2022 44th Annual International Conference of
the IEEE Engineering in Medicine & Biology Society (EMBC)
Scottish Event Campus, Glasgow, UK, July 11-15, 2022

2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) | 978-1-7281-2782-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/EMBC48229.2022.9871300

Sleep Posture Detection Using an Accelerometer Placed on the Neck

Rawan S. Abdulsadig[1], Sukhpreet Singh[1], Zaibaa Patel[1] and Esther Rodriguez-Villegas[1]

***Abstract*— Sleep position monitoring is key when attempting to address posture triggered sleep disorders. Many studies have explored sleep posture detection from a dedicated physical sensing channel exploiting optimum body locations, such as the torso; or alternatively non-contact approaches. But, little work has been done to try to detect sleep position from a body location which, whilst being suboptimal for that purpose, does however allow for better extraction of more critical biomarkers from other sensing modalities, making possible multi-modal monitoring in certain clinical applications. This work presents two different approaches, at varying levels of complexity, for detecting 4 main sleep positions (supine, prone, lateral right and lateral left) from accelerometry data obtained by a single wearable device placed on the neck. An ultra light-weight threshold-based model is presented in this work, in addition to an Extra-Trees classifier. The threshold-based model was able to achieve 95% average accuracy and 0.89 F1-score on out-of-sample data, showing that it is possible to obtain a moderately high classification performance using a simple rule-based model. The ExtraTrees classifier, on the other hand, was able to achieve 99% average accuracy and 0.99 average F1-score using only 25 base estimators with maximum depth of 20. Both models show promise in detecting sleep posture with high accuracy when collecting the signals from a neck-worn accelerometer sensor.**

I. INTRODUCTION

The severity of sleep disorders is often associated with sleep posture [1]. For example, in one study, it was shown that 55.9% of a sample of 574 adults suffering from obstructive sleep apnea (OSA) had at least twice as many apneas/hypopneas in the supine position compared to the lateral position [1]. Even worse, in the context of epilepsy, it is known that a risk factor for sudden unexpected death (SUDEP) is being at prone position during sleep, with nearly 75% of the documented SUDEP tragedies occurring in this position [2].

Many sensing modalities have been proposed to detect and monitor sleep position. A few examples include under mattress pressure sensors [3], accelerometers placed on multiple body locations [4]–[7] and a combination of under mattress pressure sensors and accelerometers [8]. But only a few studies explored using a single accelerometer for this purpose. One study investigated using an accelerometer placed on the left side of the chest [9], achieving a high level of accuracy (99.16%). However there were no occurrences of prone postures during the study. Another recent study presented a sleep monitoring mHealth app 'SleepPos' [10]. This app used the internal accelerometer of a smartphone which was then placed over the sternum using a fixation system. They reported an overall accuracy of 98.2%, but only 38.9% sensitivity for the prone position, making it inapplicable for sleep position monitoring in the context of epilepsy. Furthermore, the hardware design of their system would make it uncomfortable and unreliable for use in a daily basis.

This work presents two different methods for detecting the 4 main sleep postures, namely: Supine, Lateral Right, Lateral Left and Prone, using accelerometry data collected by a single wearable device placed on the neck. The neck was chosen as it has been previously found to be an ideal position for multi-modal unobtrusive high accuracy extraction of other important biomarkers for the purpose of SUDEP prevention [11], [12]. The first method is a threshold-based detection algorithm, where the thresholds were derived using a decision tree. The second is an Extra-Trees Classifier machine-learning algorithm. These algorithms were chosen due to their ability to segregate data into groups using interpretable conditions, which makes them easy to deploy directly on wearable monitoring devices.

II. METHODOLOGY

A. Data Acquisition

Accelerometry data was obtained using a custom PCB integrating a triaxial accelerometer (LIS2DH12, ST Electronics) with a NRF5232 microcontroller (Nordic Semiconductor) which sampled and transmitted data wirelessly via Bluetooth low energy (BLE) at 100 Hz. The PCB was housed in a 3207.23 mm^3 additive manufactured enclosure, which was then placed on the neck using a double-sided adhesive.

Signals were collected from 18 participants (12 males and 6 females) from a study approved by the Local Ethics Committee of Imperial College London (ICREC reference number: 18IC4358). The average age of those participants was 27 ± 3.2 years, and the average weight was 22.78 ± 1.87 kilograms.

