Sleep Position Detection for Closed-Loop Treatment of Sleep-Related Breathing Disorders

Breuss, A.; Vonau, N.; Ungricht, C.; Schwarz, E.; Irion, M.; Bradicich, M.; Grewe, F. A.; Liechti, S.; Thiel, S.; Kohler, M.

Published in:
2022 International Conference on Rehabilitation Robotics, ICORR 2022

DOI:
10.1109/ICORR55369.2022.9896559

2022 International Conference on Rehabilitation Robotics (ICORR)
25-29 July, 2022, Rotterdam, The Netherlands

Sleep Position Detection for Closed-Loop Treatment of Sleep-Related Breathing Disorders

A. Breuss, N. Vonau, C. Ungricht, E. Schwarz, M. Irion, M. Bradicich, F. A. Grewe, S. Liechti, S. Thiel, M. Kohler, R. Riener, and E. Wilhelm

2022 International Conference on Rehabilitation Robotics (ICORR) | 978-1-6654-8829-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICORR55369.2022.9896559

Abstract— **Reliable detection of sleep positions is essential for the development of technical aids for patients with position-dependent sleep-related breathing disorders. We compare personalized and generalizable sleeping position classifiers using unobtrusive eight-channel pressure-sensing mats. Data of six male patients with confirmed position-dependent sleep apnea was recorded during three subsequent nights. Personalized position classifiers trained using leave-one-night-out cross-validation on average reached an F1-score of 61.3% for supine/non-supine and an F1-score of 46.2% for supine/lateral-left/lateral-right classification. The generalizable classifiers reached average F1-scores of 62.1% and 49.1% for supine/non-supine and supine/lateral-left/lateral-right classification, respectively. In-bed presence ("bed occupancy") could be detected with an average F1-score of 98.1%. This work shows that personalized sleep-position classifiers trained with data from two nights have comparable performance to classifiers trained with large inter-patient datasets. Simple eight-channel sensor mattresses can be used to accurately detect in-bed presence required for closed-loop systems but their use to classify sleep-positions is limited.**

I. Introduction

With a prevalence of up to 8% in men and 1 to 5% in women, sleep apnea accompanied with daytime sleepiness is one of the most common sleep-related breathing disorders in adults [1]–[4]. Occurrence and severity of apneic events are position dependent in approximately 56% to 75% of all patients with obstructive sleep apnea [5]. In position-dependent sleep apnea, more than 50% of all apneic events occur when the patient is lying in supine position [6]. Avoiding supine position by means of positional therapy can significantly reduce the occurrence and severity of sleep apnea in this sub-group of patients. Currently, most devices that assess the body position of the users use accelerometers. These sensors need to be physically attached to the user which may induce discomfort and therefore limits acceptability. Moreover, bodyworn sensors cannot be used to detect bed occupancy which would be required for closed-loop therapeutical interventions. As an alternative, automated video-based position scoring systems have been reported to provide robust estimations of the sleeping position [7]. However, space requirements and privacy issues limit the usability of such systems in private home environments.

* A. Breuss, C. Ungricht, S. Liechti, R. Riener, and E. Wilhelm are with the Department of Health Science and Technology, Sensory-Motor Systems Lab, Institute of Robotics and Intelligent Systems, ETH Zurich, Zurich, Switzerland. Corresponding author: +41 44 632 07 14; e-mail: alexander.breuss@hest.ethz.ch

E. Schwarz, M. Irion, F. A. Grewe, S. Thiel, M. Kohler are with the Sleep Disorders Centre of the Department of Pulmonology of the University Hospital Zurich, University Hospital of Zurich (USZ), Zurich, Switzerland

R. Riener is with the Spinal Cord Injury Center, University Hospital Balgrist, Zurich, Switzerland

E. Wilhelm is with the Engineering and Technology Institute, Faculty of Science and Engineering, University of Groningen, The Netherlands

Elastic force-sensitive resistors placed underneath the posts of the bed [8] or underneath the mattress [9], pose an interesting alternative to classify body positions. These sensors can be used in the context of ballistocardiography (BCG), a non-invasive sensing method to measure forces generated by thoracic movements (e.g. respiration) and cardiac ejection of blood [10]. BCG has also been used as a non-invasive system to alert users of apneic events by turning on lights or through vibration [11]. In general, pressure-based position sensing systems can be divided in systems that apply algorithms for image processing and setups which use less sophisticated processing techniques. Systems relying on image processing techniques require high-density arrays consisting of 60 to more than 8000 sensors while low-complexity systems usually work with 1 to 100 sensors (see Table I).

Previously presented classifiers and feature extraction methods for position detection required a high number of sensors and are based on computing the earth mover's distance [12], gradients in the pressure images [13], or posture signature extraction [14]. Other approaches used with high-density pressure sensor arrays consider to cluster regions of interest [15] such as limbs [16], hip, and shoulder [17] and compare size, location, and distances between those regions. In addition, the curvature of the resulting pressure image has been suggested as a feature for classification [18]. These systems discriminate between up to nine different body positions.

