POLITEHNICA UNIVERSITY OF BUCHAREST

Doctoral School of Electronics, Telecommunications and Information Technology

Decision No. 863 **from** 28-06-2022

Ph.D. THESIS SUMMARY

Eng. Alexandra-Maria TĂUȚAN

ANALIZA AUTOMATĂ A SEMNALELOR FIZIOLOGICE PENTRU DIAGNOSTICAREA ȘI MONITORIZAREA BOLILOR NEUROLOGICE

AUTOMATIC ANALYSIS OF PHYSIOLOGICAL SIGNALS FOR THE DIAGNOSIS AND MONITORING OF NEUROLOGICAL DISORDERS

THESIS COMMITTEE

Prof. Dr. Eng. Mihai Ciuc Politehnica Univ. of Bucharest	President
Prof. Dr. Eng. Bogdan Ionescu Politehnica Univ. of Bucharest	PhD Supervisor
Sr Researcher Dr. Eng. Michael Riegler Arctic University of Norway, Norway	Referee
Prof. Dr. Eng. Liviu Goraș Technical University Gheorghe Asachi, Iași	Referee
Assoc. Prof. Dr. Eng. Dragoș Țarălungă Politehnica Univ. of Bucharest	Referee

BUCHAREST 2022

Table of contents

1 Introduction **1**

1.1 Presentation of the field of the doctoral thesis 1

1.2 Scope of the doctoral thesis 1

1.3 Content of the doctoral thesis 2

I Theoretical Concepts **3**

2 Neurological Diseases **3**

2.1 Definitions 3

2.1.1 Epilepsy 3

2.1.2 Neurodegenerative diseases 3

2.2 Physiological monitoring for automatic detection 4

2.2.1 Physiological monitoring 4

2.2.2 Taxonomy of methods 4

2.2.3 General Framework 5

2.2.4 Applications 6

2.3 Conclusions 7

II Personal Contributions **8**

3 Epileptic Seizure Detection **8**

3.1 Introduction 8

3.2 Dataset 8

3.3 Supervised Feature Extraction 8

3.3.1 Time and frequency domain features 8

3.3.2 Wavelet based feature extraction 10

3.4 Unsupervised Feature Extraction 11

3.4.1 Motivation 11

3.4.2 Proposed Approach 11

3.4.3 Evaluation 12

3.4.4 Results and Discussion 12

3.5 Conclusions . . . 13
3.6 Limitations and Future Work. . . . 13

4 Automatic Sleep Staging **14**
4.1 Introduction . . . 14
4.2 Dataset . . . 14
4.3 Single channel EEG . . . 14
4.3.1 Motivation . . . 14
4.3.2 Proposed Approach . . . 15
4.3.3 Evaluation . . . 16
4.3.4 Results and Discussion . . . 16
4.4 Study on variation of input signals . . . 16
4.4.1 Motivation . . . 16
4.4.2 Proposed Approach . . . 17
4.4.3 Evaluation . . . 17
4.4.4 Results and Discussion . . . 17
4.5 Dimensionality Reduction . . . 18
4.5.1 Motivation . . . 18
4.5.2 Proposed Approach . . . 18
4.5.3 Evaluation . . . 19
4.5.4 Results and Discussion . . . 19
4.6 Conclusions . . . 19
4.7 Limitations and Future Work . . . 19

5 Freezing of Gait Detection **20**
5.1 Introduction . . . 20
5.2 Dataset . . . 20
5.3 Deep learning . . . 20
5.3.1 Motivation . . . 20
5.3.2 Proposed Approach . . . 21
5.3.3 Evaluation . . . 22
5.3.4 Results and Discussion . . . 22
5.4 Conclusions . . . 22
5.5 Limitations and Future Work . . . 22

6 Alzheimer's disease identification **23**
6.1 Introduction . . . 23
6.2 Dataset . . . 23
6.2.1 TMS-EEG . . . 23
6.2.2 Resting state EEG . . . 24
6.3 TMS-EEG based identification . . . 24

6.3.1 Motivation 24
6.3.2 Proposed Approach 24
6.3.3 Evaluation 26
6.3.4 Results and Discussion 27
6.4 Resting state EEG based identification 27
6.4.1 Motivation 27
6.4.2 Proposed Approach 28
6.4.3 Evaluation 28
6.4.4 Results and Discussion 29
6.5 Conclusions 29
6.6 Limitations and Future Work 29

7 Conclusions **30**
7.1 Obtained results 30
7.2 Original contributions 30
7.3 List of original publications 31
7.4 Perspectives for further developments 34

References **35**

Chapter 1

Introduction

This thesis presents and investigates state-of-the-art methods in the automatic analysis of neurological disorders from physiological signal recordings. Physiological recordings are a powerful tool in the diagnosis of neurological disorders as they can be applied fast and in a non-invasive manner.Automatic analysis methods can potentially simplify and improve the medical investigation process while opening up avenues for remote patient monitoring and prodromal symptom detection in population screening studies.

In this thesis, I present a state-of-the-art literature survey on available automatic analysis techniques for two specific categories of neurological disorders: epilepsy and neurodegenerative diseases. The focus is on the bridge between the various techniques used, so that they could potentially be translated from the study of one neurological disease to another. The personal contributions focus on improvements to the state-of-the-art methods as outlined in the results and proposal of each new analysis method.

1.1 Presentation of the field of the doctoral thesis

Physiological recordings such as electroencephalography (EEG), electromiography (EMG), electrocardiography (ECG) etc. are often used in the characterization of neurological disorders along with other diagnostic techniques. By applying advanced signal processing and machine learning methods to biomedical signals, the healthcare burden of neurological diseases can be reduced by: (i) *Facilitating early prodromal detection*; (ii) *Enhancing diagnosis and differential diagnosis*; (iii) *Developing remote monitoring technologies*; (iv) *Automating steps in the diagnosis process*.

1.2 Scope of the doctoral thesis

The scope of the doctoral thesis is focused on two main categories of neurological disorders: epilepsy and neurodegenerative diseases. For epilepsy, the main focus is on the automatic detection of epileptic seizures from EEG data. For neurodegenerative

diseases, the automatic detection of sleep stages is approached as a potential aid in the diagnosis of various pathologies. Furthermore, the detection of a specific Parkinson's disease motor symptom is investigated, along with the identification of Alzheimer's disease from a novel diagnostic tool.

1.3 Content of the doctoral thesis

This thesis is structured in two parts. The first part presents the theoretical concepts related to the neurological conditions addressed: epilepsy and neurodegenerative diseases. It also summarizes and discusses the state-of-the art in the analysis of these disorders using computational methods with a focus on techniques for automatic analysis. The second part presents the personal contributions to automatic analysis of epileptic seizures, automatic sleep staging, freezing of gait and Alzheimer's disease identification. Each topic is described in a separate chapter. The motivation, methods applied, experiments performed and corresponding results are presented and discussed for each topic. Finally, the conclusions, limitations of each study and proposals for future work are described. The last chapter of the thesis presents the overall conclusions and summarizes the original contributions while highlighting perspectives for further development.

Part I

Theoretical Concepts

Chapter 2

Neurological Diseases

As one of the leading causes of disability worldwide for all age groups, neurological diseases place a great burden and cost on the healthcare system and our society. This class of diseases can be severely debilitating and their impact is not only on the patients but also on their immediate family and care takers, as well as on the healthcare workers involved in the diagnosis, treatment and monitoring of the disease. In this chapter basic concepts used in this thesis are described, along with a literature review and analysis of the state of the art focused on the targeted neurological disorders.

2.1 Definitions

Neurological diseases comprise a wide range of conditions which affect the central and peripheral nervous system. This thesis looks into a functional neurological disorder (epilepsy) and some aspects of several degenerative disorders (Alzheimer's and Parkinson's disease).

2.1.1 Epilepsy

Epilepsy is a neurological disease characterized by a predisposition of the brain to generate epileptic seizures [12]. An epileptic seizures is defined as a 'transient occurrence of signs and/or symptoms due to abnormal excessive or synchronous neuronal activity' [12].

2.1.2 Neurodegenerative diseases

Neurodegenerative diseases are a complex set of disorders that typically appear in old age and cause the progressive degeneration and death of neurons from the central nervous system. The number of known neurodegenerative diseases is in the hundreds [30].

This work analyses the computational methods applied to some of the most prevalent neurodegenerative diseases: Alzheimer's Disease (AD), Frontotemporal dementia (FTD), Lewy Body Dementia (DLB), Parkinson's Disease (PD), Huntington's Disease (HD), Amyotrophic Lateral Sclerosis (ALS) and Multiple System Atrophy (MSA). The main clinical symptoms can be divided into four types: cognitive and behavioral changes, motor impairments, speech impairments and sleep disorders. One clinical symptom can be present in multiple types of neurodegenerative diseases.

Based on the primary clinical symptoms, three main categories of neurodegenerative diseases are defined: dementing, abnormal movements and a combination of dementia and abnormal movements. The abnormal movement disorders can further be divided into hypokinetic (slowness of movement) and hyperkinetic (uncontrolled motion).

2.2 Physiological monitoring for automatic detection

Recording physiological signals is a common clinical practice in vital state assessment, disease diagnosis and monitoring. In this subsection, an overview is provided on the methods most commonly used in physiological monitoring. The general framework for automatic analysis of physiological data is provided, along with a taxonomy of computational methods and relevant applications.

2.2.1 Physiological monitoring

Physiological signals are diverse and can range from information collected with sensors attached to the human body to context information such as voice or motion tracking information to medical imaging modalities. The following physiological signals are addressed in this work: *EEG - electroencephalography; TMS-EEG - transcranial magnetic stimulation combined with EEG; EMG - electromyography; EOG - electrooculography; ECG - electrocardiography; PSG - polysomnography; actigraphy; accelerometers; gyroscopes; force sensors; audio signals for voice recordings; biomedical imaging*.

