Real-Time Detection of REM Phases with EEG Records

Valentin Bequart, Luc Fermaut, Milan Fertin,
Julie Robillart, and Rafal Kotas

***Abstract*—The article presents the development of a program capable of real-time detection of REM phases during the human sleep. For this purpose, 39 electroencephalogram (EEG) recordings from PhysioNet were used. To achieve the goal of the project authors selected following set of parameters: the average amplitude of the signal, alpha and delta power band in frequency domain and the ratio Alpha-Delta, for 30 second interval. An Artificial Intelligence (AI) has been developed with Keras and trained with those parameters for 34 patients. Finally, the AI has been tested on the last 5 patients, by simulating a true night. It reaches 62% in sensibility for REM Phase detection, and 85% in specificity. Obtained results are promising in terms of real-time REM phase detection, but the approach needs further development.**

***Index Terms*—EEG, REM sleep, NREM sleep, artificial neural network**

I. Introduction

HUMANS spend about a third of their lives sleeping. So, it is, therefore, interesting to study the effects of the sleep on humans. What happens exactly to a human during the night and why the body seems to need this rest? The answers to these questions can help to identify some pathologies during the night, help the medical staff for the treatments of the patients. For this purpose, many methods have been developed to study the sleep, like EEG (electroencephalo-graphy), EOG (electro-oculography). According to the previous studies, the existence of sleep phases had been put forward and this concept can be really helpful to better understand the importance of sleep. [1]

During sleep, body is working to support healthy brain function and maintain physical health. In children and teens, sleep also helps support growth and development. In the case of sleep deficiency, there is an increase of the risk of chronic health problem and it can affect the comportment. Sleep deficiency is linked to risk of heart disease, high blood pressure, diabetes, stroke, obesity and affects the response of immune system. [2]

The sleep can be divided into sleep phases. Those phases each have their effects, but according to [3] the fifth phase, also called REM phase, is the most important. The objective of described project is to detect REM phases in real-time.

V. Bequart, L. Fermaut, M. Fertin and J. Robillart are with the ISEN Lille Engineering School, 41 boulevard Vauban, 59046 Lille Cedex, France.

R. Kotas is with the Department of Microelectronics and Computer Science, Lodz University of Technology, ul. Wolczanska 221/223, 90-924 Lodz, Poland (e-mail: email: rkotas@dmcs.pl).

The knowledge about sleep phases came gradually, first they were defined by Loomis et al [4], then Aserinsky and Kleitman [5] discovered the REMs. After that, Dement and Kleitman introduced the cyclic patterns of sleep phases [6]. There are five phases, four non-REM and one REM. [1] Figure 1 shows the cyclic representation of a normal night, with those phases:

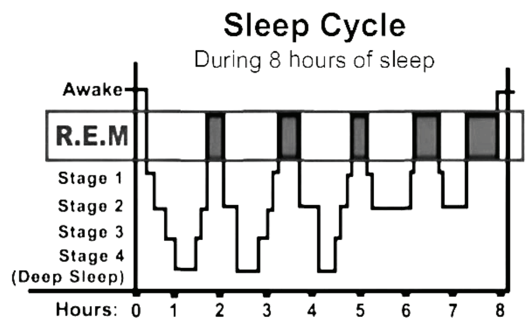

Figure 1. Classic sleep cycle

These different sleep phases were identified thanks to the categorization of human brain activity by frequency in five distinct groups. [1]

TABLE I
CATEGORIZATION OF BRAIN WAVES [1]

Unconscious		Conscious		
Delta (**Δ**)	*Theta* (θ)	*Alpha* (*α*)	*Beta* (β)	*Gamma* (γ)
0.5-4 Hz	*4-8 Hz*	*8-13 Hz*	*13-30 Hz*	*30-42 Hz*

It is possible to conclude that, if those phases can be identified, they have their own characteristics and effects.

Five sleep phases (six with the awaken state) with their consequence on the body during the sleep are as follows:

- sleep stage W: Awaken phase.
- sleep stage 1:Drowsiness: Breathing becomes slower, muscles are relaxing and consciousness decreases. During this phase, muscles can show little contractions, often with the impression of free falling.
- sleep stage 2: Slow light sleep: This phase represents 50% of the total sleeping time. During this phase, it is easy to wake up, a little bit of noise is enough, but the person remembers having slept.
- sleep stage 3 and 4: Slow deep sleep: The sleeping person is isolated from the external world by the sleep

and is recovering from fatigue during this phase. The brain emits slow and ample waves. These phases represent 20% of the total sleeping time.

