KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	Elisabeth R. M. Heremans, Trui Osselaer, Nick Seeuws, Huy Phan, Pascal Borzée, Bertien Buyse, Dries Testelmans, Maarten De Vos. (2022), **Data augmentation in semi-supervised adversarial domain adaptation for EEG-based sleep staging** 2022 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), Ioannina, Greece, 2022, pp. 1-4.
Archived version	Author manuscript: the content is identical to the content of the published paper, but without the final typesetting by the publisher
Published version	insert link to the published version of your paper https://ieeexplore.ieee.org/document/9926942
Journal homepage	insert link to the journal homepage of your paper https://ieeexplore.ieee.org/xpl/conhome/9926747/proceeding
Author contact	your email elisabeth.heremans@kuleuven.be Klik hier als u tekst wilt invoeren.
Abstract	
IR	url in Lirias

(article begins on next page)

Data augmentation in semi-supervised adversarial domain adaptation for EEG-based sleep staging

Elisabeth R. M. Heremans
KU Leuven
Leuven, Belgium
elisabeth.heremans@kuleuven.be

Trui Osselaer
Ghent University
Ghent, Belgium
trui.osselaer@ugent.be

Nick Seeuws
KU Leuven
Leuven, Belgium
nick.seeuws@kuleuven.be

Huy Phan
Queen Mary University of London
London, United Kingdom
h.phan@qmul.ac.uk

Dries Testelmans
KU Leuven
Leuven, Belgium
dries.testelmans@kuleuven.be

Maarten De Vos
KU Leuven
Leuven, Belgium
maarten.devos@kuleuven.be

***Abstract*—The upcoming era of wearable health monitoring devices has created a need for automated signal processing algorithms that can be trained with a minimal amount of labeled data. In our previous work, we showed that transfer learning techniques like semi-supervised adversarial domain adaptation can help to achieve this. We applied our method to remote sleep monitoring, by performing sleep staging on single-channel wearable EEG signals. In this work, we propose data augmentation to help in tackling this challenge. By using an artificially increased amount of labeled data, our semi-supervised adversarial domain adaptation method improves its performance on the wearable EEG data. The accuracy is increased consistently by 0.6% to 1.4% relative to the results without augmentation. As both adversarial domain adaptation and data augmentation are strategies to deal with the scarceness of data, we conclude that these methods are can effectively be combined to surpass their individual performance.**

***Index Terms*—domain adaptation, transfer learning, electroencephalography, deep learning, sleep staging, data augmentation**

I. Introduction

The recent development of wearable electroencephalography (EEG) devices is starting to enable remote monitoring of sleep disturbances. However, wearable EEG signal modalities are not easily interpreted by clinicians, because of their non-standard electrode placement. As a result, the gold standard for sleep disorder diagnosis is still based on in-hospital polysomnography (PSG). Clinicians manually perform sleep staging on these recordings by classifying each 30-second segment into a sleep stage according to developed rules [1].

In order to deploy wearable EEG devices in clinical practice, automated analysis of the captured EEG recordings is indispensable. Deep neural networks perform well on sleep staging tasks when trained on large, labeled datasets [2]–[4]. However, manual labels are hard to obtain for wearable EEG signals. Hence, these data necessitate sleep staging algorithms that do not rely on large amounts of labeled data. Previous studies have suggested transfer learning as a strategy to improve the sleep staging accuracy on wearable EEG datasets [5], [6]. This allows to exploit information from a large labeled source dataset, transferring this information to a target dataset with a smaller amount of labels. In recent years, unsupervised adversarial domain adaptation approaches have also been researched [7]–[9]. These are transfer learning techniques which only use unlabeled target data. In a recent study, we generalized these unsupervised adversarial domain adaptation methods to a semi-supervised learning framework [10]. With our general framework, we demonstrated how adversarial domain adaptation can be used to train sleep staging networks for real application in the clinic.

Transfer learning is not the only way to deal with a scarceness of data. Another strategy is to create additional training data using data augmentation techniques. When applying data augmentation to EEG signals, it is challenging to evaluate what augmentations are realistic and useful, and what augmentations result in unrealistic data or remove too much information from the signal. Despite this challenge, some recent works have successfully applied data augmentation techniques to EEG signals [11]–[14]. Some have investigated data augmentation for supervised end-to-end learning [13], [14], while others have used it for self-supervised learning [11], [12]. In the present study, we investigate the use of data augmentation in semi-supervised adversarial domain adaptation. In particular, we extend our previously developed domain adaptation framework [10] by adding data augmentation, and apply our method toward sleep staging. We demonstrate how data augmentation can further improve sleep staging performances by more efficient exploitation of the available (labeled) data.

