Classification of Physiological States through Machine Learning Algorithms Applied to Ultra-Short-Term Heart Rate and Pulse Rate Variability Indices on a single-feature basis

Marta Iovino[1], Ivan Lazic[2], Tatjana Loncar-Turukalo[2], Michal Javorka[3], Riccardo Pernice[1], and Luca Faes[1]

[1] Department of Engineering, University of Palermo, Building 9, Viale delle Scienze, Palermo, Italy
[2] Faculty of Technical Sciences, University of Novi Sad, Serbia
[3] Department of Physiology, Jessenius Faculty of Medicine, Comenius University, Martin, Slovakia

Abstract. This study investigates the feasibility of classifying physiological stress states using Machine Learning (ML) algorithms on short-term (ST, ∼5 min) and ultra-short-term (UST, < 5 min, down to 10 heartbeats) heart rate (HRV) or pulse rate variability (PRV) features computed from inter-beat interval time series. Three widely employed ML algorithms were used, i.e. Naive Bayes Classifier, Support Vector Machines, and Neural Networks, on various time-, frequency- and information domain HRV/PRV indices on a single-feature basis. Data were collected from healthy individuals during different physiological states including rest, postural and mental stress. Results highlighted comparable values using either HRV or PRV indices, and higher accuracy (> 65% for most features and all classifiers) when classifying postural than mental stress. While decreasing the time series length, time-domain indices resulted still reliable down to ∼10 s, contrary to UST frequency-domain features which reported lower accuracy below 60 heartbeats.

Keywords: Heart Rate Variability (HRV), Pulse Rate Variability (PRV), Ultra-Short-Term (UST) variability, Stress Classification, Machine Learning (ML).

1 Introduction

The American Physiological Association has defined stress as "the pattern of specific and non-specific responses an organism makes to stimuli events that disturb its equilibrium" [6]. Excess stress has been recognised as one of the most significant pathogenic factors of modern life [22]. While numerous physiological signals have been proven suitable to detect stress, the most commonly used are the electrocardiogram (ECG) and the photoplethysmogram (PPG). A typical approach for noninvasive stress assessment involves examining Heart Rate Variability (HRV), which is the beat-to-beat variation of the heart rate (HR). HRV is

typically measured by analysing the time intervals between successive heartbeats (computed from ECG R-R intervals). Analysis can be performed on 24h recordings (long-term HRV analysis), 5 minutes recordings (short-term) or even shorter recordings (Ultra-Short Term, UST). The most commonly adopted method for assessing HRV, especially for practical applications, involves the use of short-term measurements. HRV indices are considered among the most dependable indicators of mental and physical stress [20]. The complexity of cardiovascular regulation is reflected by HRV, which can indicate the body's ability to react to a variety of stimuli (such as environmental and psychological stressors) due to the inhibition of the parasympathetic and activation of the sympathetic branches of the autonomous nervous system during stress [20].
In recent times, there has been a growing interest in investigating whether and to what extent HRV can also be assessed through PPG, whose cardiovascular variability indices are usually referred as to "Pulse Rate Variability" (PRV) and are computed from pulse-to-pulse intervals [17, 20]. PPG is an optical technique used in wearable devices that can detect changes in microvascular blood volume. It is simple, low-cost, safe, and minimally invasive [12]. Although PPG and ECG are often considered interchangeable for measuring HRV, several pieces of evidence suggest that the beat-to-beat variability recorded with PPG is somewhat different from HRV [12, 17, 19]. As wearable sensors become more widespread, there is a rising demand for using UST recordings as a substitute for short-term HRV and PRV. This satisfies the increasing need for continuous and real-time monitoring of the individuals' well-being status [3]. Recently, Machine Learning (ML) algorithms have been utilised to aid in the classification of various autonomic nervous system states related to different stress types [1]. Applying ML techniques to HRV and PRV features to classify the stress level thus represents an important challenge for researchers. In particular, various studies have been focused on classifying physiological states using HRV [4, 7] or PRV [1, 15] time- or frequency-domain indices.
In a previous paper [8] we investigated the performance of several ML algorithms to classify postural and mental stress using short-term HRV and PRV indices. Another work investigated the feasibility of employing ultra-short-term time, frequency, and information-domain HRV indices to distinguish among different stress conditions, using classical statistical analyses [23]. Herein, we aim instead to apply ML algorithms to UST HRV and PRV indices computed on rest and during postural or mental stress. To the best of our knowledge, this is the first study employing single-feature ML algorithms for physiological state classification using information-domain measures alongside classical time- and frequency-domain indices. The objective is to assess the effectiveness of a single-feature classification for the identification of the features providing better discrimination of different physiological states in healthy young individuals. Moreover, the study aims to assess whether using UST HRV and PRV features allows achieving results comparable to the short-term indices.

