Convolutional Neural Network Model in Human Motion Detection Based on FMCW Radar Signals

Lazar Jugović

University of Belgrade – School of Electrical Engineering and Novelic d.o.o.
Belgrade, Serbia
lazar.jugovic@novelic.com
[0009-0003-3661-8996]

Ivan Vajs

University of Belgrade – School of Electrical Engineering and Innovation Center of the School of Electrical Engineering in Belgrade
Belgrade, Serbia
ivan.vajs@ic.etf.bg.ac.rs
[0000-0001-7039-5384]

Milica Badža Atanasijević

University of Belgrade – School of Electrical Engineering and Innovation Center of the School of Electrical Engineering in Belgrade
Belgrade, Serbia
milica.badza@ic.etf.bg.ac.rs
[0000-0002-5856-2626[

Milan Stojanović

University of Belgrade – School of Electrical Engineering and Novelic d.o.o.
Belgrade, Serbia
milan.stojanovic@novelic.com
[0000-0002-1783-1615]

Milica M. Janković

University of Belgrade – School of Electrical Engineering
Belgrade, Serbia
piperski@etf.bg.ac.rs
[0000-0002-7506-4995]

Abstract—The detection of body movements is the essential step for sleep quality analysis. Contactless approaches for sleep motion recognition are unobtrusive and are easier to use in comparison to wearable technologies. In this paper, two contactless sensors based on Frequency-Modulated Continuous Wave (FMCW) radar technology were positioned on the side of, and underneath the bed on which the participant was lying. FMCW data from 10 participants were acquired during the experiment scenario that included the following three states: resting state, movement, and cough. Magnitude-phase coherency method was applied to FMCW data for finding optimal phase signals. Finally, a one-dimensional convolutional neural network was used for the classification based on optimal phase signals. The best classification results were obtained using only FMCW data from the radar positioned underneath the bed: 72% accuracy for differentiating between the resting state, movement, and cough class, and 89% accuracy for the resting state and movement class.

Keywords - frequency-modulated continuous wave radar, movement, cough, magnitude-phase coherency, classification, convolutional neural network

I. INTRODUCTION

Sleep quality can be measured using subjective methods (retrospective questionnaires and sleep diaries) and objective methods based on physiological and behavioral parameters [1]. Polysomnography (PSG) is the gold standard for sleep disturbance detection and is widely used in sleep clinics. Full PSG includes a multimodal measurement of electroencephalography (EEG), electrooculography (EOG), chin electromyography (EMG), electrocardiography (ECG), oronasal airflow, pulse oximetry, thoracic and abdominal movements and body position. The main disadvantage of the PSG approach is its complexity which makes it impossible for home usage without additional medical staff. For that reason, the research and business community invest efforts to offer a wide range of easy-to-use technological solutions suitable for home usage that combine machine learning approaches and that are validated using the standard PSG. These solutions are based on heart rate and heart rate variability features, as well as on actigraphy (body movement classification), on respiratory rate and respiratory rate variability features and brainwave features (distinction of low and high brain frequencies) [1]. Most of the commercially available sleep quality detectors for home usage are based on wearable technology. Recently, the advancement of microwave technology enabled non-contact approaches for monitoring vital signals (heart rate and breathing rate) and body motions. This moves the limits of technical solutions for sleep quality monitoring from wearable (contact) to contactless approaches. Several published solutions for sleeping scenarios use microwave technology based on Doppler radar technology [2-5], Ultra-Wide band (UWB) radar technology [6] or frequency-modulated continuous wave (FMCW) radar technology [7].

The FMCW radar is a device that generates electromagnetic signal that has linearly increasing frequency and transmits it into the propagation space via transmitting antennas. Also, the FMCW system receives reflected signals from objects via receiving antennas and extracts information about the reflective objects. FMCW detection of motion includes estimating the magnitude and phase signals by spectral decomposition of received signals as well as choosing the optimal range bin [8]. In the literature, the optimal range bin was chosen from either magnitude or phase signals. Munoz-Ferreras et al. [9] found the optimal range bin from the maximum averaged magnitude, considering the power of reflected signal from target, while Alizadeh et al. [10] selected the optimal range bin considering the highest variation of phase. Furthermore, traditional machine learning and deep learning approaches have proven to be effective tools for detection and classification of

human motions from decomposed radar data [11-13].

