Deep Neural Network-Based Gait Classification Using Wearable Inertial Sensor Data

Dawoon Jung, Mau Dung Nguyen, Jooin Han, Mina Park, Kwanhoon Lee, Seonggeun Yoo, Jinwook Kim, Kyung-Ryoul Mun

Abstract— **Human gait has been regarded as a useful behavioral biometric trait for personal identification and authentication. This study aimed to propose an effective approach for classifying gait, measured using wearable inertial sensors, based on neural networks. The 3-axis accelerometer and 3-axis gyroscope data were acquired at the posterior pelvis, both thighs, both shanks, and both feet while 29 semi-professional athletes, 19 participants with normal foot, and 21 patients with foot deformities walked on the 20-meter straight path. The classifier based on the gait parameters and fully connected neural network was developed by applying 4-fold cross-validation to 80% of the total samples. For the test set that consisted of the remaining 20% of the total samples, this classifier showed an accuracy of 93.02% in categorizing the athlete, normal foot, and deformed foot groups. Using the same model validation and evaluation method, up to 98.19% accuracy was achieved from the convolutional neural network-based classifier. This classifier was trained with the gait spectrograms obtained from the time-frequency domain analysis of the raw acceleration and angular velocity data. The classification based only on the pelvic spectrograms exhibited an accuracy of 94.25% even without requiring a time-consuming and resource-intensive process for feature engineering. The notable performance and practicality in gait classification achieved by this study suggest potential applicability of the proposed approaches in the field of biometrics.**

I. Introduction

Human gait, defined as a personal manner of walking, is a characteristic feature of a person because it is determined by an individual's intrinsic (e.g., sex, age, etc.), physical (e.g., weight, height, limb length, physique, etc.), psychological (e.g., personality type, etc.), and pathological (e.g., trauma, neurological diseases, musculoskeletal anomalies, psychiatric disorders, etc.) factors. Compared to other biometric traits such as fingerprint, face, iris, and vein, human gait can be recognized from a distance without requiring user involvement and physical restraint. Hence, gait has been strongly supported as an effective biometric signature for human identification and authentication [1, 2].

Recently, deep learning-based techniques have produced promising results in gait recognition and classification for human identification. Zhang *et al*. proposed the Siamese neural network-based gait recognition to identify human in the surveillance camera recordings [2]. Trommel *et al*. and Seyfioğlu *et al*. devised the deep learning strategies to recognize gait for people identification using micro-Doppler systems [3, 4]. Castro *et al*. developed the convolutional neural network (CNN) model for the gait video-based gender recognition and human identification [5]. Besides the Doppler radar and video camera, abovementioned tools, wearable inertial measurement units (IMUs) have been widely used to obtain gait-related information [6]. Inertial sensors mainly including accelerometer and gyroscope can be used for reliable assessment of gait without being disturbed by appearance alternations and perceptual distortions [7].

In order to serve as a preliminary study on the feasibility of human identification and authentication using gait data collected with wearable inertial sensors, this study focused on classifying gait. The aim of this study was to develop and evaluate neural network-based classifiers for effective categorization of athlete, normal foot, and deformed foot groups' gait assessed using wearable IMUs. To achieve the aim, two approaches were proposed: gait classification based on gait parameter and fully connected neural network and that based on gait spectrogram and CNN.

II. Materials and Methods

A. Participants

This study was approved by the Institutional Review Board of Korea Institute of Science and Technology. All participants provided written informed consent. Three groups of participants were recruited. The first group consisted of 29 semi-professional athletes who had completed a triathlon or a marathon at least once in the past three months (hereafter referred to as "athlete group"). The second and third groups included 19 participants with normal foot (hereafter referred to as "normal foot group") and 21 patients with foot deformities (hereafter referred to as "deformed foot group"), respectively. The deformed foot group consisted of seven, six, two, and six patients with flat foot, plantar fasciitis, pes cavus, and deformity-related foot pain, respectively. The exclusion criteria for all groups were the presence or history of neurological disorder and age-related health issues.

B. Experimental protocol

The gait data were collected using a commercialized IMU-based motion capture system (Xsens MVN, Enschede, Overijssel, Netherland). As shown in Fig. 1, a total of seven IMUs were placed on each participant's posterior pelvis, both thighs, both shanks, and both feet. While wearing the motion capture system, the participants were required to walk on the

D. Jung, M. Park, K. Lee, J. Kim, and K. Mun are with the Center for Imaging Media Research, Korea Institute of Science and Technology, Seoul, Republic of Korea (corresponding author to provide phone: +82-2958-6785; fax: +82-2958-5769; e-mail: krmoon02@kist.re.kr).

