

# Supplementary Materials: Multi-view Self-Supervised Contrastive Learning for Multivariate Time Series

Anonymous Authors

## 1 METHOD

Here, we delve into augmentation methods within the time and frequency domains, with a specific focus on analyzing various frequency augmentation techniques. To our knowledge, this is the first direct use of phase perturbation in contrastive learning for frequency domain augmentations. The following is a detailed description of the augmentations, and Figure 1 is a visualization example corresponding to each augmentation method.

**Scale:** The time series is scaled based on its distribution and amplitude range. It enables the model to understand scale variations across different features and the impact of amplitude changes on time series, thus helping to capture crucial feature information. Each time step is  $\tilde{x}_i = \eta x_i$ ,  $\eta \sim N(0, 3)$ .

**Jitter:** Disturbing time series by adding random noise. This simulates the noise and uncertainty in the real world, which enhances the model’s ability to adapt to noisy data environments and complex data contexts.  $\tilde{x}_i = \eta x_i + \eta_i$ ,  $\eta_i \sim N(0, 3)$ .

**Dropout Mask:** This technique generates new samples by random dropout masking or discarding certain data points. It simulates scenarios where data is missing or incomplete and helps the model handle such situations more effectively. Dropout mask requires the model to make accurate predictions even when some information is missing, contributing to a better understanding of contextual consistency. In our work, the dropout rate is set at 0.2.

**Add:** Adding some new frequency components with random amplitude perturbations into the original sequence.

**Remove:** Removing some frequency components.

**Phase Perturbation:** Introducing random noise to phase spectra, resulting in irregular spectral changes.

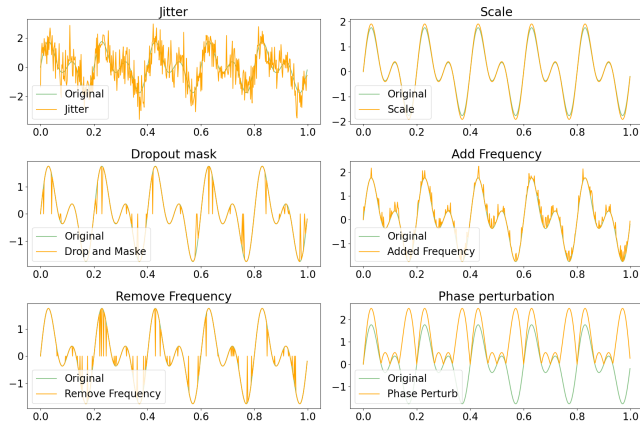


Figure 1: The visualization of the six augmentation methods

## 2 EXPERIMENTS

### 2.1 More Experiments Details

Table 1 summarizes the details of the classification dataset used in the model, i.e., the number of training samples (Train), the number of test samples (Test), the sample length (Length), the input and output channels (In\_C, Out\_C), and the number of classes (Class).

Table 2 summarizes the forecasting dataset details. The dimension indicates the number of time series (Dim.), i.e., channels, and the dataset size is organized in (training, validation, testing).

Table 1: Statistics of classification datasets used in our experiments.

Dataset	Train	Test	Length	In_C	Out_C	Class
Epilepsy	9200	2300	178	1	128	2
HAR	7352	2947	128	9	128	6
Sleep-EDF	25612	8910	3000	1	128	5
FD	8184	2728	5120	1	128	3

Table 2: Statistics of forecasting datasets used in our experiments.

Dataset	Dim.	Size	Frequency	Domain
ETTh1	7	(8545,2881,2881)	15 min	Temperature
ETTh2	7	(8545,2881,2881)	15 min	Temperature
ETTm1	7	(34465,11521,11521)	1 hour	Temperature
Weather	21	(31619,10539,10539)	10 min	Weather

### 2.2 More comparative Baselines

Classification: We compare our TFCC with recent works, i.e. SRL[1], SSL-ECG[5], TSTCC-DBPM[3]. Table 3 shows the comparison results, revealing that TFCC significantly outperforms them by a large margin across all datasets, outperforming SSL-ECG, SRL, and TSTCC-DBPM by 27.39%, 29.13%, and 4.62% in HAR, respectively.

Forecasting: We compare our model with other common representation learning and end-to-end baselines, i.e. SRL, CPC, LSTNet[2], LogTrans[4]. Table 4 shows that our proposed TFCC still exhibits significant advantages in all datasets and all lengths, and the MSE of our method decreases by 69.5%, 68.1%, 70.8%, and 58.1%; similarly, the MAE exceeds other models by 61.8%, 62.0%, 63.4%, 41.9%, which indicates the effectiveness of our framework in time series forecasting.

## REFERENCES

- [1] Jean-Yves Franceschi, Aymeric Dieuleveut, and Martin Jaggi. 2019. Unsupervised scalable representation learning for multivariate time series. *Advances in neural information processing systems* 32 (2019).