All participants were directed, via verbal cues, to perform 4 main postures: supine, prone, right and left. Participants spent at least 30 seconds in each position before moving to the next one. Fig. 1 shows a short sample of the accelerometer output signal capturing the 4 main postures.

B. Data Processing

The x, y and z accelerometer signals were filtered using a first order 0.5 Hz Butterworth low-pass filter in order to eliminate noise induced by cardiac activity and breathing motion. After the signals were filtered, a reference point was established for each subject independently at the first supine position occurrence to be considered as the relative

[1]Wearable Technologies Lab, Department of Electrical and Electronic Engineering, Imperial College London, UK

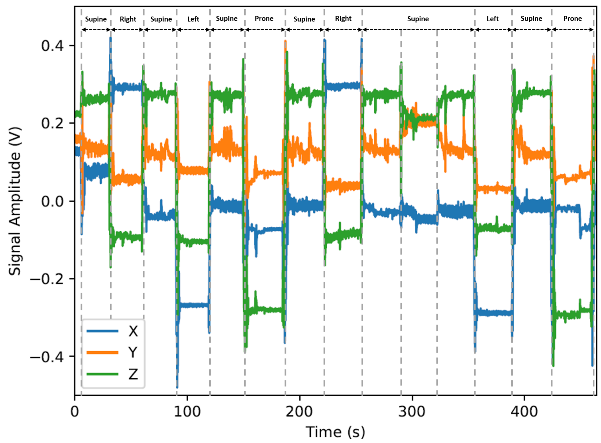

Fig. 1. Example of the raw accelerometer output signal obtained from the neck wearable and their respective annotation.

starting point, based on which the angle of acceleration was calculated. This reference point was obtained by taking the mean of the x, y and z values of a 2-second window in the middle of that first supine posture.

Once the reference point was set, the angles in the XY, XZ and YZ planes at any point were calculated using the following equation:

$$\theta_{\alpha\beta} = \arccos\frac{Ref_\alpha \times Ref_\beta}{|Ref_\alpha|\times|Ref_\beta|} - \arccos\frac{P_\alpha \times P_\beta}{|P_\alpha|\times|P_\beta|}, \quad (1)$$

where, $\alpha \in$ {X,Y}, $\beta \in$ {Y,Z}, $\alpha \neq \beta$, Ref refers to the reference point and P is the point at which the angle is calculated.

The reference point's x, y and z values were also subtracted from the values of the whole run, for each subject independently, in order to eliminate the inter-subjects differences in the orientation of their bodies and/or the orientation of the wearable which can influence the construction and the performance of the detection models.

C. Detection Models

Two main approaches were explored in this work: an ultra light-weight threshold-based method, and an ensemble-based machine-learning model, both of which utilised the decision tree algorithm.

Both approaches leveraged the time-domain information present in the mean and median values, which were found to capture most of the information needed for this type of problems [7], [13]. Those features were calculated for x, y, z, θ_{XY}, θ_{XZ} and θ_{YZ}. Transitional sections of the runs when participants moved and adjusted to the next position were eliminated. This was done by calculating the average standard deviation across the x, y and z values of a 5-second sliding window, and accepting the regions where the average standard deviation was less than 0.001.

The subjects were split into two groups, one for building the models consisting of 13 randomly-selected subjects, and the remaining subjects were used for testing the performance of that model on out-of-sample data. In the following sections of the paper, the first group will be referred to as the training-set of subjects while the second will be called the test-set.

1) Decision Tree: Threshold Search Technique: Decision Trees are well known for their ability to segregate data by finding the optimal splits across different feature spaces that result in leaves with the lowest impurity index possible [14]. The optimal splitting condition was found by searching all possible threshold values t along each of the extracted features F_i (where $i \in 1, 2, 3...N_{features}$), evaluating the level of split impurity and choosing the splitting condition resulting in the minimum split impurity value. The *Gini* index was used as the impurity criterion in this work.

The splitting process typically takes place recursively until each leaf node is pure (impurity=0) or until a regularisation condition is met, such as the maximum tree depth or the minimum samples per leaf node. Since the number of classes present in the data was 4, it was appropriate to limit the maximum tree depth to 3 levels.

2) Machine-Learning: Extra-Trees Classifier: Also known as extremely-randomised trees classifier, which is an extension of the decision tree model where an ensemble of trees work together to infer a prediction. This aggregation is often done by averaging the individual estimations of each decision tree in the ensemble in order to arrive to a unified prediction. Extensive randomisation in terms of the subset of features and the splitting conditions is key in this model to ensure low variance and high generalisation performance [15]. However, explaining this model type in depth is beyond the scope of this work.

