Contrastive-self-supervised-learning-for-climate-spatio-temporal-data

Marusov Alexander,
PhD-3

Presentation structure

1. Problem description
2. Actuality
3. Research targets
4. Method description
5. Downstream problems description
6. Results
7. Possible improvements

Problem description

Problem. Our aim is to to build universal compressed representations of spatio-temporal data using the self-supervised learning method

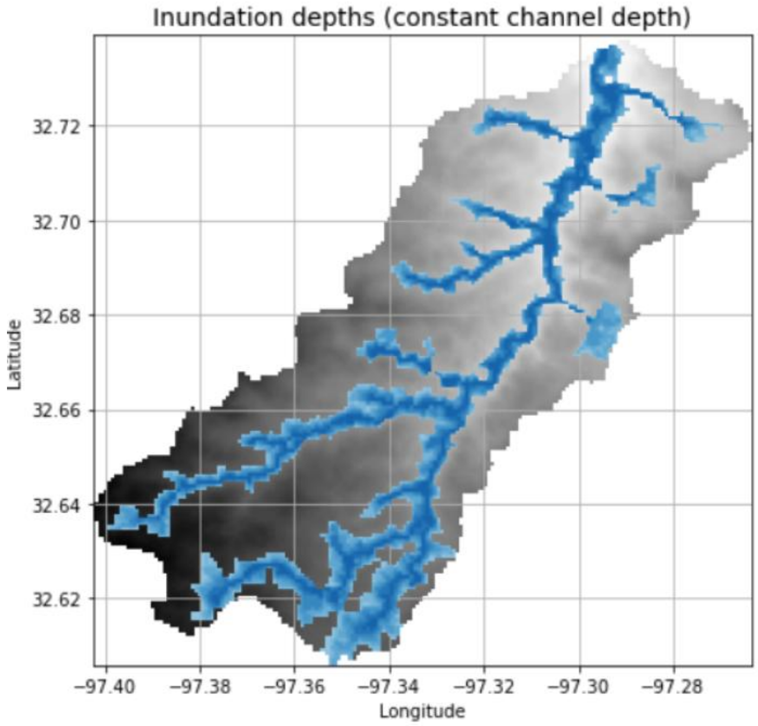

Fig.1. Spatio-temporal data example

Self-supervised learning

Why do we need SSL if we have supervised learning?

- **Neural networks are data-hungry**. However, labeled data for supervised learning are limited in volume. **Solution**: employ self-supervised learning that utilise unlabeled data.

- Quite often, we need **quick adaptation** to new forecasting problems (e.g., another drought index, another crop). **Solution**: we don't need any retraining if representations are universal.

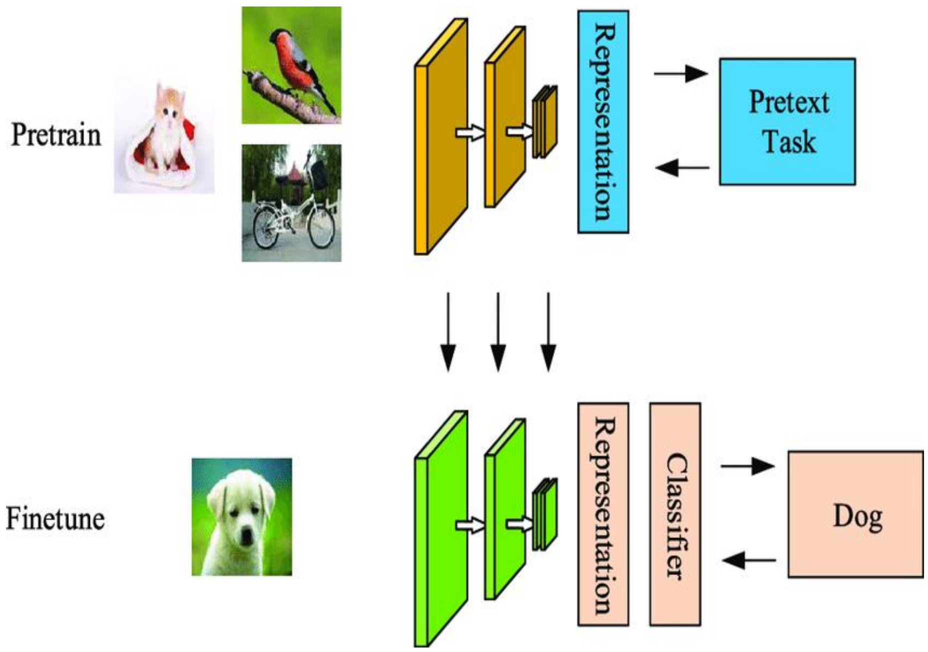

Fig.2. Illustration of self-supervised learning method

Research targets

1. Generalize the SSL-method TS2Vec[1], designed for time series, to the spatio-temporal case.
2. Using an encoder pretrained with TS2Vec, obtain representations of spatiotemporal data and use them to solve downstream problems.

Method description

1. **Spatial contrast**. Unlike to the standard TS2Vec we use contrastiveness not only via temporal dependency but also via the spatial one. To do this we use history for different regions.
2. **Encoder.** The spatial map is compressed to the embedding at each time moment via encoder.
3. **Objective function**. We use loss function, where we use both temporal and spatial contrasting.

