 DOI: 10.3217/978-3-99161-050-2-047

Transformative and Generative Data Augmentation for EEG-based BCIs

Manuel Weiss, Ricardo Chavarriaga*
Zurich University of Applied Sciences (ZHAW), Winterthur, Switzerland
*Technikumstrasse 71, 8400, Winterthur, Switzerland. E-mail: char@zhaw.ch

Introduction: The effectiveness of machine learning models used in brain-computer interfaces (BCIs) is often limited by the availability of training data [1]. In this study, we explore data augmentation methods to enhance machine learning classifier performance on imbalanced datasets, specifically in BCIs based on the Rapid Serial Visual Presentation (RSVP) paradigm [2].

Materials, Methods and Results: We considered two types of data augmentation techniques. First, *transformative* methods that generate new samples by manipulating existing datasets, including frequency shift (FS) , Fourier transform surrogate (FTS), smooth time mask (STM), and time shift (TS) [3]. The second type uses *generative* artificial intelligence approaches to produce synthetic data, including class-conditioned Wasserstein GANs (CCWGAN) [4] and a conditioned denoising diffusion probabilistic model (CDDPM) [5]. The impact of these methods for augmenting data in two RSVP public datasets [1, 6] was assessed using four different classifiers: LDA, RSVM [7], MDM [8], and EEGNet [9].

Among the *transformative* methods, STM provided the most consistent performance improvement, with gains of up to 8.36%, followed by TS with improvements of up to 5.29%.

Regarding the *generative methods*, the EEG-optimized CCWGAN excelled in generating signals with high temporal similarity and validity (c.f., Figure 1). Meanwhile, the CDDPM, adapted from computer vision, demonstrated improved training stability. However, both generative methods showed a consistent decrease in classification performance when synthetic data was used to train an MDM classifier. In turn, the EEGNet decoder maintained comparable average performance across subjects regardless of the amount of augmented data used, while some subjects exhibited improved performance.

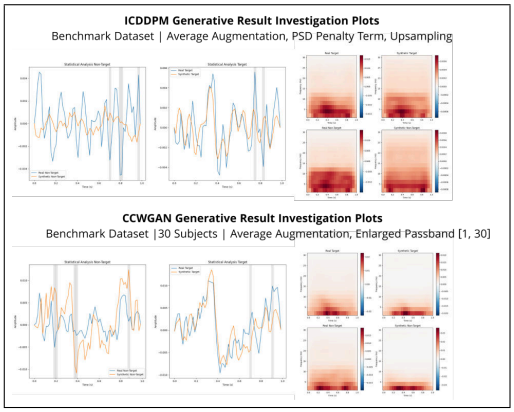

Figure 1: *Qualitative Comparison between generative Results of ICDDPM and CCWGAN.*

Conclusion: We found that *transformative* data augmentation event yield more consistent performance improvement in RSVP paradigms characterized by high class imbalance. Synthetic data created with *generative* approaches yielded signals with similar temporal an spectral characteristics. However, our results suggest that data augmentation using *generative* methods is more subject-dependent than for *transformative* methods. *Acknowledgments and Disclosures:* N/A.

References:

[1] Zhang S., et al. A benchmark dataset for RSVP- Based Brain–Computer interfaces. *Frontiers in Neuroscience*, 14, 2020. doi: 10.3389/fnins.2020.568000.

[2] Kenneth I. Forster. Visual perception of rapidly presented word sequences of varying complexity. *Perception & Psychophysics*, 8(4):215–221, July 1970. doi: 10.3758/BF03210208.

[3] Rommel C., et al. Data augmentation for learning predictive models on EEG: a systematic comparison, 2022. URL https://arxiv.org/abs/2206.14483.

[4] Panwar S., Rad P., et al. Modeling EEG data distribution with a Wasserstein generative adversarial network to predict RSVP events. *IEEE TNSRE*, 28(8):1720–1730, 2020. doi: 10.1109/TNSRE. 2020.3006180.

[5] Quinn Nichol A. and Dhariwal P. Improved denoising diffusion probabilistic models. In Marina Meila and Tong Zhang, editors, Proceedings of the 38th *Proceedings of Machine Learning Research*, 139:8162–8171. PMLR, 2021-07-18/2021-07-24.

[6] Matran-Fernandez A., Poli R. (2017) Towards the automated localisation of targets in rapid image-sifting by collaborative brain-computer interfaces. PLoS ONE 12(5): e0178498. https://doi.org/10.1371/journal.pone.0178498

[7] Barachant A. MEG decoding using Riemannian geometry and unsupervised classification. *Notes on the winner of the Kaggle "DecMeg2014 - Decoding the Human Brain" competition*, 2014.

[8] Barachant, A., et al. Riemannian geometry applied to BCI classification. *International conference on latent variable analysis and signal separation*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010.

[9] Vernon J Lawhern, et al. EEGNet: A compact convolutional neural network for EEG-based brain– computer interfaces. *Journal of Neural Engineering*, 15(5):056013, July 2018. doi: 10.1088/1741-2552/aace8c.

Published by Verlag der Technischen Universität Graz