2 Materials and methods

2.1 Subjects and experimental protocol

Analyses were carried out on a previously used to assess cardiovascular variability during orthostatic and mental stress [17, 13]. Data were collected from 76 young, healthy normotensive volunteers with normal body mass index (32 males and 44 females; age: 18 $\pm$ 2.7 years). The analysed physiological signals consisted of horizontal bipolar thoracic leads ECG recordings and arterial blood pressure acquired through the volume-clamp method (sampling rate: 1 kHz). Three different conditions of the experimental protocol have been taken into account for the analyses, in particular, a resting phase (REST), a head-up tilt state causing orthostatic stress (HUT) and a mental stress condition during which the subject is undergoing a mental arithmetic test (MA). Both physiological states have shown to be two stress conditions associated with a shift in sympathovagal balance towards sympathetic activation/parasympathetic inhibition [10]. We refer the reader to [17, 23, 8] for further details on the experimental protocol and the information on the ethical approval.

2.2 Preprocessing and features extraction

Starting from the acquired signals, time series of 300 heartbeats were extracted for each subject and condition. The R-R intervals (RRI) and the pulse-pulse intervals (PPI) time series were obtained by measuring the temporal distance between consecutive QRS complexes in the ECG and the blood pressure maxima, respectively, thus extracting the HRV and PRV time series. Indices were first computed on short-term (300 beats, i.e. $\sim$ 5 min) cardiovascular variability time series. Then, in order to perform a UST analysis, shorter time series were obtained by reducing the length of RRI and PPI series of 30 samples for each step, from 300 heartbeats down to a minimum of 60 samples [23], then of 15 samples down to a minimum of 30 heartbeats and finally of 10 samples down to 10 samples (i.e. 300, 270, 240, 210, 180, 150, 120, 90, 60, 45, 30, 20, and 10 heartbeats).

Twelve features, divided into three domains (time, frequency and information), were then calculated both on RRI and PPI series, [17, 23]. In detail, the following time domain indices were calculated for RRI and PPI series: mean (MEAN), root mean square of successive differences (RMSSD) and standard deviation of normal-to-normal intervals (SDNN). In the frequency domain, absolute spectral power was calculated in the Low Frequency (LF, 0.04-0.15 Hz) and High Frequency (HF, 0.15-0.4 Hz) bands, considering also the normalised power within these bands (LF_n and HF_n). LF-to-HF power ratio formerly used as the sympathovagal balance index (SVB) was calculated and the respiratory peak frequency f_{HF} was determined as the peak frequency in the HF band. In the information domain, three entropy measures were computed: the entropy (H), the conditional entropy (CE), and the self-entropy (SE) of the RRI and PPI time series. The last two features were estimated using the k-nearest neighbour approach, a model-free method based on nearest-neighbour metrics, as in [23]. We refer the reader to [17, 23] for further details on the computation of the features mentioned above.

2.3 Classification algorithms

Three widely used state-of-the-art ML classifiers were compared in terms of their performance in discriminating between rest and orthostatic/mental stress. The analyses were entirely conducted on MATLAB 2021b (The MathWorks, Inc., Natick, MA, USA) using the integrated "Statistics and Machine Learning Toolbox". All classifiers were trained using two input vectors: (i) the first one containing the features used for training (each column represents an observation), (ii) the second one encompassing the response variable for each observation. The following ML algorithms were used:

- **Naive Bayes Classifier (NBC)** is a supervised linear learning algorithm that requires normally distributed features, mutually exclusive or independent of each other, and divides data into K classes using K discrimination functions [11]. The algorithm used assumes that all the classes have the same not regularised diagonal covariance matrix.
- **Support Vector Machines (SVMs)** are supervised learning algorithms that work by finding the hyperplane in a high-dimensional feature space that achieves the maximum separation between the distinct classes [18]. A Gaussian kernel function and a 'one-vs-all' learner were used to perform multiclass classification. The kernel scale was set to 12.1 and data were standardised by respectively centring and scaling each column of the predictor data by the weighted column mean and standard deviation.
- **Neural Networks (NNs)** are ML models with multiple layers of interconnected nodes (resembling human neurons) [11], trained to adjust interconnection weights and biases to minimise output discrepancies. The used NN architecture consisted of a fully connected layer with one node (i.e. the number of considered features), with a rectified linear unit as an activation function. Each numeric predictor variable was centred and scaled by the corresponding column mean and standard deviation to standardise data.