TABLE I. Comparison of High-Density and Low-Complexity Systems for Position Classification found in Literature

	Reference	*Sensors*	*Postures*	*Accuracy*
High-density systems	Huang et al. [18]	60	9	46.86% [a]
	Pouyan et al. [14]	2048	8	71.1%
	Xu et al. [12]	8192	6	90.78%
	Sun et al. [16]	6144	6	97.8%
	Matar et al. [13]	1728	4	97.9%
	Ostadabbas et al. [20]	1728	3	98.4%
	Liu et al. 2014 [17], [21]	8192	6	83.0% [c]
	Mineharu et. al [15]	1768	9	77.1%
Low-complexity systems	Foubert et al. [9]	132	2	100%
	Adami et al. [8]	4	4	63.3% [b]
	Wei al. [20]	56	6	94%
	Liu et al. 2019 [19]	1	4	92.3% [b]
	Crivello et al. [22]	32	4	95%
	Hsia et al. [23]	16	3	81.43%

a. only force-sensing resistor pressure sensor mapping, b. for supine only, c. Recall

In this paper, we present personalized position classifiers that use simple 8 × 1 force sensor arrays to classify between supine and non-supine as well as between supine, lateral left, and lateral right. We compare these results to an approach in which the classifier was trained based on five participants and tested on a sixth participant in a leave-one-participant-out manner. In addition, we present a rule-based approach to detect bed occupancy for uncalibrated force sensors.

II. Materials and Study Design

A. Measurement Set-up and Ground Truth

All measurements were carried out in the context of a study that investigated the effects of controlled positional therapy on positional obstructive sleep apnea patients using ISABel I, a 90 × 200 cm single bed able of raising the backrest from 0° to 50°, and ISABel II, a 160 × 200 cm double bed capable of turning both bed halves from 0° to 30° (Figure 1).

Figure 1. The intervention beds ISABel I (left) and ISABel II (right) used for positional therapy. ISABel I can raise the backrest up to 50° and ISABel II can induce rotations about the longitudinal axes up to 30°.

Both beds were placed at ETH Zurich and were equipped with custom-made sensor mattresses (sensomative, Switzerland), each 80 × 20 cm in size. For the larger bed, two eight-channel sensor mats were put next to each other to cover the entire width of the mattress (Figure 4). For the smaller bed, a single eight-channel sensor mat was sufficient to cover the width of the entire bed (Figure 3). Eight of the 16 sensor elements of the larger bed were selected based on the estimated location of the user such that the same features could be calculated. Each sensor mat contains eight individual sensors, each 9 × 20 cm in size, and was calibrated by the manufacturer using an inflatable balloon inside a testbench with varying preset pressures (Figure 2). Data from each 8 × 1 sensor mat was sampled using an eight-channel, 24-bit ADC (AD-7771 manufactured by Analog Devices, USA) at a nominal sampling rate of 1 kHz and recorded on a Raspberry Pi 4. In case of transmission errors from the ADC to the Raspberry Pi 4, NULL Byte placeholders were stored instead.

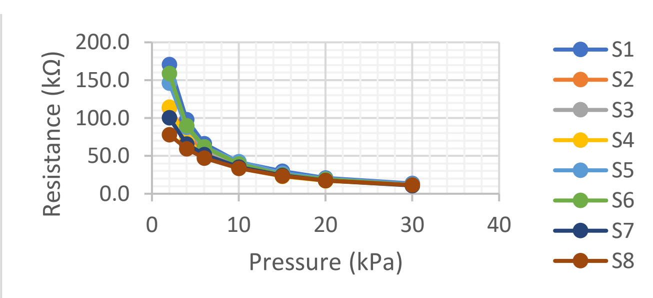

Figure 2. Calibration curve of one of the eight-channel sensor mattresses.

Ground truth positions were collected using the polysomnography system Nox A1 (Nox Medical Global, Iceland) and were manually validated by a clinical expert based on recorded video. The collected dataset (Table II) is highly imbalanced. Ground truth of the bed occupancy was extracted from the recorded video.

Figure 3. ISABel I, a standard bed with a single eight-channel force sensor.

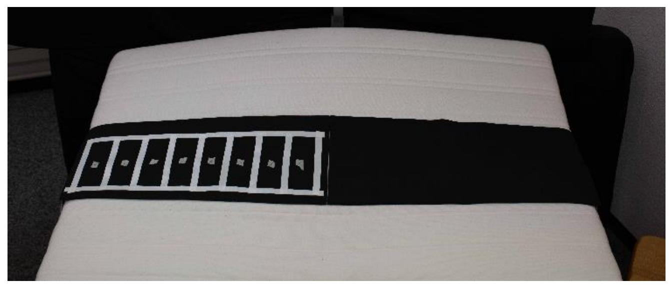

Figure 4. ISABel II, a 160 × 200 cm double bed with two sensor mattresses giving a total lateral resolution of 16 × 1 sensor elements. Only the eight sensor elements in contact with the participant are used for the position classification.