M. D. Nguyen is with the University of Science & Technology, Seoul, Republic of Korea.

J. Han and S. Yoo are with the Seoul National University of Science & Technology, Seoul, Republic of Korea.

978-1-5386-1311-5/19/$31.00 ©2019 IEEE

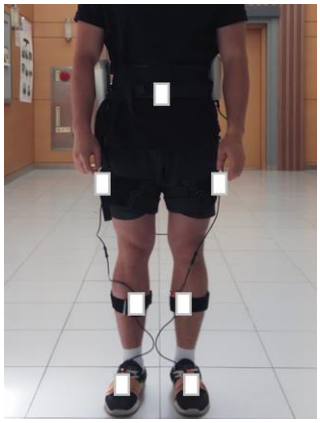

Figure 1. Positions of seven wearable inertial measurement units (white rectangles)

20-meter straight path three times at different speeds: 1) preferred, 2) fast, and 3) slow speeds. The fast and slow walking, respectively, were performed by instructing the participants to walk at 20 ± 5% faster and slower speeds than their preferred speeds.

C. *Classifier development and evaluation*

I. *Gait parameter and fully connected neural network-based classifier*

After detecting heel-strike (HS) and toe-off (TO) times of both lower limbs based on the angular velocity of both shanks, stance, swing, double limb support (DLS), single limb support (SLS), step, and stride were determined as represented in Fig. 2 [8, 9]. The following six temporal, two spatial, and one spatiotemporal gait parameters were extracted: stance time (time elapsed from a HS to a TO of the same foot), swing time (time elapsed from a TO to a HS of the same foot), DLS time (amount of time spent with both feet contacting the ground), SLS time (amount of time spent with only one foot contacting the ground), step time (time elapsed between a HS and a consecutive HS of the opposite foot), stride time (time elapsed between a HS and a consecutive HS of the same foot), step length (distance between a HS and a consecutive HS of the opposite foot), stride length (distance between a HS and a consecutive HS of the same foot), gait velocity (value calculated by dividing a stride length into a corresponding stride time) [10, 11]. The stride and step lengths were normalized with each participant's height.

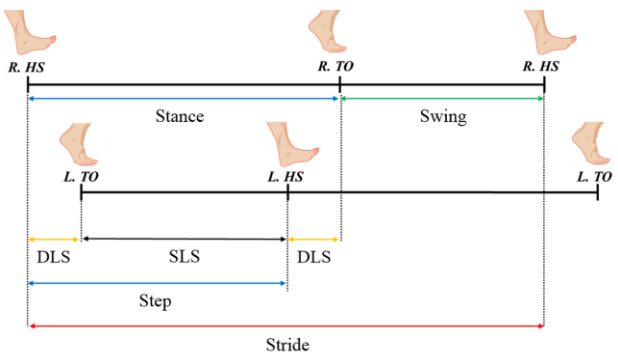

Figure 2. Definitions used to extract gait parameters

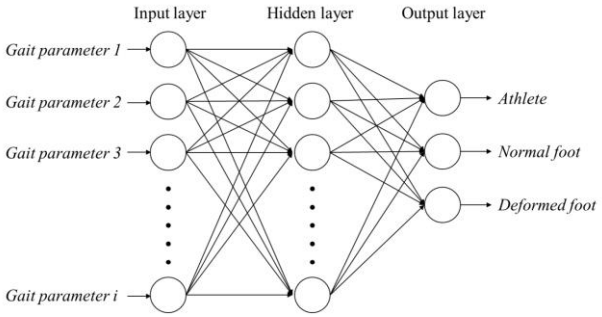

Figure 3. Architecture of the gait parameters and fully connected neural network-based classifier

Fig. 3 shows the architecture of the proposed fully connected neural network-based classifier. The average values of each of the gait parameters for each participant's four strides in the middle of the walkway were used as input variables of the classifier. The number of input neurons was the same as the number of input variables. After a grid search with a few randomly selected samples, the network with a single hidden layer was trained for 1000 iterations using the Adam optimizer and rectified linear unit (ReLU) activation function with a learning rate of 0.1. In order to avoid overfitting problems, 4-fold cross-validation was applied to 80% of the total samples. The remaining 20% were used as a test set to evaluate the classification performance.

