Finding Robust Low Dimensional Features for Sleep Detection Using EEG Data

Helen X. Mao
North Allegheny Senior High School
Wexford, PA, USA
helenmao2006@gmail.com

Joseph Widjaja
North Allegheny Intermediate High School, Wexford, PA, USA
josephwidjaja100@gmail.com

Yifan Guo
Department of Electrical and Computer Engineering, University of Pittsburgh
Pittsburgh, PA, USA
yig39@pitt.edu

Jijun Yin
Department of Electrical and Computer Engineering
University of Pittsburgh
Pittsburgh, PA, USA
davidyin96@gmail.com

Ramana Vinjamuri
Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County
Baltimore, MD, USA
rvinjam1@umbc.edu

2022 IEEE 2nd International Conference on Data Science and Computer Application (ICDSCA) | 978-1-6654-7200-5/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICDSCA56264.2022.9988155

***Abstract*—The study of sleep is crucial for understanding how our bodies function, and electroencephalogram (EEG) offers a convenient way to examine sleep. Sleep can be categorized into wakefulness, rapid eye movement (REM) sleep, and stages 1-4, with 4 being the deepest stage of sleep. We strive to study how well EEG data help classify brain waves into these stages. The goal of this paper is to construct low dimensional features that are computationally efficient, robust, and effective for sleep detection and analysis. We experiment with EEG band power analysis, principal component analysis (PCA), and autoencoder to reduce the dimensionality of the EEG data and evaluate their performances in classification. We find that, even when highly compressed, two dimensional features are still sufficient to obtain satisfactory classification accuracies: 89.3%, 88.8%, and 90% from band power analysis (using delta and alpha waves), PCA, and autoencoder, respectively.**

***Index Terms*—Autoencoder, linear discriminant analysis (LDA), PCA, sleep stages classification**

I. INTRODUCTION

Sleep is a naturally occurring state of rest where, apart from dreams, the mind resides in unconsciousness. On average, humans spend one third of their lives sleeping and cannot survive more than a few days without it. Insufficient sleep has been linked to many health problems such as type 2 diabetes, cardiovascular disease, obesity, and depression, which plague human society and reduce the quality of life [1].

A full night of rest should take us through 4 to 5 cycles of sleep stages: rapid eye movement (REM) sleep and stages 1-4, which are categorized by brain wave frequencies and used to determine the depth of sleep. Of the many ways for studying sleep, electroencephalogram (EEG) data collected through specialized electrodes and circuits can effectively display the inner levels of brain activity.

Sleep has been extensively studied in multiple fields, and researchers from the machine learning society have developed numerous approaches to classify sleep stages. For example, supervised learning methods such as support vector machines (SVM) [2]-[3] and neural networks [4]-[5] are commonly used to develop classifiers for sleep detection. Semi-supervised learning and transfer learning also offer promising approaches for sleep states classification [6]-[7]. However, these machine learning methods often require a large amount of labeled data and are computationally demanding in training and operation. This becomes especially apparent for the scenarios with constraints in computation and power, e.g., wearable computing—collecting brain wave signals and estimating brain status using compact wearable devices.

To address the above mentioned challenges, this paper studies the dimensionality reduction issue in constructing effective features from EEG data. Our goal is to find robust low dimensional features for sleep stages classification. We investigate the geometrical characteristics of EEG data in the low dimensional space and strive to gain intuition on how well the EEG data cluster in the compressed feature space. This can help us not only build classifiers directly from these low dimensional features but also incorporate clustering information of unlabeled EEG data in training the classifiers. In this paper, we experiment with three methods—EEG band power analysis, principal component analysis (PCA), and autoencoder—to search for low dimensional features. We find that the EEG band power is an informative and reliable source for construction of low dimensional EEG features: band-power related features with only two dimensions obtained by any of the three methods mentioned above are able to obtain satisfactory classification accuracies (higher than 88%).

The rest of the paper is organized as follows. Section II explains the methods, and Section III presents the experimental results and discusses the observations and insights from the results. Section IV concludes the study.