TABLE II. Relative Distribution of the Classes in the Dataset, M… Movement, U… Upright, S… Supine, LL… Lateral Left, LR… Lateral Right, P… Prone, O… Other

	M	U	S	LL	LR	P	O
P1	0.34	0.04	0.24	0.07	0.24	0	0.07
P2	0.03	0.03	0.15	0.39	0.37	0	0.03
P3	0.05	0.01	0.42	0.35	0.14	0.02	0.01
P4	0.07	0.01	0.25	0.27	0.24	0.02	0.14
P5	0.11	0.02	0.43	0.14	0.23	0	0.07
P6	0.08	0	0.36	0.29	0.24	0	0.03
Mean	0.11	0.02	0.31	0.25	0.24	0.01	0.06

B. Participants

The experimental procedure was approved by Swissmedic (10000733) and the Cantonal Ethics Committee (KEK-ZRH: 2020-01505) and was registered on clinicaltrials.gov (NCT04713267). Written informed consent was obtained from all participants. Ten participants were recruited by the Sleep Disorders Centre of the Department of Pulmonology of the University Hospital Zurich. Inclusion criteria were an apnea-hypopnea index (AHI) of at least ten events per hour, combined with a documented supine positional obstructive sleep apnea (AHI in supine position at least twice the AHI in non-supine position). Exclusion criteria were age below 18 or above 80 years, co-morbidities, and inability to follow the procedures of the study. In addition, investigators, their family members, employees, and other dependent persons were excluded. Three participants withdrew and one did not meet the inclusion criteria during the assessment. Demographics of the data sets of the remaining six patients used for training and testing of the algorithms are depicted in Table III. The AHI was calculated based on the expert scoring of the polygraphy data recorded during the baseline night. Daytime sleepiness was assessed using the standardized Epworth Sleepiness Scale (ESS) questionnaire which results in a score between 0 and 24. An ESS of more than 10 indicates increased daytime sleepiness [24].

TABLE III. DEMOGRAPHICS OF THE STUDY POPULATION (N=6)

	Mean (SD)
Age (years)	57.8 (6.7)
Height (m)	1.8 (0.1)
Weight (kg)	87.8 (13.0)
Body Mass Index (kg / m²)	28.2 (3.3)
Neck circumference (cm)	43.5 (4.7)
Apnea Hypopnea Index (/h)	30.4 (12.5)

Each participant visited the study facilities for three nights within two weeks. The baseline night was measured in the sidewards-turning bed ISABel II using the two 8 × 1 sensor mats next to each other. For the two intervention nights, each participant once slept in the sidewards-turning double bed ISABel II and once in the single bed ISABel I with the actuated backrest. Both beds provided postural interventions triggered by the clinician.

C. Evaluation Metrics and Validation

Sensitivity (Recall) is defined as the fraction of true values that were correctly identified by the classifier. Precision is defined as the fraction of true positive values among all positive predictions. The F1-score presents the harmonic mean of the precision and recall. Evaluation metrics were calculated using the sklearn.metrics package. Due to the large class imbalance (c.f. Table II), the weighted averages of these values are reported in this paper. Furthermore, we calculated the area under the receiver operating characteristic curve (AUC-ROC) to understand how well our model can distinguish between the two classes. The personalized algorithms were validated using leave-one-night-out cross validation where all possible intra-patient permutations of two nights of training and one test night were considered. After this process, the means of all evaluation metrics were calculated for all patients. To get an estimation for the inter-patient performance, we trained a second classifier for both the supine and non-supine classification and the multi-class classification. This classifier was trained on five participants and tested on the sixth participant in a leave-one-participant out cross validation.

III. DETECTION ALGORITHM

A. Time Synchronization and Drift Correction

As both the polysomnography device and the Raspberry Pi 4 have individual system clocks, time synchronization is required. To find the delay between these two devices, a random synchronization pulse captured by both systems is used. Using cross-correlation (eq. 1) between the two signals $\{X\}$ and $\{Y\}$, we determine the lag with the highest correlation and time-shift one of the signals by this amount.

$$\hat{\rho}_{i,j} = \frac{\Sigma_{t=1}^{N}(X_t-\bar{X}_t)(Y_t-\bar{Y}_t)}{\sqrt{\Sigma_{t=1}^{N}(X_t-\bar{X}_t)^2\Sigma_{t=1}^{N}(Y_t-\bar{Y}_t)^2}} \quad (1)$$

Drift compensation is required whenever the nominal frequency does not match the desired frequency. For the ADC of the sensor mattress, the nominal frequency of 1 kHz was above the actual frequency of approximately 999.998 Hz. A zero-order hold was therefore used to up-sample the actual signal.