II. *Gait spectrogram and convolutional neural network-based classifier*

Short-time Fourier transform (STFT) was used to observe the lower limb movements with the information on frequency and power that varied with time. STFT is a powerful general-purpose tool for transforming original time-domain signals to time-frequency (TF) domain signals.

$$\mathrm{STFT}\{x(t)\}(\tau,\omega) = \int_{-\infty}^{\infty} x(t)\omega(t-\tau)\,e^{-j\omega t}dt$$

Where $\omega(t-\tau)$ is a window function, which is often taken to be a Gaussian window. By multiplying the window with $x(t)$ in the time domain, the signal is windowed and the frequency is computed around the time t. The Fourier transform at the given time corresponds to the local spectrum at the specified time t [12]. In this study, the time resolution was set to 25 milliseconds. Fig. 4 displays an illustrative case of STFT application to gyroscope data obtained from a foot during a stride. Fig. 4(A) and 4(B) respectively show the foot angular velocity signal in the sagittal plane and the spectrogram generated by calculating the squared magnitude of the short-time Fourier-transformed foot angular velocity signal. Because each IMU included 3-axis accelerometer and 3-axis gyroscope, a total of 504 spectrograms were obtained from seven IMUs for each participant's 12 strides (each four strides at preferred, fast, and slow speeds). In totality, 34776 spectrograms were acquired from 69 participants.

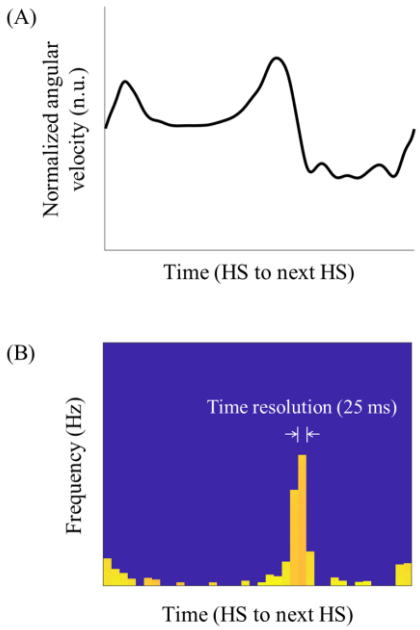

Figure 4. (A) Foot angular velocity signal in sagittal plane, (B) Spectrogram of the foot angular velocity signal

The input images for the CNN-based classification were produced through the following steps. First, all spectrograms were resized to 384 × 256 × 3 (3-channel RGB image). Second, the images for the two types of inertial sensor data, accelerometer and gyroscope data, were stacked and the new images with a size of 768 × 256 × 3 were obtained. Third, the size of the new images was adjusted to 768 × 768 × 3 by stacking the images for the 3-axis data of each inertial sensor. Finally, the size of training samples was determined with consideration for multiple IMU positions. That is, when the data measured using one, two, and all seven IMUs were taken into account, the input image sizes were 768 × 768 × 3, 768 × 768 × 6, and 768 × 768 × 21, respectively. Four-fold cross-validation was applied to 80% of the total samples and the remaining 20% were used to test the classification performance.

Fig. 5 illustrates the proposed CNN architecture that used the spectrograms of the posterior pelvis and both feet as input images. Each of the eight convolutional layers with 64, 64, 128, 128, 128, 256, 256, and 512 filters of 3 × 3 size and 1 × 1 stride, respectively, was followed by the max-pooling layer with 2 × 2 size and 2 × 2 stride. The fully connected layer was used as an output layer with three neurons corresponding to the three classification groups. Except for the softmax function at the output layer, ReLU activation function was used. Regarding the hyper-parameters, the dropout with a probability of 0.6, the batch size of 32, and the Adam optimizer for gradient update were used. The initial learning rate was set to 0.0001 and reduced by a factor of 0.5 every two epochs without improvement in the validation set accuracy. The training process was early stopped when the accuracy for the validation set was not improved within seven epochs.

III. Results

A. Participant characteristics

The demographic and anthropometric characteristics for the three groups are summarized in Table I.

