Motion Artifact Detection and Classification for Unobtrusive Cardiorespiratory Signals Using Machine Learning

Onno Linschmann[1], Carl Revander[1], Steffen Leonhardt[1], Markus Lueken[1]

[1] Medical Information Technology, Helmholtz-Institute, RWTH Aachen University, Aachen, Germany

Abstract

For personal health care applications, unobtrusive sensors, such as reflective photoplethysmography (rPPG), capacitive electrocardiography (cECG) or ballistocardiography (BCG), are used with increasing frequency. While these sensors provide more comfort for the user, they exhibit a lower signal-to-noise ratio (SNR) and especially suffer from motion artifacts (MAs). Therefore, methods for reliable detection of MAs as well as their classification in the case of, for example, sleep analysis, are researched. In this paper, support vector machines (SVM) are investigated for the detection and classification of motion artifacts. Two methods for classification with respect to eight classes of movements are presented. First, a direct multi-class classification, and second, a multi-class classification after perfect detection. Waveform-related features are created and used for the training of the SVMs. An openly available dataset (UnoVis data set) which provides nine recordings of six channels of signals with annotations for motion is used. For the binary classification, an accuracy, sensitivity and specificity of 91%, 71%, 97% (test set) and 92%, 73% and 98% (validation set) are achieved respectively. For the direct multi-class classification, the SVM's performance is rather poor with mean accuracies, sensitivities and specificities of 77%, 21% and 93% (test set) and 78%, 28% and 93% (validation set) respectively. Similar results were achieved for perfect prediction.

1. Introduction

For home care applications, the usage of conventional measuring equipment relying on adhesive electrodes or face masks is not feasible since their setup is neither simple nor comfortable. Therefore, unobtrusive sensing modalities such as reflective photoplethysmography (rPPG), capacitive electrocardiography (cECG), magnetic induction monitoring (MIM) or ballistocardiography (BCG) are promising alternatives. However, their coupling with the subject's body is not fixed and thus they suffer from a reduced signal-to-noise ratio (SNR) and motion artifacts (MA). Consequently, the robust detection and reduction of MAs is crucial to obtain reliable measurements of vital signs such as heart rate, respiratory rate and heart rate variability (HRV).

Several approaches to detect and reduce MAs exist including adaptive filtering, blind source separation such as principle component analysis and independent component analysis, wavelet denoising and empirical mode decomposition [1–3]. Most of the denoising techniques aim to compensate the MAs directly without detecting them first. However, in several applications such as sleep tracking the detection and also the classification of MAs can contain important information e.g. to differentiate periodic limp movements from other movements.

In [4–6], MAs are detected by using time measures based on statistical features of the sensor signal (cECG and BCG). Artifact segments are then either discarded or a reduction algorithm is applied. Recent methods often apply Machine Learning methods such as Deep Learning and support vector machines (SVM) [7–11]. Usually they aim to detect MAs in a single sensor setup. Furthermore, in most cases only the accuracy for the detection is provided but not the specificity and sensitivity and the algorithms are only tested on simulated MAs. In this paper, the capabilities to detect and classify MAs with SVMs using real multimodal cardiorespiratory data are investigated. In the used data set from [12], real MAs were created by controlled movements and were annotated.

The paper is structured as follows. First, the concept of SVM is briefly described followed by the introduction of the used data set. Second, the methodology and usage of SVMs is described, and finally the results are presented and discussed.

2. Method

Since SVMs are used for binary and multi-class classification, the concept of SVMs is briefly introduced. The main task for using SVMs is the creation of appropriate features such that the classes are separable in the higher

 ISSN: 2325-887X DOI: 10.22489/CinC.2022.169

dimensional feature space. Therefore, the features used are presented after describing the data set. Finally, the test scenarios for hyperparameter tuning and choosing the best feature and channel combination are presented.

2.1. Support Vector Machines

SVMs use the so-called "kernel trick", i.e. they create a linear model in the (high dimensional) feature space leading to a nonlinear model in the data space for separating the classes (usually two) [13]. The problem [14] can be described by

$$y(\mathbf{x}) = \mathbf{w^T}\phi(\mathbf{x}) + \mathbf{b}, \tag{1}$$

$\mathbf{x}$ describes the data vector, ϕ describes the feature space transformation, $\mathbf{w}$ and b describe the parameters for the linear function in feature space, and the sign of $y(\mathbf{x})$ describes the label associated with $\mathbf{x}$. SVM are optimised in a way that so-called support vectors are computed and used to maximise the margin between the two classes, since an arbitrary number of separation functions may exist [13,14]. For multi-class classification one-vs-all or one-vs-one classification schemes can be used.