II. Datasets and Methods

A. Datasets

A sleep staging network is pre-trained on a large source domain, and adapted to three different target domains. These target datasets are recorded with various wearable EEG configurations, reflecting realistic transfer learning scenarios.

Research supported by Fonds voor Wetenschappelijk Onderzoek - Vlaanderen (FWO) grant 1SC2921N and the Flanders AI Research Program.

1) Source Domain: MASS PSG: The Montreal Archive of Sleep Studies (MASS) [15] is a large, manually labeled dataset consisting of 200 PSG recordings of adult participant. The C4-A1 EEG derivation of the MASS database is used as the source domain in all experiments.

2) Target Domain 1: Surrey – cEEGrid: The Surrey - cEEGrid dataset was recorded using the cEEGrid array, a wearable device capturing the EEG around the ear using a flexible electrode strip [16]. The cEEGrid-EEG and PSG were simultaneously recorded in 12 healthy adults. The PSG signals were used for manual sleep staging. We use the right-ear front-versus-back cEEGrid channel of this dataset as our first target domain.

3) Target Domain 2: Dreem - Headband: The Dreem - Headband dataset [17] was recorded from 25 adults, using the Dreem Headband, a wearable dry-EEG device recording from five frontal and occipital channels. Simultaneous PSG measurements were also acquired and labeled based on five sleep experts' consensus. In the present study, we use the F7-F8 channel of this dataset as a second target domain.

4) Target Domain 3: Leuven – crosshead behind-the-ear: The Leuven - crosshead behind-the-ear dataset [10] consists of simultaneous recordings of crosshead behind-the-ear EEG and PSG, which was manually labeled. The recordings were obtained from 46 elderly, suspected sleep apnea patients. We select the crosshead behind-the-ear channel as a third target domain, simulating wearable behind-the-ear EEG.

B. Adversarial Domain Adaptation

In transfer learning, a model is typically pre-trained on a source domain, and this knowledge is transferred to a new dataset, the target domain. This helps to reduce the need for large amounts of labeled data of the target domain.

Adversarial domain adaptation is a popular technique for transfer learning. This technique tries to overcome the domain discrepancy between the source domain and the target domain using two competing goals. A domain discriminator is trying to identify which incoming samples are from which domain. In the meantime, a feature extractor tries to extract domain-invariant features, which make the samples from both domains indistinguishable. The framework is demonstrated in Figure 1.

To perform the given classification task, the feature extractor F and source classifier C_S are trained to classify labeled source domain samples with the categorical cross-entropy loss:

$$\mathcal{L}_S = -\mathbb{E}_{x_S \sim P_S} \sum_{c=1}^{N} y_S^c \log\ p_S^c, \tag{1}$$

with $p_S = C_S(F(x_S))$ as the class probability prediction for a source sample, and y_S as the ground truth label. c is the class number and N is the total number of classes (five in sleep staging). The domain discriminator D is trained with the binary cross-entropy loss to predict every sample's domain:

$$\mathcal{L}_D = -\mathbb{E}_{x_T \sim P_T} \log\ d_T - \mathbb{E}_{x_S \sim P_S} \log\,(1 - d_S), \tag{2}$$

with $d_S = D(F(x_S))$ and $d_T = D(F(x_T))$. The feature extractor is trained to confuse the domain discriminator with the adversarial loss function. This loss function is identical to $\mathcal{L}_D$ with inverted labels:

$$\mathcal{L}_{adv} = -\mathbb{E}_{x_T \sim P_T} \log\,(1 - d_T) - \mathbb{E}_{x_S \sim P_S} \log\ d_S. \tag{3}$$

Adversarial domain adaptation can be performed in an unsupervised way with only these three losses. However, the method can also be applied in a semi-supervised way, by using (a small amount of) labeled target domain data in addition to the unlabeled target domain data. In our previous work [10], we showed that adding such labeled target data is hugely beneficial to the performance. The semi-supervised adversarial domain adaptation approach is implemented by training the target classifier C_T on a limited number of labeled samples, using the categorical cross-entropy loss:

$$\mathcal{L}_T = -\mathbb{E}_{x_T \sim P_T} \sum_{c=1}^{N} y_T^c \log\ p_T^c. \tag{4}$$

These losses result in the following optimization framework for semi-supervised adversarial domain adaptation:

$$\min_{D} \qquad \mathcal{L}_D, \tag{5}$$

$$\min_{F, C_T, C_S} \quad \mathcal{L}_S + \mathcal{L}_T + \lambda_{adv}\mathcal{L}_{adv}. \tag{6}$$

C. Sleep Staging Network

To validate our method on the different datasets, we use a state-of-the-art sleep staging network as a feature extractor for the EEG data. The chosen network architecture is SeqSleepNet [3], which takes full advantage of the sequential information in EEG data. SeqSleepNet uses a sequence-to-sequence sleep staging strategy, taking multiple consecutive 30-second epochs as input, and concurrently predicting the corresponding sleep stages. $M = 10$ is used as a sequence length.