- sleep stage REM: Paradoxical sleep: During this phase, which represents 25% of the total sleeping time, the person shows signs of deep sleep and awakening simultaneously. Those signs are the consequence of dreaming. [7]

Table II shows EEG characteristics for all sleep phases [1].

TABLE II
CHARACTERISTICS OF SLEEP PHASES REPRESENTATION IN EEG SIGNALS [1]

Sleep Stage	Characteristics
Awake	Low voltage, mixed frequencies
NREM1	Relatively low voltage, mixed frequencies
NREM2	Low voltage, mixed frequencies, presence of K-complexes and Sleep spindles
NREM3	$20\% \leq \Delta \leq 50\%$ Variable Amplitude
NREM4	$\Delta > 50\%$ Variable Amplitude
REM	Low voltage, mixed frequencies, presence of slow alpha waves, absence of delta waves.

II. METHODS

A. EEG database

The database of 39 polysomnograms from PhysioNet [8] was used to conduct the research. The polysomnography is a type of sleep study which gather information about breathing, EEG records, and also EOG. It is performed on patients in order to detect diseases or sleep troubles.

The data were stored under the European Data Format (EDF). An EDF files contains a header (with information about the record) and data from the different signals. [9]

An advantage of the PhysioNet database is the fact that each EDF file is linked to a hypnogram for the patient, in which the different sleep phases of the patient along the night have been analysed and written by the doctors. Then, it is possible to detect some interesting parameters for each phase, because those phases are known. Furthermore, thanks to those hypnograms, it is also possible to check the results of the algorithms developed during the study. [8]

B. Reading of EEG Data

An EDF Reader from Matlab was used to extract the EEG data from EDF files [10]. At the same time, the phases were extracted of the hypnograms by a Java module - EDF4J [11]. This module was not used for data reading because some differences were detected on the values between Matlab Reader and EDF4J. Moreover the data from Matlab were more exploitable, especially for the frequency analysis.

One of the assumptions for this project was to use Java as programming language. Due to this reason an instance of Matlab is called from Java with the use of a Java API - MatlabControl [12]. When the instance is running, EDF files are read by Matlab, which stores all values in a DAT file. Those EEG values represent the voltage of the brain activity at a certain time. Sampling frequency was 100Hz. A Java algorithm was developed in order to read values from the DAT file. The objective was to make an average of the voltage each 30 second (averages out of each 3000 samples). Finally, it stores those averages in an output file, with the timestamp of this record. This approach does not cause a loss of any relevant information because the mean duration of REM phase for records from this database is 10 minutes. Given that the objective is to detect the presence of a REM phase, this interval is sufficient.

The same approach was used to read the phases. They are stored in a file with the start timestamp of each phase. Then analysis can be proceeded with the two generated files for each patient. The data read from those files are read record by record to simulate real-time analysis.

C. Amplitude analysis

The first objective of the research was to check if it is possible to identify some distinct parameters for each sleep stage. The voltage amplitude of brain activity has already been studied in previous studies. In [1] it was indicated that some stages had smaller amplitude than others.

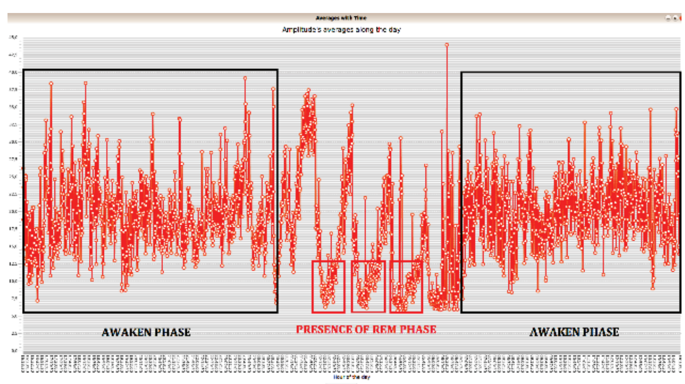

Figure 2. Visualisation of EEG values for one patient

A program to visualize the shape of the signal was developed and it is possible to observe the presence of a "cycle" during the sleeping period characterized by successive amplitude increases and decreases. By comparing the starting time of REM phases, and the time were the amplitude was relatively low, a correlation was spotted.

