A Method for Extracting Shoe-Dependent Features from Vertical GRF (Ground Reaction Force) Signals of Human Gait

Marek Parfieniuk
Faculty of Computer Science
University of Bialystok
Bialystok, Poland
marek.parfieniuk@uwb.edu.pl

Marcin Derlatka
Institute of Biomedical Engineering
Faculty of Mechanical Engineering
Bialystok University of Technology
Bialystok, Poland
m.derlatka@pb.edu.pl

***Abstract*—This paper proposes a time-domain approach to analyzing GRF (Ground Reaction Force) signals related to human gait. The method uses the basic half-band FIR filters. It is aimed at extracting features for recognizing the type of shoes worn by a person that trod on a force plate. A basic feature can be the sign of a sample of the first-order difference of the vertical GRF, and higher-level features can be defined by considering series of same-sign samples in the forepart of a force wave. The low-frequency component of the difference provides more consistent values of features of the same kind, while in the high-frequency component one can find supplementary features related to minor oscillations. This has been shown by exploring data, defining features and testing their discriminatory powers, on a large set of about 11000 GRF measurements related to stilettos, sport and patent leather shoes. The presented ideas can be used to develop classifiers able to differentiate between stilettos and other footwear, soft- and hard-sole shoes, shoes with and without heels.**

***Index Terms*—footwear, shoe, stiletto, GRF, recognition, ground reaction force, gait.**

I. INTRODUCTION

Human gait is known to carry information about a person. It can be analyzed so as to identify a walker [1], [2] or her/his conditions. As to the latter, researchers attempted to use gait observations to diagnose diseases, evaluate progress in rehabilitation and training [3], [4]. A recently stated, less investigated problem is footwear recognition [5]–[7].

A popular approach to gait evaluation is to measure ground reaction force (GRF) signals. These signals are obtained by using a force plate, placed on ground, so as humans can put feet on it when walking. Such a device is a solid plate witch sensors that convert tensions between it and ground into electrical signals. These signals are scaled and combined so as to represent forces between foot and ground [8].

Gait can also be evaluated by capturing and analyzing image sequences of a walking person or only his/her leg [4]. In surveillance cameras are only option, but in other applications they have both pros and cons compared to force plates.

The work by Marek Parfieniuk was supported by the University of Bialystok. The work by Marcin Derlatka was supported by the Bialystok University of Technology.

In particular, footwear types can be recognized from images [9], [10], but this meets obvious problems when shoe is on leg. Firstly, it is difficult to position cameras or to direct people so as to expose footwear. In particular, shoes can be hidden behind a long skirt or dress. Secondly, even when shoe is visible, it is challenging to extract features from an image when background, leg skin, tights etc. are similar in color.

This paper considers a problem of footwear recognition from only GRF measurements. We show that gait related to stilettos can be distinguished from that of sport and patent shoes. Moreover there are premises that other classes of shoes can be recognized: at least hard- and soft-sole ones, and those with and without heel.

The results of this research can be used to develop systems that prevent stiletto wearers from entering areas where floor can be damaged by these shoes. Another application can be studies of popularity of footwear types.

This work proposes a novel method for feature extraction, which uses only vertical GRF (VGRF), neglecting anterior-posterior and medio-lateral GRFs. Only the forepart of a VGRF wave is analyzed, which spans from the foot strike to the first main peak. The difference signal and its low-frequency (LF) and high-frequency (HF) components are used as sources of features. The sign of a sample of the difference is proposed as a basic feature, while higher-level features can be defined by considering series of same-sign samples. The low-frequency component of the difference provides more consistent values of features of the same kind, while in the high-frequency component one can find supplementary features related to minor oscillations.

Such decomposition of the VGRF signal is rather straightforward, but nobody has reported its use in the literature. A novelty of our work is also in revealing that it is sufficient or even advantageous to determine these signals by using the basic half-band filters.

To the best of our knowledge, nobody has reported a similar approach and so promising results. Many papers describe data sets and studies of GRF signals related to particular shoes [1], [11], [12], but do not provide insights how to recognize shoe type from gait, being only focused on explaining biomechanics behind signals.

