Preliminary assessment of the Samsung Galaxy Watch 5 accuracy for the monitoring of heart rate and heart rate variability parameters

Gianluca Rho[1,2], Francesco Di Rienzo[1], Carlotta Marinai[1], Francesca Giannetti[1], Lucia Arcarisi[1], Pasquale Bufano[3], Michele Zanoletti[3], Francesca Righetti[1], Carlo Vallati[1], Marco Laurino[3], Nicola Carbonaro[1,2], Alessandro Tognetti[1,2] and Alberto Greco[1,2]

Abstract— **In the last years, commercial smartwatches have gained popularity as non-invasive and wearable devices to be exploited for the monitoring of the cardiovascular system in daily-life settings. However, their reliability is still unclear. In this preliminary study, we evaluated the accuracy of heart rate (HR) and HR variability (HRV) estimates obtained from the Samsung Galaxy Watch 5 (SGW5) compared to a common research-grade ECG sensor, i.e., the Shimmer3 ECG unit (ShimECG), during both a resting and walking conditions. For each condition, we compared HRV features of SGW5 and ShimECG extracted in time, frequency, and non-linear domains through correlation and Bland-Altman analysis. Additionally, we compared SGW5 performance with those obtained from a research-grade PPG sensor. Our results revealed an unbiased and high-quality estimate of mean HR obtained from the SGW5. Moreover, at rest, other relevant HRV features showed a significant correlation between the SGW5 and ShimECG. Conversely, during the walking condition, we found poor performances for both PPG devices for most of the HRV features. Such preliminary results confirm the reliability of SGW5 to estimate mean HR. However, the reliability of SGW5-derived PRV to extract sympathovagal correlates is still an open question and deserves further investigation.**

I. Introduction

Heart rate (HR) and heart rate variability (HRV) are physiological parameters reflecting the general well-being of a subject [1]. Variations of such parameters can be adopted to evaluate stressful conditions, anxiety and panic [2]. Furthermore, they allow to identify autonomic imbalances associated with cardiovascular and respiratory dysregulation [3], as well as psychiatric disorders [4].

Electrocardiography (ECG) is the gold standard method adopted to estimate HR and HRV parameters in clinical and research settings. However, it is hardly used in daily-life contexts where free movements are required. To address this issue, wearable ECG sensors including shirts [5] and chest belts [6], [7] have been proposed. However, they can be obtrusive and uncomfortable over day-long recording periods. An alternative to ECG to estimate HR/HRV is the use of photoplethysmography (PPG) sensors. The pulse rate variability (PRV) estimated from the PPG signals can be used as a reliable surrogate of the HRV [8], [9]. Specifically, PPG estimates the volume variation of superficial blood vessels at either the fingertip or the ear lobe through an LED and a photo-receiver [10]. Although clinical-grade PPG monitoring devices significantly limit movements due to the constraint of wires, recently, wearable PPG sensors have rapidly spread as a key equipment of most commercial smartwatches (SW). Such large diffusion, combined with their ease of use and connectivity features with other devices, as well as the sensor positioning on the wrist, have led to the widespread use of SWs also in scientific research applications [11]–[15].

In the last years, several SW applications have been tested by comparing the SW performance against medical-grade ECG and PPG, e.g., the Apple Watch [16], Empatica E4 [17], Fitbit Charge HR [18], and Microsoft band 2 [19]. Particularly, they showed accurate HR estimates during resting conditions, with decreasing performances at the increase of the subjects' activity intensity. However, on the one hand, such devices do not always provide direct access to raw data, and HR information is extracted through black-box proprietary algorithms. On the other hand, some of them have a high cost that limits their spread. These limitations are potentially overcome by the open-source Android Wear operating system (OS), which has been recently adopted by several commercial SWs. Particularly, Android Wear OS allows developers and researchers to access sensors' data through the development of custom applications.