Correctness of the time synchronization and drift correction was ensured by visual inspection of the resulting synchronization pulses between the Raspberry Pi 4 and the polysomnography for each night.

B. Preprocessing and Filtering

Data was cut in 10-second windows with 50% overlap. Missing values caused by transmission errors of the ADC (i.e. NULL Bytes) were replaced with the median of the respective channel. For the larger bed, the eight sensors of interest were identified using the variance of the de-trended sensor signals. De-trending was performed by log-scaling with a subsequent mean subtraction. All sensors then received a ranking value according to the variance. The window with the eight neighboring sensors for which the sum of the variance ranking scores was the highest, was used for the classification. The other eight sensors were ignored. This was necessary to be able to compare the results from one bed to the other.

Windows with the label "Unknown", “Movement”, “Upright”, “Prone” or without a position label were removed from the dataset before training.

TABLE IV. FEATURES USED IN OTHER LOW-COMPLEXITY SYSTEMS. FEATURES WITH NUMBERS ARE USED IN THIS WORK

Feature Name	***No.***	***Ref.***	***Feature Name***	***No.***	***Ref.***
Sum of sensor values	1	[9]	Variance of different areas	15-19	[20]
Number of active sensors	2	[9]	Root mean square of different areas	20-24	[20]
Weighted sum of sensor values	n.a.	[9]	Mean of raw pressure values	25-32	[20], [22]
Weighted number of active sensors	n.a.	[9]	Relative amplitude H -J	n.a.	[19]
Longitudinal center of pressure	n.a.	[9]	Relative amplitude I -J	n.a.	[19]
Lateral center of pressure	3	[9]	Relative amplitude K -J	n.a.	[19]
Longitudinal variance	n.a.	[9]	Relative amplitude L -J	n.a.	[19]
Lateral variance	4	[9]	Four dimensional distribution vector	n.a.	[8]
Eigenvectors	n.a.	[20]	Kurtosis	33	[23]
Mean of different areas	5 -9	[20]	Skewness	34	[23]
Standard Deviation of different areas	10-14	[20]			

Movements were removed as they can occur in each position and most windows that were labeled with this label did not contain information on the current body position. Since most nights did not contain prone position, this class also had to be removed. This left three classes for the classification problem: supine, lateral left, and lateral right.

C. Feature Extraction

Since the mats only contain eight sensors, we implemented the features which have been reported in literature describing low-complexity systems (Table IV). Features that require two-dimensional arrays and heart rate features were not implemented and marked with n.a. in Table IV. All features were extracted using Python 3.10. Our first feature is the sum of sensor values $SSV(t) = \Sigma_{i=1}^{8} x_i(t)$ where $x_i(t)$ is the sensor value of the i'th sensor element at time t [9]. The second feature is the number of active sensors. Active sensors were defined as sensors with a high standard deviation. To be able to compare between sensors with different sensitivity, we transferred the standard deviations into z-scores. Sensors with an absolute z-score above 0.8 were defined as active sensors.

Lateral center of pressure [9] was calculated by a weighted sum across all sensor elements followed by dividing through the sum of sensor values over the same time period.

Variance, mean, standard deviation, and root square values were calculated using the Python package NumPy (version 1.22.1). To get an impression of the pressure distribution on the mattress, sensors were grouped into areas. Area 1 was defined as the first two sensors from the right. Area 2 contained the last two sensors on the left. The four sensors in the middle were grouped into Area 3. To obtain more insight, the middle area was further divided into two sub-areas each of which consists of two sensors. That resulted in 5 areas over which basic statistic features were calculated. Skewness and kurtosis were calculated using the scipy.stats package (version 1.7.3).

In addition to the features used in other low-complexity systems (c.f. Table IV), we also included the average amplitude, the difference between the areas under the curve for all possible sensor pairs, as well as the area under the curve for every sensor element. These features were used to be able to compare the behavior of sensors with an offset that was not stable between different nights.

E. Sleep-Position Detection

We used XGBoost for the classification. Personalized binary (supine / non-supine) and multiclass (supine, lateral left, lateral right) classifiers were trained separately for all patients (intra-patient) using two nights as training and one night as test data. In addition, a single binary and multiclass classifier was trained using training data from all patients (inter-patient) and was validated using leave-one-patient-out.

The personalized binary classification was performed with logistic regression with a random forest tree size of 500, a maximum tree depth of 8, a subsample ratio of 1, and a learning rate of 0.05. The inter-patient binary classification was performed with a random forest tree size of 100, a maximum tree depth of 10, a subsample ratio of 0.8, and a learning rate of 0.3.