The analysis has been oriented around several criteria and the most relevant were:

- the average amplitude of each phase, for each patient, to get one global average.
- the mean deviation between each sleep stage.

The other parameters do not provide a difference between stages. To perform this analysis, an algorithm was developed in Java. This algorithm reads and processes all the data for each patient and creates a text file which contains the results for each patient, and one last global file.

```
Averages by phase :
Average Phase W : 9.48004680019496
Average Phase 1 : 7.096899182909047
Average Phase 2 : 8.086493356256494
Average Phase 3 : 12.702757122142085
Average Phase 4 : 14.526332722125277
Average Phase R : 5.874095384196816

Averages of time (*30 seconds) :
Average Phase 1 : 2.910430331369871
Average Phase 2 : 11.614984321860282
Average Phase 3 : 3.013399780222888
Average Phase 4 : 4.757068630676995
Average Phase R : 25.502479172353773

Minimal and Maximal Values by Phase :
Min Phase 1 : 2.160473340638674 Max Phase 1 : 38.060920633845036
Min Phase 2 : 2.5585642627926726 Max Phase 2 : 44.38485470300462
Min Phase 3 : 4.006722180427331 Max Phase 3 : 32.81977337871703
Min Phase 4 : 6.136539352554007 Max Phase 4 : 34.76659356015471
Min Phase R : 2.373483595730679 Max Phase R : 40.25477900325127

Minimal and Maximal Averages by Phase :
Min Phase 1 : 3.9137266295042457 Max Phase 1 : 13.283319278684456
Min Phase 2 : 3.988166786133196 Max Phase 2 : 14.896098049752625
Min Phase 3 : 5.76323904964905 Max Phase 3 : 24.500654422210783
Min Phase 4 : 8.378148604087942 Max Phase 4 : 28.88974602281164
Min Phase R : 3.1100111295757107 Max Phase R : 15.120883007983114

Average Difference between R and 1 : 1.2228037987122313
Average Difference between R and 2 : 2.2123979720596805
Average Difference between R and 3 : 6.828661737945271
Average Difference between R and 4 : 8.98263712597782
Average Difference between R and W : 3.6059514159981485
Average Difference between 1 and 2 : 0.9895941733474484
Average Difference between 1 and 3 : 5.605857939233038
Average Difference between 1 and 4 : 7.79240718788549
Average Difference between 1 and W : 2.3831476172859167
Average Difference between 2 and W : 1.3935534439384683
Average Difference between 2 and 3 : 4.616263765885592
Average Difference between 2 and 4 : 6.874484356993778
Average Difference between 3 and W : -3.2227103219471234
Average Difference between 3 and 4 : 2.5514078871369246
Average Difference between 4 and W : -5.544563396409665
```

Figure 3. File generated by the algorithm

As it can be observed, even if the EEG signal is unstable by nature, there is a different global average between each sleep stage. Therefore, the mean deviation between all sleep stages was selected as a criterion for the identification of sleep phases, because the observed gap is generally the same for all patients, with some variations because all humans are different, for example, all REM phases average for each patient will be minor than NREM1 average, but for one patient they will be at an amplitude of 5 µV and 6 µV, and for another at 6 µV and 7 µV. The mean deviation is the same, but the amplitude is different.

It is also due to this unstable nature that a lot of research criteria around the amplitude were not conclusive, like the maximal and minimal values of each stages, that could have allowed to eliminate some stages if a value was reached. Indeed, even if the REM phase have the lower average in amplitude, it can happen that the recorded REM signal contains peaks reaching the maximum value.

This analysis leads to a conclusion that amplitude criterion can be used to distinguish REM, NREM1 and NREM2 from NREM3 and NREM4, due to their major difference in amplitude. However, this criterion is not sufficient to distinguish REM, NREM1 and NREM2. This is due to the fact that the amplitudes of these phases are very close to each other. But if a frequency analysis of EEG signals would be added to this criterion, it is possible to increase the final accuracy.