2.2. Data

The openly available data set [12, 15] used for training and evaluating the approach consists of nine recordings of six unobtrusive sensors sampled at 100 Hz (rPPG, cECG, optical BCG, EmFi BCG, and two high-frequency impedance measurements (HF)). Furthermore, a reference ECG, an impedance pneumography, and body markers for motion capturing were available. Only the motion signals were used to define segments where motion artifacts occur. Eight classes of movements were performed, i.e. torso shift left and right (TSL/TSR), torso torsion clockwise (TT)/counter-clockwise (TTc), head torsion (HT), lifting of left/right arm (LL/LR) and standing up/sitting down (S/S). Each of the movements was performed with two amplitudes except standing up/sitting down. The reference for the motion artifacts was defined by whether a movement was performed or not. The amplitudes were not differentiated.

2.3. Feature engineering

Each signal is windowed and features are computed for each window (higher-dimensional feature space). 8 features are used, namely, the variance (VAR), the mean absolute value (MAV), the waveform length (FL), the number of zero crossings (ZC), the number of slope sign changes (SSL), the mean frequency (MF), the median frequency (MDF) and discrete wavelet transform (WAV) features with a Daubechie 1 mother wavelet for one high-pass and low-pass stage. These features are then used to train an SVM where its feature space transformation is described by a kernel function

$$k(x, x') = \phi(\mathbf{x})^{\mathbf{T}}\phi(\mathbf{x}). \tag{2}$$

Note that $\mathbf{x}$ contains the previously computed features and not the raw data points.

Table 1. Full number of tuning parameters.

Channels	cECG, BCG (optical), BCG (EmFi), HF1, HF2, PPG
Window lengths [samples]	25, 50, 75, 100
Features	VAR, MAV, FL, ZC, SSC, MF, MDF, WAV

2.4. Test scenarios

Different scenarios for (hyperparameter) tuning of the SVM were created. The influence of the window length (WL), number of channels and number of features were investigated. Furthermore, the hyperparameters of the SVM were tuned using the 'auto' function of Matlab's *fitcsvm* (binary classification) or *fitcecoc* (multi-class classification) function. For the kernel function the Gaussian function was used. To validate the results a leave-one-out cross-validation was implemented since the internal cross-validation of Matlab's functions mixed the test and validation sets. For multi-class classification of the motion artifacts, a one-versus-one procedure is used. A full list of the changed parameters can be found in Tab. 1. To reduce the search space, a subset of feature/channel combinations were investigated. First, SVMs with all features and all channels, but different window lengths were trained to choose the best values for them. Second, SVMs with a single channel and a single feature were trained to investigate the influence of each channel and each feature. Finally, a reduced set of channels and features was selected for training to improve the results. The results of the binary classification were also used for the multi-class classification task. Nonetheless, SVMs using the full set of features and channels were also trained as a comparison.

3. Results

3.1. Binary classification

To assess the effect of different WLs, SVMs with all features and all channels were trained. As visible in Tab. 2, the WL does not influence the results of the classification severely. The accuracy, sensitivity and specificity lie around 0.8, 0.9 and 0.78 respectively. Therefore, for the proceeding SVMs, a WL of 50 samples was chosen. To

Table 2. Comparison of the influence of the WL on classification. Number in brackets is the mean of the leave-one-out cross validation.

WL [samples]	Accuracy	Sensitivity	Specificity
25	0.79 (0.88)	0.90 (0.76)	0.77 (0.91)
50	0.80 (0.87)	0.90 (0.77)	0.78 (0.90)
75	0.81 (0.87)	0.90 (0.76)	0.78 (0.90)
100	0.81 (0.87)	0.89 (0.75)	0.79 (0.91)

analyse which features and channels were the most important ones for classification, SVMs with only a single channel and a single feature were trained. With respect to the channels the optical BCG and EmFit BCG channels had the highest accuracy and sensitivity over all features (average of 0.85, 0.85 and 0.45 and 0.47 respectively) and the second highest specificity (average 0.97 and 0.97). The highest specificities were achieved with the rPPG (0.987) and HF1 (0.989). The features with the best results regarding accuracy, specificity and sensitivity over all channels were the VAR, MAV and WAV. An average accuracy of 0.817, 0.813 and 0.834, an average sensitivity of 0.45, 0.44 and 0.54, and an average specificity of 0.93, 0.925 and 0.926 were achieved. Therefore, a reduced set of features and channels was investigated. The SVM trained on the reduced channel and feature set (both BCG channels and VAR, MAV, WAV) achieved an accuracy of 0.92, a sensitivity of 0.72, and a specificity of 0.985. Compared with the SVM trained on the full set an increase in accuracy and specificity could be achieved while the sensitivity decreased by nearly 20 %.