In this paper, a convolutional neural network (CNN) approach was used for the movement recognition (differentiating the resting state, movement, and cough classes) based on FMCW phase data of ten lying down participants, while magnitude-phase coherency (MPC) algorithm was used for the selection of the optimal range bin.

II.METHOD

A. Theory FMCW background

FMCW radar is a device that works by frequency modulating the transmission signal. One FMCW radar contains transmitting and receiving systems. The main aim of FMCW radars is obtaining the high-frequently signal with a current frequency given by:

$$f(t) = f_c + \frac{BW}{T_{ch}} t = f_c + St \quad (1)$$

where fc represents the starting frequency for the linear modulated signal, BW bandwidth of the signal (difference between maximum and minimum frequency), Tch the period of the chirp signal, while S represents the slope of the ramp, i.e. division of bandwidth and period of the chirp signal [14].

The current phase of the transmitting signal is obtained by integrating the current angular frequency:

$$\phi(t) = \int 2\pi f(t)dt = 2\pi f_c t + \pi S t^2 \quad (2)$$

From previous relations, the final form of the transmitting signal $v_{tx}(t)$, TX in Fig. 1, is:

$$v_{TX}(t) = A_1 \cos(\phi(t)) = A_1 \cos(2\pi f_c t + \pi S t^2) \quad (3)$$

Fig. 1 illustrates that the received signal $v_{rx}(t)$, RX in Fig. 1, is a delayed version of the transmitting signal $v_{tx}(t)$and its final form is:

$$v_{RX}(t) = A_2 \cos(\phi(t-\tau)) = A_2 \cos(2\pi f_c(t-\tau) + \pi S(t-\tau)^2) \quad (4)$$

where τ is $2\frac{d}{c}$ and it represents delay time of the receiving signal vrx(t), relative to the transmitting signal vtx(t) where d is the distance from the radar to the observed target and c is the speed of light.

Using the Fourier Fast Transform, magnitude signal M(t,rk) and phase signal P(t,rk) are calculated as:

$$M(t, r_k) = 2\left|\sum_{n=0}^{N-1} v_{RX}(t) e^{-j\frac{2\pi kn}{N}}\right| \quad (5)$$

$$P(t, r_k) = 2\tan^{-1}\frac{imag(\sum_{n=0}^{N-1} v_{RX}(t) e^{-j\frac{2\pi kn}{N}})}{real(\sum_{n=0}^{N-1} v_{RX}(t) e^{-j\frac{2\pi kn}{N}})} \quad (6)$$

where $rk = \frac{c}{2BW} k$, for k = 0, ..., N-1, N is the number of observed frames. [14]

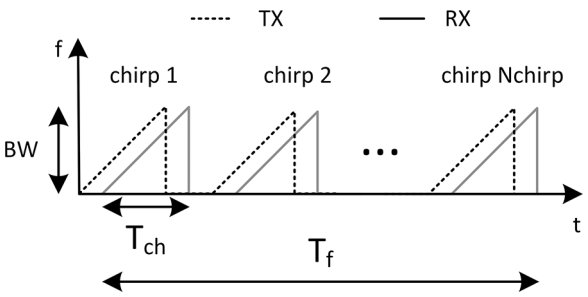

Fig. 1. One frame with duration Tf consisted of Nchirp chirps (Tch – duration of chirp signal)

B. Experiment setup and hardware description

During the experiment, two FMCW AWR1843 radars were used for data acquisition, including a board for evaluation AWR1843BOOST and a board for data collecting DCA1000EVM (Texas Instruments, USA). Number of receiving antennas, N_{rx} was 4 and number of transmitting antennas N_{tx} was 3.

The position of the radars was on the side of, and underneath the participant. Radar 1 was placed underneath the bed, directly under the chest of the participant. On the side, on the designed holder, Radar 2 was placed on the table at a height of 85 cm and 90 cm from the participants' chests, Fig. 2. The distance was measured by a laser rangefinder.