II. METHODS

A. Filtering Procedure

The EEG data came from PhysioNet [8]. We only used the first data file for consistency, and the data contained records of brain waves during sleep of a 33 year old female.

Authorized licensed use limited to: University of Maryland Baltimore Cty. Downloaded on July 06,2023 at 18:26:11 UTC from IEEE Xplore. Restrictions apply.

The collected raw data were recorded at a frequency of 100 Hz in 30 second intervals. We filtered this 3000 dimensional vector using a second-order Butterworth filter (MATLAB Signal Processing Toolbox) with cutoff frequencies of 1-4 Hz for delta wave, 5-8 Hz for theta wave, 9-12 Hz for alpha wave, and 13-25 Hz for beta wave. Their band powers can be used as 4 dimensional features representing the strength of the four brain wavebands. From here, we applied the PCA, autoencoders, and original brainwave band powers to extract low dimensional features (described in the following subsections).

B. *PCA*

PCA is a commonly used technique for dimensionality reduction. In this paper, we implemented PCA on a set of 4 dimensional vectors each containing the band powers of the filtered delta, theta, alpha, and beta waves. From our dataset (a total of 2650 data points), we performed a training/testing split in which 75% of the values were randomly chosen to be utilized for training, and the remaining 25% for testing. After normalizing all of the data, we fit PCA on the training set and then applied the mapping on both the training and testing sets. We used the linear discriminant analysis (LDA) (trained and tested with the same data used earlier) to calculate fair accuracies of classification. Specifically, we fit the LDA model with the 2 dimensional vectors (principal components) extracted by PCA from the training set, and evaluated the LDA model with the testing set.

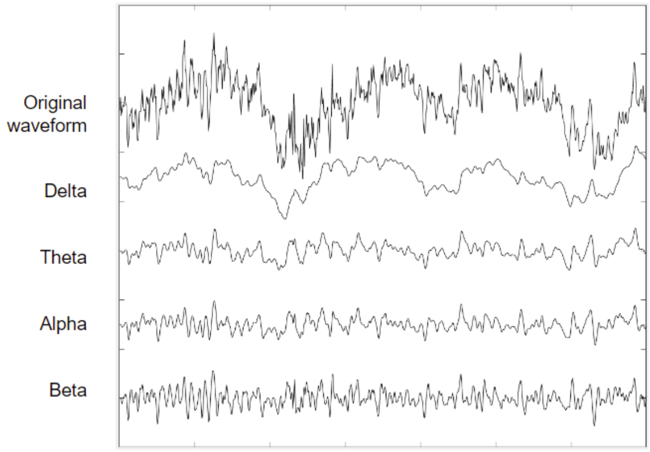

Fig. 1. Example of the waveforms of raw EEG signal and various frequency bands.

C. *Autoencoder*

An autoencoder is a specific type of neural network that is trained to produce an output which closely mirrors the input. It has a "bottle neck" in the middle where the dimensionality is reduced and the input data is compressed. We used the band powers of the filtered delta, theta, alpha, and beta waves as the 4 dimensional input to the autoencoder. We used the same 75%/25% training/testing split as used for the PCA and LDA. We experimented a set of autoencoders on Python with multiple architectures including 3, 5, and 7 layers, varying the number of nodes in each layer. The accuracies were calculated with the LDA using the same process as was used for the PCA. We fit the 2 dimensional features extracted from the autoencoder during training into the LDA model. Then, we tested it using the features from testing the autoencoder.

D. *Original Wavebands*

We also applied the LDA directly on the band powers of the four EEG waves, comparing the accuracies of using different combinations of the band powers of the EEG waves. Using the same data and training/testing split the same way as before, we fit the LDA on different wavebands and ran it on the testing data to achieve the accuracy scores.

III. Results and Discussions

A. *Raw Waveforms*

Fig. 1 shows an example of waveforms of the raw EEG signal and the four filtered frequency bands: delta, theta, alpha, and beta waves.

