Towards Better Time-series Data Augmentation for Contrastive Learning

Jang-Ho Choi, Moonyoung Chung, Taewhi Lee, Jiyong Kim
Intelligence Information Research Division
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
{janghochoi, mchung, kjy, taewhi}@etri.re.kr

***Abstract*—Contrastive learning is now a popular choice for representation learning in various domains, including image and natural language processing. However, contrastive learning for time-series data is relatively limited, due to its unrecognizable, high-dimensional temporal structures. It is still difficult to generate valid augmented views that are semantically accurate, despite the significant research advances in the field of time-series data augmentation. In this work, we survey recent works in time-series contrastive learning and propose a simple augmentation-agnostic technique that can effectively improve the fidelity of the augmented views.**

***Keywords*—Contrastive Learning, Data Augmentation, Time-series, Representation Learning**

I. Introduction

Contrastive learning has recently become very popular as it can leverage tremendous amounts of data for representation learning. In real-world scenarios, the users are often provided with only a handful of labelled data due to the high labelling cost. Thanks to its unsupervised manner, contrastive learning has widely been used in various domains such as images and natural languages. Contrastive learning learns representation that captures the similarities and differences between the data points, distinguishing positive and negative samples. Hence, the performance of contrastive learning is largely determined by the quality—such as fidelity and variety—of the positive and negative samples provided as training data [1].

The most common way to obtain positive and negative pairs is to use data augmentation. Data augmentation can be used to create two positive samples from a single data point. The positive pairs obtained from different data points within the training batch are defined as negative pairs [2]. However, as illustrated in Fig. 1, the negative pairs obtained in this way may actually belong to the same class or be very similar, so they may be incorrectly learned and have a negative impact on performance. Therefore, negative pairs are sometimes excluded in unsupervised learning [3]. The method of extracting positive pairs through data augmentation—such as cropping, replacing, cutting out, etc.—is a very efficient and effective way to achieve high fidelity, and is still widely used in image and natural language processing.

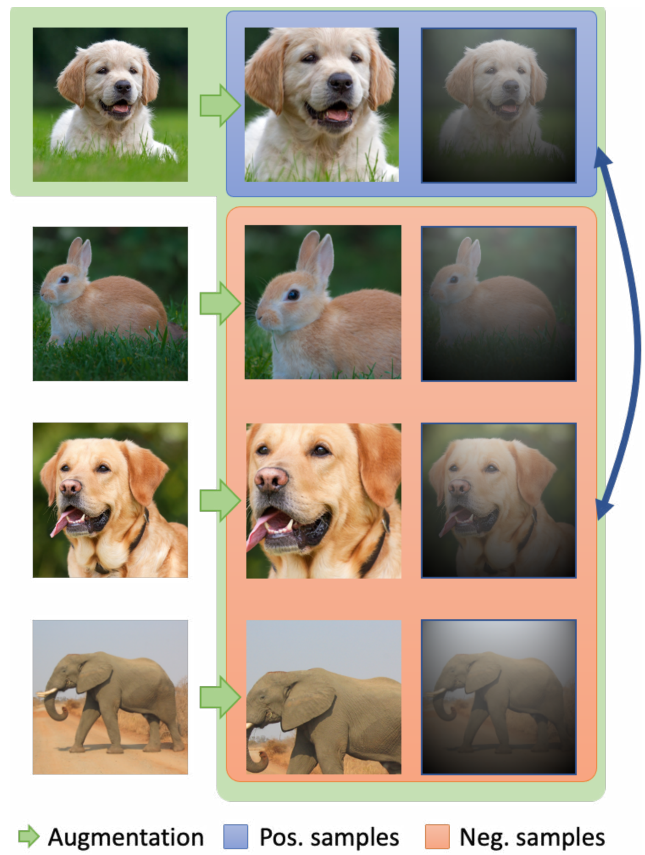

Fig. 1 Challenges in selecting positive and negative pairs using data augmentation in an unsupervised manner

For time-series data, however, the data augmentation is relatively limited due to the several characteristics: high dimensionality, noise and lack of a standard representation. In image and language domains, it is possible to generate desired augmented samples using heuristics guided by human knowledge. On the other hand, it is quite difficult to generate high fidelity positive pairs in time-series due to their human-unrecognizable temporal structure [4]. This can be crucial in contrastive learning as data augmentation in time-series can easily distort the underlying semantics.

Hence, in this paper, we analyze previous works in time-series contrastive learning. Then, we propose a data augmentation method for time-series, that can improve fidelity of the augmented data. The method is to masking portion of the time-series when applying well-known time-series augmentation techniques. The experiment show that the masking help improve the downstream task performance.

II. SURVEY

In previous work [5], the authors propose a contrastive learning framework to extract time-series representations in an unsupervised manner. The framework uses simple yet efficient augmentations to create pairs of time-series data, one that is weakly augmented and the other that is strongly augmented. The framework consists a temporal contrasting module that learns robust representations by a challenging cross-view prediction task. Furthermore, the authors adopt an additional contextual contrasting module that learns discriminative representations from the robust representations. The authors use permutation-and-jitter strategy for the strong augmentation, while jitter-and-scale strategy for the weak augmentation. The authors claim that learned representations using the proposed framework were effective for downstream tasks on three different datasets.