Only several papers discuss recognizing shoes that caused a GRF signal [2], [5] In some works, researchers analyzed data related to particular type of footwear, but this constraint was imposed so as to made signals uniform.

The only exception is [5], in which shoe type was determined by using features that are computationally demanding. The present paper is not an extension of [5]. It shows that footwear can be recognized by using less data and simpler hardware-software configuration. This has to be achieved by deeper understanding GRF time series, in the light of signal and filter theory [13].

This research was possible owing to the recent publication of a unique, suitable set of measurements [14]. It contains 6510 records related to sport shoes, 1466 measurements of stilettos and 2996 records related to patent leather shoes. So large amount of data allows for drawing credible, generalized conclusions about signals and algorithms.

It is as well important that the data set contains raw measurements, which are sampled versions of analog outputs of force plates. They have not been digitally filtered, so contain high-frequency details. These details carry information related to shoes, as it will be shown later in this paper.

II. Shoe-related information in vertical GRF

The biomechanics behind the gait cycle is well known and described in many publications. Most of such studies consider barefoot walking and symptoms of diseases, see e.g. [15]. Fewer works discuss effects of footwear [7], [16]. Exceptions are publications in which GRFs are analysed from the point of view of mathematics, information and signal theories [17], as it will be done herein.

Figure 1 shows typical waves of the VGRF. This force is called vertical because it is related to the interaction between foot and ground in the gravity direction. The VGRF is zero during the swing phase of gait, when the foot is moving though air. Its M-like waveform reflects the stance phase, during which the foot contacts the ground. This begins when one puts the heel on the floor, and ends when he/she takes toe off. Between, the VGRF increases to the first main peak, temporarily decreases during the mid-stance phase, increases again to the second main peak, and then decays. The two peaks occur because the VGRF is an effect of both gravity and muscle actions.

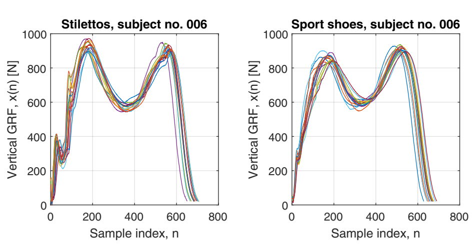

Fig. 1. Waves of vertical GRF of gait of young woman, wearing stilettos and sport shoes.

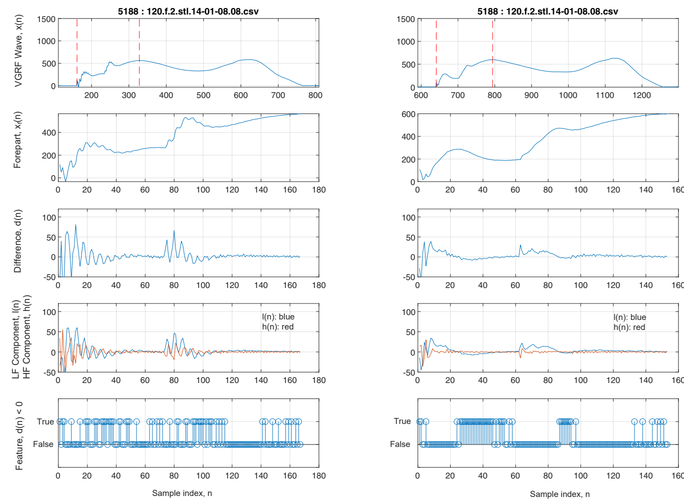

Fig. 2. Analysis of exemplary VGRF signals related to stilettos.