The personalized multiclass classification was performed using the Softmax objective, a random forest tree size of 500, a maximum depth of 8, a subsample ratio of 1, and a learning rate of 0.05. Inter-patient multiclass classification was also performed using the Softmax objective, a random forest tree size of 500, a maximum depth of 8, a subsample ratio of 0.8, and a learning rate of 0.1

F. Bed Occupancy Detection

The presence of the user on the bed can be used as input in closed-loop applications to start interventions or ensure the safety of the setup by preventing interventions while the user is not lying on the mattress. A simple decision rule to detect the presence of a user on the bed can be formulated as:

$$presence_{sensormat}(t) = \begin{cases} 1, & SMA(t) \leq thr_{sensormat} \\ & 0, else \end{cases} \quad (2)$$

where $SMA(t)$ denotes the simple moving average over a window of the past ten seconds up to timepoint t of the samples from the respective sensor mattress and $thr_{sensormat}$ some global threshold. The SMA must be smaller than the threshold as the measured values are inversely proportional to the pressure applied to the sensor elements. Setting a window length to 10 seconds ensures robustness towards false negatives that could else occur while the user is moving or briefly sitting up in the bed. The threshold is then defined for each night i as: $thr_{i,sensormat} = \bar{x}_{i,GT} - w\sigma_{i,GT}^2$ (3) where $\bar{x}_{i,GT}$ is the 10% trimmed mean and $\sigma_{i,GT}^2$ the variance of the ground truth data (i.e. the periods in the nights where the patients were not lying in the beds) of night i, and w a scalar weight parameter. The global threshold $thr_{sensormat}$ is then calculated by taking the mean of the thresholds of each night. Adjusting the threshold by a weighted variance of the data allows to better control the ratio of the sensitivity and specificity. For the computation of the global threshold, only the nights of the first five patients were used due to a partial loss of video data from patient 6. The global threshold was set to 3877158.3 AU.

IV. Results

A. Single-Class Position Classification

Each participant underwent three nights of data collection. We first report the inter-patient performance where two nights are used for training and the third night as test input. All evaluation metrics are reported as averaged means of the class performance to prevent overestimation of the classifier performance due to the highly unbalanced dataset. To see whether an algorithm would generalize to a person that was not contained in the training data, we trained and validated classifiers with the leave-one-participant out method (inter-patient). The results of can be found in Table V and VI. Personalized classifiers were able to distinguish between supine and non-supine positions with a mean F1-score of 0.61 and between supine and two lateral positions with a mean F1-score of 0.46.

TABLE V. SINGLE- AND MULTI-CLASS INTER-PATIENT PERFORMANCE IN %. FIFTEEN NIGHTS WERE USED AS INPUT FOR TRAINING, THE REMAINING 3 NIGHTS FOR EVALUATION.

P	*Single-Class*			*Multi-Class*		
	Precision	*Recall*	*F1-Score*	*Precision*	*Recall*	*F1-Score*
1	64.4	64.0	61.8	53.3	51.5	51.6
2	80.5	59.8	64.8	62.0	38.4	40.0
3	65.6	65.6	65.6	64.3	62.6	62.8
4	65.4	60.8	61.9	49.8	50.7	47.8
5	66.7	66.7	66.7	49.9	52.7	47.8
6	59.0	52.5	51.5	58.2	49.2	44.3
Mean	**66.9**	**61.6**	**62.1**	**56.3**	**50.9**	**49.1**

TABLE VI. SINGLE-CLASS INTRA-PATIENT PERFORMANCE IN % USING RANDOM AND SMOTE OVERSAMPLING. THE PERSONALIZED PATIENT-SPECIFIC CLASSIFIERS WERE TRAINED WITH TWO NIGHTS OF TRAINING DATA EACH AND VALIDATED ON THE REMAINING NIGHT. P… PATIENT ID, BED… ISABEL I AND ISABEL II, MVMT… BED MOVEMENT DURING THE NIGHT