In this study, we evaluate the performances of an Android Wear OS SW, i.e., the Samsung Galaxy Watch 5 (SGW5) [20], on a group of healthy subjects during a resting state and a walking condition. The SGW5 relative low-cost and its long-life battery [20] could make it suitable for carrying out 24h HR and HRV monitoring research studies on a consistent number of subjects. The evaluation is performed against the measurements obtained from a validated ECG device, i.e., the Shimmer3 ECG unit [21]. Furthermore, we compare the SGW5 performances with those obtained by a widely used research-grade PPG device, i.e., the Shimmer3 GSR+ unit. For each PPG signal, we derive HRV parameters in the time, frequency, and non-linear domains, respectively, and we evaluate Pearson correlation coefficients with the same

*This work has received funding from the European Union's Horizon Europe Research and Innovation Programme under grant agreement No. 101057103 – project TOLIFE

[1]G. Rho, F. Di Rienzo, C. Marinai, Francesca Giannetti, L. Arcarisi, F. Righetti, C. Vallati, N. Carbonaro, A. Tognetti, and A. Greco are with the Dipartimento di Ingegneria dell'Informazione, University of Pisa, 56122 Pisa, Italy.

[2]G. Rho, N. Carbonaro, A. Tognetti, and A. Greco are with the Research Center "E. Piaggio", University of Pisa, 56122 Pisa, Italy.

[3]P. Bufano, M. Zanoletti, and M. Laurino are with the Institute of Clinical Physiology, National Research Council, via Giuseppe Moruzzi,1, 56124 Pisa, Italy.

Corresponding author: `gianluca.rho@phd.unipi.it`

estimates obtained from ECG. Finally, we evaluate the device accuracy through a Bland-Altman analysis.

II. Materials and methods

A. Subjects

The study was conducted according to the guidelines of the Declaration of Helsinki. Twenty healthy volunteers (age 39 $\pm$ 15, 7 females) signed an informed consent to take part in the study. Subjects self-reported no history of cardiovascular diseases.

B. Experimental settings

The experimental protocol consisted of two distinct conditions: (1) 1min of resting state seated on a chair, with the arms resting on the table, and (2) 1min of walking. Subjects were asked not to talk during the experiment. Moreover, during (1), subjects were asked to minimize movements. The order of the conditions was randomized across subjects.

During each condition, we acquired the PPG signal at the wrist of the subject's non-dominant hand through the SGW5. We developed an ad-hoc Wear OS application to retrieve raw data from the PPG sensor, together with the relative universal timestamp, at the maximum available sampling frequency of 25Hz. The data was sent via Bluetooth communication to a smartphone and then stored in a computer for the processing stage.

We acquired the ECG signal through the Shimmer3 ECG unit (SHIMMER research, Dublin, Ireland) as the gold standard for this study [21]. The sensor was mounted on the subjects' chest through an elastic band, with four leads placed on the left arm, right arm, left leg, and right leg, according to the manufacturer's guidelines. We recorded ECG as the difference between the right leg and left arm leads, at the sampling frequency of 400Hz.

As an additional term of comparison, we acquired the PPG signal by using also the Shimmer3 GSR+ unit (shimPPG). The signal was acquired at the tip of the first finger of the subjects' non-dominant hand, at the sampling frequency of 25Hz.

To facilitate the following data processing, we synchronized the recordings of the SWG5, ECG, and shimPPG.

C. Data processing

The raw data were preprocessed in Matlab (Version R2021b, Mathworks, USA).

We filtered both the SGW5 and shimPPG data in the (0.7-1.8)Hz frequency range through a zero-phase band-pass IIR filter (transition band=0.1Hz). We identified pulse peaks in the signals through the multi-scale peak and trough detection (MSPDT) algorithm [22]. The outcome of the MSPDT algorithm was visually inspected, and missing peaks or peaks not properly identified were manually corrected where possible. Afterwards, we imported peak-to-peak (P-P) distance time series in Kubios HRV [1], and we derived PRV signals after uniform interpolation at 4Hz. The PRV was further corrected for the presence of artifacts (e.g., ectopic beats, abnormal P-P values) through the automatic artifact correction algorithm using a conservative threshold (*low* option in Kubios; see [1] for details). The procedure was applied to both SWG5 and shimPPG data.

