PPRL:Prototype-driven Periodic Representation Learning for Multivariate Time Series Anomaly Detection

Mingle Zhou[1,2], Shijie Miao[1,2], Yuan Gao[1,2], Min Li[1,2], and Delong Han[1,2*]

[1] *Key Laboratory of Computing Power Network and Information Security, Ministry of Education, Shandong Computer Science Center (National Supercomputer Center in Jinan), Qilu University of Technology (Shandong Academy of Sciences), Jinan, China*

[2] *Shandong Provincial Key Laboratory of Computing Power Internet and Service Computing, Shandong Fundamental Research Center for Computer Science, Jinan, China*

Email: zhouml@qlu.edu.cn, 1084281038@qq.com, yuangao@qlu.edu.cn, limin@qlu.edu.cn, handl@qlu.edu.cn*

***Abstract*—Multivariate time series anomaly detection is critical in many real-world applications, where periodic structure is pervasive and anomalies often appear as subtle violations of temporal regularities. Existing reconstruction-based methods, even when they incorporate frequency-domain modeling, typically capture periodic patterns only implicitly. As a result, they may over-reconstruct anomalous segments and wash out the residual cues needed to detect periodic anomalies. To tackle this, we propose PPRL, a self-supervised framework that makes periodicity an explicit structure through learnable periodic prototypes. PPRL builds a multi-scale periodic prototype representation and introduces two training-only objectives: prototype-driven periodic contrastive learning and prototype-driven periodic anomaly regularization. These objectives better align representation learning with periodic structure violations. The learned periodic representations are injected into a time–frequency reconstruction backbone, while inference-time scoring still relies solely on reconstruction residuals. Extensive experiments on four public benchmarks show that PPRL delivers consistent gains, especially under affiliation-based metrics, and mitigates the over-reconstruction of periodic anomalies without increasing inference-time complexity.**

***Index Terms*—Multivariate time series, anomaly detection, prototype learning,periodic modeling**

I. Introduction

Multivariate time series anomaly detection (MTS-AD) sits at the heart of many real-world systems—from industrial monitoring and data-center operations to energy management and business metric surveillance [1]–[3]. In reality, anomaly labels are rarely something we can count on: they are costly to obtain, often sparse, and sometimes simply missing. That is why self-supervised anomaly detection has become a practical and widely used default in MTS-AD [3].

A persistent challenge in time series anomaly detection is the pervasive periodic structure of real-world signals, such as daily or weekly cycles and operational rhythms. Importantly, many high-impact anomalies manifest not as isolated spikes but as subtle disruptions of periodic regularity—e.g., phase or frequency drift, harmonic distortion, or abnormal cross-variable coupling—which are difficult to capture with conventional residual-based criteria that favor large deviations [4].

*Corresponding author.

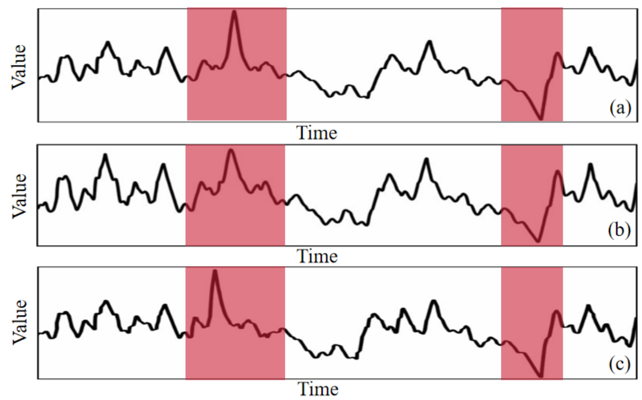

Fig. 1. Illustration of over-reconstruction and the effect of prototype-driven periodic modeling. Prototype guidance helps preserve anomalous deviations rather than smoothing them toward normal patterns.

Most existing self-supervised MTS-AD methods still follow the reconstruction or prediction paradigm: the model is trained to fit normal temporal dynamics, and anomalies are then flagged by reconstruction errors or prediction residuals [5]–[10]. More recently, some studies have started to bring in frequency-domain or time–frequency modeling, mainly to represent periodicity more directly [11], [12]. Even so, periodic anomalies are still hard to catch reliably.

A key issue is that periodicity is usually learned *implicitly* in high-dimensional representations, with no explicit structure to separate normal periodic regimes from violations. Reconstruction-oriented objectives further exacerbate this by rewarding faithful recovery, which can turn periodic anomalies into normal-looking patterns. In large-capacity models, this often results in over-reconstruction and small residuals despite clear periodic violations [7], [8], [11], [12].

Fig. 1 illustrates this phenomenon and motivates our approach. (a) shows an input window with anomalous segments highlighted. (b) depicts a reconstruction model without prototype guidance: the reconstruction tends to follow the dominant normal rhythm and smooths abnormal deviations toward normal-looking patterns, thereby reducing the residual signal available for detection. In contrast, (c) shows the behavior with prototype-driven periodic modeling. By anchoring periodic representations to a learnable prototype space, the model becomes less inclined to explain away periodic structure violations as normal variations. As a result, anomalous deviations are preserved to a greater extent, leading to more informative reconstruction residuals. This observation suggests that improving sensitivity to periodic anomalies does not necessarily require a heavier inference pipeline, but rather a more structured representation of periodic patterns and training objectives that do not reward repairing periodic violations into normality.