D. Frequency analysis

The frequency analysis is based on the recognition of brain waves which leads to the REM phases determination. First the signal is divided into records of 30 seconds (3000 samples each) and the power of the signal in a particular band frequency has been calculated. The power of the signal is based on the fast Fourier transform of a record. It enables the separation of the different frequency band of each wave. It is calculated as the sum of squares of each value in a certain range of frequency, as shown below:

$$P = \sum_{f=fmin}^{fmax} |y(f)|^2 \tag{1}$$

where:

y – the fast Fourier transform of the signal,

P – the power of the signal.

The evolution of the power for the different band shows that the Alpha waves (8-12Hz) are higher during REM phases and the Delta waves (0-4Hz) were lower. The other waves do not give relevant results to differentiate the sleep phases.

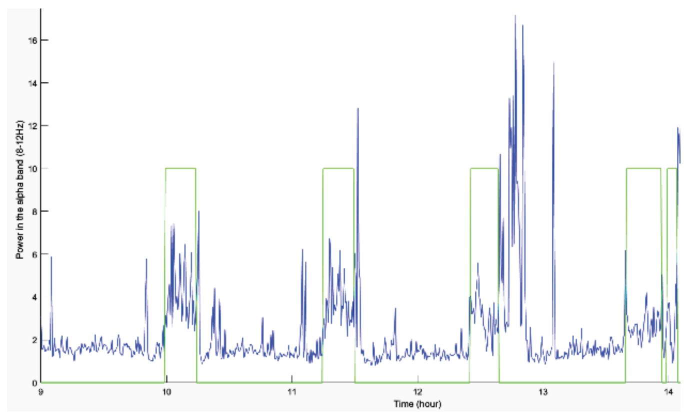

Figure 4. Blue - power in the Alpha band, green –real REM phases

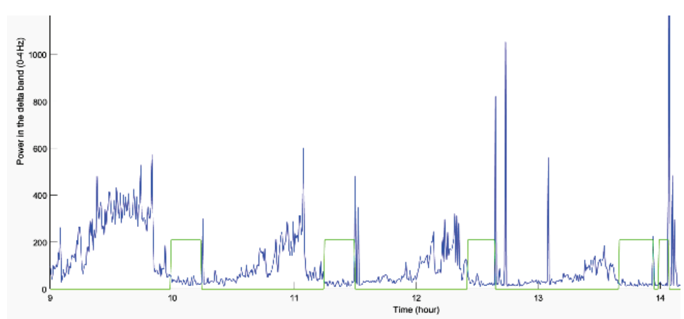

Figure 5. Blue - power in the Delta band, green – real REM phases

The next step was to combine information of both alpha waves and delta waves. Authors decided to introduce a new parameter: the ADR (Alpha-Delta Ratio), because alpha waves are high and delta waves are low during REM phases and it is the opposite during NREM phases. The ADR enables the gathering of more information in one function, and it is more universal because the main power amplitude can change from one patient to another, but the ratio will balance this difference.

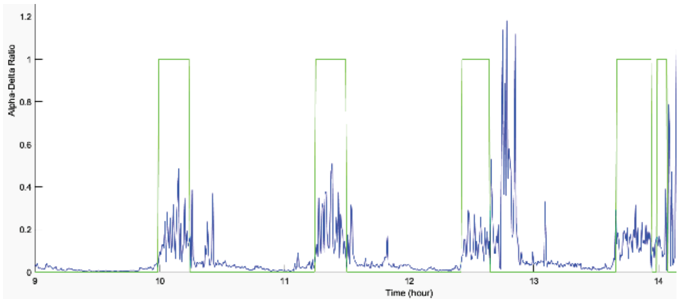

Figure 6. Blue - Alpha-Delta Ratio, green – real REM phases

To smooth the curve of the ADR function, the average of the last *n* values of the ADR was used to avoid isolated peaks that could deteriorate future analysis.

The ADR enables the authors to develop an algorithm which uses a threshold to detect REM and NREM phases. By observing the signal, there is a threshold of about 5% for most of the patients. But this value can change from one patient to another. The algorithm separates the REM and NREM phases thanks to this threshold, but it gives only the beginning of the REM phases. It can also classify mistakenly other phases as REM phases.

