

Received 26 November 2023, accepted 5 December 2023, date of publication 7 December 2023, date of current version 13 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3340620

RESEARCH ARTICLE

Underfoot Pressure-Based Left and Right Foot Classification Algorithms: The Impact of Footwear

EVE MACDONALD^{1,2}, (Graduate Student Member, IEEE),
ROBYN LARRACY¹, (Graduate Student Member, IEEE),
ANGKON PHINYOMARK¹, (Member, IEEE), AND ERIK SCHEME¹, (Senior Member, IEEE)

¹Institute of Biomedical Engineering, University of New Brunswick, Fredericton, NB E3B 5A3, Canada

²School of Computing, Queens University, Kingston, ON K7L 3N6, Canada

Corresponding author: Robyn Larracy (rlarracy@unb.ca)

This work was supported in part by the New Brunswick Innovation Foundation's Strategic Opportunities Fund, in part by the Atlantic Canada Opportunities Agency's Regional Innovation Ecosystem Program, and in part by the Natural Sciences and Engineering Research Council of Canada's Alliance Grants Program under Grant ALLRP 558340-20.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Research Ethics Board of the University of New Brunswick under Application No. REB 2022-132.

ABSTRACT High-resolution plantar pressure recordings have the potential to be used in gait biometrics, biomechanics, and clinical gait analysis. To accurately assess side-specific patterns and asymmetries, it is essential to differentiate between left and right steps, which can be challenging when manual labeling is not feasible and shoe type can vary. This research aimed to create and evaluate the performance of six distinct algorithms (two inspired by existing literature and four novel ones) that take advantage of spatial and temporal features combined with basic decision rules, machine learning, and deep learning to automatically classify left and right footsteps from underfoot pressure recordings, taking into account difficulties associated with footwear variability. A collection of more than 20,000 footsteps from 20 people and 41 different types of shoes was used to assess the six proposed classification algorithms. The results demonstrate that classification techniques based on spatial representations (peak pressure or binary images of footsteps) are more effective than those based on center-of-pressure (COP) time series. The most successful approach, which compares the area of the sole in different parts of the midfoot and forefoot, achieved an accuracy of 99.7% in determining left and right footsteps, with a convolutional neural network (CNN) algorithm at a close second (99.4%). These techniques were found to be robust to many types of footwear and may be valuable for a variety of practical, community-based gait classification tasks.

INDEX TERMS Biomechanics, biometrics, center of pressure trajectory, foot classification, footstep recognition, footwear, gait representation, gait recognition, peak pressure, plantar pressure images.

I. INTRODUCTION

The pressure exerted on the ground when walking is unique to each person and is the result of complex biomechanical, physiological, and behavioral processes. Recent technological advances in pressure-sensing tiles and mats have enabled the collection of precise, high-resolution recordings of the pressure applied to the ground. These data have a variety of

The associate editor coordinating the review of this manuscript and approving it for publication was Sotirios Goudos¹.

applications, such as gait biometrics, clinical gait analysis, and monitoring of health and rehabilitation in smart homes and facilities [1]. These systems capture the pressure exerted on the ground at multiple points in time, providing a wealth of information about a person's movement from heel to toe with each step.

To gain maximum insight into the high-dimensional, raw-pressure recordings, it is essential to extract and label individual footsteps as either left or right steps. This is important because asymmetries in gait are common in both

healthy and patient populations [2]. These asymmetries are a unique feature of an individual's gait [3], and can be used to track gait rehabilitation and disease progression [4], [5], as well as predict fall risk [6], [7]. To do this, it is necessary to analyze each side separately, uncover any distinct patterns, and compare the two sides.

Research into integrated pressure-based systems for gait analysis has been largely limited to laboratory or clinic-based barefoot walking, where left and right footsteps can be identified manually or through a straightforward protocol (e.g. asking participants to step on the sensor with a certain foot or stride pattern [8], [9], [10]). However, for practical, unsupervised applications such as long-term home or community monitoring, the large number of steps and the variety of walking paths make manual labeling impractical. Additionally, many of these applications may involve users wearing shoes, which could present additional difficulties for simple rule-based labeling techniques, especially for shoes that the system has not seen before. Consequently, there is a need for a reliable automated labeling system for left and right footsteps from pressure recordings to make the most of the technology in real-world settings.

Only a handful of studies have developed algorithms for automated classification of left and right footsteps, and they only considered barefoot samples. Certain techniques rely on the presence of multiple steps to exploit relationships such as the angle between feet [11], [12], yet they are not suitable for scenarios where users take unexpected routes or where the sensor platform only captures one footstep. Other studies have proposed techniques for classifying left and right using spatial or peak pressure features of individual footsteps, such as the number of pixels in different parts of the foot [13], the similarity of peak pressures to left and right foot template images [14], or deep transfer learning of peak pressure images [15]. Unfortunately, since these works only examined unshod (barefoot or sock-foot) samples, it is not known how these techniques will work with different types of footwear. To the best of our knowledge, no prior studies have proposed a foot classification algorithm for shod data. Therefore, the purpose of this study was to investigate the effectiveness of six different algorithms (two inspired by the existing literature and four new ones) in classifying left and right footsteps using underfoot pressure data, while taking into account the challenge of different types of footwear.

II. METHODS

A. PARTICIPANTS

The underfoot pressure data used in this work were collected as part of an ongoing pressure-based gait biometric project using a pressure-sensitive flooring system at the University of New Brunswick (UNB), Canada. This study included the footsteps of 20 subjects (10 women, 10 men) with ages ranging from 20 to 71 and shoe sizes from US women's size 5.5 to US men's size 13. Participants self-reported their race or ethnicity from a list of eight categories (Aboriginal,

Black, East/Southeast Asian, Latino, Middle Eastern, South Asian, White, or Other). Out of the group, 13 participants were White, two East/Southeast Asian, two South Asian, one Middle Eastern and two chose Other. Six participants were left-leg dominant and fourteen were right-leg dominant, where dominance was determined by the participant's answer to the question "Which leg would you normally use to kick a stationary ball straight in front of you?" [16]. All subjects provided their informed consent to participate in the study, as approved by the Research Ethics Board of the University of New Brunswick (REB 2022-132).

B. DATA COLLECTION

Dynamic walking gait footsteps were collected as participants walked back and forth on a runway consisting of a 2×6 grid of instrumented tiles (developed by Stepscan Technologies Inc.), with an additional 2×2 grid of inactive tiles at each end to allow for turning (Fig. 1). Each pressure sensing tile measures 60×60 cm with a resolution of 120×120 pressure sensitive sensors (or pixels), resulting in an active recording area of 4.32 m^2 , and a total of 172,800 pixels. The participants conducted multiple 90-s walking trials in four footwear conditions: (1) no shoes (barefoot or socks), (2) standard shoes (Grand Court 2.0 Shoes, Adidas, provided by the research team) and (3, 4) two pairs of their own personal, commonly worn shoes. These personal shoes included dress shoes (e.g., Oxfords, high heels), athletic shoes (e.g., running shoes, trainers), sandals (e.g., thong sandals, slides) and casual shoes (e.g., canvas shoes, loafers, slip-ons). Figure 2 shows a selection of pressure profiles under the foot for various footwear conditions. As part of a larger ongoing project, participants walked at four self-selected walking speeds: (1) a comfortable, regular pace, (2) a fast pace, (3) a slow pace, and (4) a slowdown (walking slower and stopping at the end of the runway) for each footwear condition. As the slowdown trials also included static footsteps, the present analysis considered only the first three walking speeds to focus on the dynamic footsteps of interest. With a sampling frequency of 100 Hz, each 90-s trial resulted in a recording of $240 \times 720 \times 9000$ (y, x, frames) and between 52 and 119 captured footsteps, depending on the participant and the walking speed.

The recordings were processed to extract individual footsteps, which were then roughly aligned by translation to their center of mass and rotation to their first principal component axis. The footsteps were normalized in time to 101 frames using nearest-neighbor interpolation. This resulted in a total of 20,083 footsteps of size $75 \times 40 \times 101$, approximately 1,000 per subject. This dataset is a subset of a larger footstep database that will be made available to the research community when collection concludes.