In order to provide the model with as many data points as possible out of each run, features were extracted using a sliding window of 5 seconds and 75% overlap. Each decision tree in the ensemble was fed with an 80% bootstrapped sample of the data. The number of estimators in the ensemble, the minimum number of samples in a leaf node and the maximum tree depth are the three main hyper-parameters that were tuned using a subject-wise leave-one-out grid search cross-validation. The specified range of values searched were [25, 200] for the number of estimators , [10, 250] for the minimum number of samples in a leaf and [5, 50] for the maximum tree depth. Any cross-validation iteration consisted of a set of validation data points belonging to 1 distinct subject out of the 13 training-set subjects, while the rest of the set was used to fit the model using the selected combination of hyper-parameters for that iteration. The performance of the model at each iteration was then evaluated using F1-score. Once the optimum values of the hyper-parameters were found, the Extra-Trees classifier was fitted to the entire training-set.

III. RESULTS AND DISCUSSION

A. Threshold Tree-Search Technique

Introducing the training-set to the decision tree algorithm resulted in the formation of the tree structure shown in Fig. 2. The fitting average accuracy and F1-score achieved by this model were 96.5% and 0.93, respectively.

Three rules were established by the fitted tree to segregate the sleep positions given the features. First, it separated the supine position from the remaining positions using the Mean (z) ≤ -0.048 condition. It then identified data instances representing the prone position using the Median (z) ≤ -0.424 condition. Finally, it was able to differentiate the lateral left from lateral right using the condition Median $(\theta_{XZ}) \leq -65.383$.

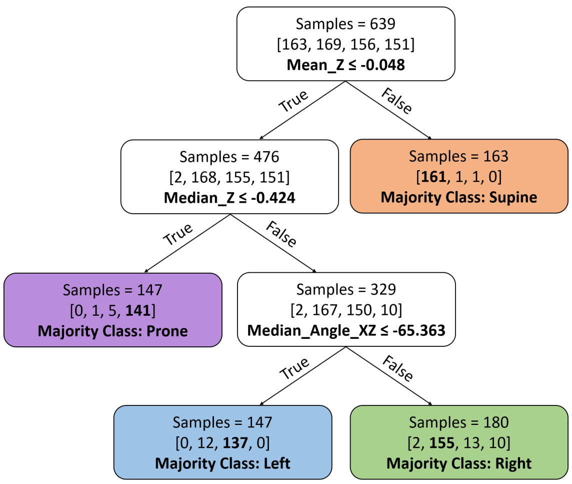

Fig. 2. A layout of the fitted decision tree, external nodes (coloured boxes) represent the resulting groups of samples while internal nodes (white branching boxes) represent the conditions. The distribution of samples at each node is shown between brackets in the order: supine, right, left and prone.

Exposing the fitted tree to the test-set in order to evaluate those rules on out-of-sample data resulted in the classification performance illustrated in Fig. 3 and summarised in Table I. The overall accuracy and F1-score achieved were 95% and 0.89, respectively.

Fig. 3. Confusion matrix presenting classification performance when evaluating the segregation thresholds on the test-set. Per-class normalised values along with the number of data points (between brackets) are provided.

The results shown in Table I illustrated that the threshold-based model was able to detect the 4 main sleep positions with 95% average accuracy despite it being made of three simple rules. The average F1-score, precision, recall (sensitivity) and specificity achieved were 0.89, 0.89 0.96 and 0.89, respectively.

TABLE I

CLASSIFICATION METRICS PER SLEEP POSITION WHEN EVALUATING THE SEGREGATION THRESHOLDS ON THE TEST-SET.

	Precision	Recall	Specificity	F1-score	Accuracy
Supine	0.90	0.90	0.97	0.90	0.95
Right	0.89	0.90	0.96	0.90	0.95
Left	0.91	0.86	0.97	0.88	0.95
Prone	0.86	0.89	0.96	0.88	0.94
Avg.	**0.89**	**0.89**	**0.96**	**0.89**	**0.95**

B. Machine-Learning: Extra-Trees Classifier

Once the leave-one-out cross-validation procedure was completed, the optimal set of values of the hyper-parameters, given the training dataset, was found. The optimum number of estimators was selected to be 25, the minimum number of leaf node samples was 40 and the maximum tree depth was 20.