Rather than taking raw EEG signals as inputs, SeqSleepNet takes time-frequency images. Every recording is thus transformed to a logarithmically scaled spectrogram, and normalized to zero mean and unit standard deviation. For further details of the sleep staging network and the pre-processing, see [3].

D. Data Augmentation

Data augmentation refers to techniques used to artificially increase the amount of data by creating modified copies of real data, or by generating new, synthetic data. Just like transfer learning, it helps coping with a shortage of labeled data. In our semi-supervised adversarial domain adaptation approach, we integrate data augmentation by augmenting the labeled target data. Labeled target data give rise to a large performance benefit in adversarial domain adaptation [10], but these data are also the most scarce. Therefore, it is in augmenting these data that we expect the largest gains in performance.

The complex character of EEG signals makes it very challenging to generate realistic synthetic EEG signals. Therefore, we focus on applying modifications and transformations to existing EEG segments. We use a combination of multiple data augmentation techniques that were proven useful in previous

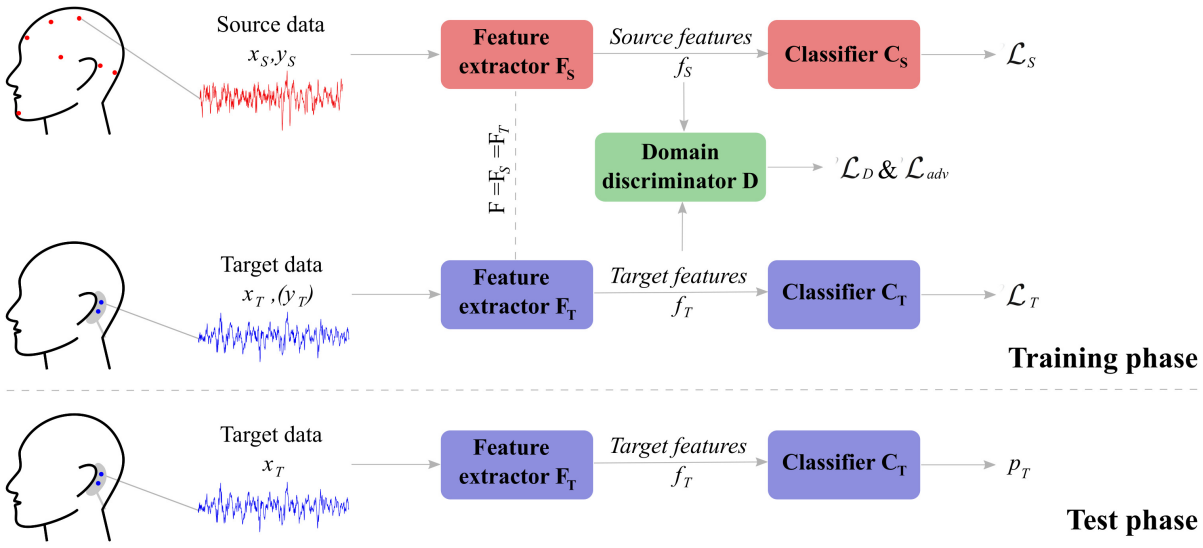


Fig. 1. Schematic illustration of adversarial domain adaptation. Subscripts S and T stand for source and target domain, respectively. The feature extractors are shared ($F_S \equiv F_T \equiv F$), but the classifiers are not ($C_S \not\equiv C_T$). Figure adapted from [10].

works on biosignals and speech signals [11]–[14], [18]. We apply the following techniques:

