Unsupervised Time-Series Representation Learning with Iterative Bilinear Temporal-Spectral Fusion

Ling Yang, Shenda Hong

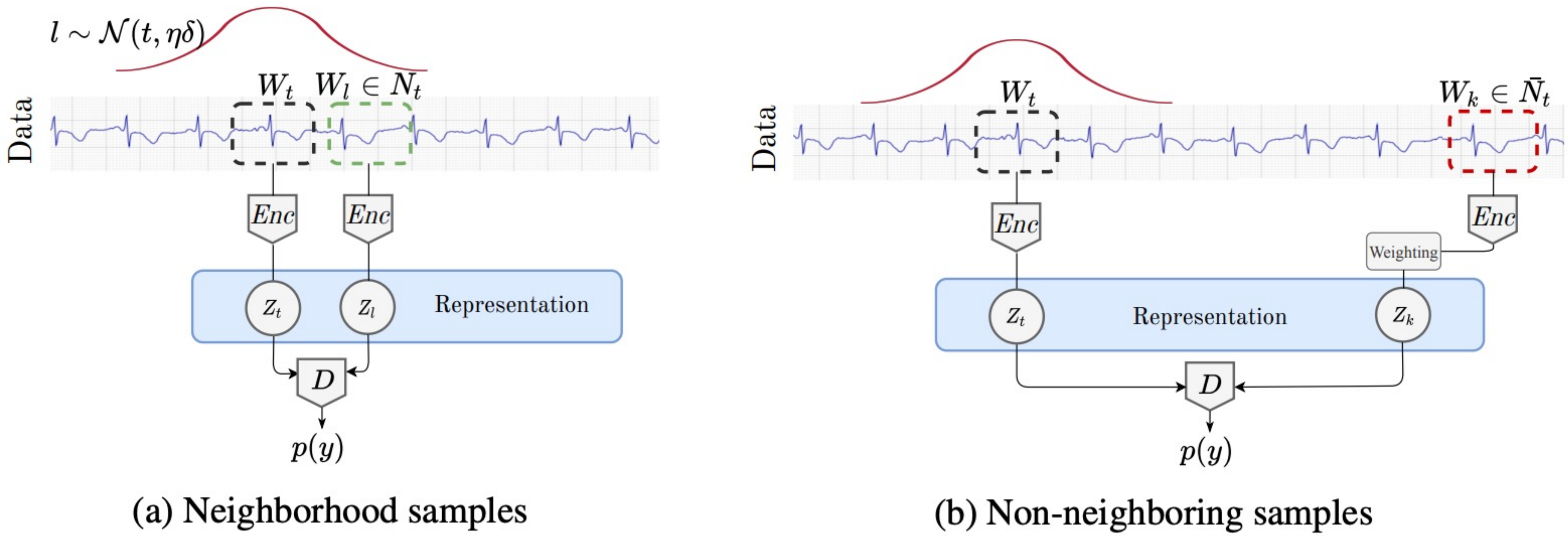
Presenter ： Ling Yang

Contact ： yangling0818@163.com

Contents

Related Work

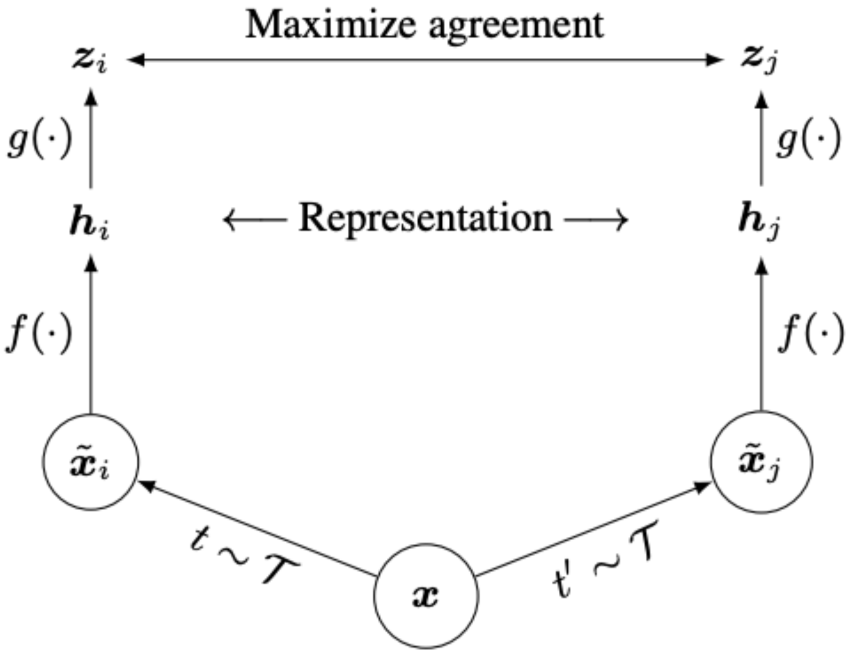
Unsupervised representation learning for time series



Tonekaboni S, Eytan D, Goldenberg A. Unsupervised Representation Learning for Time Series with Temporal Neighborhood Coding[C]//International Conference on Learning Representations. 2020.