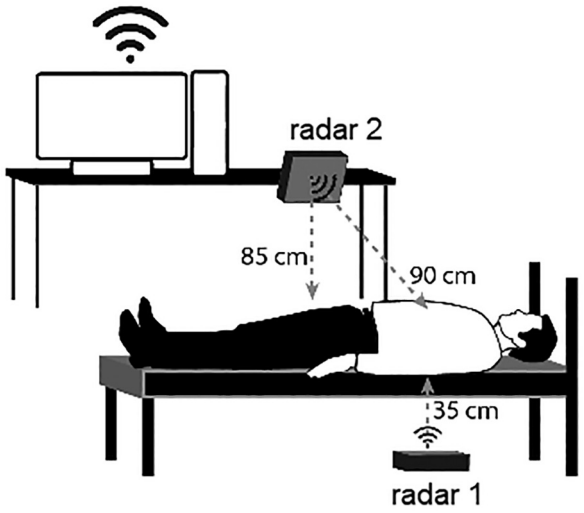

Fig. 2. Experiment setup with two FMCW radars, computer and subject in the lying position.

The configuration of both FMCW radars were set to the following values:

- Started Frequency of Chirp signal, f_c – 79Hz
- Idle Time, IT – 40 μs
- ADC sampling time, TADC – 8 μs
- Duration of Chirp signal, Tch – 75 μs
- Slope of Chirp signal, S – 30 GHz/μs
- Number of Samples per Chirp signal, N_{ADC} – 128

- The sampling frequency, f_s – 2000 kHz
- Number of Chirp signals per frame, N_{chirp} – 32
- Duration of one frame, T_f – 20 ms.

Data acquisition was performed by the mmWaveStudio software (Texas Instruments, USA).

The experiment was performed on ten healthy subjects: ages 26.7±5.23, 4 males and 6 females. They have signed the informed consent to participate in the study.

At the beginning of the experiment, the participant was lying on the back with the hands next to the body. Next, they were following the protocol of the experiment from the prerecorded audio file, when to turn over, and when to cough. The experiment had 5 phases, and the whole experiment lasts 14 minutes, Table I.

Table I. Experiment protocol

Phase No.	Phase description		
	Position	Breathing category	Duration [s]
1	Back	Normal	180
2	Right side	Normal	120
3	Stomach	Normal	120
4	Left Side	Normal	120
5	Back	Cough	5*
	Back	Normal	25*
Total duration			840 (14 min)

*This phase was repeated 10 times.

C. Algoritam description

Data was archived in .bin files with maximum file size of 1 GB. The total data SIZE per file was:

$$SIZE = N_{tx}N_{rx}N_{frames}N_{chirp}N_{ADC}4B \quad (7)$$

where Ntx is the number of transmitting antennas, N_{rx} the number of receiving antennas, N_{frames} the number of frames in one file, N_{ADC} the number of samples per chirp signal and 4B comes from the size of one sent or received data (2B for the real and 2B for the imaginary part).

Data were loaded and processed frame by frame in a 4D structure with dimensions N_{adc} x N_{chirp} x N_{rx} x N_{tx} =128x32x4x3.

The implemented algorithm includes the following steps, (Fig. 3):

- Fourier Fast Transform (1D-FFT, N_{fft}=128 points, Blackman windowing) of received signals. 1D-FFT coefficients are added to the buffer. The buffer size of 600 x N_{fft} x N_{chirp} x N_{rx} x N_{tx} was selected regarding the maximum breathing period of $T_{breating\_max}$=12 s [15] – the first dimension value 600 was calculated as $T_{breathing\_max}$ / T_{frame}.
- limiting N_{fft} dimension to the target bin range 10cm – 3m (limiting to the part where the target can be physically found)
- average buffer signals through all chirp signals, all transmitting and receiving antennas
- magnitude extraction within the buffer, M_{buffer} – applying of Eq. 5 to the averaged buffer signals
- phase extraction within the buffer, P_{buffer} – applying of Eq. 6 to the averaged buffer signals
- MPC algorithm for the optimal range bin detection (within the buffered data)(Section II C)
- creating the total phase signal, Ptotal of one .bin file
- total phase signal unwrapping of one .bin file, Ptotal_unwrapped, (Section II C2)
- CNN performance on $P_{total\_unwrapped}$ signals (Section II C3).