C. GROUND TRUTH LABELING

Ground truth labels were acquired by combining visual inspection and spatial relationships between consecutive

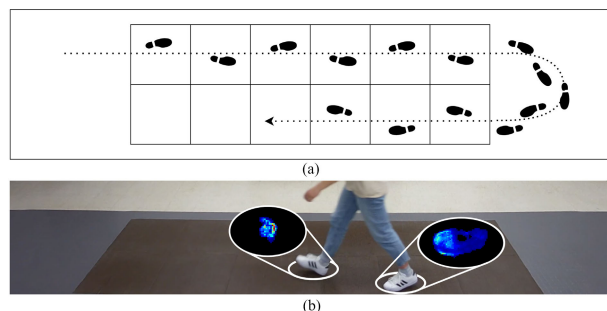


FIGURE 1. Demonstration of the tile setup and data collection protocol: (a) an example segment of a walking trajectory during a trial, where participants walked back and forth across the tiles for 90 s using the inactive regions at both ends of the runway for turning, and (b) a still from one walking trial with corresponding pressure recordings.

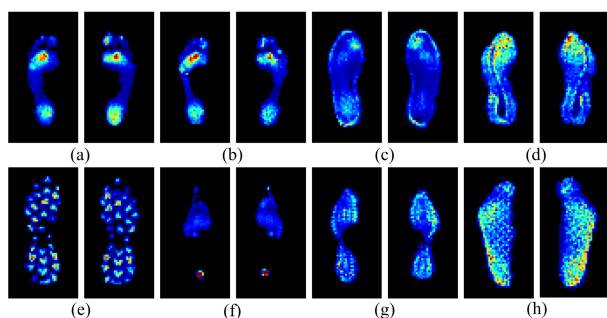


FIGURE 2. Example peak pressure images for a selection of footstep pairs with different footwear conditions: (a) barefoot, (b) sock-foot, (c) standard shoes, (d) running shoes, (e) hiking shoes, (f) high heels, (g) casual sneakers, (h) sandals.

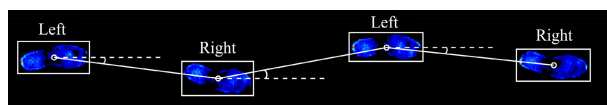


FIGURE 3. Demonstration of the angle-based method used in conjunction with visual inspection of video recordings to obtain ground truth labels.

footsteps. Video footage taken simultaneously with the pressure recordings was analyzed to determine which leg was used for the first step of each pass over the tiles and to predict an expected left/right sequence for the pass (e.g. for five recorded footsteps with the first one identified as a left foot: left, right, left, right, left). As participants walked in a consistent straight line across the tiles as per the collection protocol, the angles between successive footsteps were then used to confirm the predicted sequences. Figure 3 illustrates this process for a sequence of four steps, where the centroids of the bounding boxes were compared to calculate the angles between each step. For example, with users walking from left to right across the runway, negative angles with respect to the x -axis (i.e., the direction of walking) indicated a right step, and positive angles indicated a left step. Labels that did not match the expected walking patterns, such as two consecutive left or right footsteps, were flagged for further manual inspection from the video and corrected as necessary. In addition, all steps were examined visually to remove any

incomplete steps (e.g. participants walked partially off the tiles at the start or end of the runway) and to further confirm the labels. In total, the dataset contained 9,952 left steps (49.6%) and 10,131 right steps (50.4%).

D. LEFT AND RIGHT FOOT CLASSIFICATION

Six different techniques were implemented for automated classification of left and right footsteps. Three of them are based on temporal information (i.e., foot center of pressure (COP) time series), and the other three are based on spatial information (i.e., gait representations such as binary or peak pressure images of the footstep). In the time domain, kinetic variables obtained from the COP time series provide a thorough understanding of the forces and moments that act on the body during walking. Foot COP, which has been widely used in biometrics and biomechanics research [17], [18], [19], is generally expressed in terms of the mediolateral (ML) and anteroposterior (AP) directions. This is the path of pressures underfoot from the moment the heel touches the ground to the time the toe leaves it. It is calculated for each time point as the average coordinates of the foot weighted by the pixel intensities (i.e., pressure values). The COP was determined for each step taken, resulting in two time series of 101 points for each sample. To reduce noise, the time series were also filtered using a second-order Butterworth low-pass filter with a cutoff frequency of 20 Hz. In the spatial domain, peak pressure images, also known as maximum pressure or 100th percentile pressure (P100) images [8], are two-dimensional (2D) representations of the footstep that display the highest pressure sustained by each sensor (pixel) during the stance phase. This is the variable most commonly used in the literature on plantar pressure and is one of the most effective gait representations for gait recognition [8], [20]. The recordings of footsteps in this study capture high-resolution shape and texture data from the sole during contact with the ground (Fig. 2).

1) COP ENDPOINTS (COP EP)

The first classification approach was developed by comparing the initial and final ML coordinates of the COP time series. An example COP trajectory for one right footstep is depicted in Fig. 4(a). The triangle marker indicates the beginning of the stance (heel strike), and circle marker denotes the end of the stance (toe off). A rule-based classifier was created to compare the ML endpoints; if the toe-off's COP was higher on the ML axis than the heel strike, it was classified as a left foot, and the opposite was true for the right foot.

2) COP DYNAMIC TIME WARPING (COP DTW)

The second classification approach uses dynamic time warping (DTW) and foot COP time series. DTW is a method for measuring similarity between two time series that are not necessarily synchronized in time. The COP time series of the left and right footsteps of the training subjects were averaged to create COP templates, with the test user excluded