Supervised pre-training requires labelled datasets, which are expensive to obtain and may introduce unintended biases. To mitigate the problem, the authors propose a self-supervised contrastive pre-training method for time-series using time-frequency consistency [6]. The underlying assumption of the proposed method is that time-based and frequency-based representations of the same time-series must be embedded nearby in the joint time-frequency space. To achieve this goal, the authors apply data augmentations in each time and frequency domain and train individual encoders to produce time and frequency views separately. Then, the time and frequency contrastive losses are computed and optimized. In addition, the authors compute contrastive loss in the time-frequency space to enforce the time-frequency consistency. The authors conducted extensive experiments on eight datasets, showing that the proposed method significantly improves the baseline.

In [7], the authors propose an interesting approach to generate views for time-series, adopting a simple but effective module that learns to discover optimal views. The proposed module, called LEAVES, learns the hyper-parameters for data augmentations using adversarial training. This is similar to the work of [8], which also learns to augment time-series data in an adversarial manner, but by injecting perturbation in the latent space. LEAVES, on the other hand, consists of a series of differentiable data augmentation methods, such as permutation, time distortion, magnitude warping, scaling, and jittering. The authors claim that the proposed method can find reasonable views, which are more effective for downstream tasks.

The most common downstream task for time-series data is forecasting. The authors in [9] propose a contrastive learning framework that can be more effective for time-series forecasting as the downstream task. They argue that existing time-series contrastive learning frameworks restrict the generalization ability due to their strong dependency on specific time-series characteristics. Hence, the authors propose SimTS, a simple representation learning approach, in which the encoder learn to predict the future from the past in the latent space. Since it does not use negative pairs or data augmentation techniques that relies on specific time-series characteristics, the authors claim that SimTS can be a promising alternative for time-series contrastive learning.

Lastly, another interesting work in time-series contrastive learning is InfoTS, which leverages a meta-learner to select feasible data augmentations [4]. In their work, the authors propose a contrastive learning framework with information-aware augmentations, in which the meta-learner network selects the optimal augmentations. The authors claim that the optimal augmentations for time-series should not only be high fidelity and high variety, but also adaptive to a variety of datasets. Based on information theory, the meta-learner network is trained to minimize mutual information for high variety and cross-entropy loss using label information for high fidelity. The experiments show that the proposed framework reduces 12% in MSE on forecasting and improves 3.7% in accuracy on classification.

III. PROPOSED METHOD

The main advantage of contrastive learning is that it can learn useful representations from a large amount of data in an unsupervised manner. However, selecting or generating high fidelity and high variety positive—and/or negative—samples is a challenging task. Data augmentation is an efficient and effective method to generate positive pairs, but strong augmentation can distort the underlying semantics of the sample. Weak augmentation, on the other hand, generate augmented views with low variety. This can be fatal especially in time-series domain, where data augmentation cannot be guided using human prior due to its unrecognizable, high-dimensional, temporal structure.

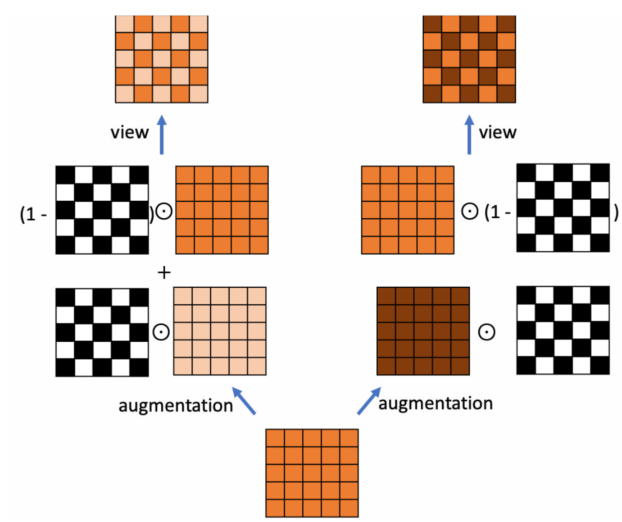

Fig. 2 A conceptual diagram of the proposed augmentation-agnostic masking technique (a fixed odd mask)

Hence, in this paper, we propose a simple, augmentation-agnostic masking technique that can be utilized with a variety of augmentations to maintain fidelity of the sample. The conceptual diagram of the proposed technique is shown in Fig. 2. The mask mixes the augmented view with the original data. The mask can be fixed, which can be randomly assigned, even/odd, or designed in custom. In addition, the fixed mask can be used with a trainable weight parameter alpha. Lastly, the mask itself can also be trained in end-to-end fashion. In this paper, we demonstrate the effectiveness of the proposed technique by using a fixed mask.