Data load, preprocessing and optimal phase detection were performed using the *Matlab* R2022a environment while neural network data preparation and experiment phase recognition were done in the *Python* environment version 3.10.11.

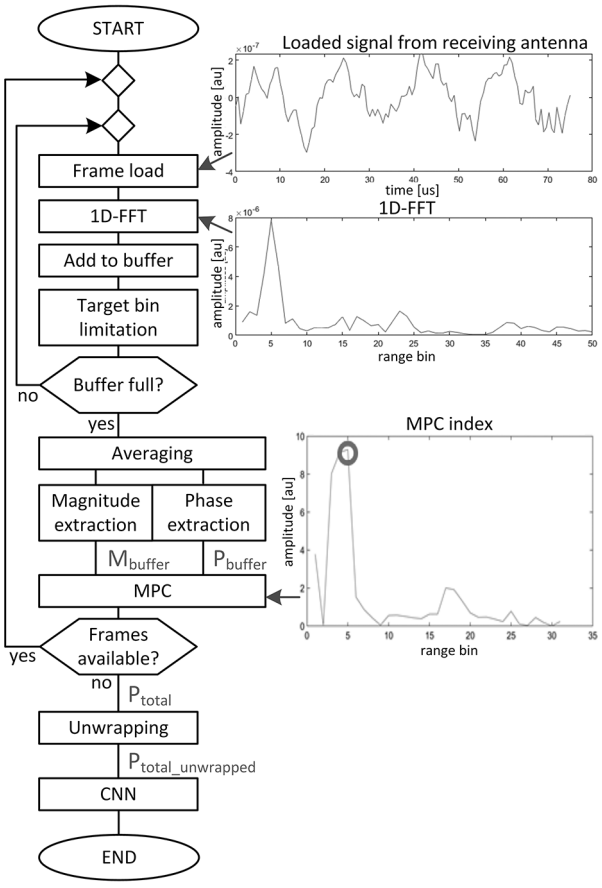

Fig. 3. Flow chart of the implemented algorithm

C1. Magnitude-phase coherency (MPC) method

An example of the magnitude signal for normal breathing and movement is presented in Fig. 4A.

It can be observed that targets existed at 25 cm to 60 cm (4-5 range bin) from the radar 1. Also, there are changes in the amplitude through bins that represent the target area because of the chest motion in those bins. The higher amplitudes represent the time when the subject inhales, so the chest is closer to the radar, and the smaller amplitudes represent the time when subject exhales, so the chest is further away from radar. Also, at the end of the figure, the amplitude is low because of the occurrence of body movement (red rectangle).

Fig. 4B, shows a magnitude signal for cough and normal breathing. A cycle of cough and normal breathing (phase 5, 5 s+25 s=30 s) is marked by a red rectangle. Higher amplitudes of signal represent cough, whereas smaller amplitudes with oscillations represent normal breathing.

MPC method uses magnitude and phase signals and considers that they are highly correlated and that the coherence between them is going to be high. Two signals are coherent if they have equal frequencies, polarizations and a constant phase ratio.

To describe the mentioned coherence, the concept of MPC index was used:

$$MPC(t,r) = \frac{\left|\sum_{s=t-t0}^{t} M(s,r)P(s,r)\right|}{\sigma_M(r)\sigma_P(r)} \quad (8)$$

where M(s,r) represents the magnitude signal, P(s,r) is the phase signal, from which the mean value is subtracted, σM(r) is the standard deviation of the magnitude signal through all frames (i.e. observed time slots), σP is the standard deviation of the phase signal through all frames. MPC(t,r) represents the obtained function – MPC index in the observed moment of time (frame) depending on the vector of range bins [16].

An example of MPC index was illustrated in Fig. 3. After forming functions of MPC indexes, it is necessary to find the position of the maximum (red circle in Fig. 3), and then declare this position for an optimal candidate for the chosen range bin. The position of the maximum in MPC indexes represents the range bin in which magnitude and phase signals are the most coherent. Observing Fig. 3, it can be obtained that the maximum MPC function is on fifth position (range bin), i.e. approximately 50 cm from the radar.