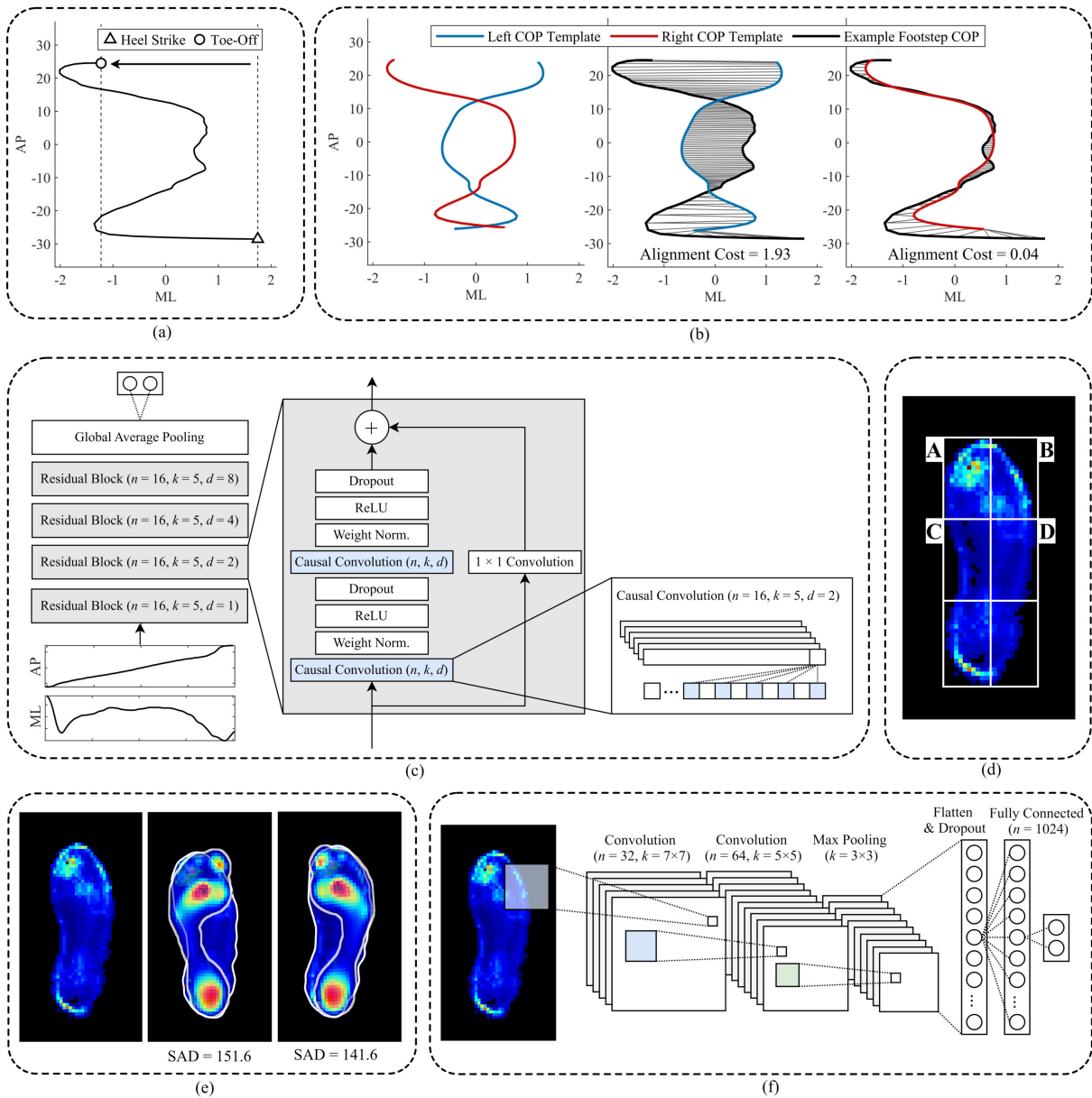


FIGURE 4. Demonstration of classification techniques for an example right footstep with the standard shoe: (a) COP EP, (b) COP DTW, (c) COP TCN, (d) P100 PC, (e) P100 TM, and (f) P100 CNN.

to prevent any information leakage. The test steps were compared to the templates by means of DTW with a cosine distance metric. The template with the lowest alignment cost (i.e., the global distance between the warped COP and the template) determined the side prediction for the footstep (Fig. 4(b)).

3) COP TEMPORAL CONVOLUTIONAL NETWORKS (COP TCN)

The third classification approach uses temporal convolutional networks (TCNs) [21] and the ML and AP COP time series. The TCN framework uses casual convolutions and dilations to adapt to sequential data, taking into account its temporal nature and necessity for large receptive fields. This study

examines an architecture that employs a sequence of four residual TCN blocks (Fig. 4(c)). Each block is made up of two 1D causal convolution layers followed by weight normalization, ReLU activation functions, and 25% dropout. All convolution layers use $n = 16$ kernels of length $k = 5$, but the dilation factors (d) increase by a power of two with each residual block: the four blocks use dilation factors of $d = 1, 2, 4,$ and 8 , respectively. An example of a causal convolution with a dilation factor of two is shown in Fig. 4(c), which demonstrates the increased range of time points that contributes to each neuron of the activation map. The TCN models were trained for 35 epochs with a learning rate of 0.001 and a batch size of 128, using Adam optimization.

4) P100 PIXEL COUNTING (P100 PC)

Li et al. [13] proposed a classification approach for barefoot pressure images that leveraged existing knowledge of the foot structure. Specifically, they noted that the midfoot tends to be more prominent on the lateral side of the sole. To implement their strategy, the midfoot region of the P100 images was segmented into two halves along the ML axis, and the number of active (non-zero) pixels in each half was used to determine whether the step belonged to the left or right foot. Inspired by this strategy, in this study a modified version was developed that also takes into account the anatomy of the forefoot, which is usually more prominent on the medial side of the sole. The P100 images were cropped to their minimum bounding rectangle and divided into six equal parts by dividing by two along the ML axis and by three along the AP axis, as shown in Fig. 4(d). The four upper segments (labeled A, B, C, and D), which typically included the forefoot and midfoot, were used to distinguish between the left and right steps. If the number of active pixels in regions A and D was higher than the number of active pixels in regions B and C, the step was classified as a right step. Otherwise, it was classified as a left step.

5) P100 TEMPLATE MATCHING (P100 TM)

Oliveira et al. [14] proposed a template matching (TM) technique to differentiate between left and right P100 pressure images of barefoot walking. To do this, a representative right footstep was manually chosen from the dataset and used as the right foot template, with a flipped version of that step serving as the left foot template. In this study, the Münster104 left and right templates [3], which are 63×27 barefoot peak pressure images averaged over 104 healthy individuals, were used as reference left and right footsteps for TM. The P100 images were normalized to the same range as the templates and then aligned to each template using a linear transformation (including rotation, scaling, and translation). The footsteps were then classified as left or right steps based on the sum of the absolute differences (SAD) between each normalized, aligned image and its respective template. The template with the lowest SAD, that is, the template that best matched the footstep, determined the predicted label (Fig. 4(e)).

6) P100 CONVOLUTIONAL NEURAL NETWORK (P100 CNN)

A convolutional neural network (CNN) was employed as the last classification approach to automatically learn features and classify peak pressure (P100) images. This research examines a lightweight network that has two convolutional layers (the first with $32 \times 7 \times 7$ filters; the second with $64 \times 5 \times 5$ filters), followed by max pooling (3×3 filter with a stride of 2) and a fully connected layer (1024 hidden neurons) (Fig. 4(f)). The convolutional layers were followed by ReLU activation functions, and the 25% dropout was applied after the max pooling layer. CNN models were trained using Adam optimization for five epochs with a learning rate of $1e-5$ and a batch size of 128.

E. EVALUATION

The three scenarios that were evaluated for each classification technique were: (1) unshod (barefoot or socks), (2) shod (wearing any type of footwear), and (3) all (both barefoot and shod) samples. A leave-one-subject-out cross-validation was used to divide the samples into training and test sets for the techniques that required training samples. Nineteen participants' footsteps were used to create the models, and one participant's samples, which were not used in the training, were used to assess the models' ability to generalize to unseen users and, for scenarios (2) and (3), unseen footwear.

A statistical analysis was conducted to compare the performance of various classification techniques and evaluation scenarios, as well as to determine the influence of factors such as sex, footwear type, dominant foot, and walking speed on classification. First, to decide which technique was the most effective, paired t -tests were conducted to compare the accuracy of the six techniques. Second, to evaluate the difference in performance between unshod and shod samples, paired t -tests were conducted to compare the performance estimates of scenario (1) with the estimates of scenario (2) for each of the six techniques. Finally, two-sample t -tests were used to compare independent subgroups, such as males and females, those aged 25 or under and those older than 25, and those who walked barefoot and those who wore socks during the unshod trials, as well as casual shoes, dress shoes, athletic shoes, and sandals. Paired t -tests were used to compare other groups, such as standard shoes and shoes owned by participants, the left and right legs, the dominant and non-dominant legs, and slow, normal, and fast walking. In addition, Cohen's effect size d was used to measure the magnitude of the difference between two groups. This was calculated by dividing the difference between the means of the two groups by the pooled standard deviation. A small effect size was defined as 0.2, a medium effect size as 0.5, and a large effect size as 0.8 or higher. Comparisons with a p value less than 0.05 and a d value greater than 0.20 were considered significant.

III. RESULTS

A. FOOT CLASSIFICATION ALGORITHMS

In this research, four novel techniques (COP EP, COP DTW, COP TCN, P100 CNN) and two modified techniques, inspired by previously proposed strategies for barefoot samples (P100 PC, P100 TM), were proposed for the classification of left and right footsteps [13], [14]. The two modifications made to previous techniques resulted in an increase in accuracy. Specifically, Li et al. [13] proposed a pixel counting strategy which only took into account the active pixels in the mid-foot region of the P100 images. This yielded accuracies of $97.7 \pm 3.7\%$ and $84.1 \pm 9.0\%$ for the unshod and shod samples, respectively. However, the proposed P100 PC technique, which also considers the forefoot, achieved 100% perfect accuracy for unshod samples

and $99.6 \pm 0.4\%$ accuracy for shod samples, representing significant improvements of $d = 0.85$ for the unshod scenario and $d = 2.29$ for the shod scenario ($p < 0.05$). Similarly, the proposed P100 TM technique demonstrated an increase in classification performance. By selecting representative right footsteps from the training set to serve as templates (the original footstep as the right template and a flipped version as the left template) as in Oliveira et al. [14], an accuracy of $99.5 \pm 1.6\%$ was found for the unshod samples and $92.5 \pm 10.3\%$ for the shod samples. By using the Münster104 healthy barefoot templates [3], the performance of the P100 TM technique was increased to 100% for unshod samples (albeit a non-significant improvement; $p = 0.17$, $d = 0.45$) and $99.0 \pm 1.0\%$ for shod samples (a significant improvement; $p < 0.05$, $d = 0.87$).