**Table 3: Classification results. Best results are highlighted in bold, while the second-best is underlined**

Methods	Epilepsy		HAR		Sleep-EDF	
	Accuracy	MF1	Accuracy	MF1	Accuracy	MF1
SSL-ECG	93.72±0.45	89.15±0.93	65.34±1.63	63.75±1.37	74.58±0.60	65.44±0.97
SRL	92.24±0.61	88.35±0.82	63.60±3.37	62.89±1.49	78.32±1.45	68.62±0.75
TSTCC-DBPM	<u>97.46±0.05</u>	<u>96.23±0.10</u>	<u>88.11±0.47</u>	<u>88.24±0.45</u>	<u>84.17±0.26</u>	<u>74.76±0.38</u>
<b>TFCC(ours)</b>	<b>98.15±0.17</b>	<b>97.92±0.35</b>	<b>92.73±0.57</b>	<b>92.52±0.59</b>	<b>84.35±0.52</b>	<b>75.60±0.45</b>

**Table 4: Multivariate forecasting results. Best results are highlighted in bold, while the second-best is underlined**

Methods		Representation Learning						Enf-to-End Forecasting			
		TFCC		SRL		CPC		LSTNet		LogTrans	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	24	<b>0.271</b>	<b>0.428</b>	0.698	0.661	0.687	0.634	1.293	0.901	<u>0.686</u>	<u>0.604</u>
	48	<b>0.417</b>	<b>0.452</b>	<u>0.758</u>	<u>0.711</u>	0.779	0.768	1.456	0.960	0.766	0.757
	168	<b>0.575</b>	<b>0.524</b>	1.341	1.178	1.282	1.083	1.997	1.214	<u>1.002</u>	<u>0.846</u>
	336	<b>0.789</b>	<b>0.677</b>	1.578	1.276	1.641	1.201	2.655	1.369	<u>1.362</u>	<u>0.952</u>
	720	<b>0.923</b>	<b>0.711</b>	1.892	1.566	1.803	1.761	2.143	1.380	<u>1.397</u>	<u>1.291</u>
ETTh2	24	<b>0.312</b>	<b>0.458</b>	1.034	0.901	0.981	0.869	2.742	1.457	<u>0.828</u>	<u>0.750</u>
	48	<b>0.618</b>	<b>0.636</b>	1.854	1.542	<u>1.732</u>	1.440	3.567	1.687	1.806	<u>1.034</u>
	168	<b>1.161</b>	<b>0.838</b>	5.062	2.167	4.591	3.126	<u>3.242</u>	2.513	4.070	<u>1.681</u>
	336	<b>1.218</b>	<b>0.862</b>	4.921	3.012	4.772	3.581	<u>2.544</u>	2.591	3.875	<u>1.763</u>
	720	<b>1.209</b>	<b>0.938</b>	5.301	3.207	4.772	2.781	4.625	3.709	<u>3.913</u>	<u>1.552</u>
ETTm1	24	<b>0.312</b>	<b>0.359</b>	0.561	0.603	0.540	0.513	1.968	1.170	<u>0.419</u>	<u>0.412</u>
	48	<b>0.331</b>	<b>0.382</b>	0.701	0.697	0.727	0.706	1.999	1.215	<u>0.507</u>	<u>0.583</u>
	168	<b>0.370</b>	<b>0.387</b>	0.901	0.836	0.851	0.793	2.762	1.542	<u>0.768</u>	<u>0.792</u>
	336	<b>0.411</b>	<b>0.514</b>	2.471	1.927	2.066	1.634	<u>1.257</u>	2.076	1.462	<u>1.320</u>
	720	<b>0.463</b>	<b>0.552</b>	2.042	1.803	1.962	1.797	1.917	2.941	<u>1.669</u>	<u>1.461</u>
Weather	24	<b>0.296</b>	<b>0.357</b>	0.688	0.701	0.647	0.652	0.615	0.545	<u>0.435</u>	<u>0.477</u>
	48	<b>0.405</b>	<b>0.462</b>	0.751	0.883	0.720	0.761	0.660	0.589	<u>0.426</u>	<u>0.495</u>
	168	<b>0.458</b>	<b>0.466</b>	1.204	1.032	1.351	1.067	0.748	<u>0.647</u>	<u>0.727</u>	0.671
	336	<b>0.462</b>	<b>0.467</b>	2.164	1.982	2.019	1.832	0.782	0.683	<u>0.754</u>	<u>0.670</u>
	720	<b>0.644</b>	<b>0.493</b>	2.281	1.994	2.109	1.861	<u>0.851</u>	<u>0.757</u>	0.885	0.773
Avg.		<b>0.582</b>	<b>0.548</b>	1.910	1.434	1.823	1.443	1.991	1.497	<u>1.388</u>	<u>0.944</u>

[2] Guokun Lai, Wei-Cheng Chang, Yiming Yang, and Hanxiao Liu. 2018. Modeling long-and short-term temporal patterns with deep neural networks. In *The 41st international ACM SIGIR conference on research & development in information retrieval*. 95–104.

[3] Xiang Lan, Hanshu Yan, Shenda Hong, and Mengling Feng. 2023. Towards Better Time Series Contrastive Learning: A Dynamic Bad Pair Mining Approach. *arXiv preprint arXiv:2302.03357* (2023).

[4] Shiyang Li, Xiaoyong Jin, Yao Xuan, Xiyu Zhou, Wenhui Chen, Yu-Xiang Wang, and Xifeng Yan. 2019. Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting. *Advances in neural information processing systems* 32 (2019).

[5] Pritam Sarkar and Ali Etemad. 2020. Self-supervised ECG representation learning for emotion recognition. *IEEE Transactions on Affective Computing* 13, 3 (2020), 1541–1554.