TIME-SERIES REPRESENTATION LEARNING VIA TEMPORAL AND CONTEXTUAL CONTRASTING (TS-TCC)

Shubham Gurjar (210260048), Atharv Tambade (210260057), Atharv Deshpande (210260012),
Siddharth Stephen (210260049), Godwin Arunim Beck (210260016)

Abstract

Learning decent representations from unlabeled time-series data with temporal dynamics is a very challenging task. Here we use an unsupervised Time-Series representation learning framework via Temporal and Contextual Contrasting (TS-TCC), to learn time-series representation from unlabeled data. First, the raw timeseries data are transformed into two different yet correlated views by using weak and strong augmentations. Second, we use a temporal contrasting module to learn robust temporal representations by designing a tough cross-view prediction task. Last, to further learn discriminative representations, a contextual contrasting module is built upon the contexts from the temporal contrasting module. It attempts to maximize the similarity among different contexts of the same sample while minimizing similarity among contexts of different samples. Experiments have been carried out on three real-world time-series datasets. TS-TCC shows high efficiency in few-labeled data and transfer learning scenarios.

1. Introduction

Time-series data are often collected without human-recognizable patterns and are challenging to label. This lack of labeled data makes it difficult to apply traditional supervised learning methods, which typically require a large amount of labeled training data. There are various limitations of labeled data, so it's in our interest to use self supervised learning. where we can train models on unlabeled data by solving pretext tasks. This approach allows for effective representation learning even with limited labeled data.

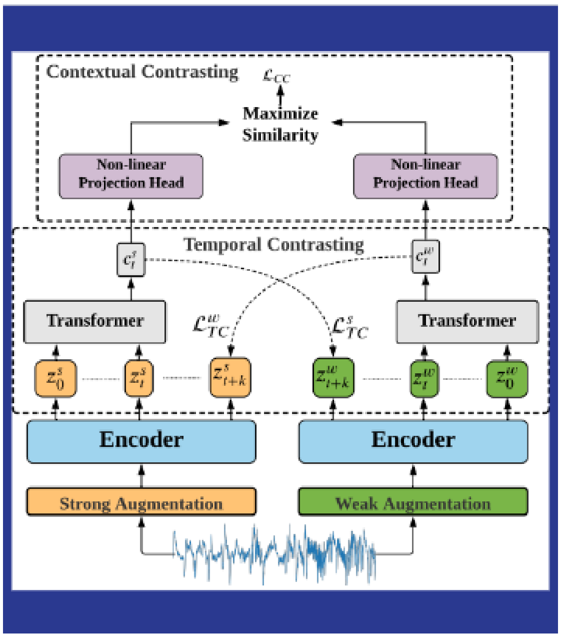


Figure 1: TS-TCC architecture

2. Problems we are addressing

- **Pretext Tasks Limitations:**Traditional self-supervised methods, which rely on pretext tasks like image rotation prediction or puzzle solving, may not be suitable for time-series data. As they won't capture the temporal dependencies
- **Contrastive Learning for Time-Series:** valuating and comparing automated approaches because most methods are compared to a single human scorer's annotations, so it may not provide an accurate assessment of the performance
- **Lack of Generalizable Methods:**While some contrastive learning methods have been proposed for specific applications in time-series data, they may not be generalizable to various types of time-series data

3. Input Sources:

3.1. Sleep EDF

- EEG signal data

- To classify into five different stages: Wake (W), Non-rapid eye movement (N1, N2, N3) and Rapid eye movement (REM).
- Includes whole-night PSG sleep recordings, with a single EEG channel (sampling rate 100Hz).

3.2. *Human Activity Recognition*

- The UCI HAR dataset
- To classify activity of a human into six categories - walking, walking upstairs, downstairs, standing, sitting, and lying down.
- Data collected from a wearable device

3.3. *Epilepsy*

- The Epilepsy Seizure Recognition dataset, again EEG recordings
- The five classes in the dataset merged into two - seizure and no seizure
- So then, binary classification

4. Implementation

We split the data into 60% , 20%, 20% for training, validation and testing, with considering subject-wise split for Sleep-EDF dataset to avoid overfitting and data leakage The pretraining and downstream tasks were done for 40 epochs, as we noticed that the performance does not improve with further training.

5. Framework of TS-TCC model

5.1. *Data Augmentation for Time-Series*

The framework starts by generating two different, yet correlated views of the input time-series data Two types of data augmentations are proposed: weak augmentation and strong augmentation Weak augmentation includes a jitter-and-scale strategy, adding random variations to the signal and scaling its magnitude. Strong augmentation includes a permutation-and-jitter strategy, which involves splitting the signal into segments and shuffling them, followed by adding random jitter. Careful consideration of augmentation hyperparameters is essential to match the nature of the time-series data.