2.2.2 Taxonomy of methods

Automatic analysis can have various purposes in the general study of diseases. Based on the literature survey performed, the following intentions have been identified in the automatic analysis of neurological disorders: (i) *disease diagnosis*; (ii) *differential diagnosis*; (iii) *identifying prodromal stages of disease*; (iv) *severity classification*; (v) *progression monitoring*; (vi) *treatment effectiveness monitoring*; (vii) *symptoms/abnormality detection*; (viii) *symptom characterization*.

Automatic methods for analysis of neurological diseases can be classified based on a multitude of criteria. In this work, the emphasis for their categorization includes: disease addressed, clinical symptoms, input signals and end goal of the analysis.

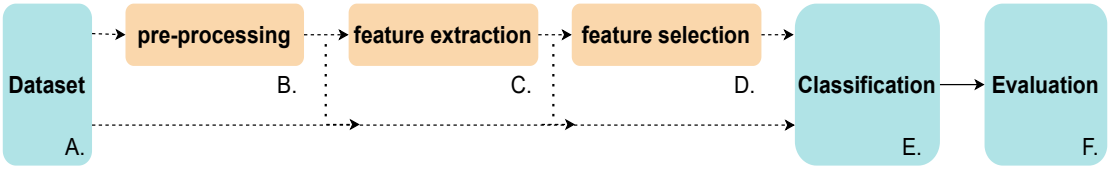

Fig. 2.1 General framework for automated analysis and classification used in problems related to neurological disorders. Blocks marked in blue (A, E, F) are always available, whereas orange blocks (B, C, D) are sometimes used depending on the problem [30]

2.2.3 General Framework

Although there is a great variety on the topics addressed with automatic analysis and neurological disorders, these computational methods have a similar structure. An overview is available in Fig. 2.1.

Datasets. The types of input data required is highly dependant on the type of problem the method aims to solve and the symptoms involved. Most often, datasets targeting a particular problem cannot be reused for another condition as the recording protocol is tailored towards the particular study aim.

Pre-processing. Feature extraction. Feature selection. In some cases, raw data can be contaminated with various types of artifacts. In order for relevant features to be extracted, adequate pre-processing methods are applied to eliminate contaminating artefacts. Feature extraction is a process highly dependent on the input data and the problem to be solved. These can be extracted either directly from raw data or after a specific domain transformation was applied.To avoid degrading the performance of the automatic analysis, feature selection methods can be applied. Dimensionality reduction techniques are used to transform the dataset in order to reduce the computational power.

Classification. In most cases, a form of model generation or classification is applied. Problems can be binary or can involve multiple classes. Most studies looking into neurological disorders make use of supervised learning techniques. These types of techniques imply that the input dataset contains both data and ground truth labels. Unsupervised learning techniques are more rarely encountered. These methods learn from the data itself and do not require a ground truth label to be provided.

Evaluation. When studies also involve machine learning methods, evaluation of the trained models is necessary. This implies splitting the input data into training and testing sets. The input dataset can be split according to the *hold out* techniques or through cross-validation. The *leave-one-subject-out* and *leave-one-record-out* methods are a particular case of cross-validation.

When evaluating the results of the model being applied on the test sets, various metrics are used. Depending on the classification problem, the choice of evaluation metric might be different. The following metrics are used throughout the work: *accuracy*, *specificity*, *sensitivity*, *area under the curve*, F_1 *score*, *Cohen's Kappa*.

2.2.4 Applications

Epileptic Seizures. **Seizure Detection and Prediction.** Features are combined with classic and neural network based algorithms. Deep neural networks are used to either extract features in an unsupervised manner or classify directly raw EEG segments. EEG signals are commonly used as a diagnosis tool for epilepsy but can also be useful in the prediction of seizures.

Cognitive & Behavioral Changes. **Cognitive Decline.** The majority of works that look into cognitive decline with machine learning methods focus on identifying mild cognitive impairement (MCI) - a prodromal symptom of different types of dementia, either from healthy controls, from more severe forms of disease or attempt to predict the conversion of the patient to a specific type of dementia [3].

Motor Impairments. **Gait, Posture &Balance.** Gait abnormalities are a clinical symptom of multiple neurodegenerative disorders, including PD, HD, MSA and ALS. Information from accelerometers, gyroscopes, force sensors or EMG sensors can be used to identify the disease from HC, give insights into the differential diagnosis or measure the severity of the disease.

Freezing of Gait (Akinesia). A common symptom in PD is the inability to follow the intention to move. Information on freezing of gait (FoG) episodes can be extracted from acccelerometer or gyroscope sensors [2].

Tremor. A common symptom in PD is uncontrollable tremor. This symptom can easily be tracked with accelerometer information. Tremor monitoring can also be used to assess the effectiveness of treatments.

Facial Muscle Rigidity. As neurodegeneration progresses into the areas of the brain controlling facial muscles, patients become unable to express the full range of emotions. This is detected through video recordings, however EMG signals can also be used.

Braykinesia. Represents a slowing down of movements which is encountered as a clinical symptom in PD. Most studies assess the severity of bradykinesia using accelerometer and gyroscopes attached to different parts of the body.

Dyskinesia. Uncontrolled movements are common symptoms in HD and can also appear in PD patients as a side effect of levodopa usage. Inertial sensors are a natural choice for tracking dyskinesia.

Speech Impairments. **Lexical Content** Signs of cognitive decline in various forms of dementia is also reflected in the content of speech of the patients. The main topics addressed include the identification of dementia from HC, the identification of the prodromal mild cognitive decline or the estimation of disease severity.

Vocalization. The control of muscles involved in speech production is also affected by neurodegeenration.Abnormal speech can be used to identify patients suffering from a neurodegenerative disorders, to enhance differential diagnosis capabilities or for the early detection of disease.

Sleep disorders. **REM Sleep Behavior Disorder.** Automatic analysis methods found in literature include the automatic detection of RBD but also methods of predicting the type of neurodegenerative disorders that would be developed later in life. The focus is on the analysis of PSG recordings as well as only EEG or EMG signals.

Restless Leg Syndrome. Automatic methods for analysis of RLS focus on the symptom characterization and its automatic detection. The characterization of RLS is performed either through the analysis of EMG or EEG signals.

Sleep/wake cycle. Alterations in the sleep architecture of patients suffering from neurodegenerative diseases is common. Both sleep macrostructure and microstructure is altered in the process of neurodegeneration.

2.3 Conclusions

This chapter provided a theoretical overview accompanied by a literature survey of the aspects presented in the next chapters of the thesis.

While conducting the general literature survey, several gaps and potential future challenges have been identified:

- Sleep in AD and in dementia in general could be better analyzed with the use of automatic analysis methods and machine learning techniques;
- Speech analysis (both lexical content and vocalization problems) could be better exploited for the differential diagnosis of PD, HD, MSA, ALS as well as for the differential diagnosis between AD and other dementia types such as FTD or DLB;
- EEG technology is not used in the study of all neurodegenerative diseases despite its great potential as a wearable tool for brain study;
- Despite the potential of TMS-EEG to be used in the study of the brain, no works have investigated its potential use in combination with machine learning techniques;
- Deep learning methods are scarcely used as the majority of the works using automatic analysis focus on classic algorithms for model creation
- Most studies focus on using hand-crafted features for clinical symptom characterization. Unsupervised feature extraction could potentially simplify the process of symptom characterization

In the next chapters of this thesis, several solutions are proposed and investigated for some of the gaps identified in this section.

Part II

Personal Contributions

Chapter 3

Epileptic Seizure Detection

3.1 Introduction

In this section, contributions to the automatic detection of epileptic seizures from EEG data are presented. Several time and frequency domain features are experimented with. Furthermore, the ability of the wavelet transform to extract more meaningful information is evaluated. Lastly, unsupervised feature extraction methods are proposed and tested.

3.2 Dataset

All contributions to the field of epileptic seizure detection have been tested on the CHB-MIT Scalp EEG database which is an open source database available on PhysioNet [14, 21]. A total of 23 recordings from 22 subjects are available (77.27% were female). All subjects had a minimum of 23 EEG signals collected at a sample rate of 256Hz and a 16-bit resolution. Annotations were provided along with the recordings.

3.3 Supervised Feature Extraction

3.3.1 Time and frequency domain features

Motivation

A significant amount of information can be extracted from the frequency content of the EEG signals. Here, several features from both the time and frequency domain are investigated to determine the most appropriate characteristic for classification of epileptic EEG signals as well as to create a benchmark algorithm for further optimization.

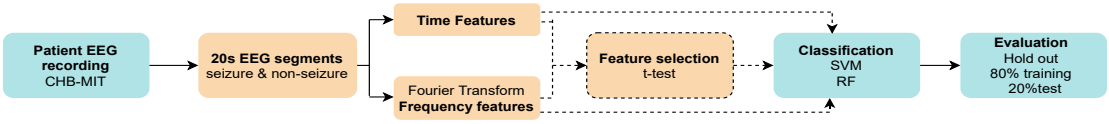

Fig. 3.1 Method for detecting epileptic seizures using time and frequency features.

Proposed Approach

The approach proposed for detecting epileptic seizures is illustrated in Fig. 3.1. A model is created for each individual patient from the dataset. For each patient the EEG data was split into non-seizure segments selected as 20 seconds prior to the annotated seizure onset and seizure segments selected as 20 seconds after the annotation. From each segment, a total of 12 features were extracted as follows:

Time Domain: *Maximum amplitude, Skewness, Kurtosis, Entropy*

Frequency Domain: EEG signals were transformed into frequency domain through Welch's method and the following characteristics were extracted: *maxPSD, maxF, meanGamma, meanBeta, meanAlpha, meanTheta, meanDelta, varPSD*

A paired t-test based feature selection is also used to identify the measures that are most helpful in clearly identifying the two states. Three types of feature combinations were used as input: (i) all 12 features from the time and frequency domain; (ii) the six selected features based on the t-test results; (iii) only maxPSD and maxF. Both an SVM and RF classification algorithm were used.

Evaluation

The evaluation of the results was performed per patient, where 80% of the data was used for training and 20% for test. The performance was quantified using accuracy, sensitivity and specificity.