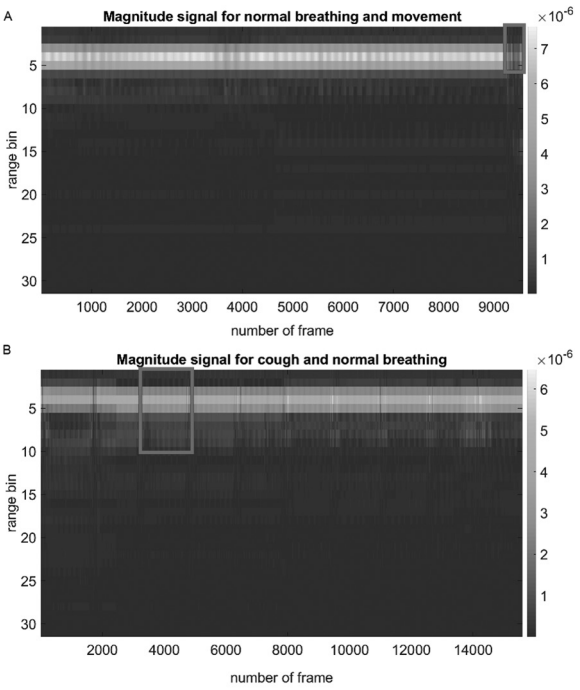

Fig. 4. A) Magnitude signal for normal breathing and movement (phase 1 of experiment protocol), B) Magnitude signal for cough and normal breathing, repeated 10 times (phase 5 of experiment protocol)

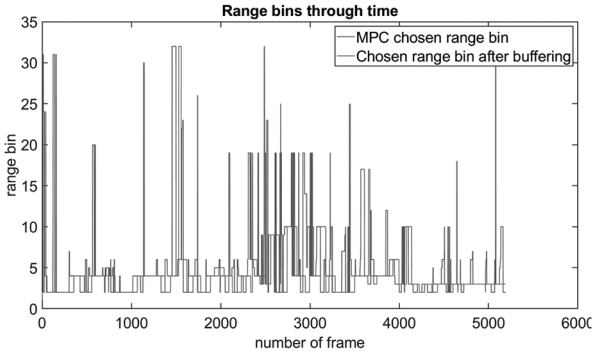

Fig. 5. An example of the range bin selection through time for normal breathing

After forming the candidate for optimal phase, it is necessary to store it in the buffer (buffer size 100). After filling the whole buffer, the most frequent range bin is found, which is going to be declared for optimal phase (i.e. range bin) in observed moment of time. Fig. 5 illustrates a graphic with obtained candidates for optimal phase through time for one .bin file, as well as the final chosen range bins after buffering which is going to be used for forming the phases.

C2. Phase unwraping

An example of selected optimal phase obtained by MPC is given in Fig. 6.

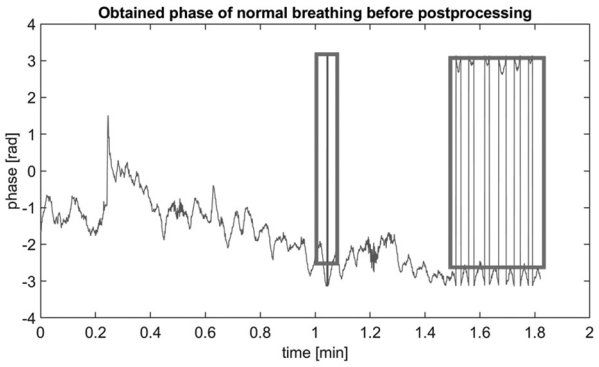

Fig. 6. An example of optimal phase signal Ptotal of normal breathing ("sudden jumps" – red rectangles)

"Sudden jumps" can be observed in the extracted optimal phase in Fig. 6 These "jumps" are a consequence of the application *matlab* function angle over magnitude signal and changes of the chosen range bin over time. It is necessary to use *matlab* function unwrap which removes "sudden jumps" from the signal (i.e. all changes in signal which are bigger than 2π). The result after the removal of "sudden jumps" is shown in Fig. 7.

