1 REM SLEEP STAGE IDENTIFICATION WITH WAVELET DECOMPOSITION AND ARTIFICIAL NEURAL NETWORK USING A SINGLE CHANNEL EEG

Gabriel Toban

2 Why Use Sleep Stages?

• Sleep stages are the most precise way to separate wakefulness from sleep state.

• Identify physiologic and pathologic events

– Paralysis in REM

– Arousals and pathology (Apnea)

– K-complexes and pathology (Epilepsy)

• Required for some sleep studies with EEG

•

3 What is an EEG?

• Electroencephalogram

• Difference in electric potentials of two electrode positions called a EEG channel that Produces “brain waves”

4 How to Identify Sleep Stages?

• American Academy of Sleep Medicine defines sleep stages and how to identify them

• Polysomnogram or Polysomnographic Record (PSG)

• Home Study

•

5 What is a PSG?

• Sleep study

• Collection of parameters EEG, EOG, chin and leg EMG, ECG, airflow, respiratory effort, Oxygen Saturation, heart rate

• Includes

– EEG Electroencephalogram (brain)

– EOG Electrooculogram (eyes)

– EMG Electromyogram (muscles chin and leg)

6 What is a Home Study?

• Typically for sleep apnea

• The AASM defines requirements for (HSAT) Home Sleep Apnea Testing

• Sleep staging required for some

• Much fewer parameters and equipment

•

7 Automated Sleep Staging

• Manual Sleep Staging Requires Tech to use equipment and study the results

• Scoring Guides from AASM

• Some use many parameters from PSG

• Some use very few parameters like a single channel EEG

8 Presentation Outline

• For the Rest of the presentation I am going to about The factors effecting the decision for

– The Single Channel EEG

– Signal Processing Algorithms

– Categorization Algorithms

• Methods (Coding) and Results

• Conclusion and Future

9

MOVING ON TO THE SINGLE EEG CHANNEL

10 The EEG

• Electrodes on the head at different position

• Two Electrodes positions make a channel

• The 10/20 International Standard defines electrode positions

• The Rechtschaffen and Kales or the R&K Standard, terminology and definitions for sleep scoring 1968

• The AASM

• Brain Waves, Activities, and Events (Markers)

• Brain waves, activities, and events (Markers) are defined by the amplitude, frequency, and identifying

factors of a single wave or grouping of waves•

11 5 EEG Frequencies

These define the scale of frequencies seen during sleep and wakefulness, but do not define the frequencies of all the markers such as Low Amplitude Mixed Frequency Waves predominately 4-7 Hz not all of theta.

12 The 10/20 System

Uses a starting position of the Nasion and moves around the head using the percentages listed here. The 10/20 name comes from the percentages around the side and over the top.

The numbers and letters that define each position are based on the brain lobe it is over the side of the head its on. Even number for the right side and odd for the left and z for the middle.

13 Brain Lobes/Regions

The locations defined by the letters are not nott all over brain lobes. central is just the line between the ears.

14 EEG Channels

The AASM recommends these 6 channels plus some backup channels not listed for an EEG in a PSG. This many channels just for an EEG would be difficult for an untrained person to place.

15 Markers to Lobes/Channels

There are a lot of “markers” for sleep staging. A lot of which cannot be found on an EEG like Rapid Eye Movement for REM sleep. For some that can be found by an EEG, the lobe or channel these markers can be maximally or adequately detected from is listed here. This is not precise. The maximal implied the wave amplitude will be maximal.. Crasto et al. [13] was able to break down stage a slow wave using channel Fpz-Cz into two sections according to the wave amplitude.

16 Sleep Stages

These are all the sleep stages defined by R&K and AASM, Notice movement time and stage 3 and 4 are combined in AASM standards. Movement time is still used today. It defines a time when there is too much movement to get an accurate recording.

17 Stages to Markers

Every sleep stage is completely defined according to the AASM using many factors

of a PSG (polysomnogram) including the EEG, EOG, and EMG [11]. Again, the rapid eye movement from an EOG to define REM, but REM

may be identified with sawtooth waves detected by an EEG. This gives an idea of the relationship between EEG markers and sleep stages.