Related Work

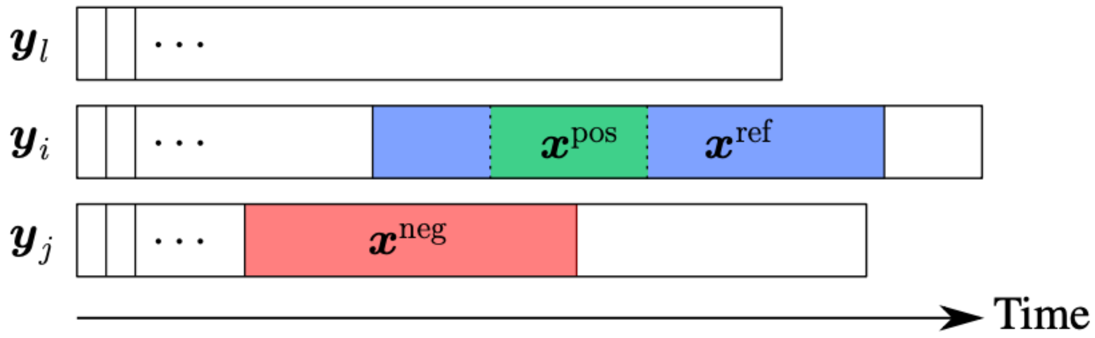
Contrastive learning



Chen T, Kornblith S, Norouzi M, et al. A simple framework for contrastive learning of visual representations[C]//International conference on machine learning. PMLR, 2020: 1597-1607.

Related Work

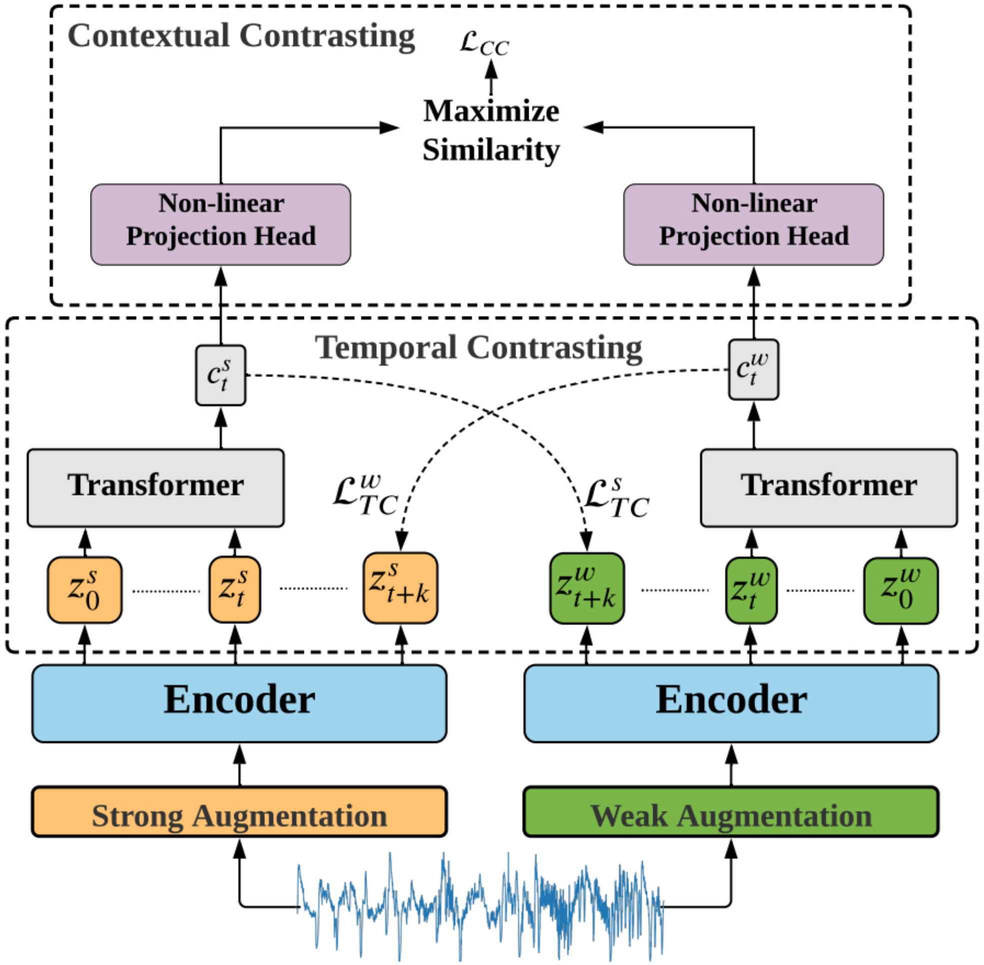
Time series contrastive learning



Franceschi J Y, Dieuleveut A, Jaggi M. Unsupervised scalable representation learning for multivariate time series[J]. Advances in neural information processing systems, 2019, 32.