Results and Discussion

The highest value for accuracy was of 69% for the experiment using a random forest classifier and only frequency features. The highest sensitivity was of 71% using all features and an SVM classifier and the highest specificity was of 67% using only frequency features and a random forest classifier. Surprisingly, the highest accuracy was obtained while using only frequency features.

Individual per patient models were created. A large variation is observed per patient model. The model accuracy reaches the highest value of approximately 90% in patient 24 and the lowest of approximately 45% in patient 15. This indicates that some of the features selected might not be effective in differentiating normal from abnormal EEG activity for some of the patients.

Fig. 3.2 Method for detecting epileptic seizures using features extracted from the detail coefficients of a 8 level Wavelet decomposition.

3.3.2 Wavelet based feature extraction

Motivation

Epileptic seizure detection from EEG data requires both time and frequency knowledge. This can be captured through the application of a time-frequency wavelet transform.

Proposed Approach

The wavelet-based method used is presented in Fig. 3.2. The EEG data is segmented in 20 second epochs. EEG segments are selected immediately prior to the start of seizure annotation for normal EEG activity and 20 seconds after for the abnormal EEG activity.

An eight level multi resolution wavelet analysis is performed based on a Daubechies mother wavelet. From each of the obtained detail level coefficients, several descriptive statistic features are extracted as follows: *maxAmp, var, std, information entropy, log entropy, skewness, kurtosis*. A total of 56 features were obtained from all detailed level coefficients. A t-test based feature selection was applied.

Classification was performed using both an SVM and RF algorithm with models being created for each patient. Experiments were performed with different combinations of input features as follows: (i) all features from all decomposition levels, (ii) all features only from decomposition levels that proved significant in the t-test selection (3, 4, 5, 6 and 7), (iii) all features except skew and kurtosis for decomposition levels 3, 4, 5, 6, 7.

Evaluation

The validation of individual models is performed using a hold-out technique. The training set represents 80% of the patient data, while the test the remaining 20%. Performance was evaluated based on the accuracy of the classification.

Results and Discussion

The highest performance of 100% accuracy was obtained for patient 9, when features selected via the t-test method (case (iii)) were used as input to a random forest classifier. However, not all patients had a similarly high performance. Overall, the performance was improved for each patient when comparing the accuracy to previous results [31].

Fig. 3.3 Method for detecting epileptic seizures using unsupervised feature extraction.

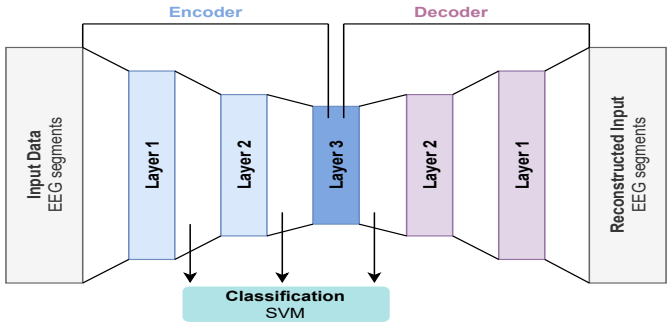

Fig. 3.4 Architecture of Autoencoder neural networks used for unsupervised feature extraction [29].

3.4 Unsupervised Feature Extraction

3.4.1 Motivation

Unsupervised learning provides an automatic method of extracting features from the data and might be helpful to build models that are more suitable for each individual patient, reducing the inter-patient variation in performance.

3.4.2 Proposed Approach

The proposed method is presented in Fig. 3.3. The EEG data is divided in equal segments prior and after the annotation provided at the start of the epileptic seizures. Experiments were performed with varying the input window size with 1s, 3s, 5s and 20s. For each segment, features are extracted using neural networks in an unsupervised manner. Two types of networks are tested: (i) based on autoencoders, (ii) based on convolutional neural networks. Individual patient models are created with a SVM classifier.

Autoencoders

The architecture of the autoencoder used for feature extraction is presented in Fig. 3.4. Dropout layers were also added to the architecture to avoid overfitting. The ReLU (rectified linear unit) activation function for all fully connected layers was used, except for the last. The last fully connected layer had a sigmoid activation function to allow the reconstruction of the input in both positive and negative values. Experiments were performed on the architecture tuning by varying:

- the input window size (therefore changing the size of the input and output layer)

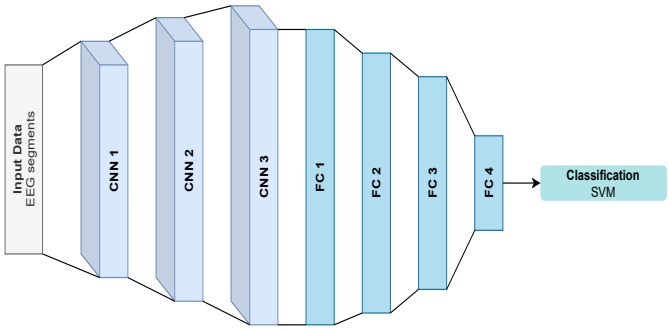

Fig. 3.5 Architecture of convolutional neural network used for unsupervised feature extraction [29].

- the addition of dropout layers with a dropout rate of 0.6 after each encoding layer
- the size of the encoding layers: (i) Layer 1 128, Layer 2: 64, Layer 3: 32 (SAE1); (ii) Layer 1: 500, Layer 2: 1000, Layer 3: 1500 (SAE2).

The values obtained after the three encoding layers are used as input to the SVM classifier.

Convolutional Neural Networks

The CNN-based network architecture for feature extraction is available in Fig. 3.5. Here, the EEG data is used as input to three convolutional layers followed by four fully connected layers. The output from the last fully connected layer is used as input to the SVM classifier. For the convolutional layers, the filter numbers are 32, 64 and 128 respectively, while the kernel size is 2 for all layers. With this type of network, the following experiments are performed:

- the input window size
- the size of the convolutional layers: (i) CNN 1: 128, CNN 2: 64, CNN 3: 32 (FC decreasing); (ii) CNN 1: 128, CNN 2: 128, CNN 3: 128 (FC constant).

3.4.3 Evaluation

The validation for each individual patient model used a leave-one-record-out cross-validation approach. The input data is split into k parts, where $k-1$ are used for training and the k^{th} for evaluation. The set is permuted until all EEG segments for one patient have been used as test. The performance was evaluated in terms of accuracy, sensitivity and specificity.

3.4.4 Results and Discussion

A summary of the best results is available in Table 3.1. For SAE2, when the hidden layer sizes are increased, the performance shows a significant improvement. The accuracy

Model	Accuracy	Sensitivity	Specificity
SAE2 - mean across models	86.21	90.10	82.31
CNN - mean across models	97.82	94.90	89.75
SAE2 - Subject 19	97.82	100	95.65

Table 3.1 Best performance obtained with unsupervised feature extraction. All presented results are reported in percentages.

value ranges from 70% when using 1s segment as input and 90% accuracy when using a 20s input window size. No significant difference in performance is observed when using features extracted from the different layers of the autoencoder model. This suggests that the quality of the unsupervised features is not dependent on the depth of the hidden layers of the Autoencoder, but on the number of points extracted from the hidden layers.

The increase in the input window length has a significant impact on performance. When using a lengthier window, the performance is higher. This could potentially be linked to the encoding of frequency information in EEG signals. The CNN based model tested for unsupervised feature extraction shows indeed an increase in classification performance. An improvement of about 10% is obtained over all input window sizes.

3.5 Conclusions

In this section, three methods of extracting features from EEG data for epileptic seizure detection were presented. All of the methods showed a good performance for the detection of epileptic EEG segments from individual patient models. The highest performance and stability in inter-patient model performance variation was obtained by using unsupervised feature extraction for classification.

3.6 Limitations and Future Work.

Although the performance obtained for individual models using the CNN based feature extraction was satisfactory, the proposed algorithm was tested only on the CHB-MIT dataset. To further validate the proposed classification methods, further testing should be conducted on several epileptic EEG datasets.

As the epileptic EEG segments can show significant variations depending on the type of seizures recorded, the features extracted in an unsupervised manner could reveal important clinical features for the characterization of epileptic episodes. Future work could focus on the relationship of the differently extracted features and the epileptic seizure types.

Chapter 4

Automatic Sleep Staging

4.1 Introduction

In this section, contributions to automatic sleep staging algorithms from polysomnographic (PSG) recordings are presented. PSG recordings are used in the assessment of sleep disorders and algorithms for their automatic scoring can help provide a faster diagnosis. Several potential improvements to automatic sleep staging algorithms are explored. The potential of using single channel EEG is investigated in in section 4.3. Section 4.4 presents the impact of several PSG signals on performance of the algorithm. Section 4.5 looks into the effects of dimensionality reduction on the model performance.

4.2 Dataset

All experiments are performed on the open source dataset available on PhysioNet "You Snooze, You Win: PhysioNet/Computing in Cardiology Challenge from 2018" [14, 13]. A total of 994 subjects (67% male) were provided. Thirteen channels were included: six EEG (F3-M2, F4-M1, C3-M2, C4-M1, O1-M2, O2-M1), EOG on the left, EMG on the chin, respiratory signals measured from a chest and abdomen belt, airflow signal, ECG and SaO2. The sample rate for all signals was of 200Hz. Seven expert annotators performed sleep stage labeling according to the latest AASM guidelines [4].

4.3 Single channel EEG

4.3.1 Motivation

PSG studies are generally cumbersome for the patient as several physiological signals need to be recorded. Simplifying the setup can make the patient more comfortable. Experiments are performed with a single channel EEG as input to investigate the potential for automatic sleep scoring. The location of the chosen EEG channel can also be relevant as different positions on the scalp record different underlying brain activity.