Regarding ECG, we applied a zero-phase band-pass IIR filter (transition band=0.1Hz) in the (0.7-25)Hz range. We estimated HRV time series using Kubios through the automatic QRS complex detection algorithm [1], followed by a uniform interpolation at the sampling frequency of 4Hz. HRV artifacts were corrected through the Kubios automatic artifact correction algorithm using the same threshold as for the PRV signals.

Using Kubios, we extracted the following features from PRV and HRV time series: (1) meanRR (i.e., the mean distance between consecutive peaks); (2) stdRR (i.e., the standard deviation of RR intervals); (3) RMSSD (i.e., root mean squared differences of successive RR intervals); (4) LF (i.e., low-frequency band; (0.04-0.15)Hz); (5) HF (i.e., high-frequency band; (0.15-0.40)Hz); (6) LF/HF ratio; (7) SD1, SD2 (i.e., the standard deviations of Poincaré plot).

D. Statistical Analysis

For each of the two experimental conditions (i.e., resting-state, walking) and for each of the features extracted (see section II-C), we computed the Pearson correlation coefficient between the SGW5 and the ECG, and between the shimPPG and ECG, to evaluate the overall degree of agreement among measurements. The resulting p-values were corrected for multiple comparisons between different devices (i.e., SGW5 vs ECG, shimPPG vs ECG) with the Bonferroni method. Moreover, we conducted an individual Bland-Altman (BA) analysis [23] on each SGW5 feature to evaluate its bias and precision with respect to the ECG estimates. The same statistical procedure was applied to the shimPPG features. The presence of a significant bias was assessed through a paired t-test on the difference between devices' measurements against the null-hypothesis of no difference. P-values were corrected for multiple comparisons.

E. Results

In Table I, we report the Pearson correlation coefficient ρ between SGW5 and ECG, and between shimPPG and ECG, for each of the estimated features (see section II-C), during both the resting state and walking condition.

TABLE I: Pearson correlation coefficient between the PRV and HRV features during the resting state and walking condition. Significant correlations are highlighted in bold and marked by * (*:p<0.05, **:p<0.01, ***p<0.001; p-values adjusted with the Bonferroni correction).

	SGW5 vs ECG		shimPPG vs ECG	
Feature	Rest	Walk	Rest	Walk
meanRR	**0.99*** **	**0.63** **	**0.99*** **	**0.69*** **
stdRR	**0.69*** **	0.28	**0.61** **	0.41
RMSSD	0.49	0.37	0.34	0.34
LF	**0.74*** **	0.05	**0.96*** **	0.12
HF	**0.64** **	0.09	**0.55* **	0.45
LF/HF	0.47	0.14	**0.61** **	0.36
SD1	0.49	0.37	0.34	0.34
SD2	**0.85*** **	0.22	**0.87*** **	0.48

