Unsupervised Deep Feature Extraction for Neonatal Sleep Stage Classification

Nick Seeuws[1], Amir Hossein Ansari[1], Sabine Van Huffel[1], and Gunnar Naulaers[2]

[1] Department of Electrical Engineering (ESAT), STADIUS, KU Leuven, Belgium
[2] Department of Development and Regeneration, University Hospitals Leuven, Neonatal Intensive Care Unit, KU Leuven, Leuven, Belgium

Preterm birth carries many risks for an infant. These infants are accomodated in a specialized unit of the hospital: the Neonatal Intensive Care Unit, NICU. The NICU utilizes many tools to provide care and assess maturation of these infants. Sleep stage monitoring is such an important tool. EEG provides a non-invasive way of performing the monitoring task. However, correctly identifying sleep stages from EEG recordings is a challenging and time-consuming task even for experts.

Machine learning offers clinicians a way of performing continuous sleep stage monitoring with a minimum of human intervention. A big drawback, however, is the need of labeled examples for training machine learning models. Due to the complexity of the task it is assumed that a substantial amount of training data is needed placing a heavy burden on the experts.

This work aims to extend the usability of machine learning models for sleep stage classification in preterm newborns. To this end, it investigates the potential of leveraging unlabeled EEG recordings as additional training data. By doing so this allows for lowering the human effort required for further improvement of classification models. The focus was on distinguishing quiet sleep from non-quiet sleep. Two approaches are considered: an unsupervised feature extraction model utilizing all data, labeled and unlabeled, followed by a supervised classifier making use of the extracted features and the corresponding labels as a first approach and a semi-supervised model jointly utilizing the labeled and unlabeled information as a second approach.

The EEG data was recorded at the Neonatal Intensive Care Unit of the University Hospitals in Leuven, Belgium. The Ethics Committee of the University Hospitals provided approval for the recordings and informed parental consent was obtained. The labeled dataset consisted of recordings of 26 preterm infants born before 32 weeks of gestation. 97 recordings taken between 27 and 42 weeks of postmenstrual age make up the dataset for a total of 492 hours of EEG data. Recordings were split up into 30 second segments and labeled as quiet or non-quiet. A full description of the dataset can be found in the work by Dereymaeker et al. [3] The unlabeled dataset consists of seven additional recordings providing an additional 109 hours of EEG recordings. This data was recorded under the same modalities as the labeled dataset. The signal is filtered and downsampled to 30Hz from 250Hz for use in this work. Altough the full dataset provides a multichannel signal, this work focused on single channel bipolar EEG data to simplify implementation and experimentation. It used the difference between the C3 and C4 electrode based on the international 10-20 system[2] as a signal.

The unsupervised approach used a Variational Auto-Encoder, VAE, as introduced by Kingma et al.[5] to extract features from the EEG segments. After training on the combined dataset of EEG recordings, the VAE defines a posterior probability distribution in latent space given an EEG segment and the mean value of this posterior was used as the extracted feature vector of a segment. A gradient boosting classifier performed the final classification using the extracted features of the labeled segments.

The semi-supervised approach made use of a Generative Adversarial Network, GAN, as introduced by Goodfellow et al.[4] The discriminator of the GAN was extended with an additional output acting as the desired classifier as proposed by Salimans et al.[7] This model can make direct use of labeled and unlabeled information and does not need an additional classifier to detect sleep stages.

The results are benchmarked against two other approaches. The first makes use of a set of features deemed relevant for sleep stage identification as proposed by Piryatinska et al.[6] mainly focused on spectral information of the EEG segments. The features are classified by a gradient boosting model. The second approach makes use of a convolutional neural network in a traditional supervised setting proposed by Ansari et al.[1] Several performance metrics were computed but models were mainly compared using Cohen's kappa coefficient. The VAE based model scored 0.47 on the test set while the GAN based model scored 0.64. Comparing these kappa values to the 0.43 scored by the classifier making use of spectral features[6] and the 0.60 reported for the supervised CNN by Ansari et al.[1] one can see the performance of a sleep stage classifier can be improved by making use of unlabeled data.

The VAE based model succeeded at extracting slightly better features compared to classical spectral features based on the improved classification performance but failed to improve upon the performance of an end-to-end deep learning approach. The GAN based model did show improved performance to supervised deep learning when making use of unlabeled information.

A further investigation of the feature space for both the VAE based model and the GAN based model leads to the conclusion that an unsupervised model struggles to separate factors of variation corresponding to sleep stage where a semi-supervised approach does identify such relevant factors.

The results show a possibility to improve sleep stage classification performance by leveraging unlabeled recordings. Increased performance is currently only observed for semi-supervised models.

References

1. Ansari, A., De Wel, O., Pillay, K., Dereymaeker, A., Jansen, K., Van Huffel, S., Naulaers, G., De Vos, M.: A convolutional neural network outperforming state-of-the-art sleep staging algorithms for both preterm and term infants. Internal Report 19-65, ESAT-STADIUS, KU Leuven (Leuven,Belgium) (2019 (submitted for publication))
2. Cherian, P.J., Swarte, R.M., Visser, G.H.: Technical standards for recording and interpretation of neonatal electroencephalogram in clinical practice. Annals of Indian Academy of Neurology **12**(1), 58 (2009)
3. Dereymaeker, A., Pillay, K., Vervisch, J., Van Huffel, S., Naulaers, G., Jansen, K., De Vos, M.: An automated quiet sleep detection approach in preterm infants as a gateway to assess brain maturation. International journal of neural systems **27**(06), 1750023 (2017)
4. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: Advances in neural information processing systems. pp. 2672–2680 (2014)
5. Kingma, D.P., Welling, M.: Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114 (2013)
6. Piryatinska, A., Terdik, G., Woyczynski, W.A., Loparo, K.A., Scher, M.S., Zlotnik, A.: Automated detection of neonate eeg sleep stages. Computer methods and programs in biomedicine **95**(1), 31–46 (2009)
7. Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X.: Improved techniques for training gans. In: Advances in neural information processing systems. pp. 2234–2242 (2016)