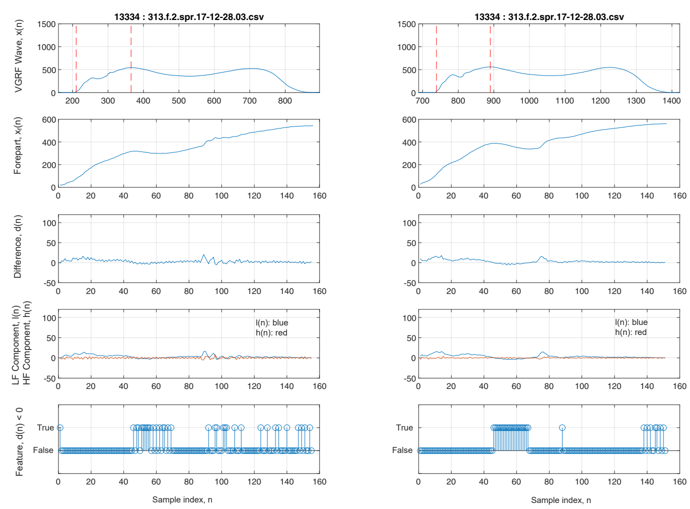

Fig. 3. Analysis of exemplary VGRF signals related to sport shoes.

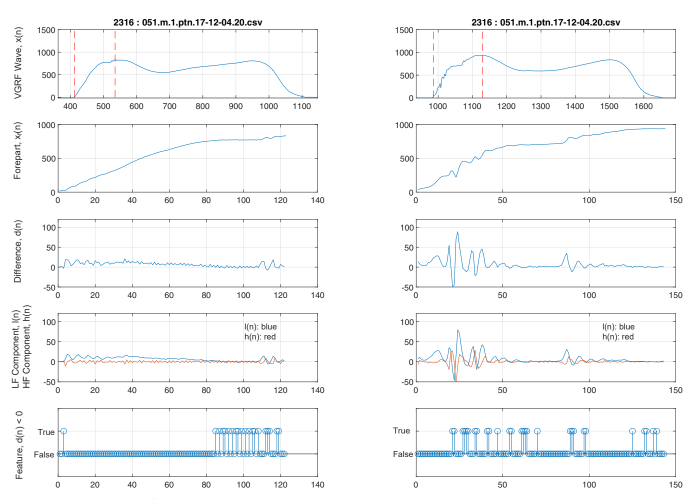

Fig. 4. Analysis of exemplary VGRF signals related to patent leather shoes.

In this research, we did not consider the two remaining GRF forces, anterior-posterior and medio-lateral ones, which act along and perpendicularly to the walking direction, respectively, in the plane parallel to ground. It seems hardly possible to extract from them features that supplement those related to the VGRF. The former force is correlated with the VGRF, while the latter is very weak and noisy.

Figure 1 shows VGRF waves of gait of one young woman. The left-hand plot presents a dozen or so signals measured when she worn stilettos. The right-hand plot is related to her sport shoes. One can conclude that an entire waveform of VGRF depends on footwear, but differences are especially evident in the forepart.

For sport shoes, VGRFs steadily increases from zero to the first peak, while stiletto-related waveforms contain sharp minor peaks before the first main one. A bend just before the second peak is also characteristics for the stilettos-related gait, but it is a rather minor deviation from a smooth curve.

The same can be concluded after inspecting other measurements from our data set. Some are shown in Figs. 2–4. In each figure, the first plot shows entire wave of the VGRF. The second plot shows the forepart of the wave.

From Figs. 2–4, a clear conclusion is that stiletto-related VGRFs contain more high-frequency contents than measurements of other shoes. However this contents can have various forms. Oscillations as well as a single pulse can occur, and their amplitudes and time extents can vary significantly.

On the other hand, high-frequency contents occur in GRF measurements related to sport and patent shoes. For these footwear, oscillations and pulses generally have lower amplitudes but not always.

One could say that information specific for stilettos is carried mainly by the high-frequency component of the VGRF. Moreover this information is located before the first main peak. The low-frequency component carries rather information related to the identity and diseases of a walker.

Undoubtedly, some information about shoes is contained also in the general shape of a VGRF wave. It seems possible to find it in relative times and levels of the main peaks. Studies of this remain as a subject for future research.

III. A METHOD FOR EXTRACTING BASIC FEATURES

The first our postulate is to analyze only the forepart of the VGRF wave. We define it as samples from the foot strike to the first peak of the VGRF wave. A common approach is to associate the foot strike with the first sample that exceeds a threshold, commonly 20 N, and is preceded by under-threshold samples, which are related to the noise produced by a force plate without a load. The first peak can be associated with the first sample that occurs after the foot strike and is followed by a considerably long series of lower values. The considerable length, 100 samples in our experiments, for the sampling rate of 1000 Hz, must be chosen so as to allow a searching algorithm to neglect minor, local peaks, which are followed by shorter series.