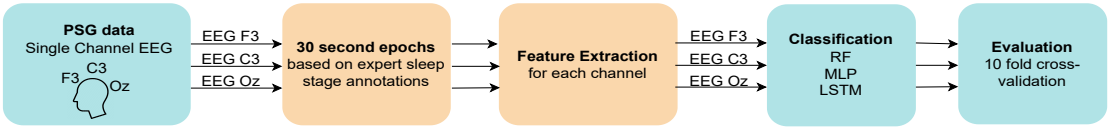

Fig. 4.1 Overview of the single channel EEG automated sleep scoring method. RF - Ranfom Forest, MLP - Multilayer perceptron, LSTM - Long Short Term Memory

4.3.2 Proposed Approach

The proposed method for creating the single channel EEG sleep scoring model is described in Figure 4.1. EEG signals from electrodes F3M2, C3M2 and O2M1 are divided into 30 second epochs according to the annotations provided for the five sleep stages. From each epoch several time and frequency domain features are extracted:

Time domain features characterize the signal statistically as follows:*mean, max, skew, kurtosis, std*

Frequency domain features analyzed five relevant EEG frequency bands [5] based on the frequency spectrum, namely *Delta 0.5-4Hz*, *Theta 4-8Hz*, *Alpha 8-12Hz*,*Sigma 12-20Hz* as follows: *mean, max, min,std, kurtosis*

Ratio of various frequency band powers were also computed. As different sleep stages are characterized by an increase in power in specific frequency bands and a decrease in others, ratios of the power spectra can help in their identification. The following ratios were computed: *delta/theta*, *theta/alpha*, *delta/alpha*.

A normal human sleep cycle contains an unequal distribution of sleep states. This implies that the database used for creating models for automatic sleep staging has an unbalanced number of classes. Unbalanced datasets can cause performance issues in classification models.To reduce the effect of the sleep stage distribution on model training, the dataset used is balanced through an undersampling technique.

After feature extraction and dataset balancing, the feature set is used as input to classification algorithms. Three algorithm types are experimented with: Random Forest (RF), Multilayer Perceptron (MLP) and Long-Short Term Memory (LSTM). After initial parameter tuning experiments, the parameters used for each network are as follows. The RF algorithm contained 10 decision trees with a minimum number of samples per leaf of 10 [32]. The MLP algorithm had an architecture comprising three hidden layers with 500 hundred units per layer. Each hidden layer perceptron had an activation *tanh* activation function [25]. The LSTM based network architecture had a single LSTM cell in a hidden layer with a design as proposed by Hochreiter and Schmidhuber [16]. A hyperbolic tangent was used as the activation function and a sigmoid function for the recurrent activation. The LSTM network was followed by a fully connected layer of size five. The

activation function for the fully connected layer was softmax. Based on experimentation, the best performance was obtained using an Adam optimizer and a binary crossentropy loss function. Furthermore, the best performance was obtained with a lookback step of 30 selected for the LSTM cell and 128 units in the hidden layer.

4.3.3 Evaluation

As the size of the dataset used for training the classification model is large, a 10-fold cross validation method was used for performance evaluation. In a k-fold cross validation, the dataset is split in *k* parts. The first *k-1* folds are used for training and the k^{th} fold is used for testing. The process is repeated until all folds have been used as a test set. The performance of the model is presented as an average of the performance metrics over all tested folds. Accuracy is used for evaluating the overall results. Precision and recall are computed for each individual class.

4.3.4 Results and Discussion

The performance obtained for the automatic sleep stage models created with EEG signals from C3, O1 and F3 was in the same range. Therefore, the results show that using a single channel EEG configuration can efficiently detected sleep stages. Either of the three electrodes can be used for this purpose.

Automatic sleep stage models were created using EEG channel F3 as input using both the unbalanced and balanced feature set. When using the balanced set, the overall accuracy improves by more than 10% and reaches 86.89%. The classification performance for individual sleep stages is improved as well, both in terms of recall and precision. The variation in performance between the different sleep stages is significantly reduced.

The best performance was obtained with the RF classifier, that showed an overall accuracy of 86.89%, whereas the lowest performance was obtained with the LSTM algorithm that presented an accuracy of 74.97%. The RF algorithm outperforms the neural network based algorithms for sleep stage classification.

4.4 Study on variation of input signals

4.4.1 Motivation

Although the classification performance using a single channel EEG is generally satisfactory for automatic sleep staging, the individual sleep stage classification performance showed room for improvement. The EEG signal records the activity of the brain. However, during sleep, other physiological signals also show differences in between sleep stages. Therefore it would be of value to investigate if adding specific signals from the PSG set would improve the performance of classifying particular sleep stages.

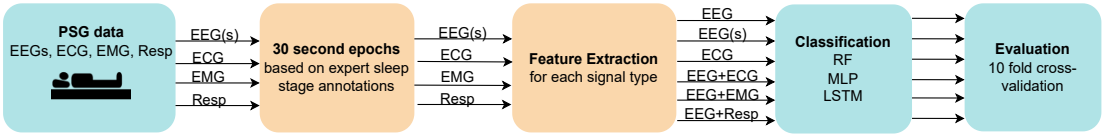

Fig. 4.2 Overview of the framework used for analyzing the impact of using different input channels or combinations of input channels for automatic sleep scoring.

4.4.2 Proposed Approach

The method used for analyzing the impact of different signal combinations as input for automatic sleep staging is available in Fig. 4.2. Similarly to the single channel EEG sleep stage classification method, PSG signals are divided into 30 second epochs according to the provided annotations. The signals considered included all six EEG channels available, ECG, EMG and all respiratory signals. Features are extracted as follows:

EEG features presented in Section 4.3.2 are extracted from all EEG channels available.

ECG *mean RR interval, BPM, TF, VLF, LF, HF, LFHF, RMSSD, SDNN, min, max, skew, kurtosis, entropy.*

EMG *mean, max, min, skew, kurtosis, variance, RMS, entropy, max freq, max PSD, mean PSD.*

Resp *mean, skew, kurtosis, variance, standard deviation, max freq, max PSD, mean PSD, nPeaks, mean distance, standard deviation of peak distance, skew of distance.*

Different variations of input signals are tested: EEG F3, all six EEG channels, ECG, EEG F3 and ECG, EEG F3 and EMG, EEG F3 and all three respiratory signals. Similar to the approach used for a single channel EEG and taking into account the results, the dataset is balanced prior to classification. Three classification algorithms are tested: RF, MLP and LSTM. The same network parameters are used as described in Section 4.3.2.

4.4.3 Evaluation

The evaluation of the methods developed with different input signals was performed using the same approach as described in Section 4.3.3.

4.4.4 Results and Discussion

RF shows higher classification accuracy for all input signal types, while LSTM shows the lowest. In terms of specific sleep stage performance, there is a variability between the different classifiers. For the RF classification, the best performance was obtained when using the EEG signal as input combined with the respiratory signals, followed by the EEG signal combined with the EMG. Adding the EMG signal improves

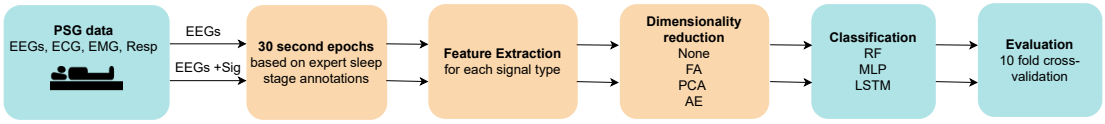

Fig. 4.3 Overview of the framework used for analyzing the impact of using dimensionality reduction techniques on the automatic sleep scoring algorithm performance. FA - Factor Analysis, PCA - Principal Component Analysis, AE -Autoencoder.

classification performance for the REM sleep stage. Adding the respiratory signals improve the precision and recall of stage N2. The precision of classifying stage W and the recall for classifying stage REM are the lowest when using only the ECG signal as input. This indicates the inability of the algorithm to detected changes in the two similar sleep stages only from heart rate information.

Interestingly, the same trends are not maintained when using the MLP and LSTM network. By extracting features, some of the temporal dependancy between the epochs might be lost which in turn will cause the LSTM lookback mechanism to not function properly. Further experiments would be needed to investigate the lower performance.

4.5 Dimensionality Reduction

4.5.1 Motivation

Datasets for sleep staging are typically large as they include hours of PSG recordings from several subjects. Furthermore, each PSG study is split into 30 second epochs which results in a significant number of instances. The high dimensionality of the data can require extensive computing power. Various dimensionality reduction methods can be applied to simplify computations.

4.5.2 Proposed Approach

The overview of the approach combining previously presented automatic sleep stage algorithms and dimensionality reduction is presented in Fig. 4.3. To study the effect of dimensionality reduction, two larger feature sets are considered. These included the features extracted from: (i) all six EEG signals - EEGs, (ii) all available PSG signals. After balancing the dataset as previously described, three different methods for reducing dimensionality are applied: (i) factor analysis (FA) [17], (ii) principal component analysis (PCA) [11], (iii) autoencoders [19].

For FA, the number of factors was varied between 3, 5, 30 and 50. For PCA,the number of principal components is also varied between 3, 5, 30 and 50. The autoencoder had only a hidden layer to compress the data with sizes between 3, 5, 30 and 50. Additionally, classification models were created when no dimensionality reduction was applied for creating a baseline.

4.5.3 Evaluation

The performance evaluation was conducted as described in Section 4.3.3.

4.5.4 Results and Discussion

Results showed that the highest number of components from either technique provided better classification results [33]. The results vary significantly depending on the type of classifier used to create the automatic sleep scoring models. When using the RF algorithm, the performance obtained is increased when using PCA with respect to the case when no dimensionality method is applied. For EEGs, the performance increases from 72.98% accuracy to 84.94% accuracy. For EEGs+Sig the performance increases from 87.97% to 91.25%. When looking at EEGs used as input for the MLP and LSTM network, FA is in the lead in terms of performance. Factor analysis also shows improvements in classification when using EEGs+Sig as input.

In summary, principal component analysis and factor analysis showed the most promising results as dimensionality reduction techniques as they not only improved the computational effort but also increased the performance. This suggests that some of the features used as input contain redundant information that is filtered out through the correlations and projections used in these two dimensionality reduction methods.

