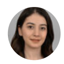
RadarSleepNet: Sleep Pose Classification via PointNet++ and 5D Radar Point Clouds

Conference Paper · December 2023

DOI: 10.1109/MAPCON58678.2023.10463831

Alessandra Fusco
Siemens Healthineers



Mervenur Akkus
Erasmus MC

Nastassia Vysotskaya
Friedrich-Alexander-University Erlangen-Nürnberg

Souvik Hazra
Infineon Technologies

RadarSleepNet: Sleep Pose Classification via PointNet++ and 5D Radar Point Clouds

Alessandra Fusco*†, Mervenur Akkus*‡, Nastassia Vysotskaya*‡, Souvik Hazra*,
Lorenzo Servadei†, Andreas Maier‡, Robert Wille†
*Infineon Technologies AG, Neubiberg, Germany
†Technical University of Munich, Munich, Germany
‡Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany

***Abstract*—Understanding sleep patterns and postures is critical for assessing overall well-being. However, traditional sleep analysis methods are often limited in their practicality due to invasive devices or complex configurations. In this study, we introduce RadarSleepNet, a non-intrusive 60 GHz Frequency-modulated Continuous Wave (FMCW) radar-based system for sleep posture monitoring that accurately infers sleep postures without compromising privacy or comfort, even in low-light conditions. Our system combines a SincNet classifier and a PointNet++ sleep pose estimation model, achieving remarkable class accuracy for each sleep posture: 98.43% for supine, 98.01% for side (chest facing radar), 97.22% for prone, and 95.72% for side (back facing radar). This demonstrates its effectiveness in accurately classifying sleep postures. This innovation offers significant potential in healthcare, providing insights into disease management and improving individual health understanding.**

***Index Terms*—FMCW radar, sleep monitoring, contactless, deep learning, SincNet, PointNet++**

I. Introduction

In the context of well-being, sleep quality extends beyond the mere duration. A comprehensive analysis of sleep patterns and positions individuals adopt during sleep can reveal much information about their overall well-being and general health. However, conventional sleep analysis methodologies do not encourage continuous monitoring, as they often rely on intrusive devices or complex setups, which prevent their broad application and ease of use. Notably, such conventional methods have been associated with patient discomfort, potentially affecting the accuracy of acquired data. In the past years, numerous approaches have been investigated to enhance sleep monitoring systems to efficiently extract patterns and derive information about overall well-being. Wearable devices and pressure-sensitive configurations, while providing fundamental information [1], alter the natural course of sleep, potentially compromising the data integrity. Similarly, vision-based systems, while being non-intrusive as they are not contact-based, still face concerns regarding privacy and dependence on ambient light [2]. In the context of continuous and contactless health monitoring, radar technology has emerged as extremely promising. Recent advancements in radar-based systems have enabled not only the monitoring of vital signs [3], [4] but also the analysis of gait [5], [6] and other applications that contribute to a global health assessment [7]. Previous research, such as [8]–[10], has explored the use of radar technology for posture classification during sleep, but only under specific conditions. The framework proposed in [9] utilizes 2D radar feature images, processed through an Inception Residual Convolutional Neural Network (CNN) architecture. Although the results appear promising, the model relies on a computationally demanding 56-layer network. Furthermore, the study lacks a detailed evaluation of the class-specific accuracy for the individual sleep postures. Similarly, the work presented in [10] exploits Doppler radar for sleep posture monitoring. The research combines analytical methods with machine learning techniques like Multinomial Logistic Regression and hybrid CNN-LSTM networks to associate radar data with sleep postures. The scope of the work is utilizing the classified sleep postures to provide a more robust heart rate estimate. Nevertheless, the experimental dataset only includes six subjects and the approach requires training a separate network for each participant, raising concerns about the generalization capabilities of the model.

Our research introduces RadarSleepNet, a pioneering approach to sleep posture classification. By integrating the capabilities of 60 GHz Frequency-Modulated Continuous Wave (FMCW) radar technology with sophisticated deep learning models, we aim to discreetly determine sleep postures. The article is structured as follows: Section II covers the basics of FMCW radar, Section III explains the RadarSleepNet deep learning architecture, Section IV presents the experimental results, and Section V provides conclusions on the effectiveness of the proposed approach.

