DOI: 10.20396/revpibic.

XXVII Congresso de Iniciação Científica Unicamp

16 a 18 de outubro de 2019 - Campinas | Brasil

Enriching synthetic data with real noise using Neural Style Transfer

Naomi Takemoto*, Lucas de M. Araújo, Tiago A. Coimbra, Martin Tygel, Sandra Avila, and Edson Borin

Abstract

Deep Learning experiments require large amounts of labeled data, but few annotated seismic datasets are available and annotation is a time-consuming, expensive activity. Synthetic modeled datasets may be a viable alternative. However, they lack the variability and intricacies of a real data signal. Moreover, methods that add colored noises are not enough to represent such variability. Thus, our goal is to produce a noise type that is characteristic of real data to create a viable synthetic dataset to train Deep Learning models. In this context, we apply the Neural StyleTransfer technique, which combines the structural content of an image with the textural style of another, to produce a synthetic data with noise characteristics extracted from a real dataset. The results show that the stylized synthetic seismic data preserves the modeled content while incorporating characteristics of some real data chosen as style, creating synthetic data with a more realistic noise profile.

Key words:

Machine Learning, Geophysics, Neural Style Transfer

Introduction

Many seismic phenomena have well understood and accepted mathematical models, and usually these models can be used to generate synthetic data labeled by construction. However, the apparent drawback of synthetic data is that it lacks the feature richness of real seismic data. In other words, if on the one hand, we may be able to perform proof-of-concept experiments, on the other quickly we may overestimate how well a Deep Learning system is adequate to solve a problem in a real-world situation.

In this work, we propose to use the Neural Style Transfer (NST) technique to enrich synthetic data with real seismic data characteristics.

In other words, given a synthetic data *S* and a real data *R*, we produce a new data that has the structure profile of *S* (what was modeled) with the feature profile of *R*.

Results and Discussion

Given some synthetic data as content and some real seismic data as style, the stylization process is as follows: i) Apply some noise to the synthetic data. This is important because the NST technique does not seem to apply the style in regions where the content is blank (see figure 2); ii) Pre-process content and style data in order to adapt the seismic data to NST's input requirement; iii) Initialize the stylized (target) data with the content data or with random noise (both approaches may yield good results, as discussed in the original paper[1]); iv) Apply NST ;4) Post-process the stylized data;

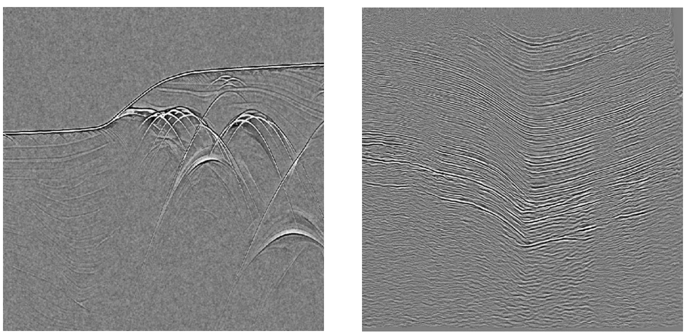

Figure 1: (left) synthetic data with Gaussian Noise. (right) real data.

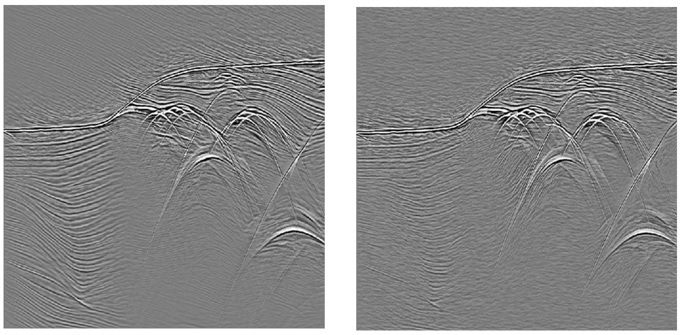

Figure 2: (left) raw output of NST process. NST applied to a seismic data with Gaussian Noise.

Conclusions

In our study, we explored the use of the concept *style* to generate a synthetic data enriched with meaningful noise, combining the structural content of a mathematically modeled data with features from real data. We also discovered that inserting an initial random noise in the synthetic input yields better style transferring. These generated data could then be used to train more robust Machine Learning systems, bypassing the scarcity of annotated data, or be used to test processing techniques.

Acknowledgement

We acknowledge support from the National Council for Scientific and Technological Development (CNPq-Brazil) and the Center for Computational Engineering and Sciences (FAPESP/CEPID #2013/08293-7-Brazil).

S. Avila is partially funded by Google Research Awards for Latin America 2018 and FAPESP (#2017/16246-0).

The authors also thank the High-Performance Geophysics (HPG) team for technical support.

---

[1] Gatys, L. A., A. S. Ecker, and M. Bethge, Image style transfer using convolutional neural networks: IEEE Conference on Computer Vision and Pattern Recognition. **2016**, 2414–2423.

This paper is a short version of the one originally prepared for presentation at the 16th International Congress of the Brazilian Geophysical Society, held in Rio de Janeiro, Brazil, August 19-22, 2019.