Motivated by this insight, we propose PPRL, a self-supervised framework that turns periodicity into an explicit structure in a learnable prototype space. PPRL follows a two-branch design: a time–frequency reconstruction backbone keeps the encoding–decoding pipeline concise and provides residual-based anomaly signals at inference, while a periodic prototype branch maps each input window into the prototype space and feeds periodic information back into the backbone representations. To capture heterogeneous periodic behaviors across variables and temporal scales, we design a multi-scale periodic prototype representation module (MPPR) that aggregates periodic cues from multiple time–frequency views [11], [12]. We further introduce two prototype-driven training objectives—prototype-driven periodic contrastive learning (PCL) and prototype-driven periodic anomaly regularization (PAR)—to shape the prototype space and better align representation learning with periodic structure violations. The prototype branch runs in the forward pass, whereas PCL and PAR are only used during training; anomaly scoring at inference still relies solely on reconstruction residuals. In this way, PPRL improves sensitivity to periodic anomalies through structured periodic modeling and objective alignment, without changing the inference-time scoring mechanism.

Extensive experiments on multiple public benchmark datasets with diverse periodic characteristics demonstrate that PPRL consistently improves detection performance, particularly for periodic structure violations, and achieves competitive results under both AUC-ROC and Affiliation-F metrics [13].

Our main contributions are summarized as follows:

- We propose PPRL, a prototype-driven periodic representation learning framework that explicitly models periodic patterns in a learnable prototype space and improves sensitivity to periodic anomalies without altering the inference-time scoring mechanism.
- We design the MPPR that leverages multi-scale time–frequency views and prototype-based aggregation to capture heterogeneous periodic regimes.
- We introduce prototype-driven training objectives, including PCL and PAR, to align representation learning with periodic structure violations and mitigate over-reconstruction.

The remainder of this paper is organized as follows. Section II reviews related work on reconstruction- and prediction-based anomaly detection, periodicity-aware time–frequency methods, and self-supervised learning for time series. Section III presents PPRL in detail. Section IV reports experimental results and ablation studies, followed by conclusions in Section V. Upon acceptance, we will release our code and the experimental configurations to facilitate reproducibility.

II. Related Work

A. Reconstruction- and Forecasting-based MTS-AD

Reconstruction- and forecasting-based methods constitute a dominant paradigm for self-supervised MTS-AD.They learn normal temporal dynamics from unlabeled data and score anomalies using reconstruction errors or prediction residuals.Representative lines of work include probabilistic generative modeling for seasonal signals (e.g., VAE-based approaches) [14], deep convolutional forecasting with residual scoring [15], and generative-adversarial frameworks that exploit both reconstruction and discrimination signals for multivariate streams [16], [17].

Despite their effectiveness, residual-based scoring can be challenged when the backbone becomes highly expressive. In such cases, the model may partially explain abnormal segments and reduce residual contrast, especially for anomalies that are not point-wise spikes but appear as subtle structural deviations over time. Moreover, pure reconstruction/forecasting objectives provide limited inductive bias to explicitly separate *typical* recurring regimes from *violations* of those regimes, which can make periodic-structure anomalies harder to expose reliably in the residual space [14], [17].

B. Modeling Multivariate Dependencies for Anomaly Detection

A key difficulty in MTS-AD is that anomalies often show up through complex cross-variable interactions, rather than as clean univariate deviations. To capture these dependencies, recent methods often turn to graph-based or hierarchical modeling. MTAD-GAT introduces graph attention to jointly model temporal dynamics and inter-sensor relations, improving robustness under correlated failures [18]. GDN further frames anomaly detection as deviation discovery on learned graph structures, enabling adaptive dependency modeling without manual graph design [19]. InterFusion proposes hierarchical inter-metric and temporal embeddings to disentangle inter-variable interactions and temporal factors, supporting both detection and interpretation [20].

These dependency-aware methods clearly strengthen multivariate modeling. Still, their representations are usually optimized for generic reconstruction/prediction fidelity or generic latent consistency, and they do not necessarily enforce an explicit organization of recurring periodic regimes that could stabilize the notion of periodic normality under multi-scale and non-stationary rhythms.

C. Self-supervised and Contrastive Representation Learning for Time Series

Self-supervised representation learning has become an effective alternative to purely residual-driven objectives in time series analysis. Temporal Neighborhood Coding (TNC) exploits local stationarity by constructing neighborhood-based positives and negatives to encourage temporally smooth representations [21]. TS2Vec further introduces a hierarchical contrastive objective that learns representations at both timestamp and subsequence levels, demonstrating strong transferability to downstream tasks such as anomaly detection [22]. From a frequency-aware perspective, CoST learns disentangled seasonal and trend representations and applies contrastive learning across time and frequency views, highlighting the importance of explicitly modeling periodic components.