3.2. Multi-class classification

For the multi-class classification two scenarios were investigated. First, direct multi-class classification and second, multi-class classification after a perfect detection of MAs. SVMs with full and reduced channel and feature sets were trained and compared. From Tab. 3, it can be seen that the accuracy and the specificity increase when using the reduced channel and feature set. However, the sensitivity stays very low at around 0.27 for the direct classification and around 0.31 for perfect detection. In case of the direct mulit-class classification it can be seen from the confusion matrix (not shown) that for the full set a slight overfitting to the no motion artifact class (NM) occurs. Furthermore, HT is often confused with NM. For the reduced set (see Fig.1), the SVM overfits severely to the NM class followed by the S/S class. In case of perfect detection, it can be seen from the confusion matrix in Fig. 2 for the full set that the SVM slightly overfits to TSL. For the reduced set the SVM overfits to HT and LA. In all cases S/S is classified the most accurate. All in all, misclassification of

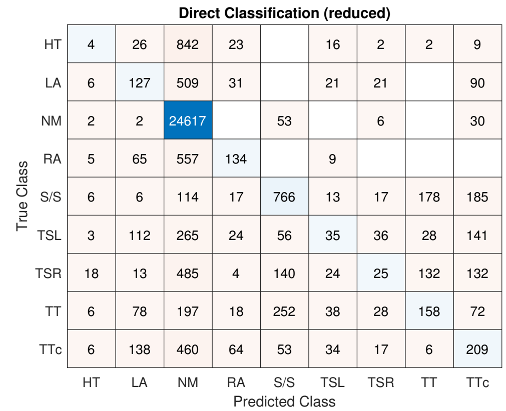

Figure 1. Confusion matrix for direct classification and the reduced channel and feature set.

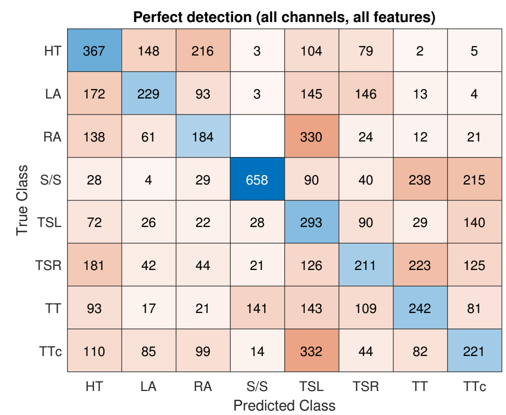

Figure 2. Confusion matrix for classification after perfect detection and all channels and all features.

movements occurs more often than correct classification.

4. Discussion and conclusion

The results of the SMVs for the binary classification are promising. However, there are certain limitations and problems with respect to the binary classification, the multi-class classification and the data set. First, the sensitivity of the binary classification, especially for the multi-class classification are still quite low. Therefore, it can be assumed that the amount of correctly classified windows is still quite low. In fact, at the beginning and end of MAs the classification is often incorrect. In the case of the multi-class classification overfitting occurs and classes are often confused. Therefore, a reliable classification of different movements could not yet be achieved. More sophisticated features and different classifiers should be investigated. The shortcomings of the classification do not only arise from the SVMs and their training, but also from some limitation of the data. MAs were defined as segments where the subject moved. However, for some movements

Table 3. Mean accuracy, sensitivity and specificity for the multi-class classification.

	Accuracy	Sensitivity	Specificity
	direct classification		
full set	0.78 (0.77)	0.28 (0.22)	0.93 (0.93)
reduced set	0.96 (0.96)	0.27 (0.27)	0.94 (0.94)
	perfect detection		
full set	0.83 (0.82)	0.31 (0.26)	0.90 (0.90)
reduced set	0.84 (0.82)	0.33 (0.28)	0.91 (0.90)

and some channels no MA was visible in the data while movement occurs, e.g. for HT. Another problem with respect to overfitting is that the data is highly imbalanced (more windows without motion than with motion). Also, the number of recordings is still limited and a larger data set would be desirable. All in all it could be shown that Machine Learning approaches can also be useful for detection and classification of real MAs in a real world application, although the sensitivity must still be improved.