B. Classification performance

Fig. 6 displays a normalized confusion matrix for the gait parameters and fully connected neural network-based classifier. In the normalized confusion matrix, the diagonal elements correspond to the normalized number of samples correctly categorized into their groups, while the off-diagonal elements indicate the normalized number of samples erroneously categorized. The gait parameter-based classification performance for the normal foot group was lower than that for the other groups, which was caused by misclassifying the participants with normal foot as athletes. The overall accuracy of the gait parameters and fully connected neural network-based classifier is summarized in the second row of Table II.

TABLE I. Demographic and Anthropometric Characteristics of the Participants

	Group		
	Athlete	***Normal foot***	***Deformed foot***
No. of participants	29	19	21
Age (years)	52.92 ± 9.60	29.18 ± 4.89	27.90 ± 3.10
Weight (kg)	69.76 ± 5.36	73.47 ± 8.09	70.19 ± 10.79
Height (cm)	171.69 ± 5.72	175.88 ± 5.05	173.07 ± 7.57

Data are presented as the mean ± SD.

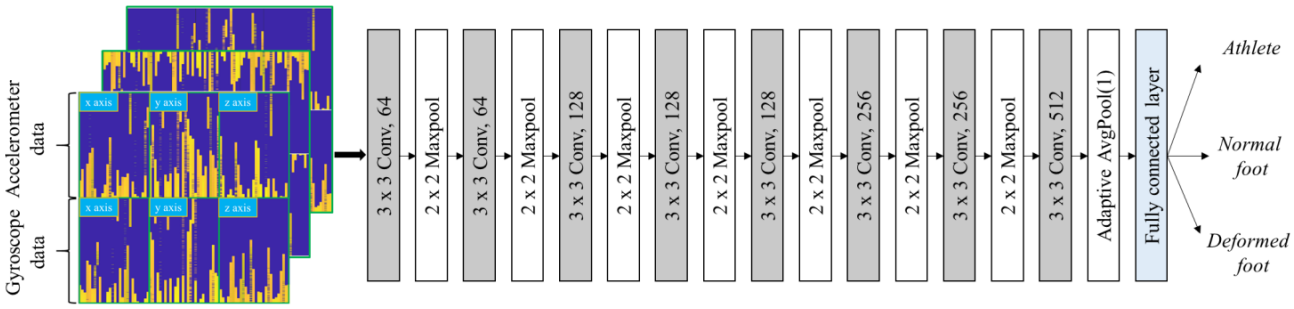

Figure 5. Architecture of the gait spectrograms and convolutional neural network-based classifier

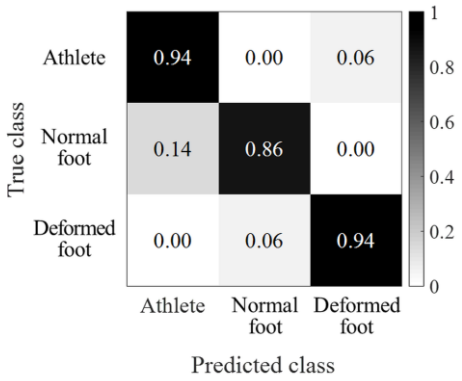

Figure 6. Normalized confusion matrix for the gait parameters and fully connected neural network-based classifier

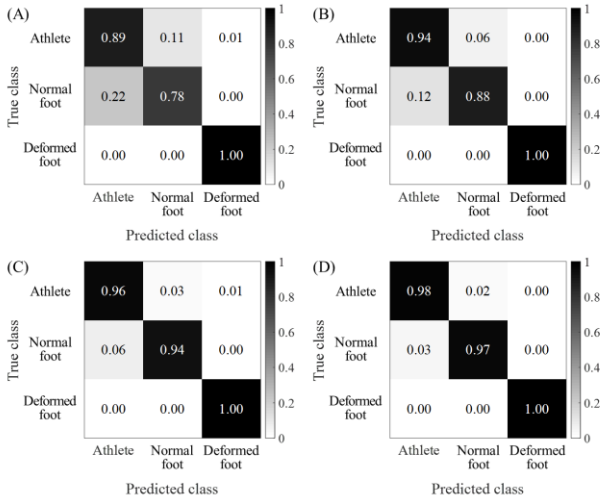

Figure 7. Normalized confusion matrices for the classifiers based on the convolutional neural networks and the gait spectrograms of (A) foot, (B) posterior pelvis, (C) foot and posterior pelvis, (D) foot, shank, thigh, and posterior pelvis