TABLE 1. Comparison of the classification accuracies of the six proposed left and right classification techniques and between the unshod samples (barefoot or sock-foot) and shod samples (in the standard footwear or the participants' shoes).

Technique	Unshod ($\mu \pm \sigma$ %)	Shod ($\mu \pm \sigma$ %)	All ($\mu \pm \sigma$ %)
COP EP	87.6 \pm 11.3	73.1 \pm 19.6	76.9 \pm 16.1
COP DTW	94.2 \pm 5.4 [*]	88.7 \pm 7.5 [*]	89.0 \pm 6.9 [*]
COP TCN	95.9 \pm 4.4 ^{*†}	95.6 \pm 4.3 ^{*†}	95.9 \pm 3.7 ^{*†}
P100 PC	100 \pm 0^{*†‡}	99.6 \pm 0.4^{*†‡§}	99.7 \pm 0.3 ^{*†‡§}
P100 TM	100 \pm 0^{*†‡}	99.0 \pm 1.0^{*†‡}	99.3 \pm 0.8 ^{*†‡}
P100 CNN	99.8 \pm 0.5 ^{*†‡}	98.5 \pm 3.6 ^{*†‡}	99.4 \pm 1.1 ^{*†‡}

Symbols (*, †, ‡, and §) represent a significant increase in performance ($p < 0.05$ and $d > 0.20$) compared to: * COP EP, † COP DTW, ‡ COP TCN, or § P100 TM using a paired t -test. Bold fonts represent a significant difference between unshod and shod samples for a given technique using a paired t -test. The classification performance is reported as the mean and standard deviation across twenty users.

The classification accuracies of the six proposed techniques for the (1) unshod, (2) shod, and (3) unshod and shod (all samples) scenarios are presented in Table 1. The P100 PC technique had the highest accuracy on average among the 20 participants, with a score of 99.7%. The P100 CNN and P100 TM techniques were close behind, with accuracies of 99.4% and 99.3%, respectively. Despite the wide range of shoe sole shapes and profiles (Fig. 2), the three classifiers based on spatial P100 image features outperformed those based on the COP time series in all cases ($p < 0.05$, $d = 1.20 - 2.14$). The COP EP technique had the worst results overall, with an average accuracy of 76.9% for all samples (ranging from 47.0% to 98.4% among 20 users). Of the three COP techniques, the TCN deep learning classifier had the highest accuracy of 95.9%.

The COP DTW and P100 TM techniques had the highest computational cost, requiring an average of 7.6 seconds to classify all of the footsteps for one user (approximately 1000 footsteps, around 7.6 ms/step) on an Intel 2.00 GHz Xeon CPU with 51 GB RAM. In comparison, the COP EP, COP TCN, P100 PC, and P100 CNN techniques were much faster, taking less than 0.3 ms/step. The CNN and TCN classifiers required a one-time training computation, which

took an average of 34 s and 67 s to complete on an NVIDIA Tesla T4 GPU.

B. THE IMPACT OF FOOTWEAR AND OTHER FACTORS

The results showed that the performance of the six techniques decreased when shoes were worn, with the COP EP, P100 PC, and P100 TM techniques being the most affected ($p < 0.05$). The average accuracy for the six techniques without shoes was 96.2%, while the average accuracy with shoes was 92.4%. However, the P100 PC technique still managed to achieve an impressive 99.6% accuracy in distinguishing between left and right footsteps for a database of 41 different types of footwear. This included potentially difficult shoes such as high heels (Fig. 2(f)), for which the P100 PC technique correctly classified all but two samples ($\approx 99\%$).

Table 2 examines how various factors affect classification performance. Unless otherwise specified (for example, when comparing different types of footwear), performance estimates include both unshod and shod samples. No differences in performance were observed between male and female participants, those younger and older than 25 years old, or walking with or without socks for any of the six classification techniques ($p > 0.05$). It was observed that the COP EP and P100 PC techniques had varying results depending on the type of shoe worn. The COP EP technique was not as successful when casual shoes (e.g. flat-soled sneakers) were worn compared to when athletic shoes (e.g. running sneakers with structured soles) were worn (85.4% > 64.7%). The P100 PC technique had a lower success rate for footsteps in dress shoes compared to sandals, yet still achieved more than 99% accuracy in both cases. It was remarkable that the three COP techniques and the P100 TM technique had significantly worse results with the standard shoes compared to the participants' own shoes ($p < 0.05$), even though the standard shoes were included as training samples from other participants. There were notable performance differences between the left and right samples for the COP EP, COP DTW, and P100 TM techniques, although the trends varied between the techniques. Furthermore, these disparities were not related to leg dominance. Lastly, the P100 CNN technique was found to perform differently depending on the walking speed; it was observed that the classification of footsteps taken at a slow pace was higher than those taken at a normal speed.

IV. DISCUSSION

The classification of left and right footsteps is a key pre-processing step for pressure-based gait recording technologies to enable further analysis and classification of footsteps. For example, footsteps can be compared regionally by registering them with side-specific templates [3], [22]. This allows the comparison of footprints from different individuals or groups, for example, to evaluate the pressure distribution of those with hallux valgus compared to healthy people [23], or for identification purposes [8]. This left and right labeling step can be done manually or with the aid of basic

TABLE 2. Effect of factors: sex, age, unshod condition, footwear category, footwear type, side, side dominance, and walking speed on the classification of left and right footsteps.

