Towards EEG signals codification using contrastive loss

Abstract

In this work we explore the use of contrastive learning to obtain vectorial representation of EEG signals. As contrastive learning is a self-supervised method, we propose the use of consecutive segments up to a certain window of time to define a property that will guide the contrastive learning. The results are promising given the small sample of EEG signals that we have access.

1 Introduction

Contrastive learning is a novel method to learn vectorial representations. It has had an impact on reaching new performance levels for different data modalities, for instance: audio (1), text (2), images (3). In this work we explore to apply it to EEG signals (electroencephalogram signals). This is the loss function for contrastive learning:

$$L(W, Y, X_1, X_2) = (1 - Y)\frac{1}{2}(Dw)^2 + (Y)\frac{1}{2}max(0, m - Dw)^2 \tag{1}$$

A common setting for contrastive learning is to use a Siamese network setting in which two raw signals s_1 and s_2 are transformed into two vectors: X_1 and X_2. These vectors are then compared by a distance function Dw and evaluated by the loss function (see equation 1) which requires a Y signal in which $Y = 1$ if the property is shared among the signals, and $Y = 0$ otherwise. The expectation is that raw signals which share a property are grouped close by a margin (m) while signals that do not share the property should be further away than the margin.

In the case of the processing of the EGG signals the goal is to infer and relate this electrical signals to specific brain activity. However it is common that these signals are noisy and their nature is not clear. When applying a self-supervised method such as, contrastive learning, our goal is able to transform EGG signals to a vectorial representation less noisy and in which relevant aspects for the signal related to brain activity are capture.

Our data come from 20 subjects that where showed three visual stimuli: pleasant, displeasant and neutral. Their activity was recorded using the standard 10-20 configured with 128 samples per second and 14 channels(4).

2 Experimental setting

Our raw input is composed of matrices of size CxL where C is the number of channels one for each EGG sensor (up to 14 in our experiments) and L is the length of the sample. The first layer of our network is a $CNN1D$ which has a kernel of size 14×1 this is done to make the rest of the processing invariant to order of the channels. The networks continue with four convolutional layers with max polling eachone, each one reduces the channels and kernels by half with batch normalization and ReLU activation's. At the end of the network we have a liner layer with 512 dimensions. We set a margin for the contrastive loss function to 0.6. We also tested Margin Ranking Loss y Cosine Embedding Loss but the results were similar to the ones showed here.

35th Conference on Neural Information Processing Systems (NeurIPS 2021), Sydney, Australia.

To train the network we use the property if two segments are close by an offset of 8 seconds. We generated a balanced dataset in which 50% of the cases are considered close and the rest not. We use 14 subjects for training, 3 for validation and 3 for testing.

Once the network was trained we translated the testing subject signals into vectors and compare how much they cluster regarding three aspects: the category, the property, the subject and the original stimulus; we measure the *silhouette* score which goes from -1 for bad clustering regarding the aspect to 1 for perfect clustering. The results can be seen in Table 1

Category	Subject	Stimulus
4.454e-05	0.0394	0.009

Table 1: Resultados de silhouette

Visually we can see the effect of the clustering in Figure 1 using T-SNE to project the vectors for the subject (coloring is assigned after projection).

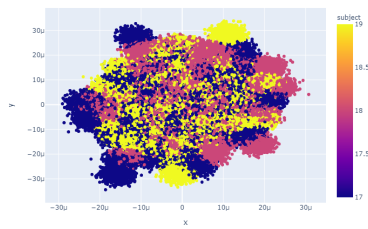

Figure 1: Proyección 2D de los vectores resultantes usando t-sne (mismo color mismo sujeto)

Acknowledgments and Disclosure of Funding

No funding was used during this research.

3 Bibliography

References

[1] A. Saeed, D. Grangier, and N. Zeghidour, "Contrastive learning of general-purpose audio representations," 2020.

[2] M. Caron, I. Misra, J. Mairal, P. Goyal, P. Bojanowski, and A. Joulin, "Unsupervised learning of visual features by contrasting cluster assignments," in *Advances in Neural Information Processing Systems* (H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, eds.), vol. 33, pp. 9912–9924, Curran Associates, Inc., 2020.

[3] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *Proceedings of the 37th International Conference on Machine Learning* (H. D. III and A. Singh, eds.), vol. 119 of *Proceedings of Machine Learning Research*, pp. 1597–1607, PMLR, 13–18 Jul 2020.

[4] M. A. Garcia, V. C. Ortiz, and R. H. Martinez, "Ritmos cerebrales relacionados a imágenes emotivas." Poster at Faculty of Psychology, UNAM.