Related Work

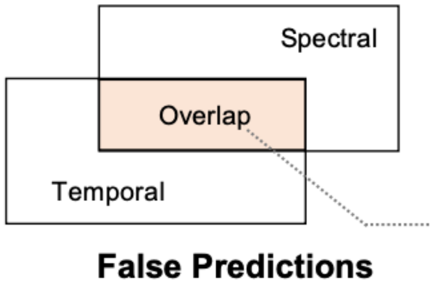
Time series contrastive learning



Eldele E, Ragab M, Chen Z, et al. Time-series representation learning via temporal and contextual contrasting[J]. arXiv preprint arXiv:2106.14112, 2021.

Limitations of Existing Approaches

1. Inconsistency between temporal and spectral representations
2. Severely based on the nearest principle when choosing positive and negative samples



Method	Overlap
SRL	32.98%
CPC	25.02%
TS-TCC	29.49%
TNC	31.08%

Figure 1. Statistics about false predictions of randomly selected evaluation samples.

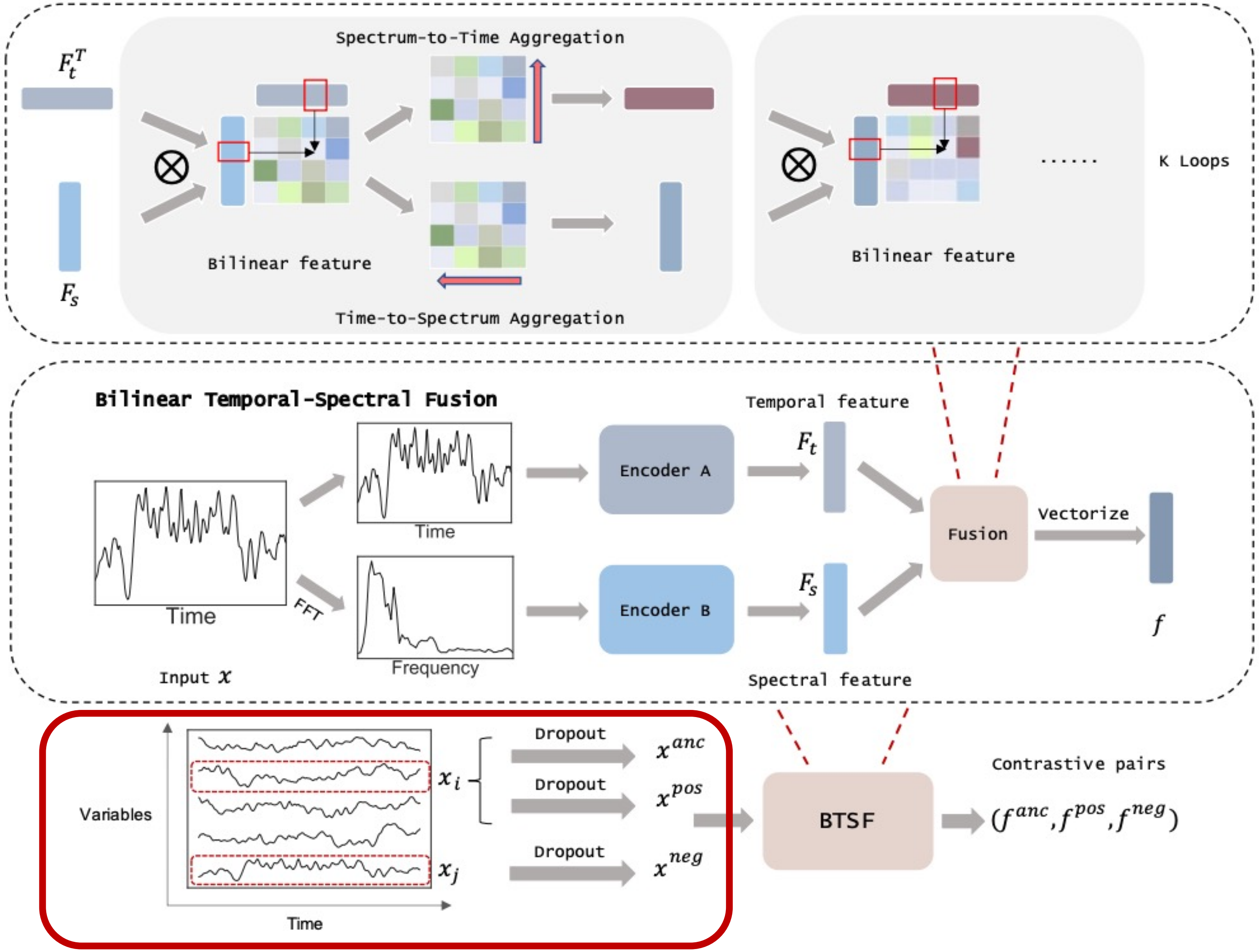
Our Proposed Framework

北京大學
PEKING UNIVERSITY

Main Contributions:

1. Instance-level augmentation technique
2. A novel iterative bilinear temporal-spectral fusion
3. Sufficient assessments including alignment and uniformity
4. Significantly outperforms previous works in downstream classification, forecasting and anomaly detection tasks