The analyses carried out in this work were aimed to compare the performances of the three mentioned ML algorithms to classify data into the different physiological states, REST phase, HUT phase and MA phase investigating as well the features which can effectively better discriminate among the distinct stress phases. Moreover, the study aims to determine whether UST HRV and PRV features can provide results comparable to the short-term indices.
For RRI and PPI data, the three classifiers were trained for each time series length after calculating the features described in Section 2.2. To carry out the training and testing of each classifier, a k-fold cross-validation (k=5), was used with a 20% of test and 80% of training. Each classifier was trained by providing the HRV or PRV features as input; the same model parameters were chosen for

both HRV and PRV features to compare the obtained results better. Since the time series had the same length in the three conditions, the classes were balanced.
The performance of the classifiers was evaluated mainly through accuracy, but also by calculating per-class metrics such as recall and specificity for each physiological state [21].

3 Results

Table 1 reports the accuracy values obtained for the three ML approaches in both the three-class and two-class classifications on a single-feature basis, to assess which state is best classified by each feature, using either HRV or PRV indices. In classification tasks, accuracy refers to the percentage of correct predictions the classifier makes. In a two-class classification problem, there are only two possible classes to which data can belong, hence the random guess accuracy is 50% (the same applies to recall and specificity). On the other hand, in a three-class classification problem, there are three possible classes to which the data can belong, thus the random guess threshold is equal to 33.33%. Values around or below the random guess threshold are reported in bold. The accuracy values of the three-class classification (REST vs. HUT vs. MA) (approximately 50%) were lower than those obtained for the two-class approach (REST vs. HUT and REST vs. MA) which achieved average values of roughly 70% and 60%, respectively. When classifying REST and MA, the only feature to achieve accuracy values higher than 65% was the MEAN, with values of 73.68%, 73.03%, and 76.32% achieved for the NBC, SVM, and NN classifiers, respectively. Higher accuracy values were obtained when distinguishing between REST and HUT classes (in grey), with multiple features exceeding an accuracy of 65% for RRI and PPI-based indices.
Table 2 illustrates the results in terms of accuracy and metrics by class, i.e. recall and specificity, for each ML algorithm considering each feature individually for the HUT class in the REST vs. HUT classification. The lowest accuracy values were obtained for SDNN, f_{HF} and LF features, with accuracy around the random guess (50%) for all three classifiers (in bold) and recall values in some cases lower than 60%. For the further analyses focusing on shorter time series, we decided to take into account only the best features in terms of accuracy, by setting a threshold of 65% allowing at least 15% improvement over a random guess [14]. Features that exceeded the 65% threshold for all the classifiers and for both HRV and PRV indices were CE, SE, HF, HF_n, LF_n, MEAN and RMSSD (in grey). The selected features achieved recall values for the HUT class in most cases higher than 80% as well.
Figure 1 depicts the accuracy of the three ML classifiers and of the most significant features, decreasing the length of RRI and PPI time series. The information-based features were computed down to 60 heartbeats while the frequency features were computed down to 20 heartbeats (black vertical lines). All features showed a decreasing trend with heartbeats, from a few, up to a maximum of 10 percentage points. Only the time domain features continued to yield higher

accuracy values for all the classifiers for both RRI and PPI-based indices till a few heartbeats, with values around 80% (RMSSD) or even 90% (MEAN).

Table 1. Total classification accuracy values using the three ML classifiers on a single-feature basis: comparison of the three-class classification (REST vs. HUT vs. MA) and the two-class classifications (REST vs. HUT or MA) using either RRI and PPI-based features.