NR: Not Required; R: Required; P: Possible; fBlankg: Not mentioned or not possible

18 This Study

• A Single Channel

There is a current push to diagnose sleep associated diseases with a cheaper at home

study. Sleep staging using a single EEG channel would be cheaper than a PSG. Ventouras et al. [29], Ebrahimi et al. [16], and Crasto et al. [13] all used a single EEG channel with accuracies between 85% and 95%. Shambroom et al. [27] tested current equipment that uses a single channel EEG to do sleep staging.

• Channel Fpz-Cz

Aside from slow waves for N3, there is not any single marker from an

EEG uniquely required to identify 1 sleep stage. The dataset used in this study has channels Fpz-Cz and Pz-Oz available. The channel to identify slow waves is not available. Since all other waves in table 4 are adequately identified in the central and frontal lobes, Fpz-Cz channel was used to for this study. Future studies may consider using other frontal lobe channels.

• 30 Second Epochs

The R&K standard [26] states 3 reason for 30

second epochs: too small epoch will be too cumbersome, too large epochs will miss out

on epoch details, and paper sizes for EEG machines typically fit 30 seconds of data. The

Physionet data set [5] used in this study has 30 second epochs marked with their sleep stage.

Current study uses 30 second epochs. BUT The AASM and R&K both mention the necessity to

consider the stage before and after the current epoch to properly score. That is part of many markers defined by the AASM. This study will use 3 epochs to stage a single epoch.

• REM Sleep Stage

REM sleep has not been the specific goal of any study found.

There is not a required unique identifier for REM sleep in the EEG.

REM sleep is almost indistinguishable from N1 sleep in an EEG. Stages W, N1, N2, and N3

are easily distinguishable from each other with an EEG (see table 6). Stages N1, N2, and

N3 all have a unique EEG marker to identify them.

19

MOVING ON TO THE SIGNAL PROCESSING

20 Signal Processing

• Why signal processing? A brainwave is a signal

Every resource I found on signal processing started with the Fourier series

. Achermann [8] stated that Dietsch [15] performed the first Fourier analysis on an EEG in 1932

• EEG Waves

• Time-Frequency Signal Analysis

• Wavelets

•

•

21 Fourier

This is the Fourier Series.

This is the Fourier Integral.

The Fourier integral theorem takes the limit as L approaches infinite of the Fourier series.

This is the Fourier Transform

It was derived from the Fourier Integral.

All of these represent harmonic or sinusoidal waves that go on

forever.

The frequency of the waves of a Fourier Transform represent the Spectral Density of the original wave [6]. The spectral density is a representation of the amplitude of a wave[7],but the sinusoidal component of the transform does not allow to capture local information[14]. In other words, abrupt changes in amplitude may not be captured. It is well-known

that Fourier Transforms cannot investigate time and frequency simultaneously [28] [16] [20]

[13] [14].

22 EEG Waves

Here are some ruff pictures of EEG waves. It is important to note that these are not harmonic or sinusoidal and they have abrupt changes in amplitude so these waves would not be captured well by the Fourier Transform.

23 Time-Frequency Analysis

Time-frequency representations can be classified as linear, quadratic, or higher-order representations.

In 1992, Hlawatsch and Boudreaux-Bartels listed at least 25 different time-frequency

representations including 2 linear representations, Short-time Fourier transform

(STFT), and Wavelet transform (WT), and 2 quadratic representations, Choi-Williams Distribution

(CWD) and Wigner-Ville Distribution (WD) [19]. Wigner-Ville Distribution (WVD) and the Choi-Williams Distribution are both generated from a more general Cohen distributions [14]. The Wigner-Ville and Wavelets have been modified many times to create many time-frequency distributions and transforms[14] [6].

A general review of the methods used to decompose EEGs found that that Short-time Fourier and Wavelets are the primary

representations used for analyzing an EEG in sleep staging [30]. Among those that used

Short-time Fourier and Wavelets, the Wavelet representations resulted in the highest accuracies [28]

[16] [20] [13] [30]. Fraiwan et al. [17] used Choi-Williams Distribution (CWD), continuous

wavelet transform (CWT), and Hilbert-Huang Transform (HHT) and found

that CWT had the highest accuracy.