B. *Waveband Powers as Low Dimensional Features*

We investigated different combinations of waveband powers, which were directly used as low dimensional features for sleep stages classification with an LDA model. The accuracy results of the LDA classifier are summarized in Table I. It can be seen that among the four wavebands, the delta wave alone was able to produce respectable classification results (with an accuracy of 86.8% ± 2.0%). Also, surprisingly, the two wavebands delta and alpha in combination formed excellent features (with an accuracy of 89.3% ± 2.0% in classification) that outperformed using three or four filtered wavebands as features. Fig. 2 visualizes the clustering and classification results using band powers of delta and alpha waves as features for the LDA.

TABLE I
ACCURACY RESULTS OF LDA USING DIFFERENT COMBINATIONS OF WAVEBAND POWERS AS FEATURES

Features	**Accuracy (standard deviation)**
Delta	86.8% (± 2.0%)
Theta	78.5% (± 1.9%)
Alpha	74.2% (± 1.4%)
Beta	73.8% (± 1.3%)
Delta and theta	89.2% (± 1.5%)
Delta and alpha	**89.3% (± 2.0%)**
Delta and beta	87.8% (± 1.7%)
Theta and alpha	86.9% (± 2.2%)
Theta and beta	86.5% (± 2.1%)
Alpha and beta	88.6% (± 1.9%)
Delta, theta, and alpha	88.2% (± 2.7%)
Delta, theta, and beta	87.4% (± 2.2%)
Delta, alpha, and beta	89.4% (± 2.4%)
Theta, alpha, and beta	87.3% (± 3.3%)
Delta, theta, alpha, and beta	88.0% (± 2.6%)

Authorized licensed use limited to: University of Maryland Baltimore Cty. Downloaded on July 06,2023 at 18:26:11 UTC from IEEE Xplore. Restrictions apply.

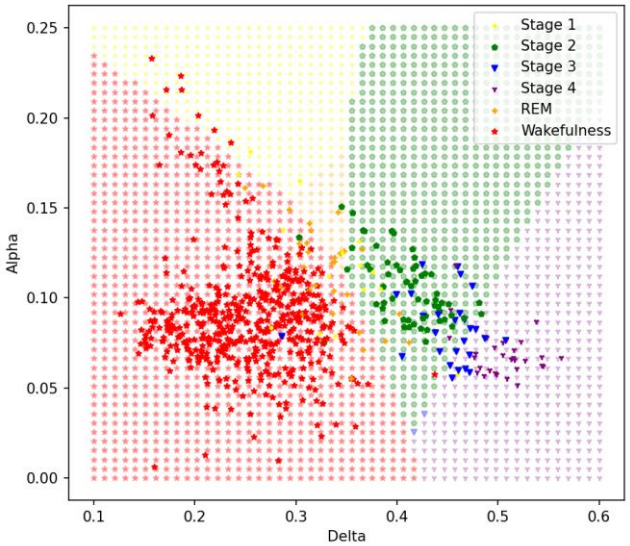

Fig. 2. Clustering and classification results using band powers of delta and alpha waves as features for the LDA.

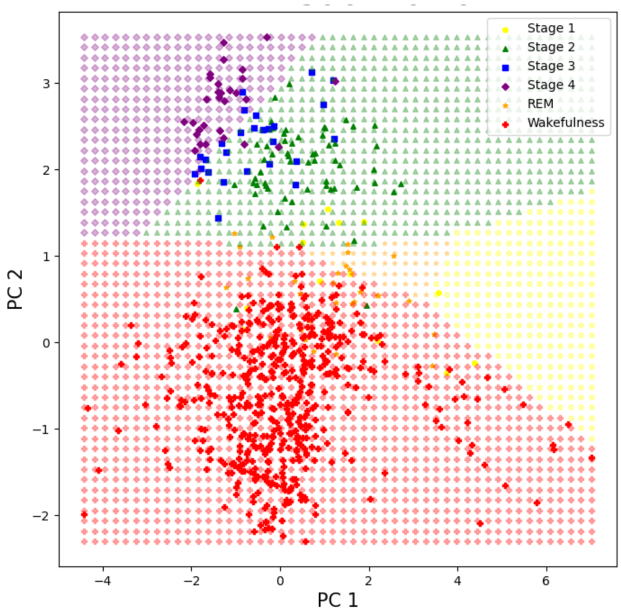

Fig. 3. Clustering and classification results using the first two PCs as features for the LDA.