Factor	Subgroup (<i>N</i>)	COP EP ($\mu \pm \sigma$ %)	COP DTW ($\mu \pm \sigma$ %)	COP TCN ($\mu \pm \sigma$ %)	P100 PC ($\mu \pm \sigma$ %)	P100 TM ($\mu \pm \sigma$ %)	P100 CNN ($\mu \pm \sigma$ %)
Sex	Male (10)	78.9 ± 19.7	90.7 ± 6.3	96.8 ± 2.9	99.6 ± 0.3	99.1 ± 0.9	99.5 ± 0.7
	Female (10)	75.0 ± 10.9	87.3 ± 7.0	95.0 ± 4.2	99.8 ± 0.3	99.4 ± 0.5	99.3 ± 1.3
Age	≤ 25 (13)	75.0 ± 16.7	87.7 ± 7.4	96.0 ± 3.9	99.7 ± 0.3	99.2 ± 0.9	99.4 ± 1.3
	> 25 (7)	80.4 ± 14.2	91.6 ± 4.9	95.9 ± 3.5	99.7 ± 0.3	99.4 ± 0.5	99.6 ± 0.4
Unshod Condition	Barefoot (4)	87.9 ± 8.2	95.3 ± 0.8	93.6 ± 8.4	100 ± 0	100 ± 0	99.9 ± 0.2
	Sock-Foot (16)	87.6 ± 11.9	94.6 ± 5.4	96.2 ± 7.0	100 ± 0	100 ± 0	99.9 ± 0.1
Footwear Category	Casual (13)	64.7 ± 19.7*	88.4 ± 8.0	96.9 ± 2.1	99.5 ± 0.7	99.4 ± 1.1	99.5 ± 0.9
	Dress (6)	76.0 ± 29.6	88.4 ± 6.2	95.9 ± 4.8	99.2 ± 0.5*	99.3 ± 0.7	98.9 ± 0.6
	Athletic (13)	85.4 ± 17.6*	91.5 ± 6.5	97.4 ± 2.7	99.7 ± 0.5	99.6 ± 0.7	98.8 ± 3.3
	Sandal (8)	76.1 ± 13.5	94.9 ± 3.9	97.4 ± 2.7	99.9 ± 0.1*	99.8 ± 0.4	99.9 ± 0.1
Footwear Type	Standard Shoe	68.2 ± 23.2*	79.6 ± 14.3*	93.9 ± 5.3*	99.7 ± 0.6	98.0 ± 2.6*	99.1 ± 1.5
	Participant Shoes	75.5 ± 18.9*	90.7 ± 6.0*	97.1 ± 2.8*	99.6 ± 0.4	99.5 ± 0.5*	99.3 ± 1.6
Side	Left Foot	87.7 ± 10.0*	86.6 ± 8.6*	95.5 ± 5.6	99.8 ± 0.3	99.5 ± 0.8*	99.4 ± 1.7
	Right Foot	66.6 ± 23.9*	91.6 ± 6.7*	96.5 ± 4.4	99.7 ± 0.4	99.0 ± 0.9*	99.4 ± 1.1
Side Dominance	Dominant Leg	73.4 ± 23.0	88.8 ± 9.3	95.8 ± 5.7	99.7 ± 0.4	99.1 ± 0.9	99.1 ± 1.9
	Non-Dominant Leg	80.9 ± 18.3	89.4 ± 6.7	96.1 ± 4.2	99.8 ± 0.3	99.4 ± 0.8	99.8 ± 0.3
Walking Speed	Slow	77.3 ± 15.6	88.2 ± 7.6	95.7 ± 4.1	99.8 ± 0.3	99.3 ± 0.6	99.7 ± 0.8*
	Fast	77.2 ± 15.4	89.8 ± 7.2	96.3 ± 3.9	99.7 ± 0.5	99.4 ± 0.7	99.4 ± 1.2
	Normal	76.0 ± 18.2	88.8 ± 7.3	95.7 ± 4.1	99.7 ± 0.4	99.1 ± 1.2	99.3 ± 1.2*

Bolded fonts and pairs of symbols (*) represent a significant difference in performance ($p < 0.05$ and $d > 0.20$) between two subgroups for a given classification technique. Two-sample t -tests were used to compare independent subgroups (for the sex, age, unshod condition, and footwear category factors) and paired t -tests were used otherwise (for the footwear type, side, side dominance, and walking speed factors).

attributes such as foot progression angle for smaller clinical examinations. However, this is not a practical approach for most real-world environments, especially when gait data is collected on a large scale over a long period of time. Additionally, classification techniques that depend on relationships between multiple steps will not be successful if sensors detect unexpected walking paths or if only a single footstep is captured. Potential environments such as care homes and facilities with gait biometric systems, like airports and offices with restricted access, are also likely to encounter footwear. This presents an additional challenge for classification, as even footsteps from the same person can vary significantly depending on the type of shoes they are wearing. This research investigated the ability to automatically classify a single footstep as a left or right step, taking into account a wide variety of footstep samples with a broad range of foot sizes and shoe types. As a result, a foot classification algorithm with near-perfect accuracy (99.7%) was successfully developed.

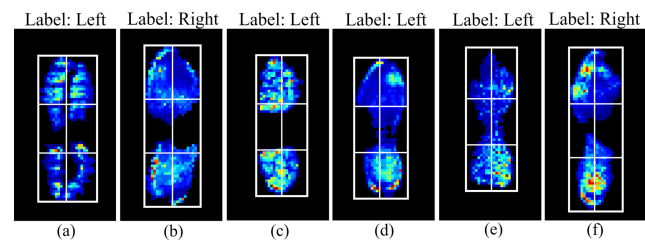


FIGURE 5. Examples of misclassified steps using the P100 PC technique and their true labels.

A. FOOT CLASSIFICATION ALGORITHMS

Existing research on foot classification techniques has been limited to barefoot samples. The current investigation revealed that previously proposed techniques [13], [14]

demonstrated relatively poor performance for shod samples in the present work (84.1 – 92.5%), despite achieving perfect classification accuracy for barefoot images in the original studies. For both the P100 PC and P100 TM methods, which were modified from the approaches in [13] and [14], the proposed changes increased the classification accuracy of the shod samples to $\geq 99\%$.

Of the six proposed classification techniques, the P100 PC approach had the best overall results, with a perfect accuracy rate for unshod samples and an average accuracy of 99.6% for shod samples. Additionally, it is a highly efficient technique in terms of computation, making it ideal for real-time processing of gait recordings. Figure 5 shows examples of some of the footsteps that were misclassified using the P100 PC technique. Notably, these misclassified steps were isolated instances; even for shoes with complex sole impressions such as those in Fig. 5(a) and 5(c), the pixel counting rule was successful in classifying the majority of samples. Only 1% of the samples from (a) and 3% of the samples from (c) were misclassified with this technique. The P100 CNN and P100 TM techniques, by contrast, were more affected by shoe type, with the CNN misclassifying more than 20% of the samples wearing the footwear from (a). However, the P100 PC approach did struggle in some cases with overpronated steps, as demonstrated by samples (d) and (f) in Fig. 5. Since these footsteps were not frequently encountered in the dataset, the P100 PC technique should be further tested in individuals with a consistently overpronated gait, as well as other potentially challenging cases, such as feet and shoes with high arches or irregular sole shapes. Interestingly, unlike the P100 TM and P100 CNN classifiers, the P100 PC technique does not take pixel intensities into account, only the shape of the sole’s contact with the ground. Although the P100 CNN and P100 TM techniques had

poorer performance overall for the dataset, these techniques do consider pressure information and did not have the same weakness to overpronated steps (i.e., both techniques were able to correctly classify the samples (d) and (f) in Fig. 5). It may, therefore, be advantageous to refine the P100 PC strategy in future studies to include the pressure distribution beneath the feet in addition to the spatial information. Moreover, to reduce the risk of misclassified steps, a confidence threshold could be established. The P100 PC classifier was more confident in correctly classified steps than in misclassified ones, given the ratio of active pixels in segments A and D to segments B and C (Fig. 4(d)) ($p < 0.05$, $d = 5.52$). By rejecting steps with a ratio > 0.85 , possibly flagging them for further evaluation, more than 90% of the misclassified steps using the P100 PC technique could be identified, with the cost of flagging approximately 10% of the otherwise correctly classified steps as well.

A deep learning-based foot classification technique, P100 CNN, was almost as successful as the leading method, P100 PC, achieving an average accuracy of 99.4% ($p < 0.05$). This research used a relatively large and diverse dataset compared to previous foot pressure studies [18], [24], with approximately 1,000 steps per subject and 41 different types of shoes among 20 participants. However, if the size of this shoe dataset were increased, the CNN might be able to gain a better understanding of left and right footstep characteristics across shoe types. These abstract, non-linear features may perform well even in cases where the basic pixel counting method is not effective. Future research should include a larger sample size and a broader range of shoe types, particularly more difficult ones, such as high heels and pointed-toe dress shoes with more symmetrical sole profiles. Additionally, this research used a computationally efficient object detection technique based on connected component labeling and SORT (Simple Online Realtime Tracking) [25] for footstep extraction as a pre-processing step. A prior study [15] applied deep learning-based object detection models to recognize left and right barefoot plantar pressure images and achieved an accuracy of 99% in classifying 974 barefoot images using YOLOv4 [26]. It would be beneficial to explore whether YOLO or other deep learning-based object detection methods could be used to combine the detection of footsteps and their labeling (left or right) for shod data.

A third approach based on spatial features (from P100 peak pressure images), P100 TM, also yielded excellent results. This technique, which evaluated each footstep's similarity to barefoot left and right templates, achieved a perfect accuracy rate for unshod samples and an average accuracy of 99.0% for shod samples. In this research, the Münster104 templates used for the P100 TM technique, although constructed from barefoot footsteps, provided a more universal standard for left and right pressure distributions than hand-picked samples from the dataset. Further research should be conducted to create templates from a wide range of shod footsteps, which could potentially lead to increased accuracy when using

the template matching (TM) technique. Additionally, rather than peak pressure images, other representations such as mean pressure images (MPI) and motion silhouette images (MSI) could be considered as input to the proposed spatial feature classification techniques [20]. Peak pressure images, which are the most prevalent image representation in plantar pressure research [8], were selected for this work in part due to the availability of the Münster104 peak pressure-based templates [3]. The University of New Brunswick is currently collecting data that could be used to create other forms of 2D templates from a variety of barefoot and shod footsteps.