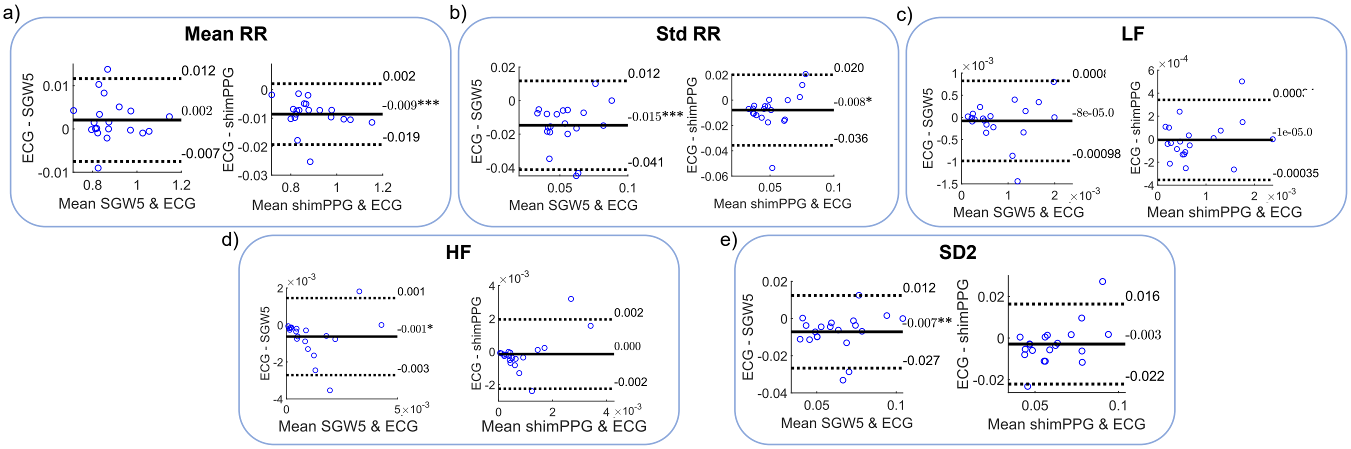

Fig. 1: BA analysis results for the SGW5vsECG and shimPPGvsECG comparisons of a)meanRR (ms), b)stdRR (ms), c)LF (ms^2), d)HF (ms^2), e)SD2 (ms) PRV/HRV features during resting-state. For each comparison and for each feature, we report the bias, computed as the average difference between the devices' measurements, and the limits of agreement, indicating the ± 1.96 standard deviation interval around the bias. Biases that significantly differed from 0 are indicated with * (*:p<0.05, **:p<0.01, ***:p<0.001; p-values adjusted with the Bonferroni correction).

We observed similar correlations for the SWG5 and shimPPG at rest. Specifically, meanRR estimates showed an almost perfect degree of agreement with ECG ones ($\rho_{SGW5vsECG} = 0.99$, $\rho_{shimPPGvsECG} = 0.99$). Moreover, SGW5 and shimPPG performances were comparable for both stdRR ($\rho_{SGW5vsECG} = 0.69$, $\rho_{shimPPGvsECG} = 0.61$) and SD2 ($\rho_{SGW5vsECG} = 0.85$, $\rho_{shimPPGvsECG} = 0.87$). In the spectral domain, we observed significant correlations for both LF and HF power estimates. Particularly, SGW5's LF measurements showed a high correlation with those obtained from ECG ($\rho_{SGW5vsECG} = 0.74$), although shimPPG correlation was even higher ($\rho_{shimPPGvsECG} = 0.96$). Conversely, SGW5's HF estimates showed a higher correlation with ECG, compared to shimPPG ($\rho_{SGW5vsECG} = 0.64$, $\rho_{shimPPGvsECG} = 0.55$). However, while we found a significant correlation between shimPPG and ECG for the LF/HF power ratio ($\rho_{shimPPGvsECG} = 0.61$), such a relationship was not observed for the SGW5. Finally, we did not find a significant correlation for both RMSSD and SD1 estimates obtained from SGW5 and shimPPG.

During the walking condition, the overall agreement between PRV and HRV features worsened with respect to the resting state for both SGW5 and shimPPG. Devices showed comparable significant correlation coefficients for meanRR, with respect to ECG ($\rho_{SGW5vsECG} = 0.63$, $\rho_{shimPPGvsECG} = 0.69$). Nevertheless, none of the other features showed a significant correlation.

In Fig.1-2, we report the results of the BA analysis for both the SGW5 and shimPPG against ECG, during the resting state and walking condition, respectively. Particularly, we report only those features for which we observed a significant correlation between SGW5 and ECG. The SGW5 estimated meanRR during resting state with no bias, while shimPPG overestimated meanRR by 9ms (Fig.1). Conversely, the SGW5 showed a higher negative bias (i.e., an overestimation), compared to shimPPG, for the estimation of stdRR (SGW5=-20ms, shimPPG=-8ms; Fig.1b), HF (SGW5=-0.001ms^2, shimPPG=0ms^2; Fig.1d), and SD2 (SGW5=-0.007, shimPPG=0; Fig.1e). Particularly, both devices showed a tendency to overestimate more HF power with the increase of its average magnitude (see Fig.1d). On the contrary, devices showed a tendency to underestimate more the LF/HF ratio with the increase of its average magnitude, with the SGW5 having a greater bias, compared to the shimPPG (SGW5=2.402, shimPPG=2.008).