II. Basics of FMCW Radar

FMCW radar sensors involve the transmission of frequency-modulated signals, known as chirps, which are reflected back by any object in their path. By mixing the transmitted chirp with the received signal, we can extract information about range, velocity and angle of the target. The transmitted signal can be represented as:

$$S_T(t) = e^{j2\pi\left(f_c t + \frac{K_s t^2}{2}\right)},$$

where t signifies the temporal progression during a chirp [11]. The central frequency is represented by f_c, and $K_s = B/T_s$ designates the chirp's frequency modulation rate, where B indicates the bandwidth and T_s the chirp duration. The received signal can be formulated as follows:

$$S_R(t) = \sigma \cdot e^{j2\pi\left((f_c+f_D)(t-\tau)+\frac{K_s(t-\tau)^2}{2}\right)},$$

where σ correlates with the radar cross-section, antenna gain, and range attenuation. The transit time is defined as $\tau = 2R/c$, with R being the target distance and c the speed of light. The Doppler frequency shift, $f_D = 2v_r/\lambda$, corresponds to the target's radial velocity, v_r.

After mixing the transmitted and the received signals and then applying low-pass filtering, the Intermediate Frequency (IF) signal is obtained:

$$S_{\text{IF}}(t) = S_T(t) \cdot S_R^*(t-\tau).$$

The IF signal samples can be arranged into a three-dimensional matrix known as the radar data cube, which can be processed using e.g., filtering, noise suppression, and normalization. The radar cube, denoted as $S_{\text{cube}}(t, R, f_D)$, is expressed as:

$$S_{\text{cube}}(t, R, f_D) = \sigma \cdot e^{j2\pi\left((f_c+f_D)(t-\tau)+\frac{K_s(t-\tau)^2}{2}\right)}.$$

From S_{cube}, we can determine the range profile using the Fourier transform across the samples per chirp:

$$R_{\text{profile}}(R, t) = \int S_{\text{cube}}(t, R, f) e^{-j2\pi f t}\, df.$$

Subsequently, we perform a Fourier transform across the chirps per frame to extract the Doppler information:

$$M_{\text{RD}}(R, f_D) = \int R_{\text{profile}}(R, t) e^{-j2\pi f_D t}\, dt.$$

The Range-Doppler map, M_{RD}, offers a two-dimensional representation, including insights into the detected targets in terms of range (distance from the radar) and Doppler shift (radial velocity towards or away from the radar). Each cell in M_{RD} represents the reflectivity strength of a target at a specific range and Doppler frequency.

III. RadarSleepNet Architecture

This section introduces RadarSleepNet, the proposed architecture for analyzing radar data to determine a user's sleep positions and movements. The architecture integrates two key components: a SincNet-based model for initial activity state classification and a PointNet++ model for detailed sleep posture identification. First, the SincNet model performs a binary classification to determine whether the subject is static or moving. If the subject is stationary, the system triggers the PointNet++ model to further process the radar data and accurately classify the user's sleep position by generating a compact spatial representation.

A. SincNet-based Classifier

The analysis of radar data for motion detection and sleep pose estimation requires robust and efficient pre-processing as well as neural network architecture. As the first step, a Fast Fourier Transform (FFT) is applied to the raw radar data to isolate the range bin with maximum intensity. In this way, we restrict the scope of the sleep monitoring system to a single subject, i.e., the closest person in bed. The range bin data that has been isolated is now given as input to the motion classifier model.

SincNet [12] distinguishes itself from traditional Convolutional Neural Networks (CNNs) by employing parameterized sinc functions for its convolutional layers. This approach simplifies the model by reducing the number of parameters to be learned. The model uses gradient optimization techniques for fine-tuning these parameters, precisely the cutoff frequencies of the bandpass filter. It also incorporates a Hamming window to enhance frequency selectivity. Due to its efficient and robust architecture, SincNet is well-suited for practical applications, including motion detection.