P	*Bed*	*Mvmt*	*Random Oversampling*			*Smote Oversampling (2nd row including own features)*		
			Precision	*Recall*	*F1-Score*	*Precision*	*Recall*	*F1-Score*
1	II		78.1	76.6	76.5	71.1 78.9	71.1 79.1	70.9 79.1
	I	x	15.3	39.0	22.0	65.9 66.5	43.7 44.1	32.9 33.5
	II	x	67.7	63.8	64.1	69.0 70.1	64.6 66.0	64.8 66.2
2	II		80.0	69.7	74.1	81.2 81.1	72.8 73.3	76.4 76.7
	I	x	59.2	72.2	61.8	71.0 69.6	73.9 73.1	65.6 62.0
	II	x	75.9	75.8	75.8	75.0 76.3	66.9 74.2	70.6 75.2
3	II		74.7	75.4	75.1	77.9 79.7	77.6 79.1	77.7 79.4
	I	x	65.8	61.2	62.8	72.4 58.2	71.2 60.9	71.7 59.4
	II	x	65.4	34.7	28.5	60.5 70.7	33.6 35.6	28.0 28.9
4	II		60.5	55.2	56.6	61.8 62.9	50.7 58.1	51.7 59.5
	II	x	63.6	60.8	61.5	59.2 64.4	53.3 61.8	53.7 62.4
	I	x	8.4	29.1	13.1	8.5 8.4	29.1 29.1	13.1 13.1
5	II		55.1	55.5	55.3	57.8 56.0	58.5 56.5	58.1 56.2
	II	x	74.8	66.1	66.5	66.2 73.3	64.3 65.4	64.9 66.0
	I	x	97.0	95.4	96.1	98.4 98.4	98.4 98.5	98.1 98.4
6	II		64.8	50.0	45.1	64.8 64.3	44.5 45.7	44.1 46.0
	II	x	43.8	50.0	45.1	48.0 44.9	54.9 50.7	47.7 46.0
	I	x	67.4	53.3	52.5	70.2 65.7	53.2 49.5	51.6 47.7
		Mean	**62.1**	**60.2**	**57.4**	**65.5** **66.1**	**60.1** **61.2**	**60.5** **61.3**

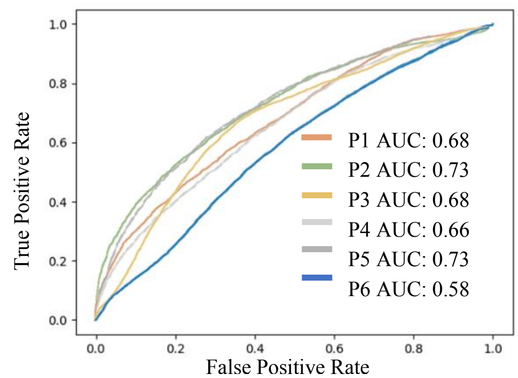

Figure 5. Obtained ROC Curve for the binary classification when trained on inter-patient data and validated using leave-one-participant out.

Multi-Class Position Classification

Results for the patient-specific (Table VII) and the inter-patient (Table VI) multi-class (supine, lateral left, lateral right) classifiers are reported. Classes are highly imbalanced, c.f. Table II. Windows labelled "Unknown", "Movement", or "Upright" were removed from the dataset prior to training. Since several participants did not sleep in prone position during the majority of nights, this class was removed. Classifiers that were trained on data from different patients, showed a mean F1-score of 0.62 and 0.49 for binary and multi-class classification, respectively.

TABLE VII. MULTI-CLASS INTRA-PATIENT PERFORMANCE IN %. THE PATIENT-SPECIFIC CLASSIFIERS WERE TRAINED WITH TWO NIGHTS OF TRAINING DATA EACH AND WERE VALIDATED ON THE THIRD NIGHT.

P	*Bed*	*Mvmt*	*Features from Literature*			*Features from Literature + own Features*		
			Precision	*Recall*	*F1-Score*	*Precision*	*Recall*	*F1-Score*
1	II		65.4	65.4	64.3	67.4	71.2	68.6
	I	x	17.1	34.1	20.5	14.6	36.2	20.8
	II	x	60.4	63.0	59.9	55.4	55.9	52.3
2	II		51.8	48.7	42.1	37.3	45.3	35.1
	I	x	12.2	35.0	18.2	39.4	36.4	21.0
	II	x	57.5	50.5	53.2	52.2	47.6	49.3
3	II		75.9	75.8	75.8	80.9	78.8	79.5
	I	x	48.9	40.0	37.8	31.1	36.7	22.1
	II	x	75.3	35.8	26.1	77.4	34.1	20.9
4	II		41.4	34.3	30.5	36.4	32.2	26.5
	II	x	40.3	40.9	37.3	42.0	44.1	40.4
	I	x	9.7	28.4	13.1	8.4	29.1	13.1
5	II		52.0	41.0	41.0	56.0	43.9	44.5
	II	x	61.1	50.5	48.2	64.5	50.7	47.6
	I	x	94.9	97.2	96.0	96.7	97.2	96.9
6	II		51.9	41.0	41.0	57.6	42.3	39.5
	II	x	61.1	50.5	48.2	44.7	52.0	45.6
	I	x	94.9	97.2	96.0	13.2	29.5	18.0
Mean			**54.0**	**51.6**	**46.2**	**48.6**	**48.0**	**42.6**

D. Bed-Occupancy Detection

Mean (SD) Sensitivity, Specificity, and F1-score were 0.963 (0.051), 0.995 (0.007), and 0.981 (0.028), respectively. All metrics are defined on ten-second windows and the weight parameter is set to one. Due to a loss of video data in the recording of the last patient, scores only represent patients 1-5.