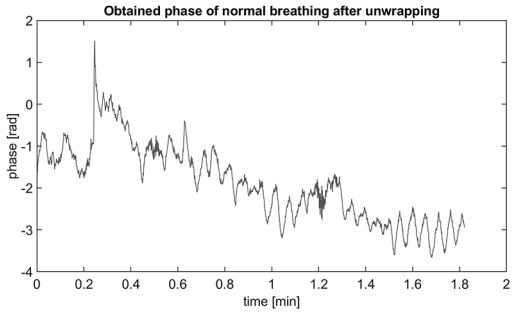

Fig. 7. Obtained phase signal of normal breathing after unwrapping, Ptotal_unwrapped

C3. CNN description

CNN was used for movement and cough detection with two evaluation scenarios. The first scenario included the classification of three states: movement, coughing and normal breathing state (three class recognition) and the second scenario included the differentiation between movement and the normal breathing state (2 classes recognition). The CNN architecture was the same for both evaluation scenarios, with the only difference being in the number of neurons in the output layer. The CNN architecture is presented in Table II. It consists of 1D convolutional (CN) and max pooling (MP) layers, dropout layers (DP), batch normalization layers (BN) and fully connected layers (FC). The network was implemented using the Python programming language version 3.10.11 with the TensorFlow library version 2.12.

For CN the strides were 1, for MP the strides were 2, DP rate was 0.5 and the activation function for all layers was the Exponential Linear Unit except for the final layer which had the Softmax activation function. All convolutional layers had l2 kernel regularizers with the l2 factor set to 0.01. The network was trained using an Adam optimizer with the initial learning rate set to 0.0001. The loss function of the network was a weighted categorical cross-entropy, with each class having a weight that is inversely proportional to the number of datapoints from that given class in the train set. The network was trained and evaluated based on the 5-fold subject-wise cross-validation (each fold contained data from two participants). Three folds were used for training, one for validation and one for testing. The maximum number of epochs was set to 300 and early stopping was implemented based on the loss function on the validation set, with the patience parameter set to 20. The network was evaluated using the F1 score, precision and recall for each class as well as the macro average for each metric and the overall classification accuracy.

III. RESULTS WITH DISCUSSION

An example of the obtained optimal phases by the MPC algorithm during the resting state, coughing and moving of subjects from the experiment are shown in Fig. 8–10.

Fig. 8–10 illustrate that it is possible, by visual inspection, to make the distinction between the resting state, cough and movement phase signals. In the cough scenario, the phase amplitude has higher values than in normal resting. Also, in the cough scenario, the higher variance of signal could be observed in comparison to the movement phase. This conclusion is expected because the subject made larger movements while switching between different sides than the movements they made during coughing.

Table III presents the CNN classification results for both classification problems: differentiating between the resting state, movement, and cough classes (three classes), and differentiating only between the resting state and movement classes (two classes recognition), for all radar

Table II. CNN architecture

	Layer number																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	1/7	18
type	CN	CN	BN	MP	DP	CN	CN	BN	MP	DP	CN	CN	BN	MP	DP	FC	DP	FC
kernel size	50	50	/	2	/	50	50	/	2	/	25	25	/	2	/	/	/	/
filter size	64	64	/	/	/	64	64	/	/	/	32	32	/	/	/	/	/	/
FC size	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	40	/	2

configurations (only radar 1, only radar 2 and both radars).

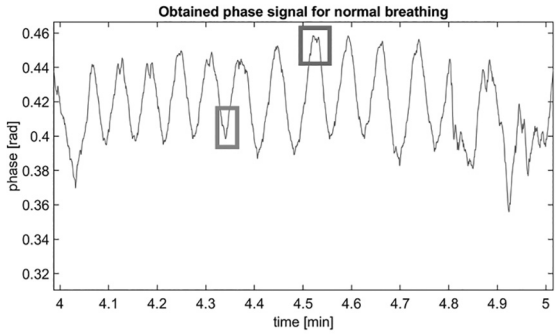

Fig. 8. An example of obtained phase signal for normal breathing – subject lying on the right side (one inhale – red rectangle, one exhale – green rectangle)