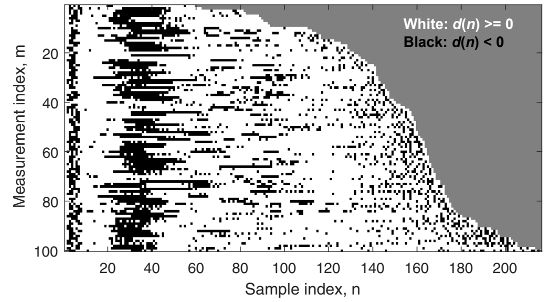

Fig. 5. Values of $(d(n) < 0)$ essential feature of 100 randomly selected VGRF measurements related to stilettos.

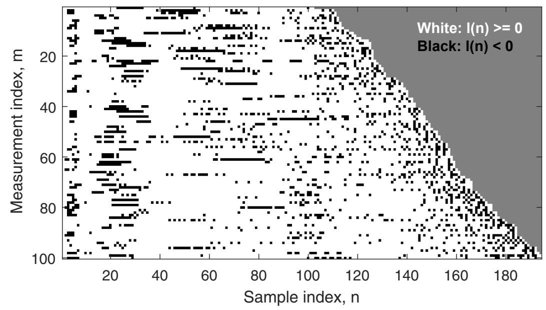

Fig. 6. Values of $(d(n) < 0)$ essential feature of 100 randomly selected VGRF measurements related to non-stiletto shoes.

The second our proposal is to focus attention on high-frequency details of a VGRF wave by analyzing its first-order difference, computed as follows:

$$d(n) = x_f(n) - x_f(n-1) \quad (1)$$

where $x_f(n)$ is the n-th sample of the finite-length auxiliary signal that represents the forepart of the VGRF signal, denoted as $x(n)$ in the plots. The difference can be interpreted as the result of high-pass filtering [13], which removes information related to the general waveform, extracting details of short-time changes, pulses and oscillations.

The difference is shown in the third plot in each of Figs. 2–4, while its sign is the fifth plot. One can easily notice that a measurement of non-stiletto shoes is usually characterized by the existence of very long series of samples for which $(d(n) > 0)$. Such a series represents an increasing slope of the VGRF wave. The existence of a series of samples for which $(d(n) < 0)$ means that the VGRF temporarily, considerably decreased, and so a notable peak occurs in the slope from the foot strike to the first main peak.

Footwear-specific patterns in the sign feature are even more visible in Figs. 5 and 6. Each image corresponds to 100 randomly selected measurements of the VGRF, of a particular type of shoes, sorted based on the length of the forepart. In a row, pixels represent values of the basic feature for samples of the forepart of a VGRF wave. A white pixel means that $d(n) >= 0$ for the corresponding sample. Black pixel means the opposite. The gray pixels represent samples after the forepart, which are excluded from consideration.

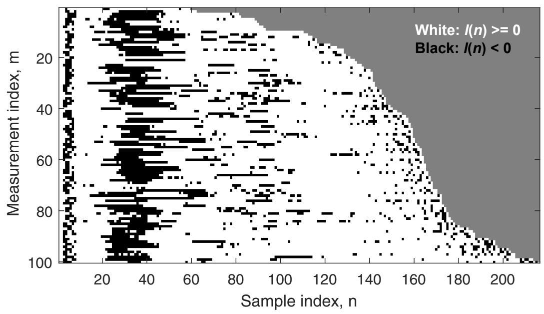

Fig. 7. Values of $(l(n) < 0)$ basic feature of 100 randomly selected VGRF measurements related to stilettos.