V. Discussion

Our work showed that sleep-position classifiers trained with data from two nights have comparable performance to classifiers trained with large inter-patient datasets. The scores are comparable to other low-complexity systems found in the literature [8]. In contrast to other studies, we used two different beds, mattresses, and sensor setups. Thereby, we were able to demonstrate transferability of the algorithms. Using only an 8-channel sensor mattress, we could show that a simple decision rule is sufficient to distinguish between in- and out-bed presence, which is required when performing closed-loop positional therapy. As sleep-related breathing disorders mainly occur in supine position, reliable detection of supine position is key. The precision, recall, and F1-score of 66.9%, 61.6%, and 62.1%, respectively, show that our model sometimes misclassified supine and non-supine. In a closed-loop setup used to treat sleep-related breathing disorders, this would result in both undertreatment and superfluous interventions. To determine whether a higher spatial resolution of the sensor mattress could improve classification accuracy, a re-run of the experiments using the same algorithms on a higher-resolution mattress would be required.

VI. Conclusions

In this work, we presented a custom-made low-resolution pressure sensor setup to detect sleep positions and bed occupancy. We could show that data from two nights is enough to obtain personalized classifiers that have comparable performance to classifiers trained with larger inter-patient datasets. Also, the classifiers were transferable between the different beds and mattresses used in our study, hence providing robustness across environments. The performances of the proposed binary- and multiclass classifiers indicate occasional misclassifications which would lead to undertreatments and superfluous interventions when transferred to a real-time closed-loop system. Finally, our presented rule-based approach can reliably detect the in-bed presence, which is an important step towards closed-loop systems performing controlled positional therapy for sleep-related breathing disorders.

Acknowledgment

We thank François Pugliese from Elite SA for his support throughout the project. In addition, we thank Peter Wolf, Michael Herold-Nadig, Lars Peter, Zsombor Kalotay, and Marco Bader from ETH Zurich for their support during the design and the development of the devices. Furthermore, we thank Sascha Steffen, Nicola Wyss, and Sebastian Küng for their help during the measurements and the six participants for taking part in the study.

References

[1] T. Young, M. Palta, J. Dempsey, J. Skatrud, S. Weber, and S. Badr, "The occurrence of sleep-disordered breathing among middle-aged adults," *N. Engl. J. Med.*, vol. 328, no. 17, pp. 1230–1235, Apr. 1993, doi: 10.1056/NEJM199304293281704.

[2] Z. F. Udwadia, A. V. Doshi, S. G. Lonkar, and C. I. Singh, "Prevalence of sleep-disordered breathing and sleep apnea in middle-aged urban Indian men," *Am. J. Respir. Crit. Care Med.*, vol. 169, no. 2, pp. 168–173, Jan. 2004, doi: 10.1164/RCCM.200302-265OC.

[3] E. O. Bixler *et al.*, "Prevalence of sleep-disordered breathing in women: effects of gender," *Am. J. Respir. Crit. Care Med.*, vol. 163, no. 3 Pt 1, pp. 608–613, 2001, doi: 10.1164/AJRCCM.163.3.9911064.

[4] H. Bearpark *et al.*, "Snoring and sleep apnea. A population study in Australian men," *Am. J. Respir. Crit. Care Med.*, vol. 151, no. 5, pp. 1459–1465, 1995, doi: 10.1164/AJRCCM.151.5.7735600.

[5] M. J. L. Ravesloot, D. White, R. Heinzer, A. Oksenberg, and J. L. Pépin, "Efficacy of the New Generation of Devices for Positional Therapy for Patients With Positional Obstructive Sleep Apnea: A Systematic Review of the Literature and Meta-Analysis," *J. Clin. Sleep Med.*, vol. 13, no. 6, pp. 813–824, 2017, doi: 10.5664/JCSM.6622.

[6] S. A. Joosten, D. M. O'Driscoll, P. J. Berger, and G. S. Hamilton, "Supine position related obstructive sleep apnea in adults: Pathogenesis and treatment," *Sleep Med. Rev.*, vol. 18, no. 1, pp. 7–17, Feb. 2014, doi: 10.1016/J.SMRV.2013.01.005.

[7] T. Grimm, M. Martinez, A. Benz, and R. Stiefelhagen, "Sleep position classification from a depth camera using Bed Aligned Maps," *Proc. - Int. Conf. Pattern Recognit.*, vol. 0, pp. 319–324, Jan. 2016, doi: 10.1109/ICPR.2016.7899653.

[8] A. M. Adami, T. L. Hayes, and M. Pavel, "Unobtrusive Monitoring of Sleep Patterns," *Annu. Int. Conf. IEEE Eng. Med. Biol. - Proc.*, vol. 2, pp. 1360–1363, 2003, doi: 10.1109/IEMBS.2003.1279555.

[9] N. Foubert, A. M. McKee, R. A. Goubran, and F. Knoefel, "Lying and sitting posture recognition and transition detection using a pressure sensor array," *MeMeA 2012 - 2012 IEEE Symp. Med. Meas. Appl. Proc.*, pp. 65–70, 2012, doi: 10.1109/MEMEA.2012.6226630.