		REST-HUT-MA			**REST-HUT**			**REST-MA**		
FEATURE		**NBC**	**SVM**	**NN**	**NBC**	**SVM**	**NN**	**NBC**	**SVM**	**NN**
$\boldsymbol{MEAN}$	$\boldsymbol{RRI}$	61.40	58.77	65.35	88.82	88.16	86.84	73.68	73.03	76.32
	$\boldsymbol{PPI}$	62.28	58.77	65.35	90.13	88.82	85.53	73.03	73.03	77.63
$\boldsymbol{RMSSD}$	$\boldsymbol{RRI}$	55.26	53.51	44.30	77.63	76.32	76.97	62.50	59.87	**50.00**
	$\boldsymbol{PPI}$	55.26	52.19	47.81	77.63	76.97	76.97	62.50	63.16	51.97
$\boldsymbol{SDNN}$	$\boldsymbol{RRI}$	46.93	42.11	35.09	66.45	66.45	61.18	61.18	61.18	53.29
	$\boldsymbol{PPI}$	44.74	39.91	37.28	65.13	65.13	65.79	61.18	61.18	55.26
$\boldsymbol{f_{HF}}$	$\boldsymbol{RRI}$	35.53	43.86	35.09	53.29	54.61	55.26	**48.68**	51.32	53.95
	$\boldsymbol{PPI}$	36.84	38.60	37.72	51.97	**50.66**	**50.66**	54.61	56.58	**46.05**
$\boldsymbol{HF}$	$\boldsymbol{RRI}$	49.12	50.44	45.18	71.05	69.74	73.68	61.18	54.61	54.61
	$\boldsymbol{PPI}$	46.05	46.05	43.42	69.08	67.76	67.11	61.84	54.61	**50.66**
$\boldsymbol{HF_n}$	$\boldsymbol{RRI}$	52.63	52.63	49.56	80.26	78.29	72.37	60.53	61.84	52.63
	$\boldsymbol{PPI}$	50.00	50.44	43.42	77.63	74.34	69.74	61.84	61.84	57.89
$\boldsymbol{LF}$	$\boldsymbol{RRI}$	36.40	**33.33**	34.65	53.29	53.29	53.95	55.92	53.95	**43.42**
	$\boldsymbol{PPI}$	41.23	**32.89**	**28.07**	**50.00**	51.97	**50.66**	55.92	51.32	**42.76**
$\boldsymbol{LF_n}$	$\boldsymbol{RRI}$	53.51	49.56	46.93	80.26	80.26	78.95	59.21	58.55	**42.76**
	$\boldsymbol{PPI}$	52.19	47.37	45.61	76.32	75.00	71.71	59.87	59.87	51.97
$\boldsymbol{SVB}$	$\boldsymbol{RRI}$	46.05	46.49	49.12	67.11	61.18	78.29	60.53	51.97	57.89
	$\boldsymbol{PPI}$	46.93	44.74	45.18	66.45	65.79	70.39	61.18	56.58	51.97
$\boldsymbol{CE}$	$\boldsymbol{RRI}$	51.32	51.32	52.19	76.32	78.29	75.66	**48.68**	51.97	51.32
	$\boldsymbol{PPI}$	52.19	49.56	45.61	74.34	74.34	71.05	53.95	**50.00**	**50.00**
$\boldsymbol{H}$	$\boldsymbol{RRI}$	45.18	39.47	41.67	67.76	67.11	60.53	61.84	61.18	51.97
	$\boldsymbol{PPI}$	42.98	38.60	41.67	65.79	66.45	57.24	59.87	60.53	**49.34**
$\boldsymbol{SE}$	$\boldsymbol{RRI}$	52.63	52.19	49.56	81.58	81.58	79.61	**50.00**	**49.34**	**50.66**
	$\boldsymbol{PPI}$	54.82	51.32	42.54	76.97	76.97	73.68	55.92	**50.00**	**44.74**

Table 2. Accuracy (ACC), recall (REC), specificity (SP), for the HUT class computed using RRI and PPI-based indices on a single-feature basis for the REST vs. HUT classification.