24 Wavelet Transforms

The wavelet transform is based on the concept of a function called the ”mother

wavelet”. This mother wavelet can be translated with b and dilated with a to fit the original wave. A and b are the coefficients of the transform. The mother wavelet replaces the exponential used in the Fourier Transform which causes it’s inability to capture local information. Daubechies was the first nonorthogonal wavelet

produced that decomposes a non-stationary wave. Ebrahimi et al. [16] and Crasto and Upadhyay [13] used

Daubechies wavelet order 2 filter with DWT. Crasto and Upadhyay [13] tested DWTs with

Daubechies wavelet order 2 to 6 to decompose a signal for input into an artificial neural

network (ANN) finding Daubechies wavelet order 2 to produce the most accurate results.

25 Mother Wavelets

For comparison, here are two mother wavelets plotted in MATLAB. One of the original mother wavelets created by Morlet and a Daubechies order 2 wavelet.

26 Wavelet Transform Modifications

The wavelet transform has many modifications to help it improve. The Continuous Wavelet Transform likely has the highest ability to capture local information. The Discrete Wavelet Transform descreases the processing time of the CWT. The multilevel Discrete Wavelet Transform takes the transform of a transform which reduces closesness of the transform to the original wave but decreases the number of coefficients. Ebrahimi et al. [16] used a wavelet packet tree (WPT) with 93% accuracy, but there are many more such as the stationary wavelet transform (SWT), complex-valued wavelet transform (CWT), dual-tree complex wavelet transform (DTCWT), etc.

27 This Study

• Multilevel Discrete Wavelet Transform with Daubechies Order 2 Mother Wavelet

28

MOVING ON TO CLASSIFICATION

29 Classification

• Classification Algorithms

• How I built my my Artificial Neural Networks (ANNs)

30 Classification Algorithms

In 2017, Alvarez et al., who completed a review of neural networks in sleep medicine, listed the following algorithms, among others, in order of highest accuracy for sleep staging: ANN, support vector machine (SVM), hidden Markov models, and discriminant analysis. In 2006, Lotte et al. [24]scored many algorithms for classification in EEG-based brain-computer interface including linear discriminant analysis, support vector machine (SVM), multilayer perceptron, others ANNs, Bayes quadratic, hidden Markov model, k nearest neighbors, Mahalanobis distance, and combinations. These classifiers were scored based on the defintions generative vs discriminative, Static vs dynamic, and Stable vs unstable. Even though Lotte et al. [24] concluded that the SVM would be best for EEG-based brain-computer interface, they also stated neural networks are the most used category of classifier.

31 Artificial Neural Networks

• Heaton [18] listed feedforward networks, deep belief

networks, deep feedforward networks, and convolution networks as the best ANNs for

classification. Studies that I have reviewed had high accuracy with feed forward back propagation

neural networks [13] [28] and multilayer perceptron neural networks [20] [29]. All of these

algorithms could be tested. This paper used a feed forward backward propagation neural network.•

32 Linear Regression

A neural network produces a non-linear decision boundary [24]. It is easy to understand

a neural network when it is described from regression which produces a linear decision

boundary.

Linear regression is based on the equation of a line. The equation of a line is y = mx+b.

Given any set of x and y coordinates, there exist a line that minimizes the distance between

the line and the coordinates [22]. In other words, there exist a red line that minimizes the

sum of the lengths of the blue dashed lines in 4. One way to do this is to minimize the least squares equation. You take the derivative of the least squares equation, set it equal to 0, and solve for the coefficients in the equation of a line. When the coordinates are multidimensional, the steps to the solution are basically the same except with matrices. When this multidimensional equation cant be solved analytically, Gradient Descent can be used.

33 Gradient Descent

Gradient descent is stepping along the gradient of a

function. Instead of solving for where the gradient equals zero by setting it to zero, gradient

descent takes small steps towards gradient value zero until it reaches it.