The left and right classification techniques based on spatial features (from P100 images) were found to be more effective than those based on COP trajectories. This could be attributed to the considerable variability in COP between individuals, footwear conditions, and even step-by-step in some cases. Figure 6 illustrates the variability in left footsteps for the unshod and standard shoe trials. The distinctness of COP profiles between individuals has been well documented in the literature [11], [27], [28], [29], and the variability of foot COP has also been shown to be affected by footwear [30]. The COP EP and COP DTW techniques were able to accurately classify samples for some of the twenty individuals with a very high degree of accuracy ($> 95\%$), yet they were entirely unsuccessful for some participants and shoe types, resulting in accuracies of around 50%, equivalent to random guessing. On average, TCN's deep-learned features were able to better capture these variable COP patterns (i.e., $> 80\%$ accuracy for all participants and shoe types).

Notably, COP-based time series techniques have an advantage over peak-pressure images in that they can be computed from relatively low-resolution sensors or even force plates. This makes them more cost-effective and allows for gait analysis and classification in a wider range of environments. The sensor system used in the present work had a very high resolution (4 sensors/cm², or 40,000 sensors/m²), which may have been a factor in the success of P100-based spatial image techniques. Vera-Rodriguez et al. [31] found that a resolution of at least 650 sensors/m² was necessary for the biometric recognition of cumulative pressure footstep images, as performance using these spatial features was significantly degraded with lower sensor densities (i.e., increases in equal error rate (EER) of 7.7% and 13.2% using subsampled sensor arrays of 430 sensors/m² and 220 sensors/m², respectively). They suggested that higher resolutions would provide even better results than their 650 sensor/m² configuration, given the observed trends in performance. Therefore, future research should investigate the effect of sensor resolution on the efficacy of spatial feature-based left and right footstep classification techniques. The combination of spatial (P100) and temporal (COP) features should also be considered, which could potentially improve the accuracy of predictions for footsteps with uncertain outcomes, especially for systems with lower sensor density.

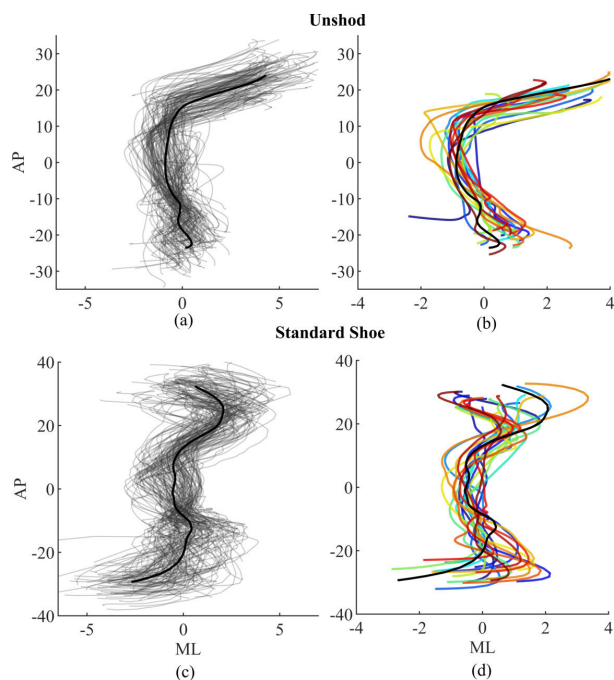


FIGURE 6. Within-user and between-user variability in COP trajectories for barefoot and standard shoe left footsteps; (a) and (c) are the COP trajectories for all footsteps from one participant, and (b) and (d) are the averaged COP trajectories for all 20 participants.

B. THE IMPACT OF FOOTWEAR AND OTHER FACTORS

The accuracy of the left and right foot classification techniques was impacted by certain factors, as is evident from Tables 1 and 2. This study discovered that the use of shoes can have a significant effect on classification performance. Surprisingly, prior studies have only focused on barefoot footsteps, and to the best of our knowledge, this is the first investigation to look into the influence of shoes on the accuracy of distinguishing between left and right feet.

Some performance differences were observed for different categories of footwear, such as significantly poorer performance in classifying dress shoes compared to sandals with the P100 PC technique. When classification techniques are developed for applications that involve shod footsteps, the influence of the type of shoe worn must be taken into account. Not only do different types of shoes leave different sole impressions and pressure distributions, they can also affect the way people walk. Bouchrika and Nixon [32] showed that recognition rates from gait video sequences were much lower when people wore flip-flops than when they were barefoot or wearing trainers or boots. Similarly, high heels have been shown to significantly impact both gait biomechanics and underfoot pressure during gait compared to lower heeled shoes [33], [34]. Footwear has also been associated with changes in walking speed [33], [35], and this study discovered that walking speed may have an effect on the accuracy of classification. The P100 CNN model was able to classify the footsteps of the slow walking trials with greater precision

than those of the normal speed trials. It is possible that the influence of shoes on walking can vary over time, both in the short-term (e.g. after an hour of walking in high heels or flat-soled shoes [36]) and in the long-term (e.g. due to musculoskeletal damage caused by wearing high heels for a long period of time [33]). Previous research with gait pressure recordings has suggested methods for reducing the influence of footwear on gait classification tasks, including person identification; for example, including examples of high heels as training data [31], [37] or using a weighting filter to eliminate irrelevant sole information and focus only on the sole of the shoe directly beneath the barefoot [38]. Future research could explore the use of similar methods to minimize the confounding effects of shoes on the identification of footsteps.

Classification performance for the standard shoes (a casual sneaker in this study, Fig. 2(c)) was observed to be worse than that of the shoes owned by the participants. Even when the footsteps in the standard shoes were used as training samples, the COP DTW and COP TCN classifiers had difficulty recognizing them. It is likely that the participants were not as comfortable and secure in these unfamiliar shoes as in their own, customary shoes, which affected their COP patterns. Melvin et al. [39] suggested that individuals should acclimatise to unfamiliar shoes by taking at least 166 steps per foot in order to stabilise peak pressure values for five different types of shoes. However, in the present study, participants only took a few steps (5-10) to test the fit of the standard shoes before recording. Moreover, as the standard shoes were new and had a relatively inflexible, flat-soled design, they were not as ‘worn-in’ as the participants’ habitual shoes. Further research is needed to explore the influence of new and unfamiliar footwear on walking patterns, particularly shoes that are significantly different from the wearer’s usual footwear, which may have a more pronounced effect on gait and posture (e.g. the effects of wearing high heels for inexperienced users [37], [40]).

For three of the classification techniques (COP EP, COP DTW and P100 TM), the performance for left and right footsteps was not the same. This is similar to the findings of Ardhiyanto et al. [15], who found that the classification performance for right footsteps was better than left footsteps. They proposed that this could be because the dataset had a larger number of participants who were dominant in the right leg. This research showed that the more successful side changed depending on the technique used and there was no evidence that leg dominance had an effect on performance. The difference in performance between left and right steps is suspected to reflect typical asymmetries in human walking, even for the healthy participants included in this work. Asymmetry may be affected by biological factors such as age, strength imbalances, and functional anomalies, as well as factors such as walking speed and external disturbances like unilateral loading (e.g., carrying a bag on one side of the body) [2], [41].