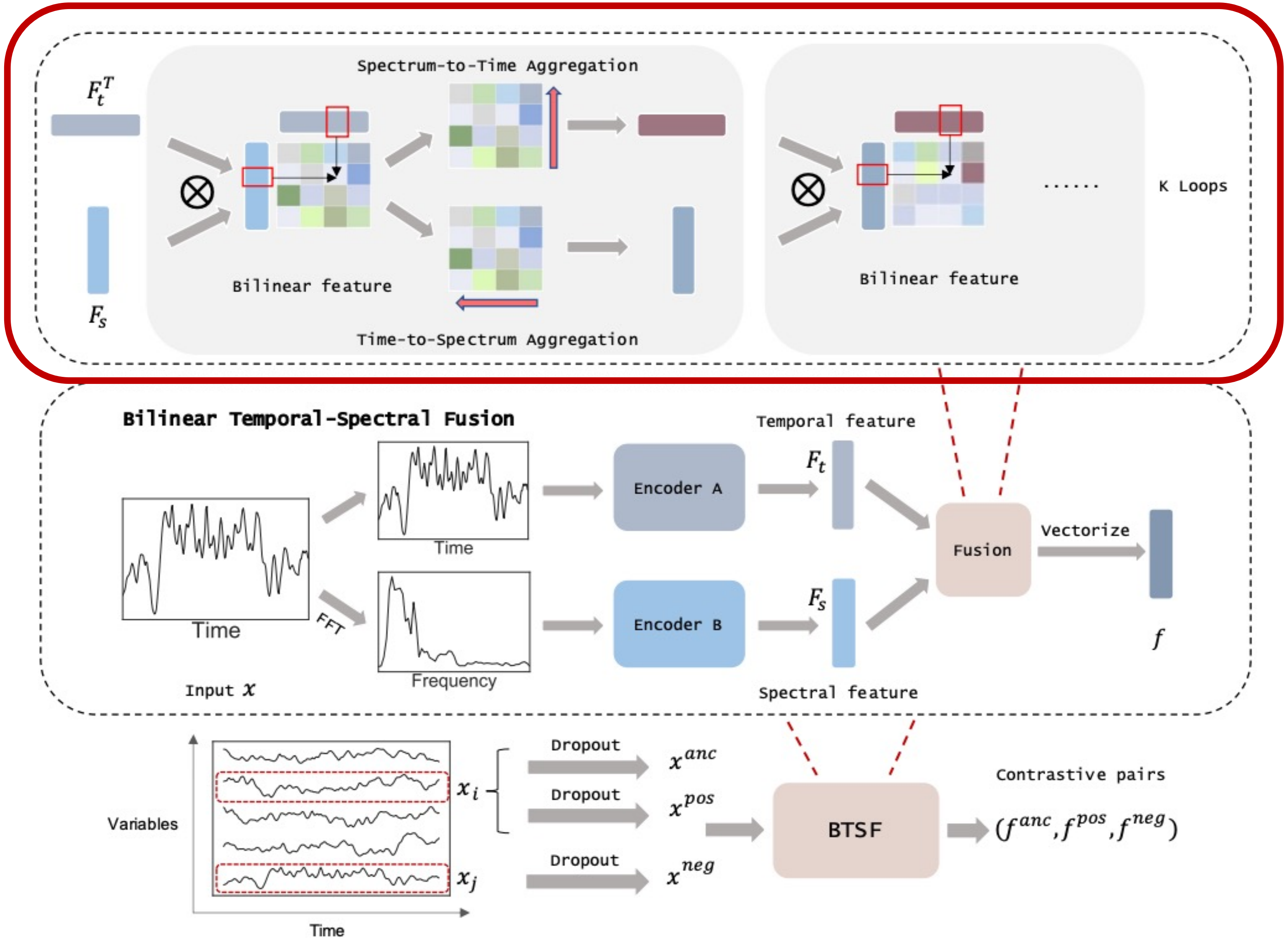
Our Proposed Framework —— BTSF



Instance-level augmentation technique

$$\boldsymbol{x}^{anc} = Dropout(\boldsymbol{x}), \quad \boldsymbol{x}^{pos} = Dropout(\boldsymbol{x}).$$