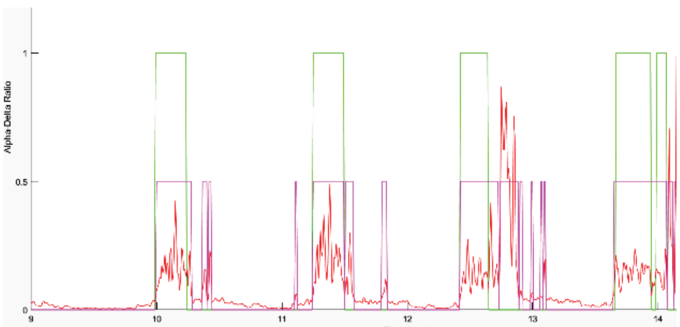

Figure 7. Red – Alpha-Delta Ratio smoothed, green – real REM phases, pink–REM phases found by the algorithm

E. Real-time simulation based on combined amplitude and frequency analysis

To check if the criteria could be effective, an algorithm which proceed to a simulated real-time analysis has been developed. The objective of this algorithm was to check the results in REM detection of combined amplitude and frequency analysis.

The input parameters for the algorithm are of two types:

- the mean deviation between each sleep stage, taken from the global conclusion file of the amplitude analysis algorithm,
- a threshold value for the ADR in frequency domain found by processing multiple analysis with various values of it (which mean if the calculated ratio is superior to this threshold, the patient is probably in a REM stage) and the number of precedent value of this ratio.

The initial values those parameters are calculated with the use of training dataset (34 patients).

In this step, all data from the night are read one by one, and an average is calculated for every 30 seconds period. At the same time, the power of α and **Δ** waves are checked within those 30 seconds, and a ratio is calculated.

The algorithm reads the new average and ratio, and determine the actual sleep stage, knowing:

- the previous values of the ratio and averages,
- the previous estimated stages,
- a global average calculated for each stage based on the previous values and the previous estimated stages,
- the mean difference between each stage, which is modified with time by the algorithm with the patient values, to better match this particular patient.

For the variation in sleep stages, the theoretical case, as described in Figure 1 was preferred, so it is considered that from the REM stage, it is only possible to go to NREM1 or stay in REM stage, in NREM1 it is only possible to go to REM stage or to NREM2 and so on. The algorithm considers a patient begin his night in NREM1, which may vary in some rare cases.

The evaluation of proposed algorithm is based on a comparison with the true stages of the night determined in the hypnograms of PhysioNet. It allows to calculate the percentage of good results for the REM stage, for the all stages in global, and the false positive REM stage detection.

F. Sleep phases detetion based on Artificial Intelligence

The other approach used to distinguish sleep phases was t use Artificial Intelligence (AI). It was decided to work with Keras, which is a high-level artificial neural networks API, using a TensorFlow backend. [13]

An Artificial Neural Network (ANN) is a computing system inspired by biological neural networks. ANNs learn to perform tasks by considering examples that they receive. In this case, authors provide them with data corresponding to different sleep phases. In general a neural network is composed of:

- an input layer: A layer composed of N neurons for N inputs (except for LSTM layers). The input layer will normalize the data for the hidden layer(s).
- hidden layer(s): One or more layers composed of many neurons depending on the needs. It transforms their inputs for inputs of the output layer.
- an output layer: A layer composed of N neurons for N outputs. It gives the final result.

At the beginning the AI had been trained with only amplitude data. Though, the results were not satisfying at all since the amplitude is not a sufficient parameter to determine sleep phases due to the unstable nature of EEG signals. Then the AI was trained with both frequency and amplitude data. Giving the AI a single value at a certain time is not relevant since what allows to visually identify the different phases is the evolution of the data.

The AI is being trained with:

- the amplitude of the 120 last records,
- the power in Alpha band of the 120 last records,
- the power in Delta band of the 120 last records,
- the Alpha-Delta Ratio of the 120 last records.

It has been decided to take the 120 last records (as soon as possible) because it was shown that usually, REM sleep happens one hour after a person fall asleep [14]. In this case, a record is equivalent to 30 seconds (3000 samples; sampling frequency is 100 Hz).

The power in Alpha band and the power in Delta band were added to the AI because they bring more information, even if the ratio gathers them. Obtained results shows that the use of those four inputs improves the AI training.

The authors tried a lot of different architectures. The final configuration of the AI was:

- three LSTM (Long Short-Term Memory) layers with 150 neurons,
- one dense layer of 7 neurons to manage all possible outputs (Awaken, NREMs, REM, and unknown),
- Adam optimizer,
- for loss function: sparse categorical cross entropy.