5.2. *Temporal Contrasting*

The Temporal Contrasting module is designed to explore the temporal features of the data using an autoregressive model, particularly a Transformer model. The Transformer model includes successive blocks of multi-headed attention (MHA) and a multi-layer perceptron (MLP). The autoregressive model summarizes previous time steps into a context vector. A tough cross-view prediction task is defined, where one view's context is used to predict future timesteps in the other view and vice versa. A contrastive loss is employed to encourage agreement between the predicted representations and the true ones for the same sample while minimizing similarity with other samples.

$$\mathcal{L}_{TC}^{w} = -\frac{1}{K}\sum_{k=1}^{K}\log\frac{\exp((\mathcal{W}_k(c_t^w))^T z_{t+k}^s)}{\sum_{n\in\mathcal{N}_{t,k}}\exp((\mathcal{W}_k(c_t^w))^T z_n^s)}$$

$$\mathcal{L}_{TC}^{s} = -\frac{1}{K}\sum_{k=1}^{K}\log\frac{\exp((\mathcal{W}_k(c_t^s))^T z_{t+k}^w)}{\sum_{n\in\mathcal{N}_{t,k}}\exp((\mathcal{W}_k(c_t^s))^T z_n^w)}$$

Figure 2: The Temporal Contrasting Loss Functions

These loss functions are implemented as shown in figure 1.

5.3. *Contextual Contrasting*

The Contextual Contrasting module is introduced to learn more discriminative representations. Non-linear transformations are applied to the context vectors using a non-linear projection head. Positive pairs are defined, consisting of the context vector and its corresponding positive sample from the other augmented view of the same input. Negative pairs are formed by the context vector and other contexts within the same batch. A contextual contrasting loss is defined to maximize the similarity between positive pairs and minimize the similarity between negative pairs.

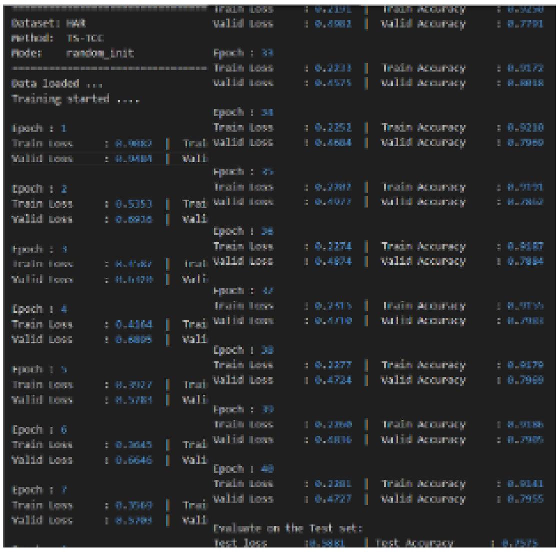
$$\mathcal{L}_{CC} = -\sum_{i=1}^{N}\log\frac{\exp\left(\operatorname{sim}\left(\boldsymbol{c}_t^i, \boldsymbol{c}_t^{i^+}\right)/\tau\right)}{\sum_{m=1}^{2N}\mathbb{1}_{[m\neq i]}\exp\left(\operatorname{sim}\left(\boldsymbol{c}_t^i, \boldsymbol{c}_t^m\right)/\tau\right)},$$

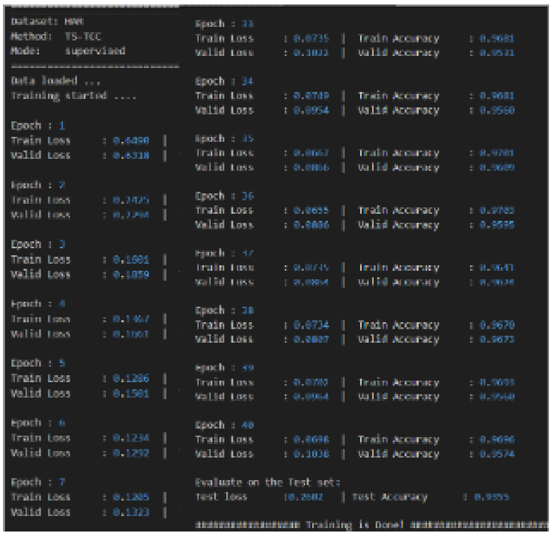
Figure 3: The Contextual Contrasting Loss