Our Proposed Framework —— BTSF



Iterative bilinear temporal-spectral fusion

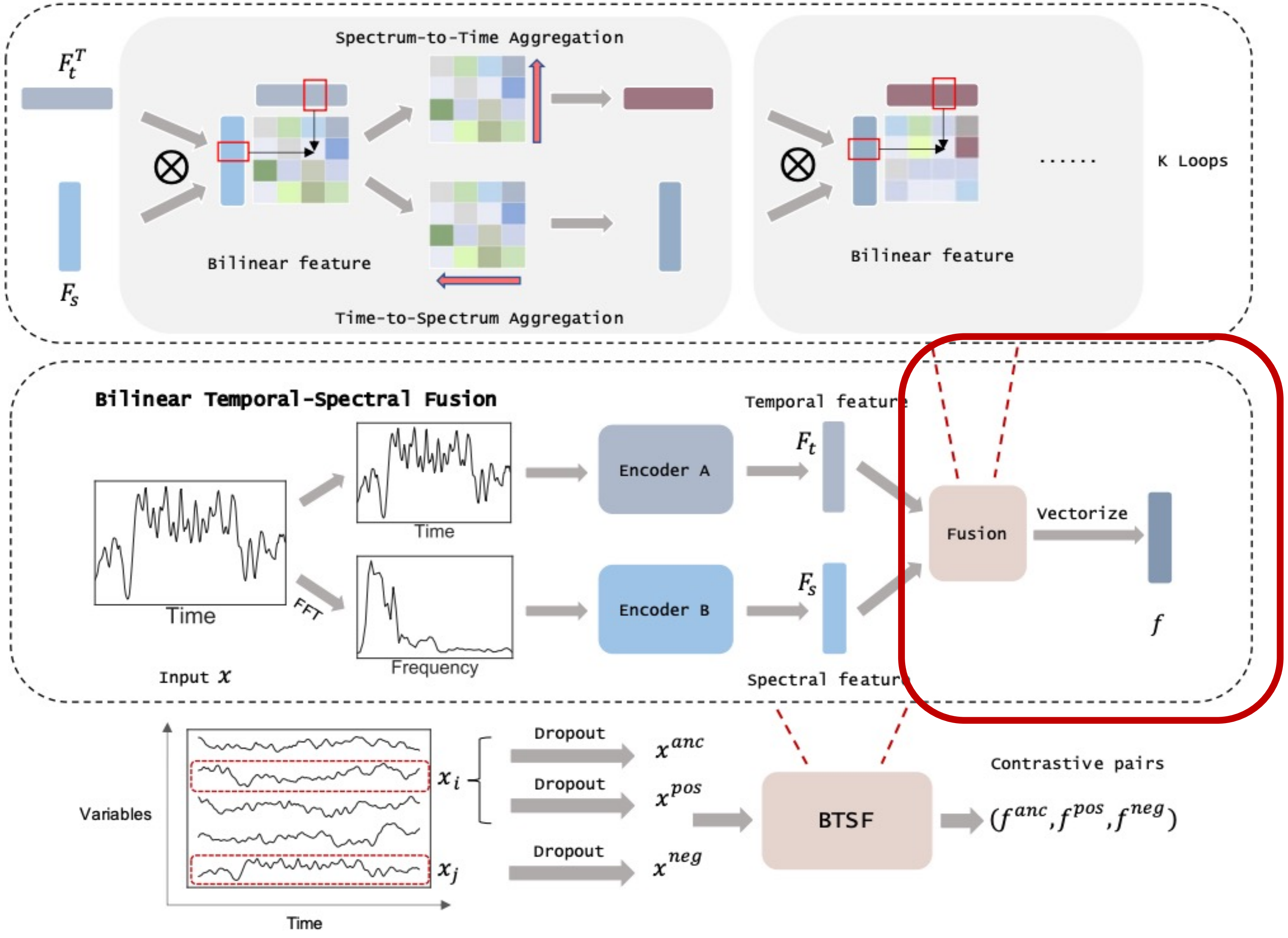
$$\boldsymbol{F}_{bilinear} = \boldsymbol{F}_t{}^T \times \boldsymbol{F}_s = \sum_{i=1}^{m}\sum_{j=1}^{n} \boldsymbol{F}(i,j)$$
$$= \sum_{i=1}^{m}\sum_{j=1}^{n} \boldsymbol{F}_t(i)^T \boldsymbol{F}_s(j)$$

$$\boldsymbol{F}_{bilinear} = \boldsymbol{F}_t{}^T \times \boldsymbol{U} \times \boldsymbol{V}^T \times \boldsymbol{F}_s$$
$$= (\boldsymbol{U}^T \times \boldsymbol{F}_t) \circ (\boldsymbol{V}^T \times \boldsymbol{F}_s)$$

$$\text{S2T}: \quad \boldsymbol{F}_t = BiCasual(Conv(\boldsymbol{F}_{bilinear}))$$
$$\text{T2S}: \quad \boldsymbol{F}_s = Conv(BiCasual(\boldsymbol{F}_{bilinear}))$$

Our Proposed Framework —— BTSF

北京大学 PEKING UNIVERSITY



Output and Loss Function

$$\boldsymbol{f} = \sigma(\boldsymbol{W_t}^T \times \boldsymbol{F_t} + \boldsymbol{W_s}^T \times \boldsymbol{F_s} + \boldsymbol{F_t}^T \times \boldsymbol{W} \times \boldsymbol{F_s})$$

$$\mathcal{L} = \mathbb{E}_{\boldsymbol{X} \sim \boldsymbol{P}_{data}}[-log(sim(\boldsymbol{f}^{anc}, \boldsymbol{f}^{pos})/\tau) + \mathbb{E}_{\boldsymbol{x}^{neg} \sim \boldsymbol{X}}[log(sim(\boldsymbol{f}^{anc}, \boldsymbol{f}^{neg})/\tau)]]$$