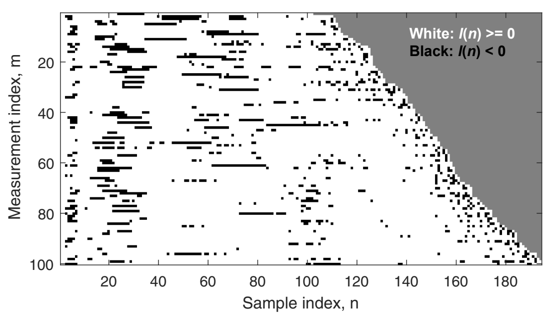

Fig. 8. Values of $(l(n) < 0)$ basic feature of 100 randomly selected VGRF measurements related to non-stiletto shoes.

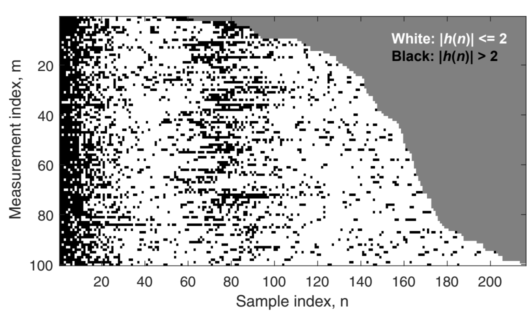

Fig. 9. Values of $(|h(n)| > 2)$ basic feature of 100 randomly selected VGRF measurements related to stilettos.

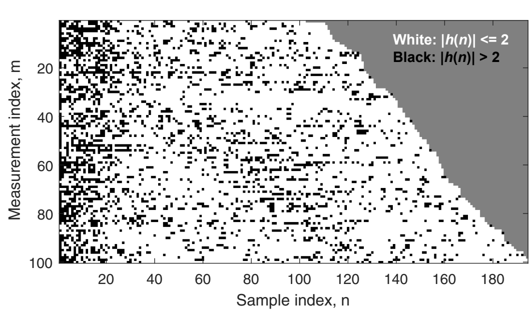

Fig. 10. Values of $(|h(n)| > 2)$ basic feature of 100 randomly selected VGRF measurements related to non-stiletto shoes.

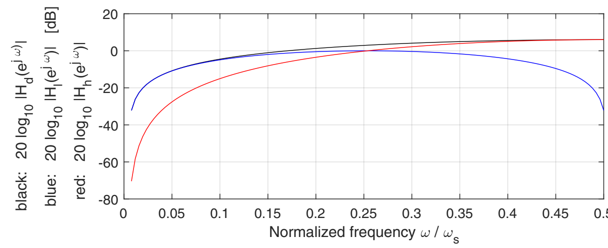

Fig. 11. Magnitude responses of filters used to analyze VGRF.

We propose to consider the sign of $d(n)$ as only one basic feature for recognizing shoes. Two more basic features can be extracted from

$$l(n) = \tfrac{1}{2}\left(d(n) + d(n-1)\right) \tag{2}$$

and

$$h(n) = \tfrac{1}{2}\left(d(n) - d(n-1)\right) \tag{3}$$

which are low- and high-frequency components of the first-order difference of the VGRF. These signals are shown in the fourth plot in Figs. 2–4.

The sign of $l(n)$ is a feature essentially equivalent to that of $d(n)$, but its equal values appear in longer series. There are fewer isolated samples whose signs are different from those of neighbors. This can be noticed in Figs. 7 and 8, which are analogous to Figs. 5 and 6.

On contrary, $h(n)$ captures minor isolated peaks and longer oscillations in $d(n)$. A basic feature that reveals them is a result of comparing the absolute value of $h(n)$ to a threshold. Figs. 9 and 10 show that distributions of values of this feature are specific for stilettos.

Equations (2) and (3) can be interpreted as a description of complementary, half-band FIR filters [13]. Their magnitude responses are shown in Fig. 11, together with that of the filter related to (1). The filters are not very selective but satisfactorily decompose signals. So simple filters have shortest possible impulse responses and minimal group delays, so that output is delayed by only one sample with respect to excitation. This is very advantageous as a GRF forepart is a relatively short, finite-length signal. For longer filters, a rapid change in VGRF would be exhibited in features related to distant samples, and boundary artifacts would contaminate features related to many samples at the beginning and ending of the VGRF forepart.