[10] I. Starr, A. J. Rawson, H. A. Schroeder, and N. R. Joseph, "STUDIES ON THE ESTIMATION OF CARDIAC OUPTUT IN MAN, AND OF ABNORMALITIES IN CARDIAC FUNCTION, FROM THE HEART'S RECOIL AND THE BLOOD'S IMPACTS; THE BALLISTOCARDIOGRAM," *Am. J. Physiol. Content*, vol. 127, no. 1, pp. 1–28, Jul. 1939, doi: 10.1152/AJPLEGACY.1939.127.1.1.

[11] Z. Zhou *et al.*, "Single-layered ultra-soft washable smart textiles for all-around ballistocardiograph, respiration, and posture monitoring during sleep," *Biosens. Bioelectron.*, vol. 155, p. 112064, May 2020, doi: 10.1016/J.BIOS.2020.112064.

[12] X. Xu, F. Lin, A. Wang, C. Song, Y. Hu, and W. Xu, "On-bed sleep posture recognition based on body-earth mover's distance," *IEEE Biomed. Circuits Syst. Conf. Eng. Heal. Minds Able Bodies, BioCAS 2015 - Proc.*, Dec. 2015, doi: 10.1109/BIOCAS.2015.7348281.

[13] G. Matar, J. M. Lina, and G. Kaddoum, "Artificial Neural Network for in-Bed Posture Classification Using Bed-Sheet Pressure Sensors," *IEEE J. Biomed. Heal. Informatics*, vol. 24, no. 1, pp. 101–110, Jan. 2020, doi: 10.1109/JBHI.2019.2899070.

[14] M. B. Pouyan, S. Ostadabbas, M. Farshbaf, R. Yousefi, M. Nourani, and M. Pompeo, "Continuous eight-posture classification for bed-bound patients," *Proc. 2013 6th Int. Conf. Biomed. Eng. Informatics, BMEI 2013*, pp. 121–126, 2013, doi: 10.1109/BMEI.2013.6746919.

[15] A. Mineharu, N. Kuwahara, and K. Morimoto, "A study of automatic classification of sleeping position by a pressure-sensitive sensor," *undefined*, Nov. 2015, doi: 10.1109/ICIEV.2015.7334059.

[16] Q. Sun, E. Gonzalez, and Y. Sun, "On bed posture recognition with pressure sensor array system," *Proc. IEEE Sensors*, Jan. 2017, doi: 10.1109/ICSENS.2016.7808623.

[17] J. J. Liu *et al.*, "A dense pressure sensitive bedsheet design for unobtrusive sleep posture monitoring," *2013 IEEE Int. Conf. Pervasive Comput. Commun. PerCom 2013*, pp. 207–215, 2013, doi: 10.1109/PERCOM.2013.6526734.

[18] W. Huang, A. A. P. Wai, S. F. Foo, J. Biswas, C. C. Hsia, and K. Liou, "Multimodal sleeping posture classification," *Proc. - Int. Conf. Pattern Recognit.*, pp. 4336–4339, 2010, doi: 10.1109/ICPR.2010.1054.

[19] M. Liu and S. Ye, "A novel body posture recognition system on bed," *2018 IEEE 3rd Int. Conf. Signal Image Process. ICSIP 2018*, pp. 38–42, Jan. 2019, doi: 10.1109/SIPROCESS.2018.8600465.

[20] A. A. Phyo Wai *et al.*, "Lying posture classification for pressure ulcer prevention," *J. Healthc. Eng.*, vol. 1, no. 2, pp. 217–238, 2010, doi: 10.1260/2040-2295.1.2.217.

[21] J. J. Liu *et al.*, "Sleep posture analysis using a dense pressure sensitive bedsheet," *Pervasive Mob. Comput.*, vol. 10, no. PART A, pp. 34–50, Feb. 2014, doi: 10.1016/J.PMCJ.2013.10.008.

[22] A. Crivello, F. Palumbo, P. Barsocchi, D. La Rosa, F. Scarselli, and M. Bianchini, "Understanding Human Sleep Behaviour by Machine Learning," pp. 227–252, 2019, doi: 10.1007/978-3-319-95996-2_11.

[23] C. C. Hsia, K. J. Liou, A. P. W. Aung, V. Foo, W. Huang, and J. Biswas, "Analysis and comparison of sleeping posture classification methods using pressure sensitive bed system," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2009, pp. 6131–6134, 2009, doi: 10.1109/IEMBS.2009.5334694.

[24] M. W. Johns, "A new method for measuring daytime sleepiness: the Epworth sleepiness scale," *Sleep*, vol. 14, no. 6, pp. 540–545, 1991, doi: 10.1093/SLEEP/14.6.540.