It is noteworthy that sex and age had no effect on the performance of the current set of 20 participants. Generally, gender and age bias is a major issue for automated decision systems such as biometric technologies [42] and healthcare screening or diagnosis systems [43]. Studies have demonstrated that gender and age can affect joint kinematics and kinetics during gait [44], [45], and gait data from video capture and wearable sensors have been used to accurately predict gender and age [46], [47]. These factors may also have an indirect effect on gait data. For example, gender and age may influence the type of footwear worn, and certain types of footwear (e.g. high heels) may be more difficult to classify using pressure-based gait data and may be disproportionately represented in one group. Although these factors did not have a significant impact on the performance of the current dataset, gender and age bias will be taken into account in future studies. The present dataset was largely composed of white participants (13 out of 20), thus making it impossible to accurately evaluate any potential bias due to ethnicity and race. To do so, future studies should include more participants and assess any potential sources of bias in the left and right classification of footsteps, as well as other components of the gait classification or analysis pipeline (e.g. during registration to templates).

V. CONCLUSION

This study developed new algorithms for automatically distinguishing between left and right feet based on floor pressure sensors. These algorithms are capable of handling both barefoot and shod footsteps. Using the proposed pixel counting technique, which is based on the spatial characteristics of the sole's contact with the sensor, it was found that the accuracy for classifying more than 20,000 footsteps, despite the wide variety of represented shoe types, was 99.7%. Furthermore, interactions between independent variables (footwear and walking speed) and dependent variables (peak pressure, center of pressure, and accuracy) were examined, indicating that these factors must be taken into account when trying to develop left and right foot classification algorithms.

VI. ACKNOWLEDGMENT

(Eve MacDonald and Robyn Larracy are co-first authors.)

REFERENCES

- [1] E. J. Harris, I.-H. Khoo, and E. Demircan, "A survey of human gait-based artificial intelligence applications," *Frontiers Robot. AI*, vol. 8, Jan. 2022, Art. no. 749274.
- [2] A. Nasirzade, H. Sadeghi, H. R. Mokhtarinia, and A. Rahimi, "A review of selected factors affecting gait symmetry," *Phys. Treatments-Specific Phys. Therapy J.*, vol. 7, no. 1, pp. 3–12, 2017.
- [3] T. C. Pataky, K. Bosch, T. Mu, N. L. W. Keijsers, V. Segers, D. Rosenbaum, and J. Y. Goulermas, "An anatomically unbiased foot template for inter-subject plantar pressure evaluation," *Gait Posture*, vol. 33, no. 3, pp. 418–422, Mar. 2011.
- [4] C. Beyaert, R. Vasa, and G. E. Frykberg, "Gait post-stroke: Pathophysiology and rehabilitation strategies," *Neurophysiologie Clinique/Clin. Neurophysiol.*, vol. 45, nos. 4–5, pp. 335–355, Nov. 2015.
- [5] A. Mirelman, P. Bonato, R. Camicioli, T. D. Ellis, N. Giladi, J. L. Hamilton, C. J. Hass, J. M. Hausdorff, E. Pelosin, and Q. J. Almeida, "Gait impairments in Parkinson's disease," *Lancet Neurol.*, vol. 18, no. 9, pp. 697–708, 2019.
- [6] K. Hill, J. Schwarz, L. Flicker, and S. Carroll, "Falls among healthy, community-dwelling, older women: A prospective study of frequency, circumstances, consequences and prediction accuracy," *Austral. New Zealand J. Public Health*, vol. 23, no. 1, pp. 41–48, Feb. 1999.
- [7] T.-S. Wei, P.-T. Liu, L.-W. Chang, and S.-Y. Liu, "Gait asymmetry, ankle spasticity, and depression as independent predictors of falls in ambulatory stroke patients," *PLoS ONE*, vol. 12, no. 5, May 2017, Art. no. e0177136.
- [8] T. C. Pataky, T. Mu, K. Bosch, D. Rosenbaum, and J. Y. Goulermas, "Gait recognition: Highly unique dynamic plantar pressure patterns among 104 individuals," *J. Roy. Soc. Interface*, vol. 9, no. 69, pp. 790–800, Apr. 2012.
- [9] R. Vera-Rodriguez, J. S. D. Mason, J. Fierrez, and J. Ortega-Garcia, "Comparative analysis and fusion of spatiotemporal information for footprint recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 4, pp. 823–834, Apr. 2013.
- [10] M. Heydarzadeh, J. Birjandtalab, M. B. Pouyan, M. Nourani, and S. Ostadabbas, "Gaits analysis using pressure image for subject identification," in *Proc. IEEE EMBS Int. Conf. Biomed. Health Informat. (BHI)*, Feb. 2017, pp. 333–336.
- [11] G. Qian, J. Zhang, and A. Kidané, "People identification using gait via floor pressure sensing and analysis," in *Proc. Eur. Conf. Smart Sens. Context. Cham, Switzerland: Springer*, 2008, pp. 83–98.
- [12] J. Zhang, G. Qian, and A. Kidane, "Footprint tracking and recognition using a pressure sensing floor," in *Proc. 16th IEEE Int. Conf. Image Process. (ICIP)*, Nov. 2009, pp. 2737–2740.
- [13] B. Li, Z. Yao, J. Wang, S. Wang, Q. Wu, P. Wang, and X. Yang, "Analysis of plantar pressure image based on flexible force-sensitive sensor array," in *Proc. 13th Int. Symp. Comput. Intell. Design (ISCID)*, Dec. 2020, pp. 326–329.
- [14] F. P. M. Oliveira, A. Sousa, R. Santos, and J. M. R. S. Tavares, "Towards an efficient and robust foot classification from pedobarographic images," *Comput. Methods Biomech. Biomed. Eng.*, vol. 15, no. 11, pp. 1181–1188, Nov. 2012.
- [15] P. Ardhiyanto, B.-Y. Liau, Y.-K. Jan, J.-Y. Tsai, F. Akhyar, C.-Y. Lin, R. B. R. Subiaktio, and C.-W. Lung, "Deep learning in left and right footprint image detection based on plantar pressure," *Appl. Sci.*, vol. 12, no. 17, p. 8885, Sep. 2022.
- [16] N. van Melick, B. M. Meddeler, T. J. Hoogeboom, M. W. G. Nijhuis-van der Sanden, and R. E. H. van Cingel, "How to determine leg dominance: The agreement between self-reported and observed performance in healthy adults," *PLoS ONE*, vol. 12, no. 12, Dec. 2017, Art. no. e0189876.
- [17] M. Roberts, D. Mongeon, and F. Prince, "Biomechanical parameters for gait analysis: A systematic review of healthy human gait," *Phys. Therapy Rehabil.*, vol. 4, no. 1, p. 6, 2017.
- [18] P. Connor and A. Ross, "Biometric recognition by gait: A survey of modalities and features," *Comput. Vis. Image Understand.*, vol. 167, pp. 1–27, Feb. 2018.
- [19] S. Mehdizadeh, K. Van Ooteghem, H. Gulka, H. Nabavi, M. Faieghi, B. Taati, and A. Iaboni, "A systematic review of center of pressure measures to quantify gait changes in older adults," *Experim. Gerontol.*, vol. 143, Jan. 2021, Art. no. 111170.
- [20] R. Larracy, A. Phinyomark, and E. Scheme, "Gait representation: From vision-based to floor sensor-based gait recognition," in *Proc. IEEE Sensors Appl. Symp. (SAS)*, Jul. 2023, pp. 1–6.
- [21] S. Bai, J. Zico Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," 2018, *arXiv:1803.01271*.
- [22] B. G. Booth, E. Hoefnagels, T. Huysmans, J. Sijbers, and N. L. W. Keijsers, "PAPPI: Personalized analysis of plantar pressure images using statistical modelling and parametric mapping," *PLoS ONE*, vol. 15, no. 2, Feb. 2020, Art. no. e0229685.
- [23] T. Hida, R. Okuda, T. Yasuda, T. Jotoku, H. Shima, and M. Neo, "Comparison of plantar pressure distribution in patients with hallux valgus and healthy matched controls," *J. Orthopaedic Sci.*, vol. 22, no. 6, pp. 1054–1059, Nov. 2017.
- [24] H. Deepashini, B. Omar, A. Paungmali, N. Amaramalar, H. Ohnmar, and J. Leonard, "An insight into the plantar pressure distribution of the foot in clinical practice: Narrative review," *Polish Ann. Med.*, vol. 21, no. 1, pp. 51–56, Jun. 2014.