IV. Higher-level Features and Their Discriminatory Powers

One can define higher-level features related to occurrence of a specific pattern of values of a basic feature. Appearance or absence of such a pattern in a measurement can be used to evaluate the probability it is related to particular type of shoes.

Table I shows our proposals of higher-level features related to the difference signal. Due to lack of space we were unable to list features related to its LF and HF components.

The last row of Table I specifies no feature. It is to show the total numbers of records related to stilettos and other shoes.

TABLE I
PROPOSED HIGHER-LEVEL FEATURES AND THEIR DISCRIMINATORY POWERS. N_F DENOTES LENGTH (IN SAMPLES) OF THE FOREPART OF VGRF WAVE.

No.	Higher-level feature of forepart of VGRF wave	Number of satisfying measurements labeled as	
		stilettos	non-stilettos
1.	Begins with a series of $\geq (10\% \cdot N_{\mathrm{F}})$ samples for which $d(n) > 0$	1.78% (26)	42.35% (4025)
2.	Begins with a series of $\geq (20\% \cdot N_{\mathrm{F}})$ samples for which $d(n) > 0$	0.14% (2)	23.05% (2191)
3.	Contains a series of $\geq (40\% \cdot N_{\mathrm{F}})$ samples for which $d(n) > 0$	15.17% (222)	61.97% (5890)
4.	Contains a series of $\geq (50\% \cdot N_{\mathrm{F}})$ samples for which $d(n) > 0$	5.13% (75)	36.57% (3476)
5.	Contains a series of $\geq (60\% \cdot N_{\mathrm{F}})$ samples for which $d(n) > 0$	2.05% (30)	23.87% (2269)
6.	Contains two nonoverlapping series of $\geq (20\% \cdot N_{\mathrm{F}})$ samples for which $d(n) > 0$	12.44% (182)	39.98% (3800)
7.	Contains two nonoverlapping series of $\geq (25\% \cdot N_{\mathrm{F}})$ samples for which $d(n) > 0$	2.94% (43)	26.19% (2489)
8.	Contains two nonoverlapping series of $\geq (30\% \cdot N_{\mathrm{F}})$ samples for which $d(n) > 0$	0.75% (11)	16.92% (1608)
9.	Contains a series of $\geq (5\% \cdot N_{\mathrm{F}})$ samples for which $d(n) < 0$	81.27% (1189)	21.35% (2029)
10.	Contains a series of $\geq (10\% \cdot N_{\mathrm{F}})$ samples for which $d(n) < 0$	51.67% (756)	4.57% (434)
11.	Contains a series of $\geq (15\% \cdot N_{\mathrm{F}})$ samples for which $d(n) < 0$	20.78% (304)	0.79% (75)
12.	Contains a sample whose for which $\|d(n)\| > 40$	79.22% (1159)	18.50% (1758)
13.	Contains a sample whose for which $\|d(n)\| > 50$	62.47% (914)	9.89% (940)
14.	Contains a sample whose for which $\|d(n)\| > 60$	41.90% (613)	5.87% (558)
15.	— (No constraints on head, so as to count all measurements of given type of shoes)	100% (1463)	100% (9504)

The rows from 12th to 14th are unrelated to the basic features. They have to show that it is not useful to test the absolute value of difference signal, $d(n)$, against a threshold. Threshold-based features well discriminate records of stilettos, but they are also satisfied by many record of other footwear.

The two right-most columns show how many records, related either to stilettos or other shoes, satisfy the condition of a given row. The exact number of matching records is given in parentheses. The primary information is the percentage of all records related to a type of footwear.

We have presented distriminatory powers in this way, because the data imbalance is a problem in footwear recognition. Measurements related to stilettos comprise 13% of the data set. In a real-world application of our method, one could expect a similar situation that only a fraction of persons wear stilettos. There are known several approaches to the imbalance, and it seems impossible to decide without experiments which of them best suits this application.

V. CONCLUSION

Our results leave no doubt that it is possible to recognize the type of shoes worn by a person by analyzing GRFs related to his/her gait. We have shown how to extract features for this purpose by using well-grounded, conceptually clear algorithms of moderate demands on computations and memory.