IV. Experiment

In this section, we demonstrate the effectiveness of the proposed augmentation-agnostic masking technique for contrastive learning. Although the experiment is conducted using time-series data, the technique can also be used in other domains such as image and tabular. In Fig 3., we show a random masking technique (p=0.5) on a simple sine wave to visualize how masking alters the augmented view.

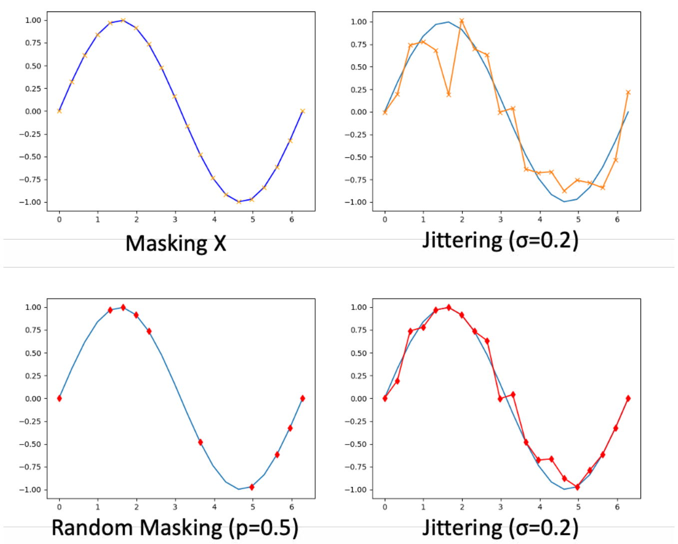

Fig. 3 An illustration of jitter and jitter with the masking (p=0.5) on a sine wave

To investigate the performance of the masking technique, we conducted an experiment under the similar setting as in [6]. We pre-trained the representation on the EEG dataset, and fine-tuned and tested on EPILEPSY dataset. The experiments were conducted in two settings: jitter for the baseline and jitter with the masking technique. The results are shown in Table 1.

TABLE I. Evaluation (Jitter vs. Jitter with masking)

	AUROC	AUPRC
Jitter	49.0251	53.2286
Masking (ours)	**52.4969**	**59.9171**

V. Conclusion

In this work, we first present a brief survey of time-series contrastive learning. Contrastive learning for time-series data is relatively limited, due to the difficulties in generating high fidelity, high variety views. Then, we introduce a simple augmentation-agnostic masking method to improve the fidelity of augmented views. For future work, we plan to conduct more extensive experiments to further investigate the effectiveness of the mask, including trainable masks. In addition, we also plan to extend the mask technique to other domains such as image and tabular data.

Acknowledgment

This work was supported by Institute of Information & Communication Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2020-0-00073, Development of Cloud-Edge based City-Traffic Brain Technology).

References

[1] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey E. Hinton, "A simple framework for contrastive learning of visual representations," Proceedings of the 37 th International Conference on Machine Learning, Vienna, Austria, PMLR 119, 2020.

[2] Chen, T., Sun, Y., Shi, Y., and Hong, L., "On sampling strategies for neural network-based collaborative filtering," In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 767–776, 2017.

[3] Jean-Bastien Grill, Florian Strub, Florent Altché, Corentin Tallec, Pierre H. Richemond, Elena Buchatskaya, Carl Doersch, Bernardo Avila Pires, Zhaohan Daniel Guo, Mohammad Gheshlaghi Azar, et al,. "Bootstrap your own latent: A new approach to self-supervised learning," 34th Conference on Neural Information Processing Systems, 2020.

[4] Dongsheng Luo, Wei Cheng, Yingheng Wang, Dongkuan Xu, Jingchao Ni, Wenchao Yu, Xuchao Zhang, Yanchi Liu, Yuncong Chen, Haifeng Chen, Xiang Zhang, "Time Series Contrastive Learning with Information-Aware Augmentations," The Thirty-Seventh AAAI Conference on Artificial Intelligence (AAAI-23)

[5] Emadeldeen Eldele, Mohamed Ragab, Zhenghua Chen, Min Wu, Chee Keong Kwoh, Xiaoli Li, Cuntai Guan, "Time-Series Representation Learning via Temporal and Contextual Contrasting," Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21)

[6] Xiang Zhang, Ziyuan Zhao, Theodoros Tsiligkaridis, Marinka Zitnik, "Self-Supervised Contrastive Pre-Training For Time Series via Time-Frequency Consistency," 36th Conference on Neural Information Processing Systems (NeurIPS 2022).

[7] Han Yu, Huiyuan Yang, Akane Sano, "LEAVES: Learning Views for Time-Series Data in Contrastive Learning," arXiv:2210.07340

[8] Beomsoo Kim, Jang-Ho Choi, Jaegul Choo, "Augmenting Imbalanced Time-series Data via Adversarial Perturbation in Latent Space," Proceedings of The 13th Asian Conference on Machine Learning, PMLR 157:1633-1644, 2021.

[9] Xiaochen Zheng, Xingyu Chen, Manuel Schürch, Amina Mollaysa, Ahmed Allam, Michael Krauthammer, "SimTS: Rethinking Contrastive Representation Learning for Time Series Forecasting," arXiv:2303.18205