Enhancing EEG Classification Performance through Generative Adversarial Networks: Investigating the Impact of Sample Sizes and Classifier Selection

Chad C. Williams[1], Daniel Weinhardt[2], Maria Wirzberger[2],
Sebastian Musslick[1]

[1]) Brown University, Carney Institute for Brain Science, 164 Angell St, Providence, RI, United States
(chad_williams, sebastian_musslick)@brown.edu
[2]) University of Stuttgart, Interchange Forum for Reflecting on Intelligent Systems, Geschwister-Scholl-Str. 24D, Stuttgart, Germany
(daniel.weinhardt, maria.wirzberger)@iris.uni-stuttgart.de

Electroencephalography (EEG)—the scalp recording of electrical activity of the brain—has played a prominent role in neuroscience to gain insights into human brain functioning. There is a major potential for using this non-invasive method in brain decoding, which has been untapped due to its low signal-to-noise ratio, and the related need for large amounts of data ([1]). Advances in machine learning have mitigated this need through data augmentation techniques, such as Generative Adversarial Networks (GANs; [2, 3, 4]).

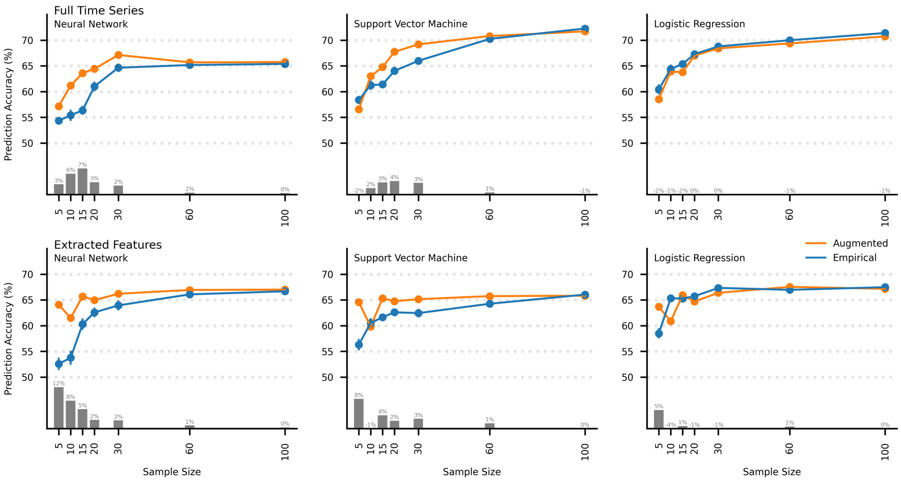

Figure 1: Classification performance using a neural network, SVM, and logistic regression, with two data formats, and across seven sample sizes for empirical (blue line) and augmented (orange line) data. Error bars reflect the standard error of the mean. The left, middle, and right panels in Figure 1 show classification performance across classifiers. The top panels show classification performance based on full EEG time series, whereas the bottom panels show classification performance for extracted EEG features. Grey bars in each panel indicate the difference between augmented and empirical performance, with positive values reflecting enhanced performance for the augmented classifications.

Here, we gauged the extent to which GANs can augment EEG data to enhance classification performance. Our objectives were to determine which classifiers benefit from GAN-augmented EEG and to estimate the impact of sample sizes on GAN-enhancements. We investigated three classifiers—neural networks, support vector machines (SVMs), and logistic regressions—across seven sample sizes ranging from 5 to 100 participants.

Our results are summarized in Figure 1. We found that augmenting EEG data enhanced classification performance for neural networks and SVMs, but not for logistic regressions. Indeed, the best classifications for sample sizes under 60 were achieved with augmented data. Further, these findings persisted across two data formats wherein the classifiers were provided inputs corresponding to the full time series data or to three extracted features—the reward positivity, delta power, and theta power. Finally, we also found that enhancements within these classifiers diminished as sample sizes increased.

Taken together, GAN-augmented EEG enhanced classification for neural networks and support vector machines, but not logistic regressions. This pattern persisted across both of the data formats (full time series and extracted features). Augmented EEG had the largest effect on neural network classifiers, reaching up to 12%, while its effect on SVMs reached up to 8%. In addition, GAN-enhanced classification was larger when using the extracted features as predictors than when using the full time series as predictors. As visible in Figure 1, GAN-enhancements diminished as sample sizes increased—suggesting it is most effective with small samples.

Given the novelty of this approach, related methodological advances can benefit ambitious research fields such as brain-computer interfacing or mobile EEG, and particularly facilitate research that is unable to collect large amounts of data, such as with clinical populations.

References

[1] Lotte, F. and Bougrain, L. and Cichocki, A. and Clerc, M. and Congedo, M. and Rakotomamonjy, A. and Yger, F. 2018. A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. *Journal of Neural Engineering*, 15:031005.

[2] Goodfellow, I. and Pouget-Abadie, J. and Mirza, M. and Xu, B. and Warde-Farley, D. and Ozair, S. and Courville, A. and Bengio, Y. 2014. Generative adversarial nets. In Ghahramani, Z. and Welling, M. and Cortes, C. and Lawrence, N. and Weinberger, K. Q., *Advances in Neural Information Processing Systems*, Curran Associates, Inc.

[3] Fahimi, F. and Dosen, S. and Ang, K. K. and Mrachacz-Kersting, N. and Guan, C. 2020. Generative adversarial networks-based data augmentation for brain–computer interface. *IEEE Transactions on Neural Networks and Learning Systems*. 32:4039–4051.

[4] Fahimi, F. and Zhang, Z. and Goh, W. B. and Ang, K. K. and Guan, C. 2019. Towards EEG generation using GANs for BCI applications. In: *2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, IEEE. pp. 1–4.