \tag{2.8}$$

where y is the target age of an input multi-angle GEI.

Once all the parameter values are learned, an age estimation is conducted by feeding a multi-angle GEI into the network to obtain the parameter values. The predicted age is then calculated from:

$$\hat{y} = \sum_{i=1}^{K} \alpha_i \mu_i \tag{2.9}$$

3. Results

3.1. Dataset

We use a gait dataset called SIIT-CN-B to assess the proposed techniques. The SIIT-CN-B dataset is a multi-viewpoint dataset collected by the staff of the CN (Cholwich-Nirattaya) lab. The data are collected by using a Kinect for Xbox 360. The Kinect is placed 235 cm from the floor and tilted down 27°. This setting makes the Kinect oversee an area of 355 cm × 250 cm. This dataset contains 393 subjects with ages ranging

Table 1. Results of experiment on SIIT-CN-B dataset

Technique	MAE
GPR [7]	6.57
Multistage-CNN [10]	7.07
Multi-task Network [12]	6.21
Our proposed with linear	5.53
Our proposed with GMM	**4.08**

from 10 to 65 years (137 females and 256 males). Each subject walked freely in front of the Kinect camera for 12 rounds. The Kinect SDK provides skeleton data, RGB, and silhouette images from each walk.

3.2. Experimental Settings

We prepare balanced training and test sets. We apply the k-means clustering algorithm to group the age values of the 393 subjects into 10 age groups. The experiments are conducted using five-fold cross validation. Therefore, we randomly split the subjects in each age group into 5 subgroups and conduct the experiments for five rounds. In each round, a subgroup from every age group is selected, combined into one set, and used as the test set. All the rest of the subgroups are combined and used as a training set.

We compare the proposed technique to three existing techniques, i.e. the GPR [7], the Multistage-CNN [10], and the multi-task network [12] techniques. Moreover, we evaluate the significance of the Gaussian mixture model by replacing the FC1, FC2 and FC3 layers with a fully-connected layer with one unit and the linear activation function. We measure the performance of each technique using the Mean Absolute Error (MAE), which is defined as the average of the absolute error between the predicted age and ground truth age. The MAE is calculated as:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{l}_i - l_i|, \tag{3.1}$$

where N is the number of test subjects, l_i is the ground truth age and $\hat{l}_i$ is the predicted age for test subject i. Table 1 shows the experimental results.

4. Discussion

The results show that the proposed method with the Gaussian mixture (GM) obtains better performance than GPR [7], Multistage-CNN [10], Multi-task network [12] and our proposed method with a linear activation.

When our proposed method is used with a linear activation function, we achieve 5.53 years of MAE. However, the proposed method with the Gaussian mixture model obtains an MAE of 4.08 years. The experimental result shows that using the Gaussian mixture achieves higher performance than using linear activation. In a general population, the age distribution is not typically linear. It may contain more than one peak of age groups. Many populations contain several unequal peaks of age groups. The gaussian mixture creates a distribution that is a weighted sum of Gaussian distributions which may fit the distribution of age data better. The linear activation is the function that generates a linear distribution to represent the data distribution, which may not suitable for the distribution of age data. This may be a reason our proposed method with Gaussian mixture model performs better than the model with a linear regression.

The MAE of a technique proposed by Xu et al. [7] is 6.57 years. Their work is based on the Gaussian process regression (GPR). Even though GPR also uses a Gaussian distribution, the MAE from GPR is significantly less than the MAE of our proposed method. The result confirms that the age distribution is not just a normal distribution with one peak. A more suitable age distribution should contain several different-height peaks. Gait features that were used in [7] are conventional GEI images of walks. Traditional GEIs are generated based on all silhouettes of all frames. The dataset used in our experiment is multi-viewpoint data where silhouettes extracted from a walk come from different observation angles and some silhouettes are not full-body silhouettes. Conventional GEIs obtained from multi-viewpoint datasets contain noise (partial silhouettes) and uneven different observation-angle silhouettes. This also may contribute to lower performance in comparison to our proposed method, since our proposed method uses 10-channel angle based GEIs as gait features. Our moving direction classification model removes partial silhouettes and classifies frames of a walk into specific angle ranges. This shows that multi-angle GEIs are more suitable for multi-viewpoint datasets than conventional GEIs.