		REST vs. HUT								
		NBC			**SVM**			**NN**		
FEATURES		**ACC**	**REC**	**SP**	**ACC**	**REC**	**SP**	**ACC**	**REC**	**SP**
$\boldsymbol{MEAN}$	***RRI***	88.82	88.16	89.47	88.16	84.21	92.11	86.84	88.16	85.53
	PPI	90.13	89.47	90.79	88.82	86.84	90.79	85.53	85.53	85.53
$\boldsymbol{RMSSD}$	***RRI***	77.63	92.11	63.16	76.32	92.11	60.53	76.97	82.89	71.05
	PPI	77.63	92.11	63.16	76.97	90.79	63.16	76.97	76.32	77.63
$\boldsymbol{SDNN}$	***RRI***	66.45	77.63	55.26	66.45	85.53	**47.37**	61.18	73.68	**48.68**
	PPI	65.13	77.63	52.63	65.13	84.21	**46.05**	65.79	77.63	53.95
$\boldsymbol{f_{HF}}$	***RRI***	**53.29**	**50.00**	56.58	**54.61**	60.53	**48.68**	55.26	60.53	**50.00**
	PPI	**51.97**	**39.47**	64.47	**50.66**	**42.11**	59.21	**50.66**	53.95	**47.37**
$\boldsymbol{HF}$	***RRI***	71.05	96.05	**46.05**	69.74	96.05	**43.42**	73.68	78.95	68.42
	PPI	69.08	94.74	**43.42**	67.76	96.05	**39.47**	67.11	72.37	61.84
$\boldsymbol{HF_n}$	***RRI***	80.26	81.58	78.95	78.29	77.63	78.95	72.37	73.68	71.05
	PPI	77.63	81.58	73.68	74.34	71.05	77.63	69.74	72.37	67.11
$\boldsymbol{LF}$	***RRI***	**53.29**	69.74	**36.84**	**53.29**	96.05	10.53	**53.95**	57.89	**50.00**
	PPI	**50.00**	63.16	**36.84**	**51.97**	93.42	10.53	**50.66**	56.58	**44.74**
$\boldsymbol{LF_n}$	***RRI***	80.26	82.89	77.63	80.26	80.26	80.26	78.95	75.00	82.89
	PPI	76.32	78.95	73.68	75.00	75.00	75.00	71.71	72.37	71.05
$\boldsymbol{SVB}$	***RRI***	67.11	**39.47**	94.74	61.18	**27.63**	94.74	78.29	73.68	82.89
	PPI	66.45	**40.79**	92.11	65.79	**35.53**	96.05	70.39	67.11	73.68
$\boldsymbol{CE}$	***RRI***	76.32	72.37	80.26	78.29	75.00	81.58	75.66	75.00	76.32
	PPI	74.34	72.37	76.32	74.34	71.05	77.63	71.05	67.11	75.00
$\boldsymbol{H}$	***RRI***	67.76	73.68	61.84	67.11	78.95	55.26	60.53	68.42	52.63
	PPI	65.79	72.37	59.21	66.45	80.26	52.63	57.24	65.79	**48.68**
$\boldsymbol{SE}$	***RRI***	81.58	77.63	85.53	81.58	81.58	81.58	79.61	80.26	78.95
	PPI	76.97	75.00	78.95	76.97	75.00	78.95	73.68	75.00	72.37