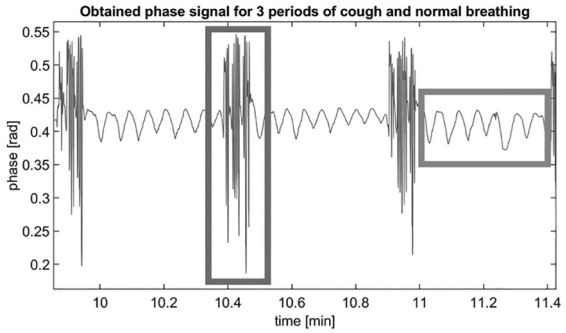

Fig. 9. An example of obtained phase signal for 3 periods of cough and normal breathing (cough – red rectangle, normal breathing – green rectangle)

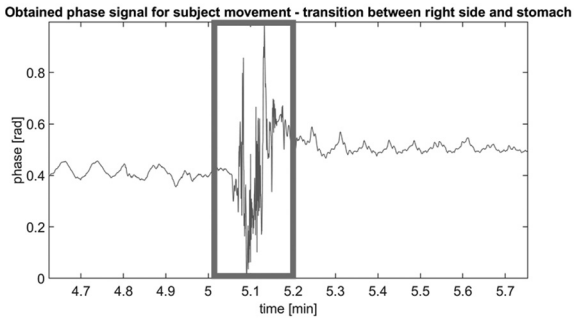

Fig. 10. An example of obtained phase signal for subject movement – transition between right side and stomach (movement – red rectangle)

It can be concluded that the maximum value of accuracy was 89 % for the two classes recognition (resting vs movement) for radar 1. Also, it can be concluded that higher values of accuracy are for radar 1 in comparison to the results for radar 2 or for both radars. The reason for these results is the small value of signal-to-noise ratio for radar 2, because of which CNN cannot detect the differences between the classes. Furthermore, observing the metric F1 Score for individual classes, it can be concluded that the best detected class was normal breathing. The imbalanced classes in combination with individual differences between the subjects and their movement and cough patterns could be the reason for the encountered accuracy detection barrier. However, with larger datasets containing more subjects, more complex CNN architecture and a combination of multiple radar phase signals from an adequately positioned sensor, further studies could further develop and improve the obtained results.

Table III. The CNN results for three classes recognition (radar 1)

Category	F1 score	Precision	Recall	Accuracy [%]
THREE CLASSES, RADAR 1				
Resting state	0.84	0.96	0.76	
Movement	0.40	0.33	0.54	
Cough	0.23	0.16	0.44	
Overall	0.49	0.48	0.58	72
THREE CLASSES, RADAR 2				
Resting state	0.80	0.96	0.69	
Movement	0.44	0.34	0.70	
Cough	0.24	0.16	0.49	
Overall	0.49	0.49	0.63	68
THREE CLASSES, BOTH RADARS				
Resting state	0.82	0.96	0.72	
Movement	0.46	0.39	0.59	
Cough	0.29	0.19	0.62	
Overall	0.53	0.25	0.65	71
TWO CLASSES, RADAR 1				
Resting state	0.94	0.99	0.89	
Movement	0.54	0.40	0.84	
Overall	0.74	0.69	0.87	89
TWO CLASSES, RADAR 2				
Resting state	0.90	0.94	0.86	
Movement	0.23	0.18	0.35	
Overall	0.56	0.56	0.61	82
TWO CLASSES, BOTH RADARS				
Resting state	0.92	0.98	0.86	
Movement	0.49	0.35	0.85	
Overall	0.70	0.67	0.86	86

IV. CONCLUSION

In this paper, an MPC method was applied to FMCW radar data for finding optimal phase signals in order to classify movement vs resting state and, in another scenario, movement, cough and resting state classes. Observing the results of the obtained phases, it can be concluded, by visual inspection, that there are distinctions between different states of the subjects. CNN results were better for radar 1 that was positioned underneath the bed than for the side radar 2 or both radars. Furthermore, CNN did not have enough data about the observed classes, because of class imbalance (resting state data being far more common than coughing or movement). For further investigations, there are possibilities for increasing the distance resolution in order to improve the implemented MPC algorithm. Also, it is necessary to record additional radar data for a larger number of subjects whereby it is expected that accuracy of CNN results would be improved.

ACKNOWLEDGMENT

This work was financially supported by the Innovation Fund of Republic of Serbia, innovation voucher, no. 760 and the Ministry of Science, Technological Development, and Innovations of the Republic of Serbia under contract number: 451-03-47/2023-01/200103 and 451-03-47/2023-01/200223.