1) Temporal cutout
   A random continuous slice of the 30-second segment is replaced with zeros. The width of this cutout window is determined by the hyperparameter τ.
2) Frequency cutout
   A random continuous frequency band of the segment is replaced by zeros. As each 30-second segment is represented in the time-frequency domain, this operation is as easy to apply as the temporal cutout. The hyperparameter ν controls the width of the cutout window.
3) Gaussian noise
   Independent Gaussian noise with mean zero is added to the signal. The standard deviation of the noise is determined by the hyperparameter σ.
4) Signal mixing
   Another randomly selected sample is added to the signal. This results in correlated noise. The amplitude of this added signal is determined with A.
5) Frequency warping
   A random continuous frequency band undergoes a deformation (stretching and squeezing) along the frequency dimension. A quadratic function is used to determine the offset for every frequency:

$$offset(f) = int(-2 * (f - f_0) * (f - f_0 - w)/w) \quad (7)$$

   in which w is a hyperparameter controlling the width of the warped frequency band, and f_0 is the random starting point.
6) Temporal recombination
   A portion of the 30-second sample is replaced by part of another sample of the same sleep stage. The point where the original sample is cut off is chosen randomly, as is the sample providing the replacement.
7) Frequency recombination
   A random portion of the frequency content is replaced by that same frequency content of another sample of the same sleep stage. Again, both the replacement sample and the point where the sample gets cut off are random.

With these data augmentation techniques, we aim to make our model robust to perturbations such as losing some frequency content or small segments in time. We also build in robustness to shifts in frequency content and to correlated or uncorrelated noise. Lastly, we create novel samples by recombining existing ones. The hyperparameters for these methods were determined based on earlier sensitivity analyses on the Leuven dataset. Table I shows the selected value for each hyperparameter. For each data sample being fed to the model, each of these operations is applied with a probability of $p = 0.3$. Multiple data augmentation strategies can thus be applied simultaneously on the same sample.

TABLE I
HYPERPARAMETER VALUES

τ	ν	σ	A	w
4	15	0.01	0.1	10

III. EXPERIMENTS AND RESULTS

A. Training Setup

SeqSleepNet was pre-trained on the MASS dataset for 10 training epochs. Then, the network was trained with adversarial domain adaptation. A cross-validation was performed for each target dataset. In each round of cross-validation, the model was trained for 20 training epochs, a validation set was

used to retain the best model, and labels were used of only two recordings of the (target domain) training set. For further details of the training setup and construction of minibatches for training, see [10].

B. Results

Table II shows the results of adversarial domain adaptation with and without data augmentation of the labeled target data. The sleep staging results are shown for every target domain, as the mean ± standard error over the cross-validation folds. The reported metrics are the accuracy (acc), Cohen's kappa (κ) and weighted F1-score (wF1). The 'direct transfer' results show the performance of the pre-trained baseline network on the target domain without any adaptation. These results allow to compare the other two methods to a lower baseline. The adversarial domain adaptation approach without data augmentation of the labeled target data consistently achieves a lower performance than the version with data augmentation.

TABLE II
PERFORMANCE OF THE ADVERSARIAL DOMAIN ADAPTATION METHOD WITHOUT AND WITH DATA AUGMENTATION (ADA AND ADA-DA RESP.)

Surrey - cEEGrid			
Method	Acc	κ	wF1
Direct transfer	51.4±3.7	0.375±0.032	50.4±3.1
ADA	66.5±4.0	0.527±0.055	64.3±4.0
ADA-DA	**67.4±4.1**	**0.529±0.057**	**64.4±4.1**
Dreem - Headband			
Method	Acc	κ	wF1
Direct transfer	62.0±3.6	0.460±0.045	61.4±3.6
ADA	68.0±3.5	0.537±0.048	67.8±3.5
ADA-DA	**68.7±3.2**	**0.544±0.045**	**68.2±3.3**
Leuven - crosshead behind-the-ear			
Method	Acc	κ	wF1
Direct transfer	57.9±1.8	0.440±0.024	58.4±1.8
ADA	64.6±2.0	0.506±0.028	64.3±2.2
ADA-DA	**65.0±2.0**	**0.512±0.028**	**64.5±2.2**

IV. DISCUSSION

In this study, we researched the use of data augmentation on wearable, single EEG-channel signals. A number of different techniques were employed and applied to the data with a predefined probability. These different data augmentation techniques were applied in a semi-supervised adversarial domain adaptation problem with limited amounts of labeled target data. The results clearly show that data augmentation of these labeled target data helps the proposed domain adaptation method to achieve better results on the target domain. It consistently improves the accuracy on the target domain by 0.6% to 1.4%. By artificially increasing the amount of data, data augmentation improves the generalizability of deep learning models when low amounts of training data are available. Therefore, this technique is suitable for transfer learning problems, characterized by a scarceness of labeled target data.