Here f is the least squares equation and lamda is the step size or the learning rate

34 Linear Regression Nodes

This is that process in two dimensions and multiple dimensions in node diagram. In here, the least squares is equation is the cost which should be reduced to find the best line. In order to make this picture represent logistic regression, the sigmoid function must be used.

35 Sigmoid Function

The sigmoid function always returns a value between 0 and 1.

36 Logistic Regression Nodes

When the sigmoid function is used the cost function must be updated to the cross entropy function. Note this cost function has become the likelihood that 1 of two values will be chosen. In order to solve for the weights and biases (the ms and bs in this picture), gradient descent is used just like in linear regression.

If there are more than 2 possible results then the output can be interpreted differently. In the picture with more than one ouput if y1 = 1 the result is the value associated with y1 if y2 is 1 then then the result is the value associated with y2. This is called one hot encoding. For one hot encoding, you need softmax.

37 SoftMax

Softmax makes all the outputs sum to equal 1. In this way, the most likely choice is the index with the highest value.

38 Artificial Neural Network

Notice that the function for a single y output, z1, and z2 is still sigmoid. For the middle layers and a single output, the sigmoid function can be replaced with 1 of many nonlinearity or activation functions including hyperbolic tangent, linear, step, and rectified linear unites (ReLU) [18]. When there are two output nodes sigmoid or softmax can be used, but when there are more than 2, only softmax can be used.

While the function for y is sigmoid, the cost function can remain the

cross-entropy function. When softmax is used, the cost function becomes the multiclass log likelihood function.

For the nodes and biases, gradient descent is used again, but it require that the negative of log likihood is used. If you do not take the negative of log likelihood then it’s called gradient ascent.

39

ON TO METHODS.

40 Methods

• Dataset

• EDFbrowser

• epochs.py

• softANN.py

41 Dataset

This study used the ”Sleep Recording and Hypnograms in European Data Format (EDF)”

data set or ”The Sleep-EDF Database [Expanded]” from physionet.org [5]. The portion of

the database used is from a study on healthy patients from 1987-1991. There are 20 patients

avaialable. There are 10 males and 10 females ranging from age 25-34 years old [5]. each patient has 2 nights of recordings, there are 40 nights available from the sleep EDF database. Of the 40 nights available, 24 were used.

Each patient had two relevant files: a polysomnogram (PSG) recording in EDF format

and a hypnogram of annotations in EDF+ [5]. EDF is a standard format for exchanging EEG

recordings [3]. EDF+ has all the capabilities of EDF and the ability to contain annotations

[3]. The PSG recording included an EEG from Fpz-Cz and Pz-Oz electrode locations, an

EOG (horizontal), a submental chin EMG, an event marker, an oral-nasal respiration, and

rectal body temperature. The hypnogram is an annotation of sleep patterns. The annotations

included W for wakefulness, R for REM, 1 for stage 1, 2 for stage 2, 3 for stage 3, 4 for

stage 4, M for movement time and ? for not scored. I skipped epochs for movement time and not scored.

Once the first annotation was reached, the rest of the annotations were 30 seconds apart creating the 30 second epochs. I used the last 30 seconds before the first annotation as the first epoch.

41 EDF BROWSER

EDFbrowser is a free EDF and EDF+ reader that can export the contents to a text file. All the contents of the PSG and hypnogram file were exported to a text file with the option to have time in seconds from the

first recording.

42 Epochs.py

The PSG and hypnogram text files were imported into a custom python 3 script. The

python 3 script, epochs.py, used the pywavelets package to complete a wavelet decomposition

of any suggested level. The output file contains the annotation and decomposition

of each epoch per line unless the annotation is for movement time or not scored. Movement

time and not scored epochs are skipped in a majority of the tests. The count of each

annotation per file is printed to standard out.

43 softAnn.py

softANN.py, accepted the output of epochs.py as input The input to the

neural network consist of the coefficients from 3 epochs: the prior epoch, the current epoch,

and the post epoch. If the current epoch is the first epoch in the file then all the coefficients to the prior epoch are set to zero. If the current epoch is the last epoch in the file then all the

coefficients of the post epoch are set to 0. The real value of the output of the neural network

is y = [1,0] if the annotation is R and [0,1] if the annotation is not REM.