Our Proposed Framework —— BTSF

Effectiveness of the Proposed BTSF

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \boldsymbol{F}_t} &= \frac{\partial \mathcal{L}}{\partial \boldsymbol{f}}\boldsymbol{W}_t + \frac{\partial \mathcal{L}}{\partial \boldsymbol{f}}\boldsymbol{W} \times \boldsymbol{F}_s, \\ \frac{\partial \mathcal{L}}{\partial \boldsymbol{F}_s} &= \frac{\partial \mathcal{L}}{\partial \boldsymbol{f}}\boldsymbol{W}_s + \frac{\partial \mathcal{L}}{\partial \boldsymbol{f}}\boldsymbol{W}^T \times \boldsymbol{F}_t\end{aligned} \tag{12}$$

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \boldsymbol{W}_t} &= \frac{\partial \mathcal{L}}{\partial \boldsymbol{f}}\boldsymbol{F}_t, \\ \frac{\partial \mathcal{L}}{\partial \boldsymbol{W}_s} &= \frac{\partial \mathcal{L}}{\partial \boldsymbol{f}}\boldsymbol{F}_s, \\ \frac{\partial \mathcal{L}}{\partial \boldsymbol{W}} &= \frac{\partial \mathcal{L}}{\partial \boldsymbol{f}}\boldsymbol{F}_t \times \boldsymbol{F}_s{}^T\end{aligned} \tag{13}$$

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}_t} &= \frac{\partial \mathcal{L}}{\partial \boldsymbol{f}}\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{F}_t}\boldsymbol{W}_t + \frac{\partial \mathcal{L}}{\partial \boldsymbol{f}}\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{F}_t}\boldsymbol{W} \times \boldsymbol{F}_s, \\ \frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}_s} &= \frac{\partial \mathcal{L}}{\partial \boldsymbol{f}}\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{F}_s}\boldsymbol{W}_s + \frac{\partial \mathcal{L}}{\partial \boldsymbol{f}}\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{F}_s}\boldsymbol{W}^T \times \boldsymbol{F}_t\end{aligned} \tag{14}$$

PEKING UNIVERSITY 北京大學

Results and Analysis

Time-Series Classification

Table 1. Comparisons of classification results.

Methods	HAR		Sleep-EDF		ECG Waveform	
	Accuracy	AUPRC	Accuracy	AUPRC	Accuracy	AUPRC
Supervised	92.03±2.48	0.98±0.00	83.41±1.44	0.78±0.52	84.81±0.28	0.67±0.01
KNN	84.85±0.84	0.75±0.01	64.87±1.73	0.75±2.88	54.76±5.46	0.38±0.06
SRL	63.60±3.37	0.71±0.01	78.32±1.45	0.71±2.83	75.51±1.26	0.47±0.00
CPC	86.43±1.41	0.93±0.01	82.82±1.68	0.73±2.15	68.64±0.49	0.42±0.01
TS-TCC	88.04±2.46	0.92±0.02	83.00±0.71	0.74±2.63	74.81±1.10	0.53±0.02
TNC	88.32±0.12	0.94±0.01	82.97±0.94	0.76±1.73	77.79±0.84	0.55±0.01
BTSF	**94.63±0.14**	**0.99±0.01**	**87.45±0.54**	**0.79±0.74**	**85.14±0.38**	**0.68±0.01**

北京大学 PEKING UNIVERSITY

Results and Analysis

Time-Series Forecasting

Table 2. Comparisons of multivariate forecasting results.

Datasets	Length	Supervised		SRL		CPC		TS-TCC		TNC		**BTSF**	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	24	0.577	0.549	0.698	0.661	0.687	0.634	0.653	0.610	0.632	0.596	**0.541**	**0.519**
	48	0.685	0.625	0.758	0.711	0.779	0.768	0.720	0.693	0.705	0.688	**0.613**	**0.524**
	168	0.931	0.752	1.341	1.178	1.282	1.083	1.129	1.044	1.097	0.993	**0.640**	**0.532**
	336	1.128	0.873	1.578	1.276	1.641	1.201	1.492	1.076	1.454	0.919	**0.864**	**0.689**
	720	1.215	0.896	1.892	1.566	1.803	1.761	1.603	1.206	1.604	1.118	**0.993**	**0.712**
ETTh2	24	0.720	0.665	1.034	0.901	0.981	0.869	0.883	0.747	0.830	0.756	**0.663**	**0.557**
	48	1.451	1.001	1.854	1.542	1.732	1.440	1.701	1.378	1.689	1.311	**1.245**	**0.897**
	168	3.389	1.515	5.062	2.167	4.591	3.126	3.956	2.301	3.792	2.029	**2.669**	**1.393**
	336	2.723	1.340	4.921	3.012	4.772	3.581	3.992	2.852	3.516	2.812	**1.954**	**1.093**
	720	3.467	1.473	5.301	3.207	5.191	2.781	4.732	2.345	4.501	2.410	**2.566**	**1.276**
ETTm1	24	0.323	0.369	0.561	0.603	0.540	0.513	0.473	0.490	0.429	0.455	**0.302**	**0.342**
	48	0.494	0.503	0.701	0.697	0.727	0.706	0.671	0.665	0.623	0.602	**0.395**	**0.387**
	96	0.678	0.614	0.901	0.836	0.851	0.793	0.803	0.724	0.749	0.731	**0.438**	**0.399**
	288	1.056	0.786	2.471	1.927	2.066	1.634	1.958	1.429	1.791	1.356	**0.675**	**0.429**
	672	1.192	0.926	2.042	1.803	1.962	1.797	1.838	1.601	1.822	1.692	**0.721**	**0.643**
Weather	24	0.335	0.381	0.688	0.701	0.647	0.652	0.572	0.603	0.484	0.513	**0.324**	**0.369**
	48	0.395	0.459	0.751	0.883	0.720	0.761	0.647	0.691	0.608	0.626	**0.366**	**0.427**
	168	0.608	0.567	1.204	1.032	1.351	1.067	1.117	0.962	1.081	0.970	**0.543**	**0.477**
	336	0.702	0.620	2.164	1.982	2.019	1.832	1.783	1.370	1.654	1.290	**0.568**	**0.487**
	720	0.831	0.731	2.281	1.994	2.109	1.861	1.850	1.566	1.401	1.193	**0.601**	**0.522**