The architecture of the proposed model is structured into four main blocks and a prediction layer. The first block is characterized by a 1D sinc-based convolution followed by 1D max-pooling with a window size of 5. Afterward, a standard CNN pipeline is applied in the second and third blocks. Each block contains a 1D convolution layer with three filters of size 15, combined with a 1D max-pooling layer and a dropout rate of 0.5. In addition, the third block includes a batch normalization and flatten layer. The fourth block comprises a dense layer with 128 units and a LeakyReLU activation. Skip connections are also utilized to facilitate learning. As the last step, the prediction layer is implemented: it includes a dense layer with a sigmoid activation function, making it suitable for a binary classification task, such as the distinction between static and non-static.

B. Sleep Position Model

To derive a compact spatial representation of the sleeping user, the radar data undergoes additional preprocessing steps, up to extracting a 5D point cloud. The radar data is integrated over multiple frames to improve the Signal-to-Noise Ratio (SNR). Then, the Moving Target Indicator (MTI) is applied through previous frame subtraction. This step helps to focus on moving targets by reducing the effect of stationary objects and background noise. Afterwards, the Fast Fourier Transform (FFT) is applied across both the sample and chirp axes. This creates the Range Doppler Image (RDI), which is essential for target detection and range estimate. Once the RDI is generated, CFAR (Constant False Alarm Rate) detection is performed to identify targets of interest within the RDI. The algorithm used for CFAR is the OS-CFAR (Ordered Statistic CFAR), and it works by identifying regions that are likely to contain our target by dynamically adjusting the detection threshold. Post CFAR detection, the DOA (Direction of Arrival) estimation is carried out using the Capon algorithm, which generates a Range-Angle Image (RAI) covering azimuth and elevation angle from -40 to +40 degrees. Finally, the processed data are transformed into a point cloud, including five features: x, y, z coordinates, intensity, and Doppler velocity for each point.

The data points now align with spherical radar coordinates $(r; \theta_a; \theta_e)$, corresponding to the range, azimuth, and elevation angle of the target. Ultimately, the transformation below finalizes the transition to standardized coordinates:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_t & \sin\theta_t \\ 0 & -\sin\theta_t & \cos\theta_t \end{bmatrix} \begin{bmatrix} r\cos\theta_e \sin\theta_a \\ r\cos\theta_e \cos\theta_a \\ r\sin\theta_e \end{bmatrix} + \begin{bmatrix} x_r \\ y_r \\ h \end{bmatrix}$$

where the computed $(x; y; z)$ represent the output Cartesian coordinates. θ_t denotes the tilt angle with respect to the horizontal plane, and $(x_r; y_r; h)$ correspond to the radar's Cartesian ground coordinates. The transformed 5D point cloud is used for the classification task.

TABLE I
POINTNET++ MODEL ARCHITECTURE SUMMARY.

Layer (Type)	Output Size	Description
Input	(B, N, 5)	Input layer with 5 features per frame, i.e., our 5D point cloud.
SA-MSG	(B, 64, N, 128)	MSG layer with 64 groups and 128 points per group.
SA-MSG	(B, 32, N, 256)	MSG layer with 32 groups and 256 points per group.
SA	(B, 1024)	PointNet abstraction layer with 1024 points.
Linear	(B, 256)	Fully connected layer with 256 output units.
BatchNorm1D	(B, 256)	Batch norm. layer for 256 feature maps.
Dropout	(B, 256)	Dropout layer with 50% dropout rate.
Linear	(B, 64)	Fully connected layer with 64 output units.
BatchNorm1D	(B, 64)	Batch norm. layer for 64 feature maps.
Dropout	(B, 64)	Dropout layer with 50% dropout rate.
Linear	(B, K)	Output fully connected layer with K output.