In conclusion, transfer learning and data augmentation are both approaches that help to improve the performance of deep learning models in low data regimes, and thus reduce labeling costs. As transfer learning methods impact the training approach, and data augmentation is fully centered around the data itself, both are complementary and can easily be combined, allowing us to make smarter use of the available data. This is extremely useful in applications characterized by data scarceness, such as remote sleep monitoring with wearable EEG systems.

REFERENCES

[1] R. Berry, R. Brooks, C. Gamaldo *et al.*, "AASM Scoring Manual Updates for 2017 (Version 2.4)," *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine*, vol. 13, 2017.

[2] I. Perez-Pozuelo, B. Zhai, J. Palotti *et al.*, "The future of sleep health: a data-driven revolution in sleep science and medicine," *npj Digital Medicine*, vol. 3, no. 1, pp. 1–15, dec 2020.

[3] H. Phan, F. Andreotti, N. Cooray, O. Y. Chén, and M. De Vos, "SeqSleepNet: End-to-End Hierarchical Recurrent Neural Network for Sequence-to-Sequence Automatic Sleep Staging," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 3, pp. 400–410, sep 2018.

[4] S. Biswal, J. Kulas, H. Sun *et al.*, "SLEEPNET: Automated Sleep Staging System via Deep Learning," jul 2017. [Online]. Available: http://arxiv.org/abs/1707.08262

[5] H. Phan, O. Y. Chén, P. Koch *et al.*, "Towards More Accurate Automatic Sleep Staging via Deep Transfer Learning," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 6, pp. 1787–1798, 2021.

[6] E. R. M. Heremans, H. Phan, A. H. Ansari *et al.*, "Feature matching as improved transfer learning technique for wearable EEG," dec 2021. [Online]. Available: https://arxiv.org/abs/2201.00644v1

[7] S. Nasiri and G. D. Clifford, "Attentive adversarial network for large-scale sleep staging," *Proceedings of Machine Learning Research*, vol. 126, pp. 1–21, 2020.

[8] E. Eldele, M. Ragab, Z. Chen *et al.*, "Adversarial domain adaptation with self-training for EEG-based sleep stage classification," jul 2021. [Online]. Available: https://arxiv.org/abs/2107.04470v1

[9] C. Yoo, H. W. Lee, and J. Kang, "Transferring structured knowledge in unsupervised domain adaptation of a sleep staging network," *IEEE Journal of Biomedical and Health Informatics*, pp. 1–1, aug 2021.

[10] E. R. M. Heremans, H. Phan, P. Borzée, B. Buyse, D. Testelmans, and M. D. Vos, "From unsupervised to semi-supervised adversarial domain adaptation in EEG-based sleep staging," *Journal of Neural Engineering*, may 2022. [Online]. Available: https://iopscience.iop.org/article/10.1088/1741-2552/ac6ca8

[11] J. Y. Cheng, H. Goh, K. Dogrusoz, O. Tuzel, and E. Azemi, "Subject-aware contrastive learning for biosignals," *arXiv*, jun 2020.

[12] X. Jiang, J. Zhao, B. Du, and Z. Yuan, "Self-supervised contrastive learning for EEG-based sleep staging," *Proceedings of the International Joint Conference on Neural Networks*, vol. 2021-July, sep 2021.

[13] E. Lashgari, D. Liang, and U. Maoz, "Data augmentation for deep-learning-based electroencephalography," *Journal of Neuroscience Methods*, vol. 346, p. 108885, dec 2020.

[14] J. Fan, C. Sun, C. Chen *et al.*, "EEG data augmentation: towards class imbalance problem in sleep staging tasks," *Journal of Neural Engineering*, vol. 17, p. 56017, 2020.

[15] C. O'Reilly, N. Gosselin, J. Carrier, and T. Nielsen, "Montreal archive of sleep studies: An open-access resource for instrument benchmarking and exploratory research," *Journal of Sleep Research*, vol. 23, no. 6, pp. 628–635, dec 2014.

[16] S. Debener, R. Emkes, M. De Vos, and M. Bleichner, "Unobtrusive ambulatory EEG using a smartphone and flexible printed electrodes around the ear," *Scientific Reports*, vol. 5, no. 1, pp. 1–11, nov 2015.

[17] P. J. Arnal, V. Thorey, E. Debellemaniere *et al.*, "The Dreem Headband compared to polysomnography for electroencephalographic signal acquisition and sleep staging," *Sleep*, vol. 43, no. 11, pp. 1–13, nov 2020.

[18] D. S. Park, W. Chan, Y. Zhang *et al.*, "SpecAugment: A Simple Data Augmentation Method for Automatic Speech Recognition," *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, vol. 2019-Septe, pp. 2613–2617, apr 2019.