References

[1] Ottenbacher J, Kirst M, Jatobá L, Huflejt M, Grossmann U, Stork W. Reliable motion artifact detection for ECG monitoring systems with dry electrodes. Annual International Conference of the IEEE Engineering in Medicine and Biology Society 2008;2008:1695–1698. ISSN 2375-7477.

[2] Sweeney KT, Ward TE, McLoone SF. Artifact removal in physiological signals–practices and possibilities. IEEE Transactions on Information Technology in Biomedicine IEEE Engineering in Medicine and Biology Society 2012; 16(3):488–500.

[3] An X, K Stylios G. Comparison of motion artefact reduction methods and the implementation of adaptive motion artefact reduction in wearable electrocardiogram monitoring. Sensors Basel Switzerland 2020;20(5).

[4] Škorić T. Reduction of artifacts in capacitive electrocardiogram signals of driving subjects. Entropy Basel Switzerland 2021;24(1).

[5] Alivar A, Carlson C, Suliman A, Warren S, Prakash P, Thompson D, Natarajan B. Motion detection in bed-based ballistocardiogram to quantify sleep quality. In 2017 IEEE Global Communications Conference. ISBN 978-1-5090-5019-2, 04.12.2017 - 08.12.2017; 1–6.

[6] Alivar A, Carlson C, Suliman A, Warren S, Prakash P, Thompson DE, Natarajan B. Motion artifact detection and reduction in bed-based ballistocardiogram. IEEE Access 2019;7:13693–13703.

[7] Stalin S, Roy V, Shukla PK, Zaguia A, Khan MM, Shukla PK, Jain A. A machine learning-based big EEG data artifact detection and wavelet-based removal: An empirical approach. Mathematical Problems in Engineering 2021; 2021:1–11. ISSN 1024-123X.

[8] Enayati M, Farahani NZ, Skubic M. Machine learning approach for motion artifact detection in ballistocardiogram signals. In Munson SA, Schueller SM, Arriaga R (eds.), Proceedings of the 14th EAI International Conference on Pervasive Computing Technologies for Healthcare. New York, NY, USA: ACM. ISBN 9781450375320, 05182020; 406–410.

[9] Sivanantham A. Measurement of heartbeat, respiration and movements detection using smart bed. In 2015 IEEE Recent Advances in Intelligent Computational Systems (RAICS). ISBN 978-1-4673-6670-0, 10.12.2015 - 12.12.2015; 105–109.

[10] Zargari AHA, Aqajari SAH, Khodabandeh H, Rahmani AM, Kurdahi F. An accurate non-accelerometer-based PPG motion artifact removal technique using CycleGAN. ArXiv 2021;abs/2106.11512.

[11] Goh CH, Tan LK, Lovell NH, Ng SC, Tan MP, Lim E. Robust PPG motion artifact detection using a 1-d convolution neural network. Computer Methods and Programs in Biomedicine 2020;196:105596.

[12] Hoog Antink C, Schulz F, Leonhardt S, Walter M. Motion artifact quantification and sensor fusion for unobtrusive health monitoring. Sensors Basel Switzerland 2017;18(1).

[13] Efron B, Hastie TJ. Computer age statistical inference: Algorithms, evidence, and data science, volume 5 of Institute of Mathematical Statistics Monographs. Cambridge: Cambridge University Press, 2019. ISBN 978-1107149892.

[14] Bishop CM. Pattern recognition and machine learning. Information Science and Statistics. New York: Springer, 2006. ISBN 978-0-387-31073-2.

[15] Hoog Antink C, Leonhardt S, Schulz F, Walter M. MuSeSe - a multisensor armchair for unobtrusive vital sign estimation and motion artifact analysis. Annual International Conference of the IEEE Engineering in Medicine and Biology Society 2017;2017:857–860. ISSN 2375-7477.

Address for correspondence:

Onno Linschmann
Medical Information Technology
Helmholtz Institute, RWTH Aachen University
Pauwelsstr. 20, D-52074 Aachen
linschmann@hia.rwth-aachen.de