4.6 Conclusions

In this chapter, several methods for automatic sleep stage detection were presented along with techniques of improving performance and advancing state of the art. Automatic sleep scoring algorithms using a single channel EEG showed good prospects in terms of performance.Several techniques for improving the training classification performance were investigated. Balancing the data from the training set with an undersampling technique improved the per class performance. Using single channel EEG with various other input PSG signals was also investigated as a method of improving specific sleep stage performance. Finally, dimensionality reduction methods can be successfully used to decrease the computational cost of the training phase.

4.7 Limitations and Future Work

The proposed methods for automatic sleep staging were trained and tested on one single dataset. Although the dataset used was large, data from all subjects was collected in similar circumstances. Training the algorithms and improving performance using a single dataset might not be representative for other real-world scenarios where different recording equipment and controls are used.

As the manual selection of features might lead to a poor characterization of the input signals in terms of sleep stages, unsupervised feature extraction can be used for classification with deep neural networks using raw physiological recordings directly as input. Annotations were provided along with the datasets however little information was available on the number of annotators and the annotation process used was provided. More investigations should be performed on the impact of scorer performance.

Chapter 5

Freezing of Gait Detection

5.1 Introduction

In this section, contributions to the automatic identification of Parkinson's Disease freezing of gait moments from accelerometer data are presented. Freezing of Gait (FoG) episodes are sometimes reported by PD patients. During these events, patients exhibit an inability to move. FoG episodes can be detected from various accelerometer sensors placed on the body of the patient.

5.2 Dataset

Experiments on the detection of FoG from accelerometer data were performed on the *Daphnet Freezing of Gait Data Set* [2]. Three wearable sensors incorporating accelerometers were placed on the shank (ankle sensor), on the thigh (leg sensor) and one attached to the belt (waist sensor). The data was collected with a sample frequency of 64Hz. A total of 10 Parkinson's disease patients were included in the study (30% female) [2]. Annotations were provided for regular walking and freezing of gait episodes.

5.3 Deep learning

5.3.1 Motivation

The detection of FoG events using machine learning is useful to characterize PD progression as well as to help the patient overcome these episodes. In order to provide support in coming out of FoG episodes, a fast reaction time is needed. Therefore, the possibility of detecting FoG events from different input time windows while maintaining classification performance is investigated.

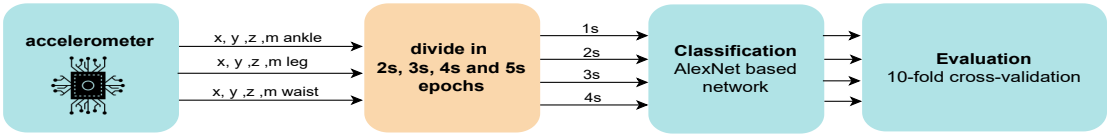

Fig. 5.1 Overview of the framework used for the classification of FoG detected with accelerometer data (x, y, z, m - magnitude) from the ankle, leg and waist sensors based on deep learning.

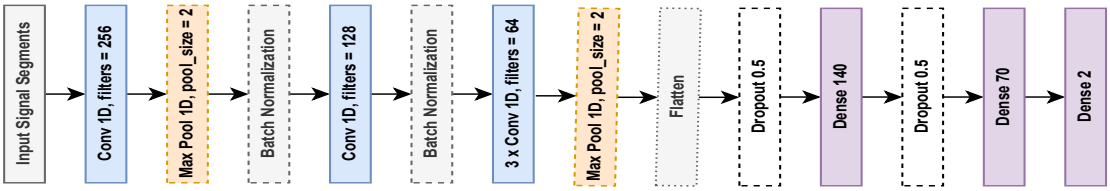

Fig. 5.2 Architecture of the AlexNet inspired deep neural network for identification of Freezing of Gait episodes from accelerometer data [26]

5.3.2 Proposed Approach

The overview of the approach used to detect FoG events from accelerometer data is presented in Fig. 5.1. All accelerometer input signals available in the dataset are considered. The three accelerometer sensors placed on the waist, leg and ankle have three axes representing the direction of motion x, y and z. The magnitude m is also computed as the square root of the sum of squares of the three axes signals [26].

In order to eliminate unwanted interference and unnecessary content the signals are filtered with a butterworth bandpass filter between 0.1 and 15Hz [**?**]. In order to remove inter-subject variability in the signal amplitude, all input signals are normalized to the maximum signal value.

All of the signals are divided into epochs by using a sliding window. Experiments were performed with windows of 2s, 3s, 4s and 5s lengths. Each segment is labeled either as normal walking or an FoG event. Since the number of normal walking segments is significantly different than that of FoG events, an undersampling dataset balancing technique is applied. The obtained segments are fed directly into a deep neural network for classification. The architecture of the classification network is presented in Fig. 5.2. The network is an adaptation of the AlexNet network [1]. An Adam optimizer with mean squared logarithmic error loss were used for model compilation. Batches of 32 samples and 50 epochs were used for training the model. The input layer size depended on the size of the input segment.

5.3.3 Evaluation

The performance of the classification model is evaluated using a 10-fold cross-validation method. The metrics quantifying performance were accuracy, sensitivty and specificity.

5.3.4 Results and Discussion

Overall, the performance was high for all input signals and sensors with accuracy values between 80.5% and 91.73%. Compared to previous work, the performance is higher or within the same range. The original work of Bachlin et al. [2] used a simple threshold method and obtained a sensitivity and specificity of 69% and 94% respectively.

With respect to input signal used, there is no trend identified. In terms of sensors that show a better performance, the leg and ankle sensors consistently show an increased accuracy when compared to the waist sensor. In terms of input window size, the best performance is obtained generally with a 5s window. However, the difference in performance is within 2%. Therefore, it is also acceptable to use a 2s window as input for classification.

5.4 Conclusions

In this chapter, a method for classifying freezing of gait events from accelerometer data was proposed. A neural network inspired by the AlexNet was adapted to accelerometer data. The leg and ankle sensors showed a slightly better performance. Using an input window size of 2s showed satisfactory results in the identification of FoG events.

5.5 Limitations and Future Work

Although the performance obtained when classifying FoG directly from raw accelerometer data was good, there is still room for improvement. The dataset used for training the classification models contains a small number of subjects. In order to gain more statistical significance and to have a better representation of real-world scenarios a larger dataset should be collected for training.

Lastly, the classification algorithm was adapted from AlexNet [1]. However, the classification model parameters should be optimized to the new dataset. Future work will focus on further optimizing the proposed model as well as implementing other deep learning algorithm architectures such as VGGNet, LeNet etc.

Chapter 6

Alzheimer's disease identification

6.1 Introduction

Transcranial Magnetic Stimulation combined with EEG (TMS-EEG) has proven a useful tool in the study of the brain and its response to external, electric stimulation both in healthy and pathological states [18]. By applying a focused magnetic pulse on specific regions of the brain, a targeted electrical stimulation is provided. The propagation of the electric impulse through the brain as recorded through EEG can reveal information on reactivity of the brain to external perturbation. In this chapter, the potential of TMS-EEG as an AD diagnostic tool is investigated. Section 6.3 describes a method of creating machine learning models to identify AD patients. Section 6.4 describes a similar algorithm but applied to resting state EEG for comparison.

6.2 Dataset

The dataset used in the AD identification experiments was a private dataset collected at the Santa Lucia Foundation (Rome, Italy) between January 2014 and June 2020 [24]. The study was approved by the ethics committee of the Santa Lucia Foundation according to the principles defined in the Declaration of Helsinki. Participants were screened according to the latest AD criteria. A total of 76 patients chose to participate and were eligible. Both TMS-EEG and resting state EEG data was collected at the time of enrolment. Due to the quality of the data, some of the recordings were excluded from further analysis.

6.2.1 TMS-EEG

A magnetic biphasic stimulator connected to a figure-of-eight coil with a diameter of 70mm that can generate 2.2T as maximum output was used to apply transcranial magnetic stimulation. The left-dorso lateral pre-frontal cortex (DLPFC) was targeted by the stimulation.The intensity of the stimulation of single pulse TMS was set at 90% of

the adjusted motor threshold. A block of 120 single-pulse TMS with an inter-stimulus interval of 1 to 4 seconds was applied. EEG was recorded during and after the TMS protocol with 62 electrodes. EEG data collected during TMS stimulation is contaminated by many sources of interference and therefore several pre-processing steps are applied.

Based on the availability of specific electrode data and the quality of the EEG, a total of 38 AD patients (60% female, average mini-mental state score 19) and 17 healthy controls (58% female, average mini-mental state score 29) were included in the experiments for automatic identification of AD participants.

6.2.2 Resting state EEG

For some of the participants, resting state EEG was also collected prior to the application of the TMS protocol. Two minutes of EEG recordings were obtained for each of the two recording conditions: the participants had their eyes open and closed. Although resting state EEG is susceptible to less artefact sources than TMS, it can still be affected by muscle movements or blinks. Several pre-processing steps were applied.

For resting state EEG, less healthy participants were recorded. In the end, only 37 AD subjects and 10 HC were included for both open eyes and closed eyes conditions of the protocol.

6.3 TMS-EEG based identification

6.3.1 Motivation

The use of TMS-EEG on Alzheimer's disease patients can be useful in providing further information for differential diagnosis with other dementia conditions, identifying the disease at incipient stages or further characterizing the disease and extracting useful biomarkers for disease progression. TMS-EEG as a physiological response has not been previously combined with machine learning as an analysis technique for the study of Alzheimer's disease. Therefore, this section presents a method of identifying AD patients from HC subjects using TMS-EEG and creating a machine learning model.

6.3.2 Proposed Approach

The proposed approach for identifying AD patients from TMS-EEG data is presented in Fig. 6.1. Here, the baseline is considered 500ms to 200ms prior to the TMS pulse. Due to the removal of TMS artefacts and subsequent filtering, 200ms prior to the pulse are discarded to eliminate the influence of the the final filtering steps from the pre-processing. The TMS-EEG response is considered 1000ms after the delivery of the TMS pulse.