If the current epoch is the first epoch in the file then all the coefficients to the prior epoch are set to zero. If the current epoch is the last epoch in the file then all the coefficients of the post epoch are set to 0.

softANN.py does the training and prediction testing of the artificial neural network.

The outputs for training are 4 CSV files and current statistics to standard out at specified

intervals. The 4 CSV files are the weights and biases between the input and hidden layers

and the hidden and output layers. The current statistics include the cost and the accuracy

of the weights and biases. The stopping condition was split between 100% accuracy and

cost less than 10^-11. The output for prediction testing includes the accuracy, sensitivity, and

specificity [10].

44 Results: Train

there were an average of 1589 epochs per night. Of those epochs, 89% were

not REM and 10% were REM.

Using one of the nights, some preliminary test were run on the softANN.py to determine

the optimal wavelet decomposition level, the optimal number of hidden nodes, and the

optimal learning rate. It appears any level is capable of achieving 100% accuracy on the test

night. Lower levels achieve 100% accuracy while reducing the Cost more than higher levels.

The difference in time to 100% between levels is negligible. 10e-5 is the optimal learning

rate for all levels. It was found that learning rates 10e-3 and larger caused an overflow.

Crasto et al. [13] used principal component analysis to reduce the number of coefficients

from approx 3000 to 400. These wavelet levels can reduce the number of components to 14,

but level 3 reduces the number of coefficients to 377 which is closest to 400. Since 3 epochs

are used, there are 1131 input nodes with 3 levels of decomposition. Various numbers of

hidden nodes were tested, but the results were not stored. The number of hidden nodes were

set to half the number of input nodes, equal the number of input nodes, and one and a half the number of input nodes with no difference noticed. The test for training and predictions

all used the same number of hidden nodes as input nodes.

The training data sets all used 1, 2, or 4 nights of data. Of the 8 trainings, 1 was

stopped by the user, 1 resulted in overflow, and 1 had the weights lost as a result of a copying

error. The training stopped by the user was stopped after 3 days. No other training took

longer than 12 hours. All of the training was stopped in the program when the accuracy

of the current data set reached 100% except 1 which was stopped after the cost reached

10e-1. Each training set was given an id based on the level of decomposition and the number of iterations of training. The level of decomposition is the first number and the number of iterations is the second number in the id.

45 Results: Predict

Each training session produced weights and biases that were used to make test predictions.

the test predictions were completed with all of the remaining nights. The average accuracy, sensitivity, and specificity are listed

46

47 Conclusion

There are many factors left formerly untested: Over-fitting as a stopping condition in

training, the number of nodes in the hidden layers, the number of hidden layers, more than 2

nights as input, level 2 decomposition, etc.

The best result is that the ability to pick REM sleep is not good. The percentage of

correctly chosen REM was less than 50% a majority of the time. Sensitivity increases when

the number of nights included in the input increased. There are not enough test to know if

lowering the decomposition level had an effect.

A positive result is the specificity which was consistently near or above 90%.

The results support the idea that all sleep stages besides REM are easier to find in a

single channel EEG.

48 Future

In future studies, every aspect of this study could be changed to improve necessary

results. The choice of EEG channel. The choice of sleep channel. The choice of sleep stage

could be changed. The choice of epoch size could be changed. Of time-frequency techniques,

the wavelet transform is the best choice, but the modification of wavelet transform should

be changed to possibly a continuous wavelet transform or a new modification of the discrete

wavelet transform. The support vector machine, multilayer perceptron, convolution neural

networks, deep feed forward neural networks, and the hidden Markov model should all be

tested for classification.

This study was limited to a choice between two sleep channels. For at home studies,

there are different devices that use different channels. One device uses a two electrodes

around the Fpz electrode position. This position is basically on the forehead of a person and

would be one of the easiest positions to locate for an untrained person.