6. Results

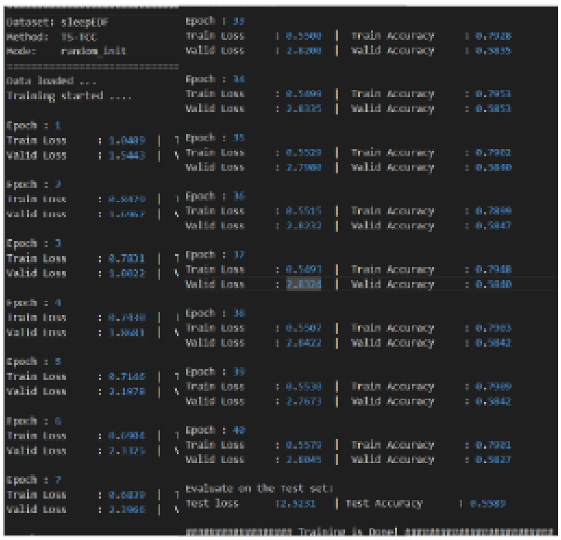
	HAR		Sleep-EDF		Epilepsy	
Baseline	ACC	MF1	ACC	MF1	ACC	MF1
Random Initialization	57.89±5.13	55.45±5.49	35.61±6.96	23.80±7.96	90.26±1.77	81.12±4.22
Supervised	90.14±2.49	90.31±2.24	**83.41±1.44**	**74.78±0.86**	96.66±0.24	94.52±0.43
SSL-ECG [P. Sarkar, 2020]	65.34±1.63	63.75±1.37	74.58±0.60	65.44±0.97	93.72±0.45	89.15±0.93
CPC [Oord *et al.*, 2018]	83.85±1.51	83.27±1.66	82.82±1.68	73.94±1.75	96.61±0.43	94.44±0.69
SimCLR [Chen *et al.*, 2020]	80.97±2.46	80.19±2.64	78.91±3.11	68.60±2.71	96.05±0.34	93.53±0.63
TS-TCC *(ours)*	**90.37±0.34**	**90.38±0.39**	83.00±0.71	73.57±0.74	**97.23±0.10**	**95.54±0.08**

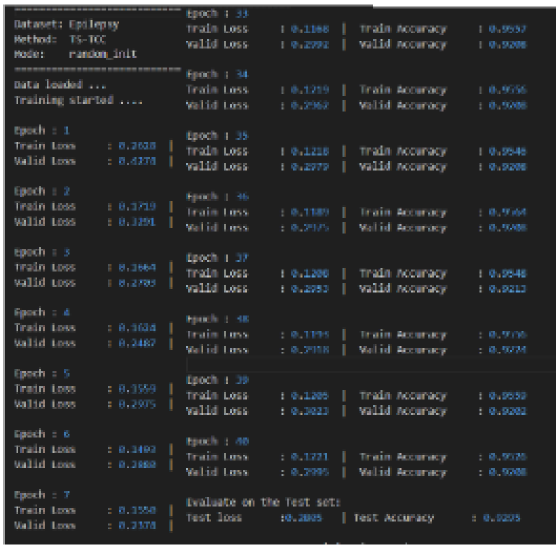
Figure 4: Comparison between our proposal TS-TCC model against baselines using linear classifier evaluation experiment





These above two are for the random initialization and supervised learning approaches on the HAR dataset.

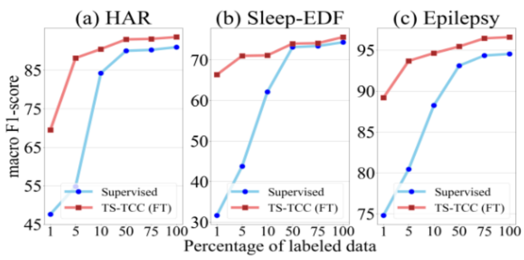




These two are for the Random Initialization on the sleepEDF and Epilepsy datasets, respectively.

7. Enlightening Observations

TS-TCC is highly effective in semi-supervised settings. Even with just 1% of labeled data, TS-TCC outperforms supervised training, showcasing its ability to leverage unlabeled data for improved performance. A comprehensive performance comparison of the proposed TS-TCC framework against several baseline methods, including Random Initialization, Supervised training, SSL-ECG, CPC, and SimCLR. It is observed that TS-TCC outperforms all the baseline methods, showcasing the effectiveness of the proposed framework for self-supervised time-series representation learning. Temporal features are more important than general features in time-series data, as CPC, which focuses on temporal information, performs better than SimCLR, which was originally designed for images.



8. Ways to the future

To advance this field further and build on the findings of this study, there are several directions and opportunities for future research:

1. Enhanced Architectures : Future work can explore more advanced neural network architectures tailored to time-series data. Investigating modifications or novel architectures for temporal feature extraction and contrastive learning can lead to even

more powerful self-supervised models
2. Generalization to Diverse Domains : While the study evaluates the performance of TS-TCC in various domains, future research can extend this evaluation to an even broader range of applications, including finance, healthcare, and environmental monitoring. Generalization across diverse domains is an important step for real-world applicability.
Future research should continue to push the boundaries of self-supervised learning for time-series data, making it more accessible, interpretable, and applicable to a wide array of real-world problems.

9. Final Words

In conclusion, the TS-TCC framework proves to be a powerful and efficient approach for self-supervised time-series representation learning. It is effective in various training settings, including linear evaluation, semi-supervised training, and transfer learning. We have developed a framework useful not just in sleep data classification, but highly important to all of artificial intelligence, increasing accuracy working with time-series data. The research findings highlight the importance of leveraging both temporal features and discriminative learning in the context of time-series data, offering a promising approach for real-world applications involving time-series data.

Acknowledgements

1. https://paperswithcode.com/task/automatic-sleep-stage-classification
2. https://www.ijcai.org/proceedings/2021/0324.pdf
Special thanks to Prof. Kshitij Jadhav and teaching assistant Raja Kumar for assisting us throughout the project.