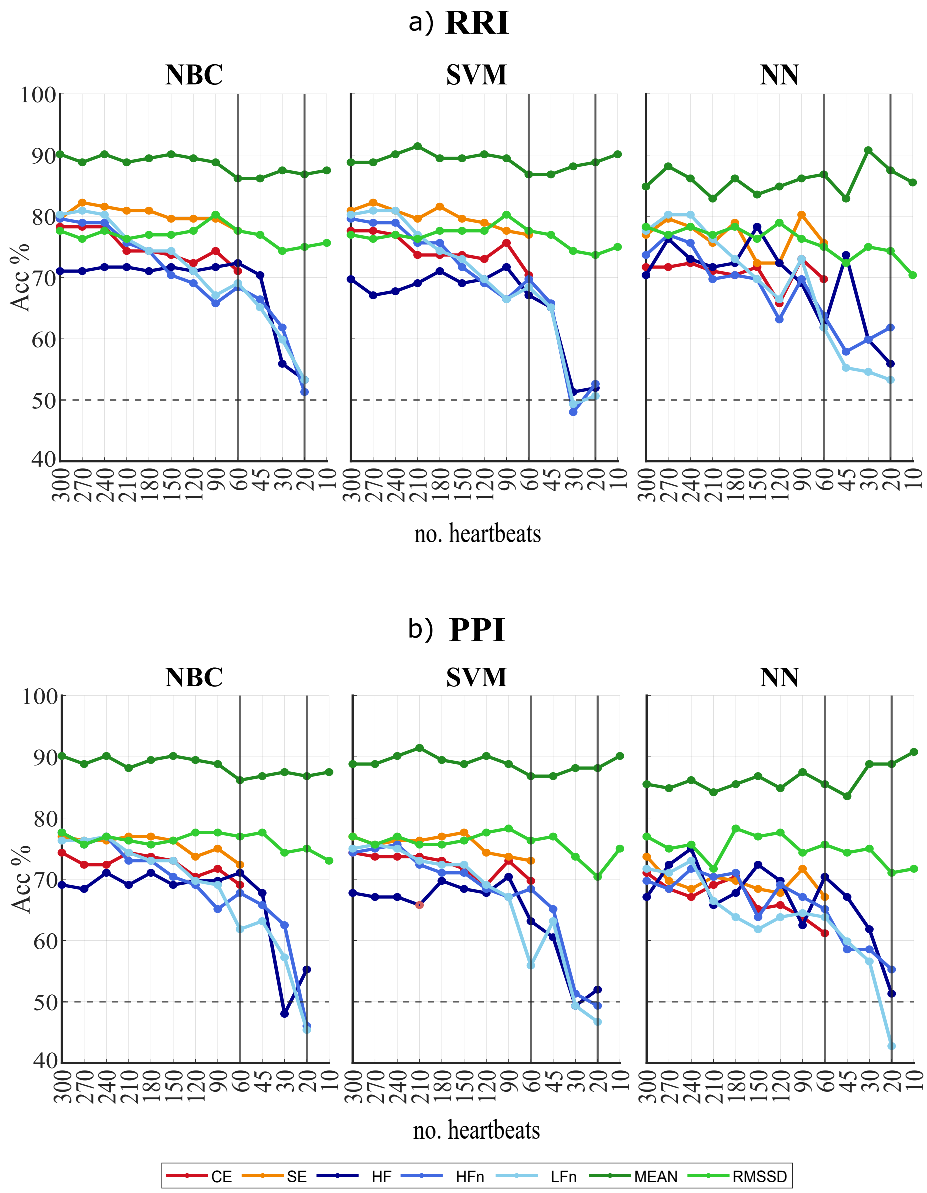

Fig. 1. Accuracy of the 3 ML classifiers on a single-feature basis at decreasing time series length (heartbeats) for (a) HRV and (b) PRV indices.

4 Discussion

The results described in Section 3 will be herein discussed. Firstly, in Section 4.1, our attention will be devoted to examining the viability of classifying physiological states by utilising various machine learning (ML) algorithms based on individual features of standard short-term HRV or PRV. Both three-stage (REST vs. HUT vs. MA) and two-stage classification (REST vs. HUT or vs. MA) will be taken into account. Subsequently, the viability of employing shorter time series (i.e. UST analysis), as a substitute for the short-term standard will be discussed in Section 4.2, focusing only on REST-HUT classification.

4.1 Classification using short-term HRV and PRV features

The accuracy values obtained with the three-class classification (Table 1) were indicative of an arbitrary classification primarily attributable to the classifiers' inability to effectively distinguish between the REST and MA classes. This observation is confirmed by the classification results obtained for the two-class REST vs. MA classification approach, where very low accuracy values were obtained. The classification accuracy was higher when distinguishing between the REST and HUT classes, for either RRI or PPI-based indices. This finding is deemed reliable as it has been previously demonstrated that the orthostatic stress condition is more marked than MA stress [23], given the stronger sympathetic activation and the concurrent parasympathetic withdrawal [16, 17].
In light of these results, the focus was shifted to REST vs. HUT classification, taking in this case also into account per-class metrics for a more detailed assessment of the effectiveness of the classification model. The results shown in Table 2 highlighted recall values consistently above 60%, reinforcing the effectiveness of the classification model for discriminating between the REST and HUT classes. On the other hand, only the frequency-domain HF and SVB features, strongly related to each other, exhibited low values across all three classifiers and for both RRI and PPI data.
Finally, similar accuracy and per-class metrics results were obtained using either HRV or PRV indices (cf. Tables 1 and 2).