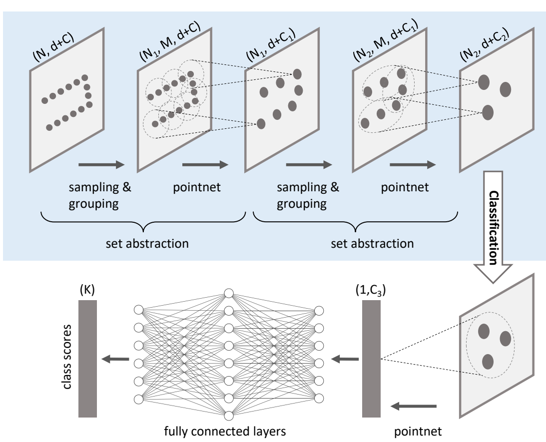


Fig. 1. PointNet++ architecture for sleep pose classification task.

We have selected a PointNet++ architecture [13] to classify the processed 5D point cloud into distinct sleep postures. PointNet++ is an extension of the original PointNet [14], designed to process point cloud data with varying resolutions and densities. Next, we will elaborate on the architecture and specifications of the model. Table I provides a detailed breakdown of the model architecture. We refer to B as the batch size and N as the number of input points. The total number of features is five, given by the sum of two quantities d and C in Figure 1: the spatial coordinates (d = 3), namely x, y, and z, and the additional features (C = 2), namely intensity and Doppler velocity.

TABLE II
OPERATING PARAMETERS OF THE RADAR

Parameters	Value	Units
Ramp start frequency, $f_{\min}$	57.5	GHz
Ramp stop frequency, $f_{\max}$	58.5	GHz
Frame rate, f	20	fps
Number of samples per chirp, N_{TS}	64	-
Sampling frequency, f_s	2	MHz
Chirp time, T_c	64	μs
Number of chirps, P_N	128	-
Number of Tx antennas, N_{Tx}	1	-
Number of Rx antennas, N_{Rx}	3	-

The hierarchical architecture includes Set Abstraction (SA) and Multi-Scale Grouping (MSG). The goal is to capture local features by grouping the M nearest neighbors of each point into local regions at varying scales using SA-MSG layers. The Farthest Point Sampling (FPS) algorithm is implemented within the SA layer to select key points, which are then processed to produce a more comprehensive set of features. These features are transformed and combined using fully connected layers to capture underlying patterns. Batch Normalization is applied to standardize the features and Dropout to prevent over-fitting. The final output layer has K units, representing the number of unique sleep positions targeted for classification. The model is trained using a batch size of 64, a learning rate of 0.001, and spans 150 epochs. Loss minimization is conducted via the Adam optimizer, with an adjustable learning rate every 20 epochs and a decay rate of 1e-4.

IV. EXPERIMENTAL RESULTS

This research utilizes the XENSIV$^{\mathrm{TM}}$ BGT60TR13C FMCW radar chipset from Infineon Technologies to capture data for the RadarSleepNet

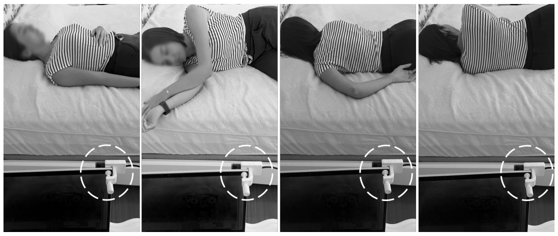


Fig. 2. Depictions of sleep postures in the experimental set-up. Sequentially from left to right: supine, side (chest facing the radar), prone and side (back oriented toward the radar) positions. In each image, the radar sensor is highlighted with a white dashed circle.