This model was trained with the data of 34 patients and tested on the last 5 patients.

III. Results

A. Results of the approach based on combined amplitude and frequency analysis algorithm

The algorithm was launched with different values of n (the number of last values of the ratio to calculate an average to smooth the ADR function) and different threshold for the average to detect a REM stage. The NREM stages were mainly detected with amplitude analysis, and REM stage with the frequency analysis.

Table III presents the results obtained with the variations of the described parameters for REM phase detection in terms of Sensitivity and Specificity.

TABLE III
DETECTION OF REM PHASE WITH SENSITIVITY AND (SPECIFICITY)

		Threshold to reach REM stage		
		0.03	0.05	0.07
Number of previous values	2	0.9142 (0.4347)	**0.8068 (0.621)**	0.6252 (0.7432)
	5	0.9169 (0.4177)	0.8226 (0.5985)	0.6443 (0.7231)
	10	0.9245 (0.4018)	0.8226 (0.5772)	0.6342 (0.7067)

The algorithm also detects the NREM phase. Table IV, with the same parameters as in Table III, presents the percentage of all stages (REM and NREM) correctly detected.

TABLE IV
DETECTION OF ALL SLEEP STAGES IN PERCENTAGE

		Threshold to reach REM stage		
		0.03	0.05	0.07
Number of previous values	2	34%	**40%**	43%
	5	34%	39%	42%
	10	34%	40%	43%

B. Results of the approach based on AI

The combination of the 120 last records for amplitude, power in Alpha and Delta band, Alpha-Delta Ratio and the AI structure is used for REM phase detection. The results are presented in Table V. The sleep phases were determined record by record – the algorithm compares the true REM phases with the detected REM phases.

TABLE V
RESULTS OF REM PHASE DETECTION FOR 5 PATIENTS IN TERMS OF SENSIBILITY AND SPECIFICITY

Patient	Sensibility	Specificity
1	0.7200	0.8574
2	0.8186	0.8323
3	0.6800	0.8900
4	0.2784	0.8273
5	0.6073	0.8780
Global:	0.6209	0.8550

62.09% of exact REM phases were detected. Furthermore, the number of false positives has considerably decreased compared to the first algorithm, with 85.50% of specificity.

IV. Discussion

A. Combined amplitude and frequency algorithm

Obtained results show that, even if it is possible to have a good detection of REM stage, particularly with a threshold of 0.03 or 0.05, there is a significant risk of false positive in the detection. This can be explained by the fact that this threshold which was fixed with the definition of the ratio is different for each patient, for example the 0.05 can be efficient for most of the patients, but in some cases, it was spotted that this limit was too low, and a majority of the EEG signal for the concerned patient was detected as a REM stage.

This problem could be fixed by adjusting the limit for each patient. It could be done by the AI, that will learn to adapt this ratio to the patient data.

B. Approach based on AI

The AI approach reached 62.09% of REM phase detection. Part of the REM phases was not detected due to the gap between the beginning of the true REM phase and its detection by the AI. Figures 8 and 9 show graphs for patient no. 1 and no. 4 respectively, with the true (orange) and detected (blue) REM phases.

It can be observed, with the example of the first patient, that even if his statistic result is only 72% of sensibility, each of his REM phases was detected. Each REM phase was detected with short delay and was longer than the true one.

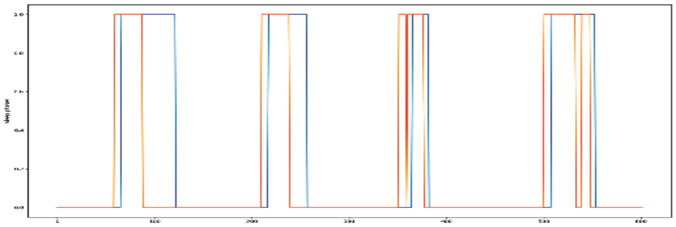

Figure 8. Patient no. 1. Orange: true REM phases. Blue: detected REM phases.