Fig 4 gives an insight into the contribution of each of the features in the construction of the resulting model. This contribution was represented by the average *Gini* impurity reduction each feature provided to the decision trees of the ensemble.

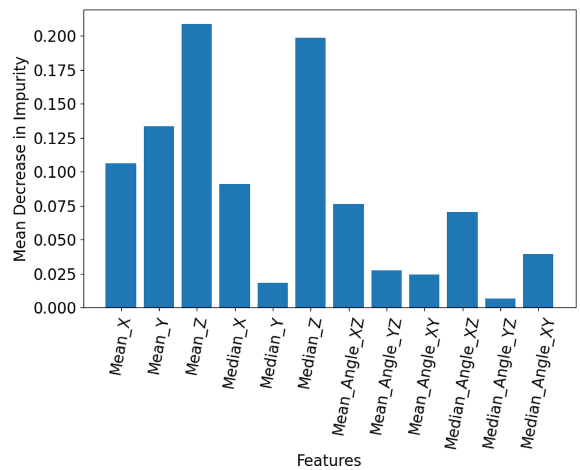

Fig. 4. *Gini* feature importance scores when fitting the Extra-Trees classifier to the training-set. The Y-axes represents the mean decrease in the impurity score granted by each feature, ranging from 0 to 0.2. The X-axis shows the 12 different features, being the mean and median values of x, y, z, θ_{XY}, θ_{XZ} and θ_{YZ}.

It is shown in Fig. 4 that the highest impurity reduction when fitting the Extra-Trees classifier was granted by the mean and the median values of z, followed by the mean and median values of x and θ_{XZ} in addition to the mean of y, all of which at scores above 0.05. The rest of the features appear to provide very little contribution towards separating the different positions given the data. Those results were also supported by Fig. 2 where the three selected thresholds were built along the mean and median values of z and the median values of θ_{XZ}.

Testing the optimised model on the test-set resulted in an average accuracy and F1-score of 99% and 0.99, respectively. Fig. 5 shows the classification performance achieved, which is also summaries in Table II.

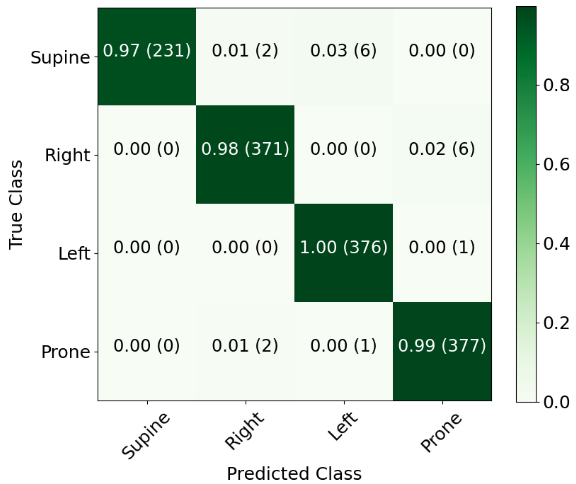

Fig. 5. Confusion matrix produced when evaluating the classifier performance on the test-set. Per-class normalised values along with the number of data points (between brackets) are provided.

TABLE II

CLASSIFICATION METRICS PER SLEEP POSITION WHEN EVALUATING THE CLASSIFIER PERFORMANCE ON THE TEST-SET.

	Precision	**Recall**	**Specificity**	**F1-score**	**Accuracy**
Supine	1.00	0.97	1.00	0.98	0.99
Right	0.99	0.98	1.00	0.99	0.99
Left	0.98	1.00	0.99	0.99	0.99
Prone	0.98	0.99	0.99	0.99	0.99
Avg.	**0.99**	**0.98**	**0.99**	**0.99**	**0.99**

The Extra-Trees classifier was able to exceed the performance of the threshold based approach as expected. Table II shows that the model was able to construct a good representation of the data allowing it to achieve high performance on the test-set with only 25 estimators of 20 maximum depth. The table also shows that the model performed almost equally well in detecting all positions including the prone position, with 0.99 sensitivity of detecting it. The achieved average accuracy, F1-score, specificity, recall (sensitivity) and precision were 99%, 0.99, 0.99, 0.98 and 0.99, respectively.