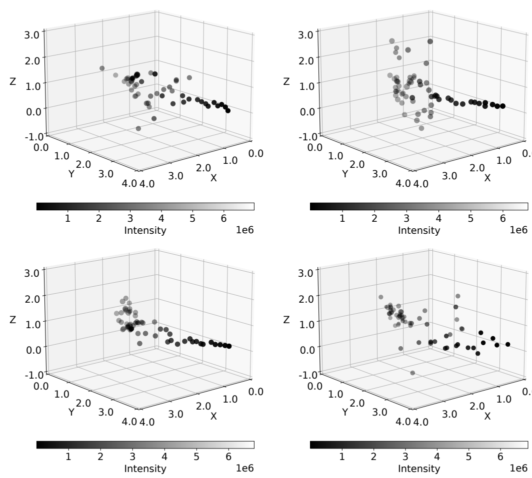


Fig. 3. Exemplary representations of 5D point cloud for different sleep postures arranged in a 2x2 grid. The first row, from left to right, shows the supine and side (chest facing the radar) positions. The second row displays the prone and side (back oriented toward the radar) positions.

model, which is designed to classify sleep poses into one of four categories: supine, prone, side (facing the radar), and side (back to the radar). The key operating parameters of the radar are summarized in Table II.

A total of 32 subjects participated in the experimental protocol, designed to collect a comprehensive dataset for model training and testing. Participants were instructed to lie on a bed and adopt each of the four target sleep postures, visually detailed in Figure 2, while data was recorded during both motionless and movement periods.

TABLE III
NORMALIZED CONFUSION MATRIX OF SINCNET MODEL

Non - Static	97.79%	2.21%
Static	3.35%	96.65%

TABLE IV
NORMALIZED CONFUSION MATRIX OF POINTNET++ MODEL

Supine	98.43%	0.82%	0.41%	0.34%
Side (chest)	0.83%	98.01%	0.50%	0.66%
Prone	1.42%	0.75%	97.22%	0.62%
Side (back)	1.11%	1.73%	1.44%	95.72%

The study employed a total of 1.512.000 frames (corresponding to 1.260 minutes of recordings for a frame repetition rate of 20 fps). This ensures an exhaustive dataset containing a balanced number of recordings for each of the four investigated sleep positions. Figure 3 provides exemplary 5D point cloud representations for the sleep postures. In these plots, the spatial coordinates x, y, and z are presented along the respective axes, the magnitude is illustrated through a grayscale intensity map, and the velocity is conveyed by the size of the points.

In summary, the proposed RadarSleepNet model utilizes a SincNet classifier to discern whether the target was static or non-static, and a PointNet++ model for sleep pose estimation. During the SincNet model's training phase, the cross-entropy loss decreased from 0.8 to 0.3 across 150 epochs, suggesting effective feature learning. Table III showcases the confusion matrix for the motion classification model based on SincNet architecture. The high performance of the SincNet model, with a class accuracy of 97.79% for moving targets and 96.65% for static targets, indicates its suitability for extracting frequency-related features from radar data for motion determination. In addition, the model shows 97% precision and recall values, while its F1-score is 97%. Table IV reports the confusion matrix for the sleep postures classification model, i.e., the PointNet++. The matrix reveals remarkable class accuracy for each sleep posture: 98.43% for supine, 98.01% for side (chest facing radar), 97.22% for prone, and 95.72% for side (back facing radar). The results highlight the efficacy of the proposed learning approach in identifying various positions during sleep. In addition to the class accuracy, the model's precision and recall values are 97.34% and 97.36%, respectively, and the F1 score 97%. The high accuracy of the proposed approach confirms high robustness and ability to generalize effectively to new data, making it highly suitable for real-world applications involving frequency-related feature extraction from radar data for motion detection and sleep pose estimation.

V. Conclusions

This work introduces RadarSleepNet, a pioneer approach to detect motion and identify sleep positions based on 60 GHz FMCW radar sensor. The experimental results, based on a comprehensive dataset including 32 subjects and 21 hours of recordings, demonstrate high accuracy and robustness. The proposed approach revealed to be effective in extracting meaningful features from radar data for potential applications in healthcare, where monitoring sleep quality and detecting sleep-related disorders is critical.

In the future, the models will be fine-tuned using an even more heterogeneous dataset collected from real-world environments like sleep clinics. Furthermore, the framework will be expanded by incorporating radar-based breathing rate estimate [15] and apnea detection [16], but also auxiliary sensors for the detection of snoring and coughing [17]. Such integrations open the door to a multi-faceted approach to sleep monitoring, offering a holistic view of an individual's sleep health and potentially uncovering health-related issues.