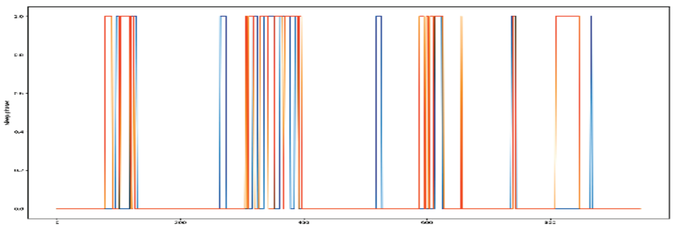

Figure 9. Patient no. 5. Orange: true REM phases. Blue: detected REM phases.

Furthermore, for the fifth patient, which has the lowest result, the AI algorithm detects most of the moments where the patient reaches the REM phase. However the result for this patient is so low due to the fact that REM phases registered in the hypnogram are a succession of many short, frequently changing REM phases instead of the long, typical REM phase.

V. Conclusion

In conclusion, the goal of conducted research was achieved. Even though, the final result reaches only 62%, it has to be stated that it is because of the fact that the algorithm tried to determine each 30 second period separately. However, in general, presented algorithm found all REM phases during sleep and can indicate in real-time whether the patient is in the REM phase or not with a false positives rate relatively low (specificity at 85.50%). Further work will be carried out to improve the algorithm and increase the detection sensibility.

References

[1] Branco J, Paiva T, Martins R. Data Acquisition System for Sleep Stage Detection: Signal Processing

[2] National Heart, Lung, and Blood Institute: Why Is Sleep Important? - Sleep Deprivation and Deficiency, Retrieved from https://www.nhlbi.nih.gov/node/4605 (2018, July 23)

[3] Sound Sleep Health: What Stage of Sleep Is Most Important? NREM vs REM Sleep, Retrieved from https://www.soundsleephealth.com/blog/what-stage-of-sleep-is-most-important-nrem-vs-rem-sleep (2018, July 23)

[4] Loomis AL, Harvey EN, Hobart GA., Cerebral states during sleep, as studied by human brain potentials. J Exp Psychol 1937; 21: 127-144.

[5] Aserinsky E, Kleitman N., Regularly occurring periods of eye motility and concomitant phenomena during sleep. Science 1953; 118: 273-274

[6] Dement WC, Kleitman N., Cyclic variations in EEG during sleep and their relation to eye movements, body motility and dreaming. Electroencephalogr clin Neurophysiol 1957; 9: 673-690

[7] Estrada E, Nazeran H, Nava P, Behbehani K, Burk J, Lucas E. EEG feature extraction for classification of sleep stages, The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Francisco, CA, 2004, pp. 196-199

[8] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 101(23):e215-e220; 2000

[9] Kemp B, Olivan J. European data format 'plus' (EDF+), an EDF alike standard format for the exchange of physiological data. Clin Neurophysiol. 2003 Sep;114(9):1755-61.

[10] DennisDean, EDF Reader in Matlab: Retrieved from https://github.com/DennisDean/BlockEdfLoad (2018, July 23)

[11] EDF Reader in Java (used for reading hypnograms): Retrieved from https://github.com/MIOB/EDF4J (2018, July 23)

[12] Matlab Control from Java: Retrieved from https://code.google.com/archive/p/matlabcontrol/ (2018, July 23)

[13] Keras Documentation, Retrieved from https://keras.io/ (2018, July 23)

[14] Lund HG, et al. Sleep patterns and predictors of disturbed sleep in a large population of college students. Journal of adolescent health 46.2 (2010): 124-132.

Valentin Bequart, is a student at ISEN Lille (Institute of Electronics and Digital Engineering, France). His experience is related to software development in the field of web applications, cryptography and artificial intelligence.

Luc Fermaut is a student at ISEN Lille (Institute of Electronics and Digital Engineering, France). His experience is related to software development in the field of signal processing and artificial intelligence.

Milan Fertin is a student ISEN Lille (Institute of Electronics and Digital Engineering, France). His experience is related to project management and software development in the field of cybersecurity and web applications.

Julie Robillart is a student at ISEN Lille (Institute of Electronics and Digital Engineering, France). Her experience is related to software development in the field of web and mobile applications and voice recognition systems.

Rafal Kotas received the M.Sc. and Ph.D. degrees from the Lodz University of Technology in 2009 and 2014 respectively. Since November 2014 until now he is with the Department of Microelectronics and Computer Science, Lodz University of Technology. His main field of study is processing and analysis of bioelectric signals for the purpose of medical diagnosis, algorithms development.