C. Extracting Low Dimensional Features Using PCA

Fitting PCA on the training set (each data point being a four dimensional vector of EEG waveband powers) achieved variances of 0.5822 and 0.4042 for the first and second principal components (PCs), respectively, and these two PCs in total explained variance of approximately 0.9864. After applying these two PCs as the low dimensional features, the LDA classifier achieved an accuracy of 88.8% (± 2.0%). Fig. 3 visualizes the clustering and classification results using the first two PCs as features for the LDA.

D. Autoencoder

The autoencoder was built and tested with varying architectures. Table II lists the accuracy results of LDA using the two dimensional features obtained by the autoencoder (using the outputs of the two nodes in the middle layer of the autoencoder as features). Each number under the "structure" column represents the number of nodes in a layer of the autoencoder. For example, "4_2_4" represents a 3 layer autoencoder with 4 nodes in the input layer, 2 nodes in the middle layer, and 4 nodes in the output layer. The highest accuracy was 90.0% (± 2.0%), which was obtained with the simplest autoencoder of three layers. Fig. 4 shows the clustering and classification results using the two features (denoted by x_1 and x_2) extracted by the "best" autoencoder.

TABLE II
ACCURACY RESULTS OF LDA USING FEATURES EXTRACTED FROM AUTOENCODER

Structure of autoencoder	**Accuracy (standard deviation)**
4_2_4	**90.0% (± 2.0%)**
4_3_2_3_4	88.7% (± 1.4%)
4_4_2_4_4	89.6% (± 1.5%)
4_5_2_5_4	88.8% (± 1.6%)
4_6_2_6_4	89.0% (± 1.5%)
4_7_2_7_2	88.4% (± 1.8%)
4_8_2_8_4	88.3% (± 1.6%)
4_9_2_9_4	88.6% (± 1.4%)
4_10_2_10_4	88.7% (± 1.6%)
4_11_2_11_4	88.7% (± 1.8%)
4_12_2_12_4	88.7% (± 1.9%
4_13_2_13_4	88.6% (± 1.5%)
4_3_3_2_3_3_4	88.6% (± 1.8%)
4_3_4_2_4_3_2	88.5% (± 1.4%)
4_4_3_2_3_4_4	88.9% (±1.8%)
4_4_4_2_4_4_4	88.7% (± 1.4%)
4_5_3_2_3_5_4	88.4% (± 1.5%)
4_6_3_2_3_6_4	88.2% (± 1.7%)
4_5_4_2_4_5_4	89.0% (± 1.9%)
4_6_4_2_4_6_4	88.3% (± 1.5%)

Authorized licensed use limited to: University of Maryland Baltimore Cty. Downloaded on July 06,2023 at 18:26:11 UTC from IEEE Xplore. Restrictions apply.

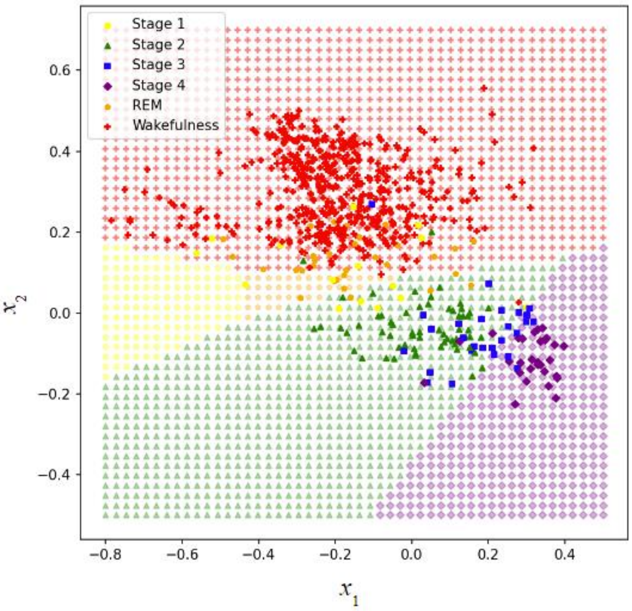

Fig. 4. Clustering and classification results using the two features extracted by the autoencoder.