Four confusion matrices in Fig. 7 show the performance of the gait spectrograms and CNN-based classifiers. Among the classifications using each of the spectrograms of posterior pelvis, thigh, shank, and foot, the worst and the best performance were respectively resulted from the foot spectrograms and from the pelvic spectrograms (Fig. 7(A) and 7(B)). Nevertheless, the foot spectrograms could improve the classification performance by being used together with the pelvic spectrograms (Fig. 7(C)). The most outstanding performance was achieved by using all spectrograms of posterior pelvis, thigh, shank, and foot (Fig. 7(D)). No participants in the athlete and normal foot groups were erroneously categorized into the deformed foot group by the classifier trained with the pelvic spectrograms and by that trained with all spectrograms (Fig. 7(B) and 7(D)). The third to sixth rows of Table II show the overall accuracies in the gait spectrograms and CNN-based classification.

IV. Discussion

This study proposed the neural network-based approaches to classify the athlete, normal foot, and deformed foot groups' gait using the wearable inertial sensor data. The gait classifier trained with the temporal, spatial, and spatiotemporal gait parameters exhibited an accuracy of 93.02%. The Kruskal-Wallis test, performed using SPSS statistics software (v.21.0, SPSS Inc., Chicago, Illinois, USA), revealed the significance of differences in all of the gait parameters among the three groups as shown in Table III (all $P < 0.05$). These statistical results suggest the potential effectiveness of the gait parameters as biometric identifiers, and this effectiveness could be improved by utilizing a neural network model as proposed in this study. The spectrograms as well as the gait parameters were used to assess gait. The accuracy was 98.19% for the gait classification based on the CNN and gait spectrograms obtained from the TF domain analysis of the raw acceleration and angular velocity data. This result is associated with the reliability of the gait spectrogram as an image effectively encoding temporal and spatial gait characteristics and with the suitability of CNN to be used for gait spectrogram-based classification. The time-normalization effect, achieved by resizing and stacking the spectrograms to make them suitable to be used as inputs of the CNN model, is also related to the outstanding performance of the gait spectrograms and CNN-based classifier.

The use of CNN is noteworthy in that a lot of labor, empirical knowledge, and time-consuming and resource-intensive process required for feature engineering can be avoided. In terms of real-world applicability and usability, it is also notable that the athlete, normal foot, and deformed foot

TABLE II. Gait classification performance

Neural network input		Accuracy (%)
Gait parameters		93.02
Gait spectrograms of	Foot	88.82
	Posterior pelvis	94.25
	Foot and posterior pelvis	96.68
	Foot, shank, thigh, and posterior pelvis	98.19

TABLE III. Difference in gait parameters among the three groups

	Group		
	Athlete	***Normal foot***	***Deformed foot***
Stance time (s)	0.60 ± 0.10	0.63 ± 0.11*	0.75 ± 0.16*†
Swing time (s)	0.48 ± 0.05	0.50 ± 0.08*	0.62 ± 0.10*†
DLS time (s)	0.06 ± 0.03	0.13 ± 0.02*	0.08 ± 0.06*†
SLS time (s)	0.48 ± 0.05	0.50 ± 0.10*	0.61 ± 0.10*†
Step time (s)	0.54 ± 0.07	0.59 ± 0.11*	0.68 ± 0.12*†
Stride time (s)	1.08 ± 0.14	1.13 ± 0.19*	1.37 ± 0.23*†
Step length (m)	0.43 ± 0.06	0.45 ± 0.05*	0.46 ± 0.07*†
Stride length (m)	0.86 ± 0.12	0.85 ± 0.09*	0.91 ± 0.12*†
Gait velocity (m/s)	1.42 ± 0.34	1.36 ± 0.32*	1.51 ± 0.40*†

Data are presented as the mean ± SD.

*$P < 0.05$ in comparison to the athlete group; †$P < 0.05$ in comparison to the normal foot group (Kruskal-Wallis test).

groups were able to be classified with 94.25% accuracy only using the raw accelerometer and gyroscope data measured at the posterior pelvis during a single stride, regardless of walking speed. A deeper analysis on the pelvic movements including hike, drop, tilt, and rotation during gait, reflected in each axis of the inertial sensors, could be helpful to improve the practicality of the proposed approach.

This study is expected to serve as an inspiration to researches on relationships between individual's gait and intrinsic, physical, psychological, and pathological characteristics in the field of lower limb biomechanics, and these researches could provide clinical applicability of gait data to health monitoring.