Results and Analysis

Time-Series Anomaly Detection

Table 3. Comparisons of multivariate anomaly detection.

Datasets	Metric	Supervised	SRL	CPC	TS-TCC	TNC	**BTSF**
SAaT	F1	0.901	0.710	0.738	0.775	0.799	**0.914**
WADI	F1	0.649	0.340	0.382	0.427	0.440	**0.653**
SMD	F1	0.958	0.768	0.732	0.794	0.817	**0.972**
SMAP	F1	0.842	0.598	0.620	0.679	0.693	**0.863**
MSL	F1	0.945	0.788	0.813	0.795	0.833	**0.957**

北京大学
PEKING UNIVERSITY

Results and Analysis

Alignment and Uniformity

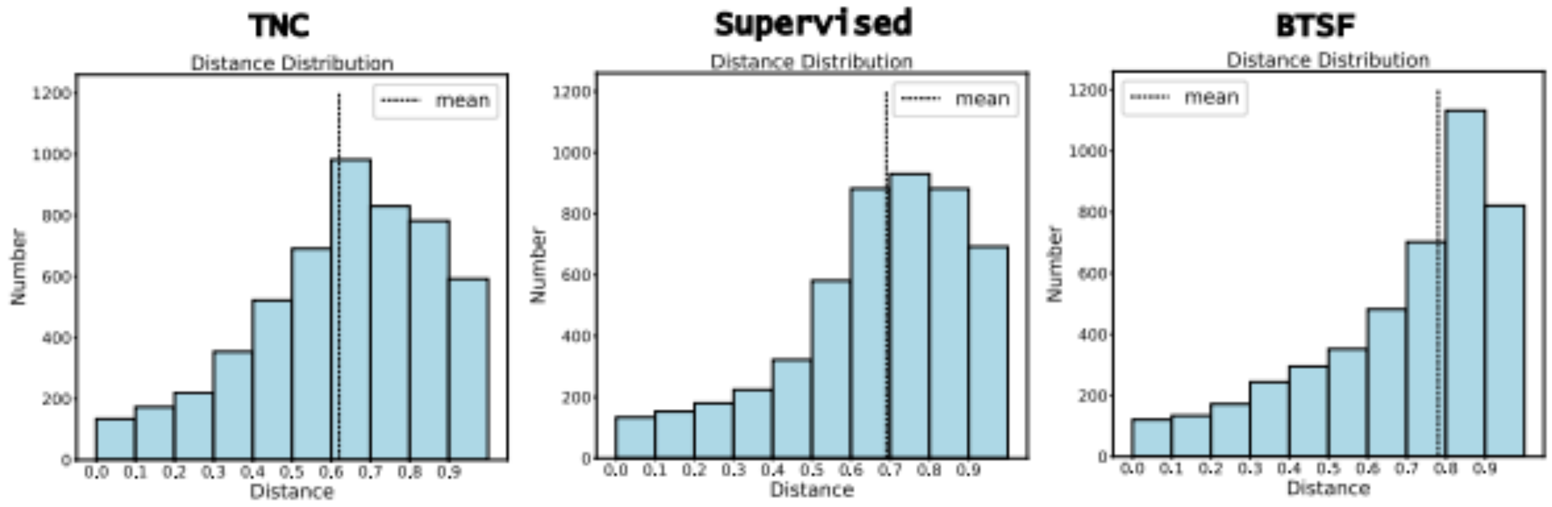


Figure 6. Distance distribution of positive pairs for assessing alignment. Our BTSF is well aligned.