Concerning the walking condition, both PPG sensors estimated meanRR with no bias with respect to ECG. Nevertheless, the limits of agreement highlighted a wide range of variability among measurements (SGW5: [-78ms, 84ms]; shimPPG: [-80ms, 78ms]; see Fig.2). Additionally, we observed a significant overestimation and broad limits of agreement for all the other HRV features considered. Particularly, the overestimation trend increased at the increase of the measurements' magnitude.

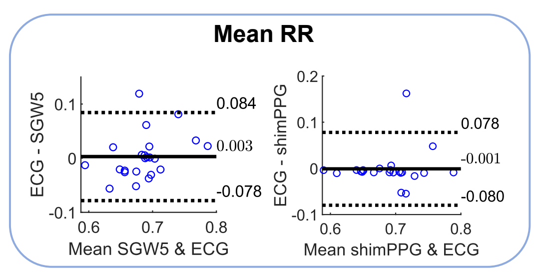

Fig. 2: BA analysis results for the SGW5vsECG and shimPPGvsECG comparisons of meanRR (ms) estimates during walking. For each comparison we report the bias and the limits of agreement (± 1.96 standard deviations interval around the bias). Biases that significantly differed from 0 are indicated with * (*:p<0.05, **:p<0.01, ***:p<0.001; p-values adjusted with the Bonferroni correction).

F. Discussion

In this study, we investigated the accuracy of the commercially-available SGW5 to monitor HR dynamics in a group of healthy volunteers during a resting state and a walking condition. To this aim, we estimated the PRV and its main features from the SGW5's PPG built-in sensor, and we compared them with the HRV features obtained from the Shimmer3 ECG wearable unit [21]. Additionally, the outcome of the comparison between the SGW5-PRV features and the ShimmerECG-HRV ones was further compared with that obtained by replacing the SGW5 with a research-grade PPG sensor, i.e., the Shimmer3 GSR+ unit.

At resting state, our preliminary results show the SGW5 as a reliable wearable tool to provide unbiased estimates of the average HR (from PRV time series). Moreover, the correlation analysis of SGW5 HRV features against ECG highlighted significant correlation coefficients with those observed for shimPPG for relevant features such as stdRR, LF, HF, and SD2. Nevertheless, Bland Altman's analysis showed a tendency for the SGW5 to estimate HRV features with a higher bias compared to shimPPG. Given the important informative content of these features (especially the frequency ones) linked to the autonomic nervous system, such bias could limit the reliability of the inference on sympathovagal balance. It is worthwhile noting that, while ECG is characterized by sharp R peaks, PPG has a smoother sinusoidal nature which makes peak detection intrinsically more difficult and unprecise. On the other hand, the comparison with shimPPG could be affected by the acquisition site. Indeed, previous studies reported more accurate HRV parameters extracted from the finger, with respect to the wrist [8]. In this light, we cannot exclude an effect of the sensor position on the differences observed between SGW5 and shimPPG.

Concerning the walking condition, both the SGW5 and shimPPG confirmed good reliability for the estimation of the mean RR, although BA analysis indicated a significant range of variability among measurements. The PRV-related features performed poorly compared to the resting state. These results were not totally unexpected, as they are in line with previous validation studies on different PPG wearable sensors [16]–[19]. Indeed, such devices are known to be particularly susceptible to motion artifacts [10]. Accordingly, deriving accurate PRV time series in such a context may be tricky.