References

[1] Abdulsadig, R.S. and Rodriguez-Villegas, E., 2023. Sleep Posture Monitoring Using a Single Neck-Situated Accelerometer: A Proof-of-Concept. IEEE Access, 11, pp.17693-17706.

[2] Tam, A.Y.C., So, B.P.H., Chan, T.T.C., Cheung, A.K.Y., Wong, D.W.C. and Cheung, J.C.W., 2021. A blanket accommodative sleep posture classification system using an infrared depth camera: A deep learning approach with synthetic augmentation of blanket conditions. Sensors, 21(16), p.5553.

[3] Hazra, S., Fusco, A., Kiprit, G.N., Stadelmayer, T., Habib, S., Servadei, L., Wille, R., Weigel, R. and Santra, A., 2023. Robust Radar-based Vital Sensing With Adaptive Sinc Filtering and Random Body Motion Rejections. IEEE Sensors Letters.

[4] Alizadeh, M., Shaker, G., De Almeida, J.C.M., Morita, P.P. and Safavi-Naeini, S., 2019. Remote monitoring of human vital signs using mm-wave FMCW radar. IEEE Access, 7, pp.54958-54968.

[5] Niazi, U., Hazra, S., Santra, A. and Weigel, R., 2021, May. Radar-based efficient gait classification using Gaussian prototypical networks. In 2021 IEEE Radar Conference (RadarConf21) (pp. 1-5). IEEE.

[6] Addabbo, P., Bernardi, M.L., Biondi, F., Cimitile, M., Clemente, C. and Orlando, D., 2020, June. Gait recognition using FMCW radar and temporal convolutional deep neural networks. In 2020 IEEE 7th International Workshop on Metrology for AeroSpace (MetroAeroSpace) (pp. 171-175). IEEE.

[7] Abedi, H., Ansariyan, A., Lehman, C., Morita, P.P., Boger, J., Wong, A. and Shaker, G., 2022, October. Non-Visual and Contactless Wellness Monitoring for Long Term Care Facilities Using mm-Wave Radar Sensors. In 2022 IEEE Sensors (pp. 1-4). IEEE.

[8] Yue, S., Yang, Y., Wang, H., Rahul, H. and Katabi, D., 2020. BodyCompass: Monitoring sleep posture with wireless signals. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 4(2), pp.1-25.

[9] Zhou, T., Xia, Z., Wang, X. and Xu, F., 2021, September. Human sleep posture recognition based on millimeter-wave radar. In 2021 Signal Processing Symposium (SPSympo) (pp. 316-321). IEEE.

[10] Higashi, K., Sun, G. and Ishibashi, K., 2019, July. Precise heart rate measurement using non-contact Doppler radar assisted by machine-learning-based sleep posture estimation. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 788-791). IEEE.

[11] Jankiraman, M., 2018. FMCW radar design. Artech House.

[12] Mirco Ravanelli, *SincNet*, 2018, PyTorch Implementation, https://github.com/mravanelli/SincNet.

[13] Qi, C.R., Yi, L., Su, H. and Guibas, L.J., 2017. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. Advances in neural information processing systems.

[14] Qi, C.R., Su, H., Mo, K. and Guibas, L.J., 2017. Pointnet: Deep learning on point sets for 3d classification and segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 652-660).

[15] Choi, H.I., Song, W.J., Song, H. and Shin, H.C., 2021. Selecting target range with accurate vital sign using spatial phase coherency of FMCW radar. Applied Sciences, 11(10), p.4514.

[16] Zhuang, Z., Wang, F., Yang, X., Zhang, L., Fu, C.H., Xu, J., Li, C. and Hong, H., 2022. Accurate contactless sleep apnea detection framework with signal processing and machine learning methods. Methods, 205, pp.167-178.

[17] Dafna, E., Tarasiuk, A. and Zigel, Y., 2013. Automatic detection of whole night snoring events using non-contact microphone. PloS one, 8(12), p.e84139.