4.2 REST vs HUT classification using Ultra-Short-Term features

To assess the feasibility of using UST analysis, only the features that yielded accuracy values exceeding 65% for all three classifiers were selected, i.e. MEAN and RMSSD w.r.t. time-domain, HF, HF_n and LF w.r.t. frequency-domain and finally CE and SE w.r.t. information-domain features. Using the above-mentioned seven features, the three classifiers were trained and tested by varying the length of the series from the initial 300 heartbeats down to 10 heartbeats. Although according to [3, 20] a minimum of 1 minute is needed to reliably estimate HF and at least 2 minutes for the LF component with regard to frequency-domain HRV/PRV analysis, we have adopted the approach followed by several other studies which have instead computed such features on even shorter time series

down to 10 s [2, 5, 9]. In our case, it was feasible to compute the information domain features using 60 heartbeats as a minimum. On the contrary, the frequency features were only up to 20 beats, while they started already to lose their discriminating capability from 60 beats downwards.
Our results highlight that the accuracy values are almost constant when considering time- and information-domain UST features (until they can be computed). On the other hand, frequency features accuracy values exhibited a decreasing trend, particularly when going below 60 heartbeats. Moreover, the three classifiers exhibited similar accuracy levels, and this was also observed when considering either HRV or PRV features. As already demonstrated by [3], not all UST indices are good short-term HRV and PRV surrogates, and generally, time domain features (e.g. MEAN and RMSSD) proved more effective in classifying the two physiological states. Although employing different methods and datasets, these results are in agreement with ours [2, 3, 5]. Herein, MEAN and RMSSD time domain features yielded higher accuracy values for RRI and PPI in all the classifiers, underscoring their high potential for classifying postural stress.

5 Conclusion

This work presented a comparison of the performances of various ML algorithms for classifying physiological states using either HRV or PRV indices on a single-feature basis. Our results indicated comparable results for RRI or PPI-based indices and confirmed that it was easier to discriminate postural than mental stress. Our findings identified some reliable UST HRV and PRV features (e.g. MEAN and RMSSD) that can be employed for detecting postural stress even for very few heartbeats ($\sim$ 10 s). On the other hand, UST frequency-domain HRV and PRV indices were worse short-term surrogates, with decreased accuracy for shorter ($<$ 60 s) time series. Future activities may envisage the use of the entire feature set and the application of other widely used classifiers (e.g. Random Forest [7]).

Acknowledgement

This work was supported by "Sensoristica intelligente, infrastrutture e modelli gestionali per la sicurezza di soggetti fragili" (4FRAILTY) project, funded by Italian MIUR (PON R&I 2014-20, grant no. ARS01_00345). R.P. is partially supported by the European Social Fund (ESF) Complementary Operational Programme (POC) 2014/2020 of the Sicily Region.

References

1. Awasthi, K., Nanda, P., Suma, K.: Performance analysis of machine learning techniques for classification of stress levels using ppg signals. In: 2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), pp. 1–6. IEEE (2020)