This study has some limitations. The first is related to the use of spectrogram generated by STFT. In order to avoid problems with poor frequency resolution caused by a short-time window, other TF domain analysis methods such as Stockwell transform, continuous wavelet transform, and extended modified B-distribution will be used in future work. The second limitation is that this study was not able to avoid the use of black-box model which provides little understanding of the results produced by the model. Further study on the mechanism of the classifiers will be performed by visualizing the features after the CNN layers. The other limitations are associated with generalizability and reproducibility. Applying the proposed approaches to larger groups with diverse demographic, anthropometric, and pathological characteristics will be conducted along with inertial data acquisition using other wearable devices. Collecting each participant's gait data through repeated experiments over multiple days will be also carried out to ensure within-day and day-to-day reproducibility.

V. Conclusion

In this study, novel approaches for gait classification using wearable inertial sensor data were proposed. The gait parameters and spectrograms, trained in the neural network framework after being acquired from the acceleration and angular velocity data at the posterior pelvis and lower limbs, were effectively used to classify the athlete, normal foot, and deformed foot groups. This study has a potential to be applied for reliable and practical human identification and authentication based on gait assessed using wearable inertial sensors.

Acknowledgment

This research was supported by the High-tech based national athletic performance improvement (Winter), Korea Sports Promotion Foundation, and the Korea Institute of Science and Technology (KIST) Institutional Program (Project No. 2E29450).

References

[1] L. Wang, T. Tan, H. Ning, and W. Hu, "Silhouette analysis-based gait recognition for human identification," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 25, no. 12, pp. 1505-1518, 2003.

[2] C. Zhang, W. Liu, H. Ma, and H. Fu, "Siamese neural network based gait recognition for human identification," in *IEEE Int. Conf. Acoust. Speech Signal Process.*, Shanghai, 2016, pp. 2832-2836.

[3] M. S. Seyfioğlu, S. Z. Gürbüz, A. M. Özbayoğlu, and M. Yüksel, "Deep learning of micro-Doppler features for aided and unaided gait recognition," in *2017 IEEE Radar Conf.*, Seattle, 2017, pp. 1125-1130.

[4] R. P. Trommel, R. I. A. Harmanny, L. Cifola, and J. N. Driessen, "Multi-target human gait classification using deep convolutional neural networks on micro-Doppler spectrograms," in *2016 European Radar Conf.*, London, 2016, pp. 81-84.

[5] F. M. Castro, M. J. Marín-Jiménez, N. Guil, and N. Pérez de la Blanca, "Automatic learning of gait signatures for people identification," in *14th International Work-Conference on Artificial Neural Networks*, Cádiz, 2017, pp. 257-270.

[6] A. Mannini, D. Trojaniello, A. Cereatti, and A. M. Sabatini, "A machine learning framework for gait classification using inertial sensors: Application to elderly, post-stroke and huntington's disease patients," *Sensors,* vol. 16, no. 1, p. 134, 2016.

[7] Y. Zhang, G. Pan, K. Jia, M. Lu, Y. Wang, and Z. Wu, "Accelerometer-based gait recognition by sparse representation of signature points with clusters," *IEEE Trans. Cybern.,* vol. 45, no. 9, pp. 1864-1875, 2015.

[8] K. Aminian, B. Najafi, C. Büla, P. F. Leyvraz, and P. Robert, "Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes," *J. Biomech.,* vol. 35, no. 5, pp. 689-699, 2002.

[9] J. Kamruzzaman and R. K. Begg, "Support vector machines and other pattern recognition approaches to the diagnosis of cerebral palsy gait," *IEEE Trans. Biomed. Eng.,* vol. 53, no. 12, pp. 2479-2490, 2006.

[10] J. H. Hollman, E. M. McDade, and R. C. Petersen, "Normative spatiotemporal gait parameters in older adults," *Gait Posture,* vol. 34, no. 1, pp. 111-118, 2011.

[11] K.-R. Mun, G. Song, S. Chun, and J. Kim, "Gait Estimation from Anatomical Foot Parameters Measured by a Foot Feature Measurement System using a Deep Neural Network Model," *Sci. Rep.,* vol. 8, no. 1, pp. 1-10, 2018.

[12] W.-K. Lu and Q. Zhang, "Deconvolutive short-time Fourier transform spectrogram," *IEEE Signal Proc. Let.,* vol. 16, no. 7, pp. 576-579, 2009.