After segmenting the signal according to the various TMS-EEG trials, features are extracted for each individual trial in the time and frequency domain as follows:

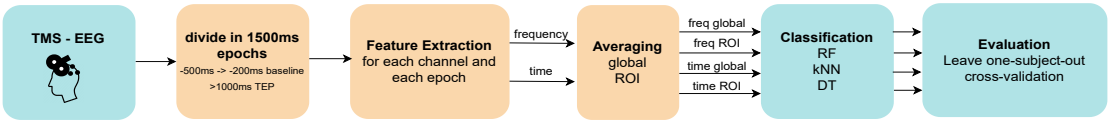

Fig. 6.1 Overview of the framework used for the identification of Alzheimer's disease patients from healthy control subjects from TMS-EEG data. ROI - region of interest, RF - random forest, kNN - k-neareest neighbours, DT - decision tree.

Time domain - Prior to extracting the time domain features, the signals (baseline and TEP) are normalized to have values between -1 and 1. The descriptive statistics, Hjorth parameters and signal energy features are computed over the TEP, but they can also be normalized to baseline. The extracted features are as follows:

Descriptive statistics - characterizing the TEP signal after the TMS pulse : *maximum, minimum, mean, skew, kurtosis*.

Hjorth parameters - statistical features based on the variation of the signal to provide insights into the frequency content of the signal without applying additional frequency transforms [15]: *activity, mobility, complexity*.

Signal energy - the signal changes after the TMS pulse can also be characterized through the mean signal energy [27].

TEP peaks - Due to the high inter-individual variability, the peaks were computed in specific time windows that were also investigated in prior literature for characterizing AD. The maximum value in the selected window is detected and a short window around it is used to calculate the average around this maximum [20]. The following peaks are detected: *P1 - peak in between 25-40ms, P2 - peak in between 45-80ms, P3 - peak in between 85-150ms, P4 - peak in between 160-250ms*

Mean Field Power - is a measure of power in between different electrodes. It can be computed globally over all electrodes or for specific regions of interest, including only particular electrodes according to equation 6.1 [23].

$$MFP(t) = \sqrt{\frac{\sum_j^M \left(x_j(t) - x_{mean}(t)\right)^2}{M}} \tag{6.1}$$

where t - time, x_j - signal from the $j^{th} channel$,x_{mean} - mean signal over the selected channels, M - total number of channels. For the global mean field potential (GMFP) M equals to all electrodes available, for the local mean field potential (LMFP) M equals the number of the electrodes selected in a particular region of interest. For each GMFP and LMFP signal the area under the curve is extracted.

Frequency domain - the power spectrum is computed for each individual trial [10] and the z-score is extracted from the TEP with respect to baseline. The following were extracted from each trial:

- mean value of the power spectrum after the TMS pulse in the delta (δ - 0.5-3Hz), theta(θ - 3-6Hz), alpha (α - 7-12Hz), beta (β - 13-30Hz), low gamma (γ_L - 31-60Hz) and high gamma (γ_H - 61-100Hz).
- the ratio between mean power spectral powers according to equation 6.2 [28]:

$$r = \frac{\alpha + \beta}{\delta + \theta} \tag{6.2}$$

- the mean value of the power spectrum in the delta, theta, alpha, beta, low gamma and high gamma in specific windows: 10-80ms (W1), 80-150ms (W2), 150-500ms (W3), 500-1000ms (W4).

All features extracted per trial are averaged per channel. The values obtained for each channel are averaged globally over the entire electrode set or are averaged over specific regions of interest (ROI). The following regions of interest are defined: FL - frontal left (F1, F3, F5. FC1, FC3, FC5), FR - frontal right (F2, F4, F6, FC2, FC4, FC6), CL - central left (C1, C3, C5), CR - central right (C2, C4, C6), CPL - central parietal left (CP1, CP3, CP5, P1, P3, P5), CPR - central parietal right (CP2, CP4, CP6, P2, P4, P6), POL - parietal occipital left (PO3, PO7, O1), POR - parietal occipital right (PO4, PO8, O2), TL - temporal left (T7, TP7, P7), TR - temporal right (T8, TP8, P8).

Separate models are created for the time and frequency features. Different combinations of features are used as input for the classification models between AD and HC: only global, only ROI based, both global and ROI based.

The dataset used for the experiments is unbalanced: more AD patients are available than healthy controls. This can cause the classifier to create a model with a classification preference towards the majority AD class. The Synthetic Minority Oversampling Technique (SMOTE) [7] was applied for data balancing.

Classification models are created with three different types of classic algorithms: random forest (RF), k-nearest neighbour (kNN) and decision tree (DT). The RF algorithm was created using 100 trees in the forest with a minimum number of one sample per leaf. kNN was used with 7 neighbours and a ball tree distance computation. The DT algorithm used entropy as a criterion for the branch split. To avoid the variability introduced with some of the classification algorithms, the models are run 100 times.

6.3.3 Evaluation

Since the dataset is relatively small, a leave-one-subject-out classification method was used. The k^{th} subject is kept for testing while the rest of the $k-1$ are used for training. Subjects are rotated until all of them have been used as a test. The performance is evaluated by calculating the accuracy, sensitivity, specificity and F_1 score. The algorithms are run 100 times and the metrics presented are computed as an average.

Model	Accuracy	Sensitivity	Specificity	F_1 score
Time Global	92.05	96.15	87.94	92.03
Time ROI	88.34	95.73	80.94	88.26

Table 6.1 Best performance obtained on time domain models for identifying AD from TMS-EEG data. All presented results are reported in percentages.

6.3.4 Results and Discussion

The best performance for the time domain models is presented in Table 6.1. The highest values were obtained using the RF classifier on global features. This could potentially indicate that AD is a disorder that affects the entire brain. Although particular regions are more affected and can help in its identification and characterization, better results are obtained when analyzing the entire electrode set [24, 27].

For the frequency models, the highest accuracy of 90.78% is obtained when combining global and ROI based features with a DT classifier. The highest sensitivity of 100% is obtained with a kNN algorithm applied on ROI based features. The performance of the different classification algorithms when using frequency-based input features varies within the range of 66.2% and 90.78% accuracy. There is no conclusive result regarding the best performance in terms of classification method.

For the frequency models, the results obtained with the globally averaged and ROI based averages are within the same range. Alzheimer's disease is known as a disconnection syndrome: long range network connections in the brain are lost first [9], but local neuronal malfunctions can also occur due to hyperexcitability and altered neuronal communication [8]. These altered connectivity patterns might be reflected in the different frequencies of the neuronal oscillations as recorded by the EEG sensors. Hence, both globally averaged as well as ROI averaged EEG frequency can potentially be a biomarker of neurodegeneration in Alzheimer's disease [28].

6.4 Resting state EEG based identification

6.4.1 Motivation

The combination of TMS-EEG and machine learning has not been previously used in the study of Alzheimer. Therefore, the performance and results obtained and presented in section 6.3 are difficult to compare with available literature. As resting state EEG was also collected for some of the participants, creating classification models is possible and can serve as a reference for the classification of AD patients using TMS-EEG.

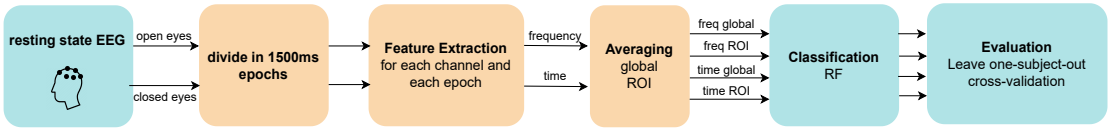

Fig. 6.2 Overview of the framework used for the identification of Alzheimer's disease patients from healthy control subjects from resting state EEG data. ROI - region of interest, RF - random forest, kNN - k-neareest neighbours, DT - decision tree.

6.4.2 Proposed Approach

The method used for AD identification from HC is described in Fig. 6.2. Resting state EEG is divided in epochs of 1500ms to resemble the division of the TMS-EEG data. Different effects are expected on resting state EEG data collected while the subject had open eyes than when the subject kept the eyes open. Therefore, experiments are performed for both recording conditions.

From each epoch, several time and frequency domain features are extracted. The same features are used as for TMS-EEG. In this case, specific TMS-EEG features such as the TEP peaks are excluded as there is no evoked response to be analyzed. The features included in the analysis are as follows:

Time domain - computed for each epoch:

Descriptive statistics - *maximum, minimum, mean, skew, kurtosis*.
Hjorth parameters - *activity, mobility, complexity*.
Signal energy - over the selected epoch.
Mean Field Power - is calculated as described in equation 6.1 both globally and locally [23]. For each GMFP and LMFP signal the area under the curve is extracted.

Frequency domain - the power spectrum is computed for each individual trial [10]. The same features as in Section 6.3.2 are extracted.

As in the case of TMS-EEG data, all features are extracted per epoch and then averaged per channel and subject. The values obtained per channel are averaged both on a global level as well as on specific regions of interest. The ROIs are defined in Section 6.3.2.

Different models are created for the frequency and time features. Similarly, features used as input can be global, ROI averaged or a combination of the two. In the case of resting state EEG, the dataset also presents an unequal number of subjects per class. Therefore, here as well the SMOTE [7] techniques was used for dataset balancing (see Section 6.3.2). As better results were obtained with the RF classifier on TMS-EEG data, here only the RF classifier is applied for comparison purposes.

6.4.3 Evaluation

The same evaluation methods as described in Section 6.3.3 were used here as well.

6.4.4 Results and Discussion

The best performance was obtained with resting state EEG collected during closed eyes using a combination of globally and ROI averaged feature from time domain. The classification performance was of 94.66%, 97.58%, 91.75% and 94.66% for accuracy, sensitivity, specificity and F_1 score respectively. The performance obtained with resting state EEG is similar to that obtained with TMS-EEG data. This confirms that both methods are relevant in the problem of AD identification.

The best performance is obtained when combining globally and ROI averaged features. This is unlike the case of TMS-EEG data classification, when averaging globally showed a better performance. In the case of TMS-EEG, the generated TMS pulse travels though different brain networks that might be affected by AD. Therefore, having a view over the entire brain could potentially be better in capturing these altered neuronal communication paths. In the case of resting state EEG, the communication between different brain hubs might not be captured as there is no electrical signal induced and transmitted. Therefore local alterations in connectivity and neuronal activity might be a better predictor of AD when looking at resting state EEG data.