The goal of this paper is to support an at home sleep study. A sleep study needs to be

able to detect all stages, but it is very important for a sleep study to identify sleep from

wakefulness. If REM can not be identified, all stages can not be identified with a single

channel EEG.

The epoch was originally defined based on the ease of use. The complication of paper

size and time required for an analysis can be eliminated in choice of epoch size. A training

dataset may be difficult to find for a small epoch size, but each individual wave can be

classified based on the definitions provided by the AASM. This could result in epochs being

scored literally by which waves make up the majority of the epoch.

Time-frequency analysis has many new modifications for discrete wavelet transforms.

These modifications should be tested for improvements to the current discrete wavelet

transforms. Also, the continuous wavelet transform (CWT)can contain more detail than the

discrete wavelet transform (DWT). The CWT should be test against the new modifications

of the DWT.

49 References

• [1]“10/20 system | Polysomnography Study Guide.” .

• [2]10/20 System Positioning Manual. Trans Cranial Technologies.

• [3]R. B. Berry, MD (Chair) et al., AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications, vol. 2.4. American Academy of Sleep Medicine.

• [4] van S. B, K. B, K. Ha, and V. der V. Ea, “Alternative electrode placement in (automatic) sleep scoring (Fpz-Cz/Pz-Oz versus C4-A1).,” Sleep, vol. 13, no. 3, pp. 279–283, Jun. 1990.

• [5]A. Rechtschaffen and A. Kales, A manual of Standardized Terminology, Techniques and Scoring

System for Sleep Stages of Human Subjects. .

• [6]E. A. Wolpert, “A Manual of Standardized Terminology, Techniques and Scoring System for

Sleep Stages of Human Subjects.,” Archives of General Psychiatry, vol. 20, no. 2, p. 246, Feb. 1969.

• [7]F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, “A review of classification algorithms for EEG-based brain–computer interfaces,” J. Neural Eng., vol. 4, no. 2, p. R1, 2007.

• [8]J. Heaton, Artificial Intelligence for Humans Volume 3: Deep Learning and Neural Networks.

2015.

• [9]R. K. Sinha, “Artificial Neural Network and Wavelet Based Automated Detection of Sleep

Spindles, REM Sleep and Wake States,” J Med Syst, vol. 32, no. 4, pp. 291–299, Aug. 2008.

• [10]L. Fraiwan, K. Lweesy, N. Khasawneh, H. Wenz, and H. Dickhaus, “Automated sleep stage identification system based on time–frequency analysis of a single EEG channel and random forest classifier,” Computer Methods and Programs in Biomedicine, vol. 108, no. 1, pp. 10–19, Oct. 2012.

• [11]F. Ebrahimi, M. Mikaeili, E. Estrada, and H. Nazeran, “Automatic sleep stage classification based on EEG signals by using neural networks and wavelet packet coefficients,” in 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2008, pp. 1151–

1154.

• [12]L. Parrino, R. Ferri, O. Bruni, and M. G. Terzano, “Cyclic alternating pattern (CAP): The marker of sleep instability,” Sleep Medicine Reviews, vol. 16, no. 1, pp. 27–45, Feb. 2012.

•

• [13]LazyProgrammer.me, “Data Science: Deep Learning in Python.”

• [14]LazyProgrammer.me, “Data Science: Linear Regression in Python.”

• [15]LazyProgrammer.me, “Data Science: Logistic Regression in Python.”

• [16]T. Rao and D. D. Vishwanath, “Detecting sleep disorders based on EEG signals by using discrete wavelet transform,” in Green Computing Communication and Electrical Engineering (ICGCCEE), 2014 International Conference on, 2014, pp. 1–5.

• [17]D. Sundararajan, Discrete Wavelet Transform: A Signal Processing Approach. John Wiley & Sons, 2016.

• [18]P. Achermann, “EEG analysis applied to sleep,” Epileptologie, vol. 26, pp. 28–33, 2009.

• [19]“European Data Format (EDF).” [Online]. Available: <http://www.edfplus.info/>. [Accessed: 19- Jun-2017].