The MAEs of Multistage-CNN [10] and Multi-task Network[12] are 7.07 and 6.21 years, respectively. These two techniques use linear activation to estimate ages which may not be suitable for age distributions, causing lower performances. The Multistage-CNN uses three-stage CNN models to predict a person's age, whereas the Multi-task Network is based on ResNet-18. ResNet is designed to connect a previous layer's output after skipping a few layers. This attempts to deal with the vanishing gradient. Although ResNet typically enjoys high performance in the field of image classification, the experimental result shows that ResNet alone is not enough to achieve high performances in the prediction of a person's age. Both Multistage-CNN [10] and Multi-task networks [12] use conventional GEIs as their gait features which, as mentioned before, are not suitable for multi-viewpoint datasets.

5. Conclusion

In this paper, we proposed a new age estimation model from gait using multi-angle GEIs and a combined convolution neural network with a Gaussian mixture (GM). The proposed model consists of two main parts: the moving direction classification model [13] that generates angle-based GEIs and the newly proposed age estimation model. The moving direction classification model classifies the image of a walk into a negative sample (partial silhouette) and 10 groups of observation angles. Multi-angle GEIs are 10-channel GEIs that are created from these different observation angle ranges. Multi-angle GEIs are used as inputs for the age regression model. Our age estimation model is a modified convolutional neural network combined with three sub-networks, that calculate the three parameters of the Gaussian mixture. Our proposed method has an MAE of 4.08 years, which outperforms existing methods on a multi-viewpoint dataset (SIIT-CN-B). The experimental result suggests that multi-angle GEIs and the Gaussian mixtures are more suitable for age distributions.

Acknowledgments

This work was supported by Thammasat University Research Unit in Gait Analysis and Intelligent Technology (GaitTech).

References

[1] H. Han, C. Otto, A.K. Jain, Age estimation from face images: Human vs. machine performance, International Conference on Biometrics (ICB) (2013), 1–8.

[2] R. Jana, D. Datta, R. Saha, Age estimation from face image using wrinkle features, Procedia Computer Science 46 (2015) 175–1761.

[3] M. Xia, X. Zhang, W. Liu, L. Weng, Y. Xu, Multi-stage feature constraints learning for age estimation, IEEE Transactions on Information Forensics and Security 15 (2020) 2417–2428.

[4] A. Othmani, A. Rahman Taleb, H. Abdelkawy, A. Hadid, Age estimation from faces using deep learning: A comparative analysis, Computer Vision and Image Understanding 196 (2020) Article ID 102961.

[5] J. Han, B. Bhanu, Individual recognition using gait energy image, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (2) (2006) 316–322.

[6] J. Lu, Y.P. Tan, Ordinary preserving manifold analysis for human age and head pose estimation, IEEE Transactions on Human-Machine Systems 43 (2013) 249–258.

[7] C. Xu, Y. Makihara, G. Ogi, X. Li, Y. Yagi, J. Lu, The OU-ISIR gait database comprising the large population dataset with age and performance evaluation of age estimation, IPSJ Transactions on Computer Vision and Applications 9 (24) (2017) 1–14.

[8] N. Mansouri, M.A. Issa, Y.B. Jemaa, Gait features fusion for efficient automatic age classification, IET Computer Vision 12 (1) (2018) 69–75.

[9] M. Hema, S. Pitta, Human age classification based on gait parameters using a gait energy image projection model, The 3rd International Conference on Trends in Electronics and Informatics (ICOEI) (2019), 1163–1168.

[10] A. Sakata, N. Takemura, Y. Yagi, Gait-based age estimation using multi-stage convolutional neural network, IPSJ Transactions on Computer Vision and Applications 11 (2019) Article No. 4.

[11] A. Sakata, Y. Makihara, N. Takemura, D. Muramatsu, Y. Yagi, Gait-based age estimation using a DenseNet, In Computer Vision - ACCV 2018 Workshops (2019), 55–63.

[12] S. Zhang, Y. Wang, A. Li, Gait-based age estimation with deep convolutional neural network, International Conference on Biometrics (ICB) (2019), 1–8.

[13] K. Kitchat, N. Khamsemanan, C. Nattee, Gender classification from gait silhouette using observation angle-based geis, IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM), Bangkok, Thailand (2019), 485–490.