6.5 Conclusions

In this chapter, methods of identifying AD patients from HC subjects through TMS-EEG data was investigated. Input features from both time and frequency domain were explored, with averaging applied globally or on specific regions of interest. The performance of the classification models created using various features and methods of combining them has been satisfactory. As a comparison, AD classification models were created from resting state EEG data as well. The performance of models created from TMS-EEG and resting state EEG was comparable. Therefore TMS-EEG is a good method of identifying Alzheimer's disease patients.

6.6 Limitations and Future Work

The TMS-EEG dataset used contained 38 AD patients and 17 HC patients. The resting state EEG data contained 37 AD patients and only 10 HC. Although a data augmentation balancing technique was applied, this method is not necessarily representative for a real-world scenario. More healthy control subjects should be used in the classification to match the number of Alzheimer's disease patients.

Alzheimer's disease patients present altered network connectivity patterns. Network connectivity metrics could potentially be used as input features to classification models of AD. Future work would involve using connectivity metrics for AD identification.

Chapter 7

Conclusions

This thesis addresses various methods of using automatic detection tools for the diagnosis, treatment and monitoring of neurological disorders, more specifically for epilepsy and several other neurodegenerative disorders. Several approaches were proposed and investigated with the aim of improving the state-of-the-art. Focus was placed on algorithms for epileptic seizure detection from EEG data, automatic sleep staging from multiple polysomnographic recordings, freezing of gait detection from raw accelerometer data and Alzheimer's disease identification from healthy control based on TMS-EEG data.

7.1 Obtained results

For each of the topics presented in this thesis, the summarized findings are as follows. Chapter 2 presents the general concepts used in this thesis along with an analysis of the state-of-the-art. Several gaps and opportunities for research are identified. The second part of the thesis describes personal contributions and addresses some of the gaps identified in the first part. Chapter 3 describes several methods for person-specific classification of epileptic seizures from EEG data based on different methods of feature extraction. Chapter 4 describes several methods applied to the automatic classification of sleep stages and investigations carried out for performance improvements. Chapter 5 presents a deep learning method applied directly to raw accelerometer data for the detection of freezing of gait in Parkinson's disease patients. Chapter 6 describes a method of using TMS-EEG for Alzheimer's disease identification and characterization. To sum up, the results obtained are promising and show improvements or alternatives to the state-of-the art in the automatic analysis of epilepsy and neurodegenerative disorders.

7.2 Original contributions

The original contributions brought to the state-of-the-art are as follows:

1. In (C1) and (C3) various classic feature based algorithms for classification of epileptic seizures from EEG data are proposed and investigated. In (C4), the state-of-the-art is ex-

tended through the use of unsupervised features extracted from various neural network architectures. The obtained results showed a good classification performance for individualized models compared to the state of the art. Furthermore, better inter-individual model performance is obtained with the used of the proposed feature extraction methods as compared to using classic techniques.

2. In (C5), (C7) and (J3) experiments are performed with various physiological signal combinations used as input for automatic sleep staging algorithms. Using a single channel EEG as input showed a similar performance or even higher performance depending on the classifier used when compared to using multiple signals as input and compared to the results found in literature.
3. In (J2) several dimensionality reduction methods were experimented with when using a large dataset and various physiological recordings as input. The use of factor analysis helped not only for reducing the dimensionality of the data, but also improved the performance of the classification with more than 10% in some cases.
4. In (C6) a classification algorithm that can detected Freezing of Gait events directly from raw accelerometer data was proposed. The obtained performance was high especially when using as input sensor data collected from the limbs. The results were stable even when the input window size had a length of only 2 seconds.
5. In (C10), (J4) and (J5) a novel method of identifying and characterizing Alzheimer's disease pathology from TMS-EEG data was proposed. For all classification algorithms using various combinations of features extracted, the classification performance was above 70% and is comparable to that obtained with similar techniques on resting state EEG and the performance described in the state-of-the-art.
6. In (C11) a new method of characterizing the duration of the TMS evoked response on the EEG is proposed. The algorithm for obtaining the target engagement duration index is presented and exemplified on the problem of classifying Alzheimer's disease patients from healthy controls. The TMS induced perturbation on the EEG signal is analyzed based on the mean signal energy. No previous works attempt the characterization of TMS responses by evaluating the variation in terms of signal energy with respect to baseline.

7.3 List of original publications

In this section, a list of all published papers and works submitted and/or under review is presented.

Journal papers :

(J1) **A.M. Tăuțan**, A.C. Rossi, R. De Francisco, B. Ionescu: Dimensionality reduction for EEG-based sleep stage detection: Comparison of autoencoders, principal

component analysis and factor analysis. Biomedical Engineering/Biomedizinische Technik 66 (2), 125-136, ISSN:1862-278X, DOI:https://doi.org/10.1515/bmt-2020-0139, October 2020. (Q3 journal, **Impact Factor: 1.404**, WOS: 000634935600002)

(J2) **A.M. Tăuțan**,B. Ionescu, E. Santarnecchi: Artificial intelligence in neurodegenerative diseases: A review of available tools with a focus on machine learning techniques. Artificial Intelligence in Medicine 117, 102081, ISSN:0933-3657, DOI:https://doi.org/10.1016/j.artmed.2021.102081, July 2021. (Q1 journal, **Impact Factor: 5.326**, WOS: 000661230700001)

(J3) **A.M. Tăuțan**, A.C. Rossi, B. Ionescu: Automatic sleep scoring with LSTM networks: impact of time granularity and input signals. Biomedical Engineering/ Biomedizinische Technik, ISSN:1862-278X, DOI: http://dx.doi.org/ 10.1515/bmt-2021-0408, June 2022. (Q3 journal, **Impact Factor: 1.404**, WOS:000806164300001)

(J4) (Under review) **A.M. Tăuțan**, E.P. Casula, M.C. Pellicciari, I. Borghi, M. Maiella, S. Bonni, M. Minei, M. Assogna, A. Palmisano, C. Smeralda, S.M. Romanella, B. Ionescu, G. Koch, E. Santarnecchi. TMS-EEG perturbation biomarkers for Alzheimer's disease identification. Scientific Reports, ISSN:2-45-2322. (Q1 journal, **Impact Factor: 4.379**, WOS pending)

(J5) (In preparation) **A.M. Tăuțan**, E.P. Casula, M.C. Pellicciari, I. Borghi, M. Maiella, S. Bonni, M. Minei, M. Assogna, A. Palmisano, C. Smeralda, S.M. Romanella, B. Ionescu, G. Koch, E. Santarnecchi. Classification of Alzheimer's Disease patients using frequency content of TMS-EEG data.

Conference papers :

(C1) **A.M. Tăuțan**, A.I. Munteanu, D.D. Țarălungă, R. Strungaru, G.M. Ungureanu: Automated Classification of Epileptiform Discharges in EEG Signals Using the Wavelet Transform. International Conference and Exposition on Electrical And Power Engineering, ISBN:978-1-5386-5062-2, DOI:10.1109/ICEPE.2018.8559773, Iași, Romania, October 18-19, 2018 (WOS:000458752200171, IEEE Xplore).

(C2) D.D. Țarălungă, **A.M. Tăuțan**, G.M. Ungureanu: An Efficient Method for Fetal Heart Sounds Detection Based on Hilbert Transform. International Conference and Exposition on Electrical And Power Engineering, ISBN:978-1-5386-5062-2, DOI:10.1109/ICEPE.2018.8559893, Iași, Romania, October 18-19, 2018 (WOS:000458752200179, IEEE Xplore).

(C3) **A.M. Tăuțan**, I. Mândruță, O.A. Băjenaru, R. Strungaru, D.D. Țarălungă, B. Hurezeanu, G.M. Ungureanu: The automatic detection of epileptic seizures based on EEG signals processing: investigation of different features and classification algorithms. World Congress on Medical Physics and Biomedical Engineering, IFMBE Proceedings, ISSN:1680-0737, DOI:10.1007/978-981-10-9038-7_74, 2019 (WOS:000449742700074, Springer).

(C4) **A.M. Tăuțan**, M. Dogariu, B. Ionescu: Detection of epileptic seizures using unsupervised learning techniques for feature extraction. 41st Annual International

Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), ISSN:1558-4615, DOI:10.1109/EMBC.2019.8856315, Berlin, Germany, July 23-27, 2019 (WOS:000557295302184, IEEE Website).

(C5) **A.M. Tăuțan**, A.C. Rossi, R. De Francisco, B. Ionescu: Automatic sleep stage detection using a single channel frontal EEG. E-Health and Bioengineering Conference (EHB). ISSN: 2575-5145, DOI:10.1109/EHB47216.2019.-8969973, Iași, Romania, November 21-23, 2019 (WOS:000558648300105, IEEE Website).

(C6) A.G. Andrei, **A.M. Tăuțan**, B. Ionescu: Parkinson's disease detection from gait patterns. E-Health and Bioengineering Conference (EHB). ISSN: 2575-5145, DOI:10.1109/EHB47216.2019.8969942, Iași, Romania, November 21-23, 2019 (WOS:000558648300074, IEEE Website).

(C7) **A.M. Tăuțan**, A.C. Rossi, R. De Francisco, B. Ionescu: Automatic sleep stage detection: a study on the influence of various PSG input signals. 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), ISSN: 2694-0604, DOI:10.1109/EMBC44109.2020.-9175628, Montreal, QC, Canada, July 20-24, 2020 (WOS:000621592205159, IEEE Website).

(C8) **A.M. Tăuțan**, A.G. Andrei, B. Ionescu: Freezing of gait detection for parkinson's disease patients using accelerometer data: Case study. International Conference on e-Health and Bioengineering (EHB), ISBN: 978-1-7281-8804-1, DOI:10.1109/EHB50910.2020.9280223, Iași, Romania, October 29-30, 2020 (WOS:000646194100095, IEEE Website).