- [25] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, "Simple online and realtime tracking," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 3464–3468.
- [26] A. Bochkovskiy, C.-Y. Wang, and H.-Y. Mark Liao, "YOLOv4: Optimal speed and accuracy of object detection," 2020, *arXiv:2004.10934*.
- [27] P. C. Connor, "Comparing and combining underfoot pressure features for shod and unshod gait biometrics," in *Proc. IEEE Int. Symp. Technol. Homeland Secur. (HST)*, Apr. 2015, pp. 1–7.
- [28] G. Qian, J. Zhang, and A. Kidané, "People identification using floor pressure sensing and analysis," *IEEE Sensors J.*, vol. 10, no. 9, pp. 1447–1460, Sep. 2010.
- [29] T. Takeda, K. Kuramoto, S. Kobashi, and Y. Hata, "On optimal operator for combining left and right sole pressure data in biometrics security," *Adv. Fuzzy Syst.*, vol. 2013, pp. 1–10, Jan. 2013, doi: [10.1155/2013/620312](https://doi.org/10.1155/2013/620312).
- [30] Z. B. Barrons and G. D. Heise, "The impact of shod vs unshod walking on center of pressure variability," *Gait Posture*, vol. 81, pp. 116–119, Sep. 2020.
- [31] R. Vera-Rodriguez, J. S. D. Mason, J. Fierrez, and J. Ortega-Garcia, "Analysis of spatial domain information for footstep recognition," *IET Comput. Vis.*, vol. 5, no. 6, p. 380, 2011.
- [32] I. Bouchrika and M. S. Nixon, "Exploratory factor analysis of gait recognition," in *Proc. 8th IEEE Int. Conf. Autom. Face Gesture Recognit.*, Sep. 2008, pp. 1–6.
- [33] N. J. Cronin, "The effects of high heeled shoes on female gait: A review," *J. Electromyogr. Kinesiol.*, vol. 24, no. 2, pp. 258–263, Apr. 2014.
- [34] S.-H. Hyun, Y.-P. Kim, and C.-C. Ryew, "Effect on the parameters of the high-heel shoe and transfer time of ground reaction force during walking," *J. Exercise Rehabil.*, vol. 12, no. 5, pp. 451–455, Oct. 2016.
- [35] R. Dobbs, A. Charlett, S. Bowes, C. O'Neill, C. Weller, J. Hughes, and S. Dobbs, "Is this walk normal?" *Age Ageing*, vol. 22, no. 1, pp. 27–30, 1993.
- [36] D. Y. Ko and H. S. Lee, "The changes of COP and foot pressure after one hour's walking wearing high-heeled and flat shoes," *J. Phys. Therapy Sci.*, vol. 25, no. 10, pp. 1309–1312, 2013.
- [37] D. Marcin, "Human gait recognition based on ground reaction forces in case of sport shoes and high heels," in *Proc. IEEE Int. Conf. Innov. Intell. Syst. Appl. (INISTA)*, Jul. 2017, pp. 247–252.
- [38] Y.-O. Cho, B.-C. So, and J.-W. Jung, "User recognition using sequential footprints under shoes based on mat-type floor pressure sensor," *Adv. Sci. Lett.*, vol. 9, no. 1, pp. 591–596, Apr. 2012.
- [39] J. M. A. Melvin, S. Preece, C. J. Nester, and D. Howard, "An investigation into plantar pressure measurement protocols for footwear research," *Gait Posture*, vol. 40, no. 4, pp. 682–687, Sep. 2014.
- [40] P. A. de Oliveira Pezzan, S. M. A. João, A. P. Ribeiro, and E. F. Manfio, "Postural assessment of lumbar lordosis and pelvic alignment angles in adolescent users and nonusers of high-heeled shoes," *J. Manipulative Physiol. Therapeutics*, vol. 34, no. 9, pp. 614–621, Nov. 2011.
- [41] X. A. Zhang, M. Ye, and C. T. Wang, "Effect of unilateral load carriage on postures and gait symmetry in ground reaction force during walking," *Comput. Methods Biomech. Biomed. Eng.*, vol. 13, no. 3, pp. 339–344, Jun. 2010.
- [42] P. Drozdzowski, C. Rathgeb, A. Dantcheva, N. Damer, and C. Busch, "Demographic bias in biometrics: A survey on an emerging challenge," *IEEE Trans. Technol. Soc.*, vol. 1, no. 2, pp. 89–103, Jun. 2020.
- [43] R. R. Fletcher, A. Nakeshimana, and O. Olubeko, "Addressing fairness, bias, and appropriate use of artificial intelligence and machine learning in global health," *Frontiers Artif. Intell.*, vol. 3, Apr. 2021, Art. no. 561802.
- [44] S. H. Cho, J. M. Park, and O. Y. Kwon, "Gender differences in three dimensional gait analysis data from 98 healthy Korean adults," *Clin. Biomech.*, vol. 19, no. 2, pp. 145–152, Feb. 2004.
- [45] A. G. Thakurta, "Human gait with reference to age, gender and impact of load: A review," *Adv. Appl. Physiol.*, vol. 1, no. 2, p. 24, 2016.
- [46] J. P. Singh, S. Jain, S. Arora, and U. P. Singh, "Vision-based gait recognition: A survey," *IEEE Access*, vol. 6, pp. 70497–70527, 2018.
- [47] M. A. R. Ahad, T. T. Ngo, A. D. Antar, M. Ahmed, T. Hossain, D. Muramatsu, Y. Makihara, S. Inoue, and Y. Yagi, "Wearable sensor-based gait analysis for age and gender estimation," *Sensors*, vol. 20, no. 8, p. 2424, Apr. 2020.



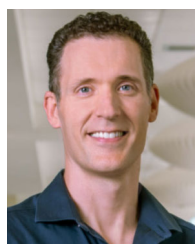
EVE MACDONALD (Graduate Student Member, IEEE) received the B.Sc. degree in electrical engineering from the University of New Brunswick, Fredericton, NB, Canada, in 2023. She is currently pursuing the M.Sc. degree with the School of Computing, Queen's University, Kingston, ON, Canada. Her research interests include gait recognition, artificial intelligence, and cybersecurity.



ROBYN LARRACY (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from the University of New Brunswick (UNB), Fredericton, NB, Canada, in 2019 and 2021, respectively, where she is currently pursuing the Ph.D. degree in electrical engineering. She is a member of the Health Technologies Laboratory, UNB. Her research interests include gait analysis, gait biometrics, and machine learning.



ANGKON PHINYOMARK (Member, IEEE) received the B.Eng. degree (Hons.) in computer engineering and the Ph.D. degree in electrical engineering from the Prince of Songkla University, Hat Yai, Thailand, in 2008 and 2012, respectively. His dissertation research was awarded the "Best Ph.D. Thesis Award" from the National Research Council of Thailand. He was a Postdoctoral Research Fellow with GIPSA and LIG Laboratories, Université Grenoble Alpes, Grenoble, France (2012–2013), and the Human Performance Laboratory, University of Calgary, AB, Canada (2013–2016). From 2016 to 2017, he was a Researcher with the ISI Foundation, Turin, Italy. He has been a part of the Institute of Biomedical Engineering, University of New Brunswick, Fredericton, NB, Canada, since 2017, where he is currently a Senior Research Scientist. His research interests include biomedical signal processing and machine learning, wearable sensors, and gait biomechanics and biometrics.



ERIK SCHEME (Senior Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the University of New Brunswick (UNB), Fredericton, NB, Canada, in 2003, 2005, and 2013, respectively. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, the Associate Director of the Institute of Biomedical Engineering, and the Director of the Health Technologies Laboratory, UNB. His research interests include the application of signal processing and machine learning for human-machine interfaces, mobility and rehabilitation, digital health and diagnostics, and biometrics. He is a Registered Member of the Association of Professional Engineers and Geoscientists of New Brunswick (APEGNB).

...