IV. CONCLUSIONS

This paper presented, for the first time, two methods for estimating sleep posture given accelerometry data obtained from a single wearable placed on the neck. The first model was an ultra light-weight threshold-based model which was able to achieve 95% average accuracy on out-of-sample data, proving that it is possible to obtain a moderately high classification performance using a simple yet well-optimised model. A relatively complex machine-learning ensemble model (Extra-Trees classifier) was also investigated. This model was able to achieve 99% average accuracy and 0.99 average F1-score using only 25 base estimators with maximum depth of 20; which further proves that a well optimised and regularised model is able to have the right bias-variance balance needed for this type of problems. Furthermore, given their light-weight quality, either models could be feasibly deployed and integrated into the hardware of a wearable system for on-line detection of sleep posture. In that case, the accuracy-complexity trade off would be the deciding factor on which model to use for the specific system with respect to the final application of concern.

ACKNOWLEDGMENT

This research was supported by the European Research Council (ERC), grant agreement no. 724334.

REFERENCES

[1] A. Oksenberg, D. S. Silverberg, E. Arons, and H. Radwan, "Positional vs nonpositional obstructive sleep apnea patients: Anthropomorphic, nocturnal polysomnographic and multiple sleep latency test data," *Chest*, vol. 112, pp. 629–639, 9 1997.

[2] J. A. Liebenthal, S. Wu, S. Rose, J. S. Ebersole, and J. X. Tao, "Association of prone position with sudden unexpected death in epilepsy," *Neurology*, vol. 84, pp. 703–709, 2 2015.

[3] M. Enayati, M. Skubic, J. M. Keller, M. Popescu, and N. Z. Farahani, "Sleep posture classification using bed sensor data and neural networks," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2018-July, pp. 461–465, 10 2018.

[4] S. Jeon, T. Park, A. Paul, Y. S. Lee, and S. H. Son, "A wearable sleep position tracking system based on dynamic state transition framework," *IEEE Access*, vol. 7, pp. 135742–135756, 2019.

[5] P. Jiang and R. Zhu, "Dual tri-axis accelerometers for monitoring physiological parameters of human body in sleep; dual tri-axis accelerometers for monitoring physiological parameters of human body in sleep," 2016.

[6] E. P. Doheny, M. M. Lowery, A. Russell, and S. Ryan, "Estimation of respiration rate and sleeping position using a wearable accelerometer," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2020-July, pp. 4668–4671, 7 2020.

[7] R. M. Kwasnicki, G. W. Cross, L. Geoghegan, Z. Zhang, P. Reilly, A. Darzi, G. Z. Yang, and R. Emery, "A lightweight sensing platform for monitoring sleep quality and posture: a simulated validation study," *European Journal of Medical Research*, vol. 23, p. 28, 5 2018.

[8] Y. Nam, Y. Kim, and J. Lee, "Sleep monitoring based on a tri-axial accelerometer and a pressure sensor," *Sensors 2016, Vol. 16, Page 750*, vol. 16, p. 750, 5 2016.

[9] H. Yoon, S. Hwang, D. Jung, S. Choi, K. Joo, J. Choi, Y. Lee, D.-U. Jeong, and K. Park, "Estimation of sleep posture using a patch-type accelerometer based device," vol. 2015-November, pp. 4942–4945, IEEE, 8 2015.

[10] I. Ferrer-Lluis, Y. Castillo-Escario, J. M. Montserrat, and R. Jané, "Enhanced monitoring of sleep position in sleep apnea patients: Smartphone triaxial accelerometry compared with video-validated position from polysomnography," *Sensors*, vol. 21, no. 11, p. 3689, 2021.

[11] E. Rodriguez-Villegas, G. Chen, J. Radcliffe, and J. Duncan, "A pilot study of a wearable apnoea detection device," *BMJ open*, vol. 4, no. 10, p. e005299, 2014.

[12] S. Singh, M. Kozłowski, I. García-López, Z. Jiang, and E. Rodriguez-Villegas, "Proof-of-concept of a novel neck-situated wearable ppg system for continuous physiological monitoring," *IEEE Transactions on Instrumentation and Measurement*, 2021.

[13] A. Abdallah, R. S. Abdulsadig, and M. Amien, "A comparative study on human loco-motor activity recognition using wearable sensors," *Proceedings of the International Conference on Computer, Control, Electrical, and Electronics Engineering 2019, ICCCEEE 2019*, 9 2019.

[14] J. Han, J. Pei, and M. Kamber, *Data mining: concepts and techniques*. Elsevier, 2011.

[15] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine learning*, vol. 63, no. 1, pp. 3–42, 2006.