• [20]G. Dietsch, “Fourier-Analyse von Elektrencephalogrammen des Menschen,” Pflügers Arch., vol.

230, no. 1, pp. 106–112, Dec. 1932.

50 References

• [21]F. Hlawatsch and G. F. Boudreaux-Bartels, “Linear and quadratic time-frequency signal representations,” IEEE Signal Processing Magazine, vol. 9, no. 2, pp. 21–67, Apr. 1992.

• [22]“Machine Learning FAQ,” Sebastian Raschka’s Website. [Online]. Available:

sebastianraschka.com/faq/docs/closed-form-vs-gd. [Accessed: 10-Aug-2017].

• [23]“MethodsEEGMeasurement.pdf.” .

• [24]A. Baratloo, M. Hosseini, A. Negida, and G. El Ashal, “Part 1: Simple Definition and Calculation of Accuracy, Sensitivity and Specificity,” Emerg (Tehran), vol. 3, no. 2, pp. 48–49, 2015.

• [25]C. L. Byrne, Signal Processing, 2nd Edition, 2nd ed. CRC Press, 2014.

• [26]E. M. Ventouras et al., “Sleep spindle detection using artificial neural networks trained with filtered time-domain EEG: A feasibility study,” Computer Methods and Programs in Biomedicine, vol. 78, no. 3, pp. 191–207, Jun. 2005.

• [27]“Spectral density,” Wikipedia. 31-Jul-2017.

• [28]M. B. Kurt, N. Sezgin, M. Akin, G. Kirbas, and M. Bayram, “The ANN-based computing of drowsy level,” Expert Systems with Applications, vol. 36, no. 2, pp. 2534–2542, Mar. 2009.

• [29]S. F. Quan et al., “The association between obstructive sleep apnea and neurocognitive performance—the Apnea Positive Pressure Long-term Efficacy Study (APPLES),” Sleep, vol. 34, no.

3, p. 303–314B, 2011.

• [30]“The CAP Sleep Database.” [Online]. Available: https://physionet.org/pn6/capslpdb/. [Accessed: 24-May-2017].

• [31]M. G. Terzano, D. Mancia, M. R. Salati, G. Costani, A. Decembrino, and L. Parrino, “The Cyclic

Alternating Pattern as a Physiologic Component of Normal NREM Sleep,” Sleep, vol. 8, no. 2, pp.

137–145, Jun. 1985.

• [32]P. S. Addison, The Illustrated Wavelet Transform Handbook: Introductory Theory and

Applications in Science, Engineering, Medicine and Finance, Second Edition. CRC Press, 2017.

• [33]“The Sleep-EDF Database [Expanded].” [Online]. Available:

https:/[/www.physionet.org/physiobank/database/sleep-edfx/](http://www.physionet.org/physiobank/database/sleep-edfx/). [Accessed: 24-May-2017].

• [34]“Time-frequency Signal Analysis with Applications.” [Online]. Available: <http://eds.b.ebscohost.com.ezproxy.mtsu.edu/eds/ebookviewer/ebook/bmxlYmtfXzc1MzU4OF9fQ> U41?sid[=8234de8b-f2a6-417d-93de-597fe37750c7@sessionmgr102&vid](mailto:8234de8b-f2a6-417d-93de-597fe37750c7@sessionmgr102&vid)=6&format=EB&rid=8. [Accessed: 21-Jun-2017].

• [35]D. Álvarez et al., “Usefulness of Artificial Neural Networks in the Diagnosis and Treatment of

Sleep Apnea-Hypopnea Syndrome,” 2017.

• [36]J. R. Shambroom, S. E. Fábregas, and J. Johnstone, “Validation of an automated wireless system to monitor sleep in healthy adults,” Journal of Sleep Research, vol. 21, no. 2, pp. 221–230, Apr.

2012.

• [37]N. Crasto and R. Upadhyay, “Wavelet Decomposition Based Automatic Sleep Stage

Classification Using EEG,” in Bioinformatics and Biomedical Engineering, 2017, pp. 508–516.

• [38]L. Debnath and F. Shah, Wavelet Transforms and Their Applications. Springer, 2014.

•