Results and Analysis

Alignment and Uniformity

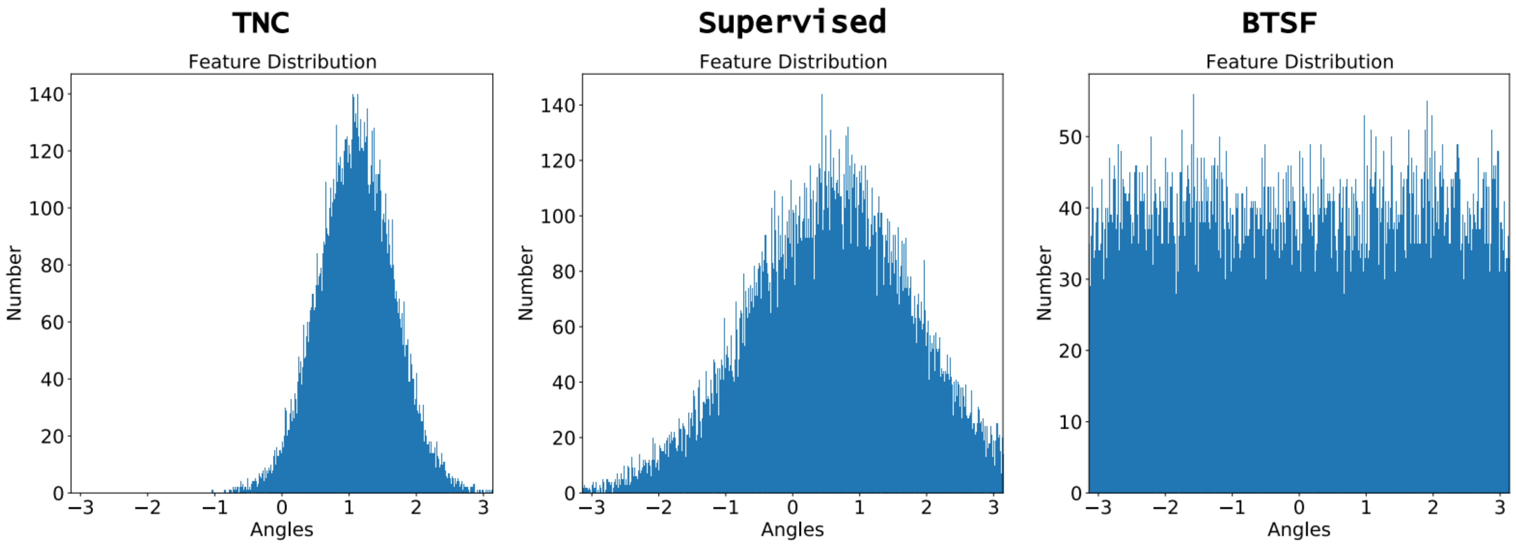


Figure 7. Feature distribution of samples in different classes on the normalized surface area for assessing uniformity. Features extracted by BTSF are evenly distributed.

Results and Analysis

Visualization of Learned Representations

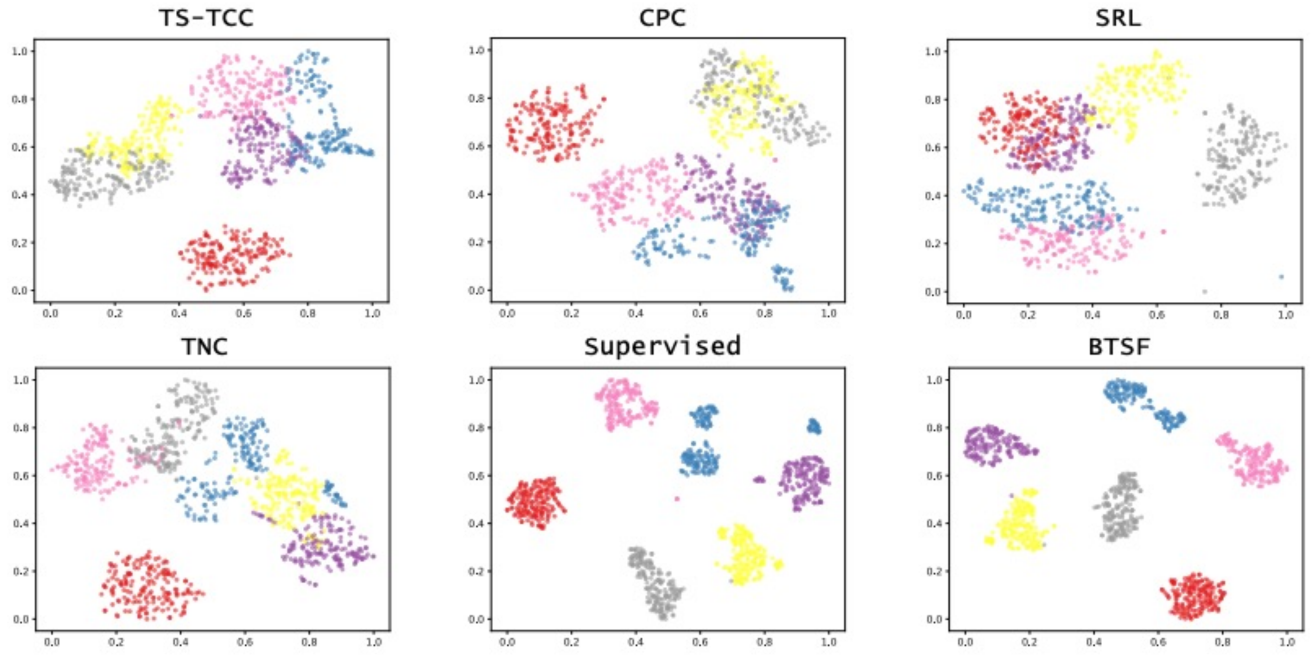


Figure 4. T-SNE visualization of signal representations for HAR dataset.

Thank You!