REFERENCES

[1] Crivello, A. Barsocchi, P. Girolami, M. and F. Palumbo, “The meaning of sleep quality: a survey of available technologies,“ IEEE Access, vol. 7, pp. 167374-167390, 2019.

[2] Hong, H. Zhang, L. Gu, C. Li, Y. Zhou, G. and X. Zhu, “Noncontact sleep stage estimation using a CW doppler radar,“ IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 8, pp. 260-270, 2018.

[3] Zakrzewski, M. Vehkaoja, A. Joutsen, A. S. Palovuori, K. T. and J. J. Vanhala, “Noncontact respiration monitoring during sleep with microwave Doppler radar,“ IEEE Sensors Journal, vol. 15, pp. 5683-5693, 2015.

[4] De Chazal, P. Fox, N. O’HARE, E. M. E. R., Heneghan, C. Zaffaroni, A. Boyle, P. McNicholas and W. T. “Sleep/wake measurement using a non-contact biomotion sensor. Journal of sleep research,“ vol. 20, pp. 356-366, 2011.

[5] Chung, K. Y. Song, K. Shin, K. Sohn, J. Cho, S. H. and J. H. Chang, “Noncontact sleep study by multi-modal sensor fusion. Sensors,“ vol. 17(7), pp. 1685, 2017.

[6] R. De Goederen, S. Pu, M. Silos Viu, D. Doan, S. Overeem, W.A. Serdijin, K. F. M. Joosten, X. Long, J. Dudink, “Radar-based sleep stage classification in children undergoing polysomnography: a pilot-study. Sleep Medicine, “ vol 81, pp. 1-8, 2021

[7] Turppa, E. Kortelainen, J. M. Antropov, O. and T.Kiuru, “Vital Sign Monitoring Using FMCW Radar in Various Sleeping Scenarios. Sensors,“ vol. 20, pp. 6505, 2020.

[8] M. He, Y. Nian and Y. Gong, „Novel signal processing method for vital sign monitoring using FMCW radar,“ School of Biomedical Engineering, Third Military Medical University, Chongqing, 400038, China, vol. 33, pp. 335-345, 2017.

[9] J. -M. Muñoz-Ferreras, J. Wang, Z. Peng, C. Li and R. Gómez-García, "FMCW-Radar-Based Vital-Sign Monitoring of Multiple Patients," 2019 IEEE MTT-S International Microwave Biomedical Conference (IMBioC), Nanjing, People’s Republic of China, pp. 1-3, 2019.

[10] M. Alizadeh, G. Shaker, J. C. M. D. Almeida, P. P. Morita and S. Safavi-Naeini, "Remote Monitoring of Human Vital Signs Using mm-Wave FMCW Radar," IEEE Access, vol. 7, pp. 54958-54968, 2019.

[11] Y. Ge, A. Taha, SA. Shah, K. Dashtipour, S. Zhu, J. Cooper, QH. Abbasi and MA. Imran "Contactless WiFi Sensing and Monitoring for Future Healthcare-Emerging Trends, Challenges, and Opportunities. IEEE Reviews in Biomedical Engineering,“ pp. 171-91, 2023.

[12] Y. Song, W. Taylor, Y. Ge, M. Usman, Imran MA, QH. Abbasi, „Evaluation of deep learning models in contactless human motion detection system for next generation healthcare. Scientific Reports,“ vol. 12, pp. 21592, 2022.

[13] A. Shrestha, H. Li, J. Le Kernec and F. Fioranelli, „Continuous human activity classification from FMCW radar with Bi-LSTM networks,“ IEEE Sensors Journal, vol. 20, pp. 13607-19, 2020.

[14] Skolnik I. Merril, „Introduction to Radar Systems,“ MacGraw-Hill, New York City, United States, 2001.

[15] A. De Groote, M. Wantier, G. Cheron, M. Estenne, and M. Paiva, ‘‘Chest wall motion during tidal breathing,’’ J. Appl. Physiol., vol. 83, pp. 1531–1537, 1997.

[16] Choi, H. Song and H. -C. Shin, "Target Range Selection of FMCW Radar for Accurate Vital Information Extraction," IEEE Access, vol. 9, pp. 1261-1270, 2021.