E. Discussions

The highest accuracy result obtained by the autoencoder was 90.0% (± 2.0%). By comparison, we see that regular EEG waveband powers directly can serve as effective features for sleep stages classification, especially the delta and alpha wavebands, with an accuracy of 89.3% (± 2.0%). Even the PCA achieved close results of 88.8% (± 2.0%).

This contradicted our original expectation. For the autoencoder, we surmised that the more complex 5 or 7 layer architectures would have greater success, but this was not the case. Also, we did not expect for the two wavebands alpha and delta to be such excellent low dimensional features. Compared to the autoencoder, which is computationally heavier, the delta and alpha waves are robust indicators for sleep detection. This may be due to the fact that delta waves are present during deep sleep, while alpha waves occur more when we are awake or relaxed.

IV. Conclusion

Sufficient sleep is crucial for us to function properly and maintain health. Professionals in many fields have studied it using a range of methods. One such way to analyze sleep patterns is through EEG. During sleep, the brain cycles through 5 important stages (REM and stages 1-4) multiple times. We seek to study the classification of brain waves into these stages using EEG data. To do this, we attempt to construct robust and computationally efficient low dimensional features for sleep detection. We use machine learning with EEG band power analysis, PCA, and autoencoder to extract low dimensional features from the EEG data. This allows us to evaluate their classification performances. With classification accuracies of 89.3% (band power analysis using delta and alpha waves), 88.8% (PCA), and 90% (autoencoder), we discover that the highly compressed two dimensional features are robust indicators and obtain respectable accuracies.

References

[1] Centers for Disease Control and Prevention, "Sleep and chronic disease," https://www.cdc.gov/sleep/about_sleep/chronic_disease.html#:~:text=Notably%2C%20insufficient%20sleep%20has%20been,disease%2C%20obesity%2C%20and%20depression

[2] M. Diykh, Y. Li, and P. Wen, "EEG sleep stages classification based on time domain features and structural graph similarity," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 11, pp. 1159–1168, 2016.

[3] E. Alickovic and A. Subasi, "Ensemble SVM method for automatic sleep stage classification," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 6, pp. 1258–1265, 2018.

[4] M. Sokolovsky, F. Guerrero, S. Paisarnsrisomsuk, C. Ruiz, and S. A. Alvarez, "Deep learning for automated feature discovery and classification of sleep stages," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 17, no. 6, pp. 1835–1845, 2019.

[5] P. Krauss, C. Metzner, N. Joshi, H. Schulze, M. Traxdorf, A. Maier, and A. Schilling, "Analysis and visualization of sleep stages based on deep neural networks," *Neurobiology of Sleep and Circadian Rhythms*, vol. 10, pp. 1-9, May 2021.

[6] X. Wuzheng, S. Zuo, L. Yao, and X. Zhao, "Semi-supervised sparse representation classification for sleep EEG recognition with imbalanced sample sets," *Journal of Mechanics in Medicine and Biology*, vol. 21, no. 5, Article no. 2140006 (13 pages), 2021.

[7] M. Abdollahpour, T. Y. Rezaii, A. Farzamnia, and I. Saad, "Transfer learning convolutional neural network for sleep stage classification using two-stage data fusion framework," *IEEE Access*, vol. 8, pp. 180618-180632, 2020.

[8] PhysioNet, https://physionet.org/content/sleep-edfx/1.0.0/#files-panel

Authorized licensed use limited to: University of Maryland Baltimore Cty. Downloaded on July 06,2023 at 18:26:11 UTC from IEEE Xplore. Restrictions apply.