In noisy scenarios, several processing techniques in the frequency domain have been proposed to accurately estimate the mean HR from PPG recordings [24]–[27]. However, to the best of our knowledge, such approaches do not provide an alternative means to estimate HRV time series and extract HRV features. Hence, our work aimed at evaluating the reliability of the PRV time series derived from the SGW5.

In conclusion, our results, although preliminary, highlight the reliability of the SGW5 as an open-source, non-invasive and low-cost wearable device to monitor mean HR. On the other hand, the reliability of PRV dynamics is still an open question and deserves further investigation. Particularly, for those parameters providing a window on sympathovagal regulation such as HF power, which is a reliable parasympathetic correlate to investigate stress, anxiety, and fatigue conditions [2], future studies should evaluate whether such HRV features extracted from the SGW5 can still be used to distinguish different psychophysiological states.

References

[1] M. P. Tarvainen, J.-P. Niskanen, J. A. Lipponen, P. O. Ranta-Aho, and P. A. Karjalainen, "Kubios hrv–heart rate variability analysis software," *Computer methods and programs in biomedicine*, vol. 113, no. 1, pp. 210–220, 2014.

[2] F. Shaffer and J. P. Ginsberg, "An overview of heart rate variability metrics and norms," *Frontiers in public health*, p. 258, 2017.

[3] A. L. Roque, V. E. Valenti, T. Massetti, T. D. Da Silva, C. B. d. M. Monteiro, F. R. Oliveira, Á. D. de Almeida Junior, S. N. B. Lacerda, G. C. Pinasco, V. G. Nascimento, *et al.*, "Chronic obstructive pulmonary disease and heart rate variability: a literature update," *International archives of medicine*, vol. 7, no. 1, pp. 1–8, 2014.

[4] A. Birkhofer, G. Schmidt, and H. Förstl, "Heart and brain–the influence of psychiatric disorders and their therapy on the heart rate variability," *Fortschritte der Neurologie-psychiatrie*, vol. 73, no. 4, pp. 192–205, 2005.

[5] R. Paradiso, G. Loriga, and N. Taccini, "A wearable health care system based on knitted integrated sensors," *IEEE transactions on Information Technology in biomedicine*, vol. 9, no. 3, pp. 337–344, 2005.

[6] M. Schaffarczyk, B. Rogers, R. Reer, and T. Gronwald, "Validity of the polar h10 sensor for heart rate variability analysis during resting state and incremental exercise in recreational men and women," *Sensors*, vol. 22, no. 17, p. 6536, 2022.

[7] S. Hartikainen, J. A. Lipponen, P. Hiltunen, T. T. Rissanen, I. Kolk, M. P. Tarvainen, T. J. Martikainen, M. Castren, E.-S. Väliaho, and H. Jäntti, "Effectiveness of the chest strap electrocardiogram to detect atrial fibrillation," *The American journal of cardiology*, vol. 123, no. 10, pp. 1643–1648, 2019.

[8] M. Nardelli, N. Vanello, G. Galperti, A. Greco, and E. P. Scilingo, "Assessing the quality of heart rate variability estimated from wrist and finger ppg: A novel approach based on cross-mapping method," *Sensors*, vol. 20, no. 11, p. 3156, 2020.

[9] N. Selvaraj, A. Jaryal, J. Santhosh, K. K. Deepak, and S. Anand, "Assessment of heart rate variability derived from finger-tip photoplethysmography as compared to electrocardiography," *Journal of medical engineering & technology*, vol. 32, no. 6, pp. 479–484, 2008.

[10] D. Castaneda, A. Esparza, M. Ghamari, C. Soltanpur, and H. Nazeran, "A review on wearable photoplethysmography sensors and their potential future applications in health care," *International journal of biosensors & bioelectronics*, vol. 4, no. 4, p. 195, 2018.

[11] F. Z. Tajrishi, M. Chitsazan, M. Chitsazan, F. Shojaei, V. Gunnam, and G. Chi, "Smartwatch for the detection of atrial fibrillation," *Critical Pathways in Cardiology*, vol. 18, no. 4, pp. 176–184, 2019.