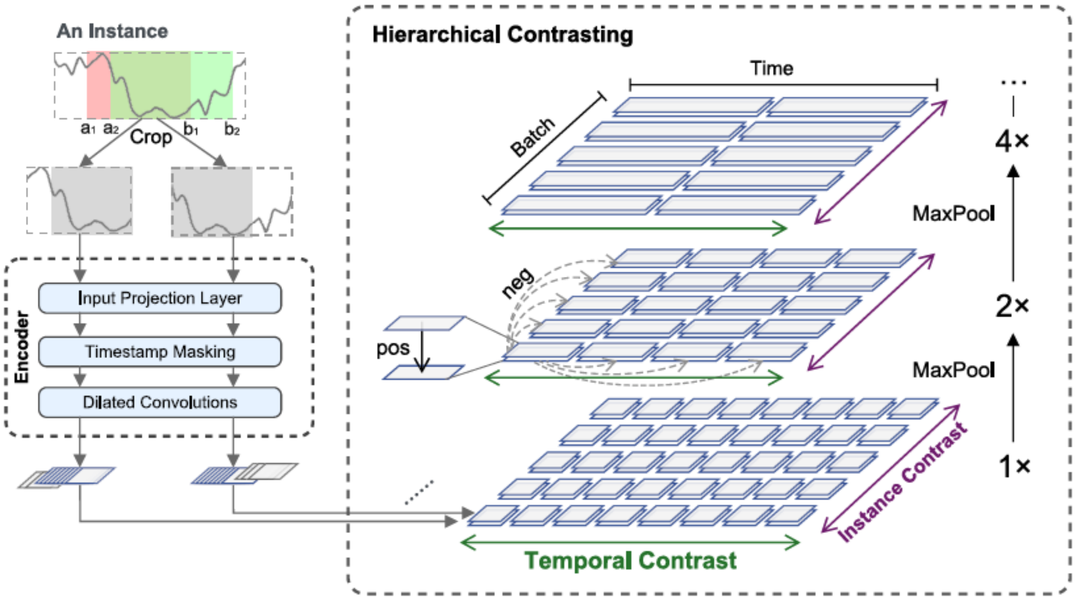

Fig.3. TS2Vec architecture [1]

Downstream problems description

- **Drought prediction**. We need to fit an encoder that takes as input the drought history for a given region and as output predicts the probability of drought for several months in advance.
- **Temperature forecasting**. Our aim is to predict spatial distribution of the temperature for one week ahead.

- **Approaches**
  - **Supervised method.** For the drought prediction we used gradient boosting, while for the temperature forecasting we used linear regression.
  - **SSL-based apporach**. We use gradient boosting and linear regression correspondingly for drought and temperature prediction but on the embeddings from the SSL.

Results

We can see that currently SSL-based results are worse than supervised data. However we see several steps described on the next slide to improve the results.

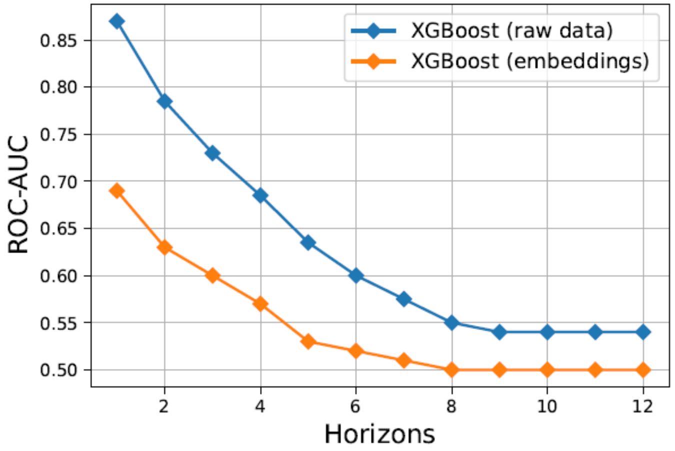

Fig.4. Drought prediction

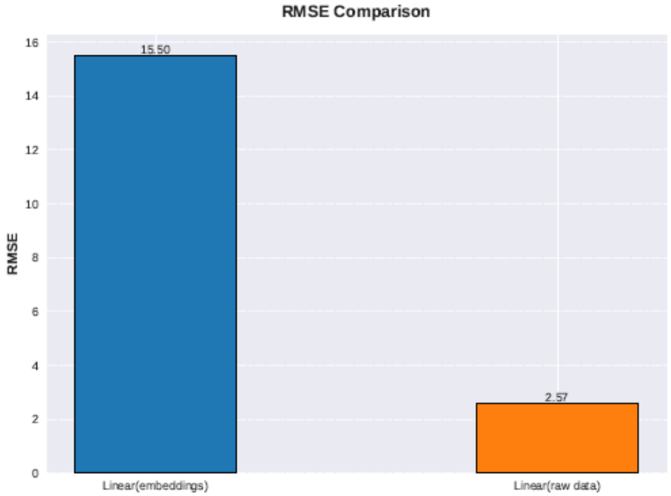

Fig.5. Temperature prediction

Possible improvements

Below we describe several steps to improve the results:

- **More appropriate accounting of temporal information**. Despite that TS2Vec positions itself as a universal SSL-framework for time series authors claim that two nearby observations should be treated as a negative pair. However this claim is wrong for specific type of data (e.g. climate), where nearby observations for sure can be treated as a positive pair. For example, temperature indicators for 2 p.m. and 3 p.m. are definitely correlate with each other. Hence we need a specific treatment for such type of data.

- **Accounting for data periodicity**. The potential periodicity of the data (like in the climate case) should be mentioned and processed separately due to its importance.

References:

1. Zhihan Yue, Yujing Wang, Juanyong Duan, et al. “Ts2vec: Towards universal representation of time series”. In: Proceedings of the AAAI conference on artificial intelligence. Vol. 36. 8. 2022, pp. 8980–8987.