Simulating Brain Signals: Creating Synthetic EEG Data via Neural-Based Generative Models for Improved SSVEP Classification

Nik Khadijah Nik Aznan, Amir Atapour-Abarghouei, Stephen Bonner, Jason D. Connolly, Noura Al Moubayed and Toby P. Breckon

Durham University

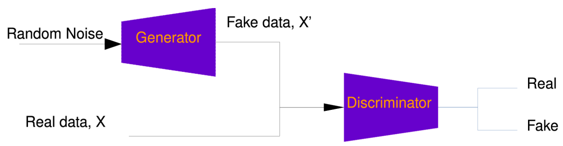
Issue:

Low reliability and usability of BCI caused by **limitations** such as **difficulty** of collecting **high quality EEG data**, requirement for per-subject and per-session **calibration** and **highly subject and session variant**.

Approach:

Generation of new synthetic dry-EEG data using a selection of neural-based generative models.

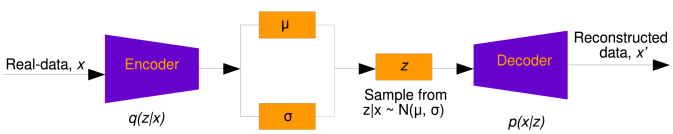
a. Deep Convolutional Generative Adversarial Network



Generative Adversarial Network

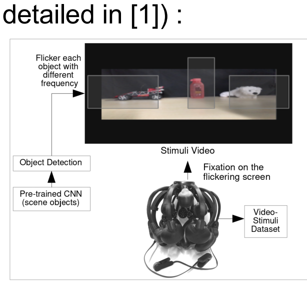
b. Improved Wasserstein Generative Adversarial Network

c. Variational Autoencoder

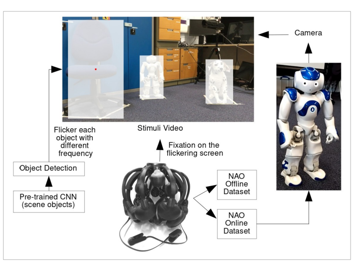


Variational Autoencoder

Empirical SSVEP Dry-EEG Datasets (collected and detailed in [1]) :



Video-Stimuli Dataset: train the generative model



NAO Dataset: experimental evaluation

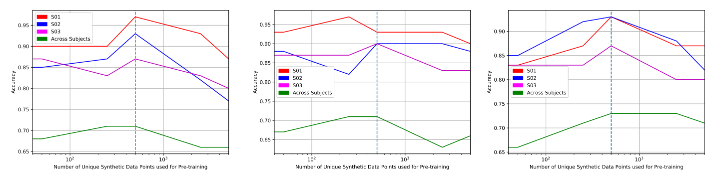
Contribution:

(1) Classification of synthetic and real data as a single training set [Baseline - the classification of only real data] :

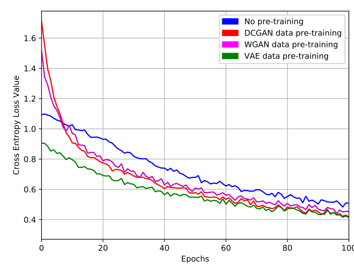
Method	S01	S02	S03	Across Subjects
Baseline	0.91 ± 0.04	0.87 ± 0.10	0.84 ± 0.03	0.69 ± 0.03
DCGAN	0.86 ± 0.01	0.80± 0.04	**0.89** ± 0.03	**0.72 ± 0.03**
WGAN	**0.93 ± 0.04**	0.90 ± 0.04	0.87 ± 0.02	0.71 ± 0.04
VAE	0.92 ± 0.03	**0.90 ± 0.03**	0.79 ± 0.03	0.67 ± 0.02

(2) Pre-training classification:

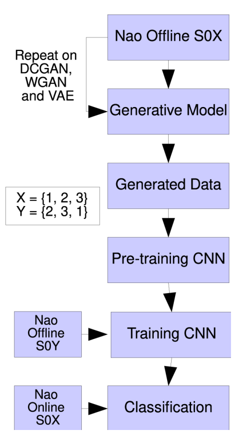
Method	S01	S02	S03	Across Subjects
Baseline	0.91 ± 0.04	0.87 ± 0.10	0.84 ± 0.03	0.69 ± 0.03
DCGAN	**0.97 ± 0.03**	**0.93 ± 0.03**	**0.87 ± 0.01**	0.70 ± 0.02
WGAN	0.93 ± 0.03	0.90 ± 0.04	**0.87 ± 0.03**	0.72 ± 0.03
VAE	0.93 ± 0.03	0.92 ± 0.06	**0.87 ± 0.05**	**0.73 ± 0.03**



Test accuracy when varying the volume of synthetic data used for pre-training.



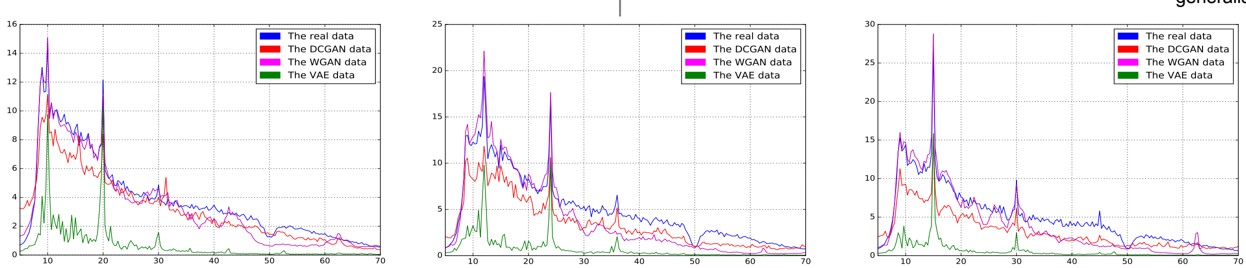
Convergence of the Cross-Entropy value plotted over training epochs.



Cross-subject generalisation

(3) Cross-subject generalisation:

Method	S01	S02	S03	Mean
Baseline	0.35	0.54	0.42	0.45 ± 0.08
DCGAN	0.42	0.71	0.57	0.57 ± 0.12
WGAN	0.40	0.60	0.50	0.56 ± 0.13
VAE	0.70	0.90	0.85	**0.82 ± 0.08**



Comparing real and synthetic data from the generative models for SSVEP signal at 10Hz, 12Hz and 15Hz. Synthetic data clearly displays the characteristic SSVEP frequency peaks at the same frequencies as those observed in the real data.

Conclusion:

- **generation of synthetic dry-EEG data** containing SSVEP signals using **DCGAN, WGAN and VAE**
- **improvement of performance** in term of **classification for pre-training and dataset augmentation** and **convergence rate** of classification models.

[1] N. K. N. Aznan, J. D. Connolly, N. A. Moubayed, and T. P. Breckon, "Using variable natural environment brain-computer interface stimuli for real-time humanoid robot navigation," in IEEE International Conference on Robotics and Automation. IEEE, 2019.