2. Bernardes, A., Couceiro, R., Medeiros, J., Henriques, J., Teixeira, C., Simões, M., Durães, J., Barbosa, R., Madeira, H., Carvalho, P.: How reliable are ultra-short-term hrv measurements during cognitively demanding tasks? Sensors **22**(17), 6528 (2022)
3. Castaldo, R., Montesinos, L., Melillo, P., James, C., Pecchia, L.: Ultra-short term hrv features as surrogates of short term hrv: A case study on mental stress detection in real life. BMC medical informatics and decision making **19**(1), 1–13 (2019)
4. Dalmeida, K.M., Masala, G.L.: Hrv features as viable physiological markers for stress detection using wearable devices. Sensors **21**(8), 2873 (2021)
5. Finžgar, M., Podržaj, P.: Feasibility of assessing ultra-short-term pulse rate variability from video recordings. PeerJ **8**, e8342 (2020)
6. Gerrig, R.J., Zimbardo, P.G.: American Psychological Association: Glossary of Psychological Terms. Pearson Education, Education, Incorporated (COR) (2002)
7. Giannakakis, G., Marias, K., Tsiknakis, M.: A stress recognition system using hrv parameters and machine learning techniques. In: 2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos(ACIIW), pp. 269–272. IEEE (2019)
8. Iovino, M., Javorka, M., Faes, L., Pernice, R.: Comparison of machine learning approaches for physiological states classification using heart rate and pulse rate variability indices. In: 8th National Congress of Bioengineering – GNB 2023 (submitted, 2023)
9. Kim, J.W., Seok, H.S., Shin, H.: Is ultra-short-term heart rate variability valid in non-static conditions? Frontiers in Physiology **12**, 596,060 (2021)
10. Lacey, K., Zaharia, M., Griffiths, J., Ravindran, A., Merali, Z., Anisman, H.: A prospective study of neuroendocrine and immune alterations associated with the stress of an oral academic examination among graduate students. Psychoneuroendocrinology **25**(4), 339–356 (2000)
11. Lim, H.W., Hau, Y.W., Lim, C.W., Othman, M.A.: Artificial intelligence classification methods of atrial fibrillation with implementation technology. Computer Assisted Surgery **21**, 154–161 (2016). DOI 10.1080/24699322.2016.1240303
12. Mejía-Mejía, E., May, J.M., Torres, R., Kyriacou, P.A.: Pulse rate variability in cardiovascular health: A review on its applications and relationship with heart rate variability. Physiological Measurement **41**(7), 07TR01 (2020)
13. Mijatovic, G., Pernice, R., Perinelli, A., Antonacci, Y., Busacca, A., Javorka, M., Ricci, L., Faes, L.: Measuring the rate of information exchange in point-process data with application to cardiovascular variability. Frontiers in Network Physiology **1**, 16 (2022)
14. Noulas, A., Mascolo, C., Frias-Martinez, E.: Exploiting foursquare and cellular data to infer user activity in urban environments. In: 2013 IEEE 14th International Conference on Mobile Data Management, vol. 1, pp. 167–176 (2013). DOI 10.1109/MDM.2013.27
15. Panganiban, F.C., de Leon, F.A.: Stress detection using smartphone extracted photoplethysmography. In: 2021 IEEE Region 10 Symposium (TENSYMP), pp. 1–7. IEEE (2021)
16. Paton, J., Boscan, P., Pickering, A., Nalivaiko, E.: The yin and yang of cardiac autonomic control: vago-sympathetic interactions revisited. Brain research reviews **49**(3), 555–565 (2005)
17. Pernice, R., Javorka, M., Krohova, J., Czippelova, B., Turianikova, Z., Busacca, A., Faes, L.: Comparison of short-term heart rate variability indexes evaluated through electrocardiographic and continuous blood pressure monitoring. Medical & biological engineering & computing **57**(6), 1247–1263 (2019)

18. Rani, S., Shelyag, S., Karmakar, C., Zhu, Y., Fossion, R., Ellis, J., Drummond, S., Angelova, M.: Differentiating acute from chronic insomnia with machine learning from actigraphy time series data. Frontiers in Network Physiology **2**, 1036,832 (2022)
19. Rinella, S., Massimino, S., Fallica, P., Giacobbe, A., Donato, N., Coco, M., Neri, G., Parenti, R., Perciavalle, V., Conoci, S.: Emotion recognition: Photoplethysmography and electrocardiography in comparison. Biosensors **12**(10) (2022). DOI 10.3390/bios12100811
20. Shaffer, F., Ginsberg, J.P.: An overview of heart rate variability metrics and norms. Frontiers in public health p. 258 (2017)
21. Shah, S.A., Ren, A., Fan, D., Zhang, Z., Zhao, N., Yang, X., Luo, M., Wang, W., Hu, F., Rehman, M.U., et al.: Internet of things for sensing: A case study in the healthcare system. Applied sciences **8**(4), 508 (2018)
22. Umair, M., Chalabianloo, N., Sas, C., Ersoy, C.: Hrv and stress: A mixed-methods approach for comparison of wearable heart rate sensors for biofeedback. IEEE Access **9**, 14,005–14,024 (2021). DOI 10.1109/ACCESS.2021.3052131
23. Volpes, G., Barà, C., Busacca, A., Stivala, S., Javorka, M., Faes, L., Pernice, R.: Feasibility of ultra-short-term analysis of heart rate and systolic arterial pressure variability at rest and during stress via time-domain and entropy-based measures. Sensors **22**(23), 9149 (2022)