(C9) A. Ciurea, C.P. Manoila, **A.M. Tăuțan**, B. Ionescu: Low latency automated epileptic seizure detection: Individualized vs. Global approaches.International Conference on e-Health and Bioengineering (EHB), ISBN: 978-1-7281-8804-1, DOI:10.1109/EHB50910.2020.9280267, Iași, Romania, October 29-30, 2020 (WOS:000646194100135, IEEE Website).

(C10) **A.M. Tăuțan**, E.P. Casula, I. Borghi, M. Maiella, S. Bonni, M. Minei, M. Assogna, B. Ionescu, G. Koch, E. Santarnecchi: Preliminary study on the impact of EEG density on TMS-EEG classification in Alzheimer's disease. 44th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Glasgow, Scotland, United Kingdom, July 11-15, 2022 (WOS pending, IEEE Website).

(C11) **A.M. Tăuțan**, E.P. Casula, I. Borghi, M. Maiella, S. Bonni, M. Minei, M. Assogna, B. Ionescu, G. Koch, E. Santarnecchi: Characterizing TMS-EEG perturbation indexes using signal energy: initial study on Alzheimer's Disease classification. 44th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Glasgow, Scotland, United Kingdom, July 11-15, 2022 (WOS pending, IEEE Website).

7.4 Perspectives for further developments

Future work should focus on further developing the proposed algorithms into forms that allow their translation to different problems. For instance, unsupervised feature extraction could be applied to other classification problems involving physiological data. Furthermore, the current work focused mostly on disease or symptom detection. The use of the same input signals combined with various methods should be applied to differential diagnosis problems.

References

[1] Alex, K., Ilya, S., and Geoffrey, H. (2017). ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, 60:84–90.

[2] Bachlin, M., Plotnik, M., Roggen, D., Maidan, I., Hausdorff, J. M., Giladi, N., and Tr, G. (2010). Wearable Assistant for Parkinson's Disease Patients With the Freezing of Gait Symptom. *IEEE Transactions on InformationTechnology in Biomedicine*, 14(2):436–446.

[3] Benussi, A., Grassi, M., Palluzzi, F., Cantoni, V., Cotelli, M. S., Premi, E., Di Lorenzo, F., Pellicciari, M. C., Ranieri, F., Musumeci, G., Marra, C., Manganotti, P., Nardone, R., Di Lazzaro, V., Koch, G., and Borroni, B. (2021). Classification accuracy of TMS for the diagnosis of mild cognitive impairment. *Brain Stimulation*, 14(2):241–249.

[4] Berry, R. B., Brooks, R., Gamaldo, C., Harding, S., Lloyd, R., Quan, S. F., Troester, M., and Vaughnh, B. V. (2017). The AASM Manual for the Scoring of Sleep and Associated Events. Version 2.4. *American Academy of Sleep Medicine*.

[5] Biswal, S., Kulas, J., Sun, H., Goparaju, B., Westover, M. B., Bianchi, M. T., and Sun, J. (2017). SLEEPNET: Automated Sleep Staging System via Deep Learning. *arXiv*, pages 1–17.

[6] Casula, E. P., Bertoldo, A., Tarantino, V., Maiella, M., Koch, G., Rothwell, J. C., Toffolo, G. M., and Bisiacchi, P. S. (2017). TMS-evoked long-lasting artefacts: A new adaptive algorithm for EEG signal correction. *Clinical Neurophysiology*, 128(9):1563–1574.

[7] Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Technical report.

[8] D'Amelio, M. and Rossini, P. M. (2012). Brain excitability and connectivity of neuronal assemblies in Alzheimer's disease: From animal models to human findings. *Progress in Neurobiology*, 99(1):42–60.

[9] Delbeuck, X., Van Der Linden, M., and Collette, F. (2003). Alzheimer's Disease as a Disconnection Syndrome? Technical Report 2.

[10] Delorme, A. and Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Technical report.

[11] Fan, Y. (2018). Research on Feature Extraction of EEG Signals using MSE-PCA and Sleep Staging. *2018 IEEE International Conference on Signal Processing, Communications and Computing, ICSPCC 2018*, (2):1–5.

[12] Fisher, R. S., Van Emde Boas, W., Blume, W., Elger, C., Genton, P., Lee, P., and Engel, J. (2005). Epileptic seizures and epilepsy: Definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia*, 46(4):470–472.

[13] Ghassemi, M. M., Moody, B. E., Lehman, L. W. H., Song, C., Li, Q., Sun, H., Mark, R. G., Westover, M. B., and Clifford, G. D. (2018). You Snooze, You Win: The PhysioNet/Computing in Cardiology Challenge 2018. *Computing in Cardiology*, 2018-Septe:20–23.

[14] Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., and Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet. *Circulation*, 101(23):E215–20.

[15] Hjorth, B. (1970). EEG analysis based on time domain properties. *Electroencephalography and Clinical Neurophysiology*, 29(3):306–310.

[16] Hochreiter, S. and Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8):1735–1780.

[17] Khosla, N. (2004). Dimensionality reduction using factor analysis. *Components*.

[18] Koch, G., Bonnì, S., Pellicciari, M. C., Casula, E. P., Mancini, M., Esposito, R., Ponzo, V., Picazio, S., Di Lorenzo, F., Serra, L., Motta, C., Maiella, M., Marra, C., Cercignani, M., Martorana, A., Caltagirone, C., and Bozzali, M. (2018). Transcranial magnetic stimulation of the precuneus enhances memory and neural activity in prodromal Alzheimer's disease. *NeuroImage*, 169:302–311.

[19] Najdi, S., Gharbali, A. A., and Fonseca, J. M. (2017). Feature Transformation Based on Stacked Sparse Autoencoders for Sleep Stage Classsification. *IFIP International Federation for Information Processing*, 499:144–153.

[20] Rogasch, N. C., Zipser, C., Darmani, G., Mutanen, T. P., Biabani, M., Zrenner, C., Desideri, D., Belardinelli, P., Müller-Dahlhaus, F., and Ziemann, U. (2020). The effects of NMDA receptor blockade on TMS-evoked EEG potentials from prefrontal and parietal cortex. *Scientific reports*, 10(1):3168–3168.

[21] Shoeb, A. (2009). Application of machine learning to epileptic seizure onset detection and treatment. *Diss. Massachusetts Institute of Technology*, pages 157–162.

[22] St Louis, E. K. and Boeve, B. F. (2017). REM Sleep Behavior Disorder: Diagnosis, Clinical Implications, and Future Directions. *Mayo Clinic Proceedings*, 92(11):1723–1736.

[23] Stokes, M. G., Chambers, C. D., Gould, I. C., Henderson, T. R., Janko, N. E., Allen, N. B., and Mattingley, J. B. (2005). Simple metric for scaling motor threshold based on scalp-cortex distance: Application to studies using transcranial magnetic stimulation. *Journal of Neurophysiology*, 94(6):4520–4527.

[24] Tautan, A., Casula, E., Pellicciari, M. C., Borghi, I., Maiella, M., Bonnì, S., Minei, M., Assogna, M., Palmisano, A., Smeralda, C., Romanella, S., Ionescu, B., Koch, G., and Emiliano, S. (2022a). (UNDER REVIEW) TMS-EEG perturbation biomarkers for Alzheimer's Disease classification. *Scientific Reports*, XX(XX):XX.

[25] Tautan, A., Rossi, A. C., De Francisco, R., and Ionescu, B. (2020). Automatic Sleep Stage Detection: A Study on the Influence of Various PSG Input Signals. In *42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) in Conjunction with the 43rd Annual Conference of the Canadian Medical and Biological Engineering Society.*

[26] Tăuţan, A.-M., Andrei, A.-G., and Ionescu, B. (2020). Freezing of Gait Detection for Parkinson's Disease Patients using Accelerometer Data: Case Study. In *2020 International Conference on E-Health and Bioengineering (EHB)*, pages 1–4.

[27] Tautan, A.-M., Casula, E., Borghi, I., Maiella, M., Bonnì, S., Minei, M., Assogna, M., Ionescu, B., Koch, G., and Emiliano, S. (2022b). (UNDER REVIEW) Preliminary study on the impact of EEG density on TMS-EEG classification in Alzheimer's disease. In *44th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Glasgow, Scotland, UK.

[28] Tautan, A.-M., Casula, E., Pellicciari, M. C., Borghi, I., Maiella, M., Bonnì, S., Minei, M., Assogna, M., Palmisano, A., Smeralda, C., Romanella, S., Ionescu, B., Koch, G., and Emiliano, S. (2022c). (UNDER REVIEW) Classification of Alzheimer's Disease patients using frequency content of TMS-EEG data. *Brain Stimulation*, XX(XX):XX.

[29] Tautan, A. M., Dogariu, M., and Ionescu, B. (2019a). Detection of Epileptic Seizures using Unsupervised Learning Techniques for Feature Extraction. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pages 2377–2381.

[30] Tautan, A.-M., Ionescu, B., and Santarnecchi, E. (20021). Artificial intelligence in neurodegenerative diseases: A review of available tools with a focus on machine learning techniques. *Artificial Intelligence in Medicine*, 117.

[31] Tautan, A.-M., Mandruta, I., Bajenaru, O.-A., Strungaru, R., Taralunga, D., Hurezeanu, B., and Neagu, G. M. (2018). The Automatic Detection of Epileptic Seizures Based on EEG Signals Processing: Investigation of Different Features and Classification Algorithms. In *World Congress on Medical Physics and Biomedical Engineering*, pages 393–397.

[32] Tautan, A. M., Rossi, A. C., De Francisco, R., and Ionescu, B. (2019b). Automatic sleep stage detection using a single channel frontal EEG. *2019 7th E-Health and Bioengineering Conference, EHB 2019*, 1:4–7.

[33] Tautan, A.-M., Rossi, A. C., De Francisco, R., and Ionescu, B. (2020). Dimensionality reduction for EEG-based sleep stage detection: Comparison of autoencoders, principal component analysis and factor analysis. *Biomedizinische Technik*, 66(2):125–136.