[12] B.-G. Lee, B.-L. Lee, and W.-Y. Chung, "Smartwatch-based driver alertness monitoring with wearable motion and physiological sensor," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 6126–6129, IEEE, 2015.

[13] J. He, J. Ou, A. He, L. Shu, T. Liu, R. Qu, X. Xu, Z. Chen, and Y. Yan, "A new approach for daily life blood-pressure estimation using smart watch," *Biomedical Signal Processing and Control*, vol. 75, p. 103616, 2022.

[14] M. Boukhechba, L. Cai, C. Wu, and L. E. Barnes, "Actippg: using deep neural networks for activity recognition from wrist-worn photoplethysmography (ppg) sensors," *Smart Health*, vol. 14, p. 100082, 2019.

[15] W.-K. Beh, Y.-H. Wu, and A.-Y. A. Wu, "Robust ppg-based mental workload assessment system using wearable devices," *IEEE Journal of Biomedical and Health Informatics*, 2021.

[16] D. Hernando, S. Roca, J. Sancho, Á. Alesanco, and R. Bailón, "Validation of the apple watch for heart rate variability measurements during relax and mental stress in healthy subjects," *Sensors*, vol. 18, no. 8, p. 2619, 2018.

[17] N. Milstein and I. Gordon, "Validating measures of electrodermal activity and heart rate variability derived from the empatica e4 utilized in research settings that involve interactive dyadic states," *Frontiers in Behavioral Neuroscience*, vol. 14, p. 148, 2020.

[18] D. T. Weiler, S. O. Villajuan, L. Edkins, S. Cleary, and J. J. Saleem, "Wearable heart rate monitor technology accuracy in research: a comparative study between ppg and ecg technology," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 61, pp. 1292–1296, SAGE Publications Sage CA: Los Angeles, CA, 2017.

[19] D. Morelli, L. Bartoloni, M. Colombo, D. Plans, and D. A. Clifton, "Profiling the propagation of error from ppg to hrv features in a wearable physiological-monitoring device," *Healthcare technology letters*, vol. 5, no. 2, pp. 59–64, 2018.

[20] "Samsung galaxy watch 5." `https://www.samsung.com/it/watches/galaxy-watch/galaxy-watch5-40mm-graphite-bluetooth-sm-r900nzaaitv/`. Accessed: 2023-02-02.

[21] A. Burns, B. R. Greene, M. J. McGrath, T. J. O'Shea, B. Kuris, S. M. Ayer, F. Stroiescu, and V. Cionca, "Shimmer™–a wireless sensor platform for noninvasive biomedical research," *IEEE Sensors Journal*, vol. 10, no. 9, pp. 1527–1534, 2010.

[22] S. M. Bishop and A. Ercole, "Multi-scale peak and trough detection optimised for periodic and quasi-periodic neuroscience data," in *Intracranial Pressure & Neuromonitoring XVI*, pp. 189–195, Springer, 2018.

[23] J. M. Bland and D. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *The lancet*, vol. 327, no. 8476, pp. 307–310, 1986.

[24] Z. Zhang, Z. Pi, and B. Liu, "Troika: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise," *IEEE Transactions on biomedical engineering*, vol. 62, no. 2, pp. 522–531, 2014.

[25] Z. Zhang, "Photoplethysmography-based heart rate monitoring in physical activities via joint sparse spectrum reconstruction," *IEEE transactions on biomedical engineering*, vol. 62, no. 8, pp. 1902–1910, 2015.

[26] Y. Zhang, B. Liu, and Z. Zhang, "Combining ensemble empirical mode decomposition with spectrum subtraction technique for heart rate monitoring using wrist-type photoplethysmography," *Biomedical Signal Processing and Control*, vol. 21, pp. 119–125, 2015.

[27] B. Sun and Z. Zhang, "Photoplethysmography-based heart rate monitoring using asymmetric least squares spectrum subtraction and bayesian decision theory," *IEEE Sensors Journal*, vol. 15, no. 12, pp. 7161–7168, 2015.