It seems possible to achieve more then only to recognize stilettos. Figs. 3 and 4 suggest that features and patterns can be found that allow for distinguishing sport shoes from patent leather shoes. On the other hand, Figs. 5–10 suggest that subclasses of footwear can be determined that has not been labeled: soft- and hard-sole shoes, those with and without heel, various heels.

More higher-level features can be defined by considering subpatterns and their relative positions in $d(n)$, $l(n)$, and $h(n)$. Alternatively, one can resample the forepart, so as to classify signals that have as many samples, without considering the relative position of a pattern.

Depending on feature properties, various types of classifier can be used. Equal-length vectors of samples can be processed by a nearest-neighbors classifier, while high-level features well suit decision trees. Both kinds of data can be used to learn a neural network.

Because of so wide perspectives of future research, we confine ourselves to only presenting features in this paper.

REFERENCES

[1] K. A. Duncanson *et al.*, "Modelling individual variation in human walking gait across populations and walking conditions via gait recognition," *Journal of The Royal Society Interface*, vol. 21, no. 221, p. 20240565, 2024.

[2] M. Derlatka *et al.*, "Recognition of human gait based on ground reaction forces and combined data from two gait laboratories," *Acta Mechanica et Automatica*, vol. 18, pp. 361–366, 06 2024.

[3] M. Bonanno *et al.*, "Gait analysis in neurorehabilitation: From research to clinical practice," *Bioengineering*, vol. 10, 2023.

[4] D. Sethi, S. Bharti, and C. Prakash, "A comprehensive survey on gait analysis: History, parameters, approaches, pose estimation, and future work," *Artificial Intelligence in Medicine*, vol. 129, p. 102314, 2022.

[5] M. Derlatka and M. Bogdan, "Recognition of a person wearing sport shoes or high heels through gait using two types of sensors," *Sensors*, vol. 18, no. 5, 2018.

[6] M. Pavlou and N. M. Allinson, *Encyclopedia of Biometrics*. Boston, MA: Springer, 2009, ch. Footwear Recognition, pp. 557–562.

[7] E. Macdonald *et al.*, "Underfoot pressure-based left and right foot classification algorithms: The impact of footwear," *IEEE Access*, vol. 11, pp. 137 937–137 947, 2023.

[8] *Kistler Force Plate Formulae*, Kistler, 1999, ver. 03.02.99 V/CI.

[9] F. G. Bloedorn and C. G. Webber, "Transfer learning applied to a classification task: a case study in the footwear industry," *IEEE Latin America Transactions*, vol. 21, pp. 427–433, 2023.

[10] H. Zhan, B. Shi, and A. C. Kot, "Cross-domain shoe retrieval with a semantic hierarchy of attribute classification network," *IEEE Transactions on Image Processing*, vol. 26, pp. 5867–5881, 2017.

[11] S. Jandova *et al.*, "Gait cycle of slow walking in high heels," *Health Problems of Civilization*, vol. 19, no. 1, pp. 47–55, 2025.

[12] M. Wiedemeijer and E. Otten, "Effects of high heeled shoes on gait. a review," *Gait & Posture*, vol. 61, pp. 423–430, 2018.

[13] G. Strang and T. Q. Nguyen, *Wavelets and Filter Banks*. Wellesley, MA: Wellesley-Cambridge Press, 1996.

[14] M. Derlatka and M. Parfieniuk, "Real-world measurements of ground reaction forces of normal gait of young adults wearing various footwear," *Scientific Data*, vol. 10, no. 1, p. 60, Jan 2023.

[15] K. Shah, M. Solan, and E. Dawe, "The gait cycle and its variations with disease and injury," *Orthopaedics and Trauma*, vol. 34, no. 3, pp. 153–160, 2020.

[16] K. F. E. Nadège *et al.*, "Wearing high heel shoes during gait: Kinematics impact and determination of comfort height," *American Journal of Life Sciences*, vol. 3, no. 2, pp. 56–61, apr 2015.

[17] A. Tongen and R. E. Wunderlich, *Biomechanics of Running and Walking*, ser. Dolciani Mathematical Expositions. Mathematical Association of America, 2010, pp. 315–328.