Despite their effectiveness, most existing contrastive formulations primarily focus on learning general invariances and discriminative representations. For anomaly detection under periodic structure violations, it remains challenging to construct a representation space in which periodic regimes are explicitly anchored and deviations from these regimes remain consistently detectable.

III. Method

A. Problem Definition

Let $\mathbf{X} \in \mathbb{R}^{T \times D}$ denote a multivariate time series with D variables and length T. We segment $\mathbf{X}$ into overlapping windows $\mathbf{X}^{(i)} \in \mathbb{R}^{L \times D}$ using a sliding window of length L and stride S. PPRL operates in a *self-supervised* setting without anomaly annotations and learns an anomaly scoring function that assigns a score to each window (and subsequently to each time step via overlap aggregation), where larger scores indicate higher abnormality. This work focuses on anomalies arising from *violations of periodic structures* (e.g., phase/frequency drift, harmonic distortion, and abnormal periodic relations across variables).

B. Overall Framework

PPRL adopts a two-branch architecture, as illustrated in Fig. 2. The first branch is a time–frequency reconstruction backbone, which reconstructs each input window and provides reconstruction residuals as the sole source of anomaly scores. The second branch is a periodic prototype branch built upon the proposed MPPR module, which explicitly models periodic structures from the frequency domain.

Specifically, MPPR extracts *per-variable periodic features* and injects them into the backbone feature stream to guide reconstruction toward regular periodic patterns. The prototype branch itself does not output anomaly scores. Instead, PCL and PAR act as *training-only* objectives (dashed links in Fig. 2) to shape periodic representations, while inference relies exclusively on reconstruction residuals from the backbone (solid links).

C. Time–Frequency Reconstruction Backbone

Given an input window $\mathbf{X}^{(i)}$, we first apply reversible instance normalization (RevIN) to alleviate non-stationarity:

$$\tilde{\mathbf{X}}^{(i)} = \mathrm{RevIN}\Big(\mathbf{X}^{(i)}\Big). \tag{1}$$

The normalized signal is then transformed into the frequency domain via FFT,

$$\mathbf{Z}^{(i)} = \mathcal{F}\Big(\tilde{\mathbf{X}}^{(i)}\Big), \tag{2}$$

and processed by a frequency-domain encoder that operates on patchified real and imaginary components to capture cross-variable dependencies:

$$\mathbf{F}_{\mathrm{bb}}^{(i)} = \mathrm{Enc}\Big(\mathrm{Patch}\Big(\Re(\mathbf{Z}^{(i)}), \Im(\mathbf{Z}^{(i)})\Big)\Big). \tag{3}$$

The encoded features are mapped back to complex coefficients and transformed to the time domain through inverse FFT and inverse RevIN,

$$\hat{\mathbf{X}}^{(i)} = \mathrm{RevIN}^{-1}\Big(\mathcal{F}^{-1}\Big(\hat{\mathbf{Z}}^{(i)}\Big)\Big). \tag{4}$$

During inference, anomaly scores are computed solely from the reconstruction residuals.

D. Multi-scale Periodic Prototype Representation

MPPR makes periodicity explicit by extracting per-variable periodic features anchored to a *shared* bank of learnable periodic prototypes (periodic templates). Given the normalized window $\tilde{\mathbf{X}}^{(i)}$, for each variable d we compute M time–frequency views via multi-scale STFT:

$$\mathbf{S}_d^{(i,m)} = \mathrm{STFT}_m\Big(\tilde{\mathbf{X}}_{:,d}^{(i)}\Big), \quad m = 1, \ldots, M. \tag{5}$$

At each scale, we identify dominant periodic components by mining the top-k non-DC frequency bins with the largest average magnitude:

$$\mathcal{K}_d^{(i,m)} = \mathrm{TopK}\Big(\mathrm{Avg}_t \Big|\mathbf{S}_d^{(i,m)}(t,\cdot)\Big|\Big), \tag{6}$$

and summarize the selected spectral components into a query embedding $\mathbf{q}_d^{(i,m)}$ through a lightweight projection head.

We then softly match $\mathbf{q}_d^{(i,m)}$ to a learnable prototype bank $\{\mathbf{p}_k^{(m)}\}_{k=1}^{K}$ and aggregate prototypes to obtain a prototype-aware representation:

$$\begin{aligned} \mathbf{h}_d^{(i,m)} &= \sum_{k=1}^{K} \alpha_{d,k}^{(i,m)}\, \mathbf{p}_k^{(m)}, \\ \alpha_{d,k}^{(i,m)} &= \mathrm{Softmax}_k\Big(\mathrm{sim}(\mathbf{q}_d^{(i,m)}, \mathbf{p}_k^{(m)})/\tau\Big). \end{aligned} \tag{7}$$

Finally, we fuse $\{\mathbf{h}_d^{(i,m)}\}_{m=1}^{M}$ to form the periodic feature $\mathbf{p}_d^{(i)}$ for variable d, and stack all variables as $\mathbf{P}^{(i)} = [\mathbf{p}_1^{(i)}, \ldots, \mathbf{p}_D^{(i)}]^\top$.

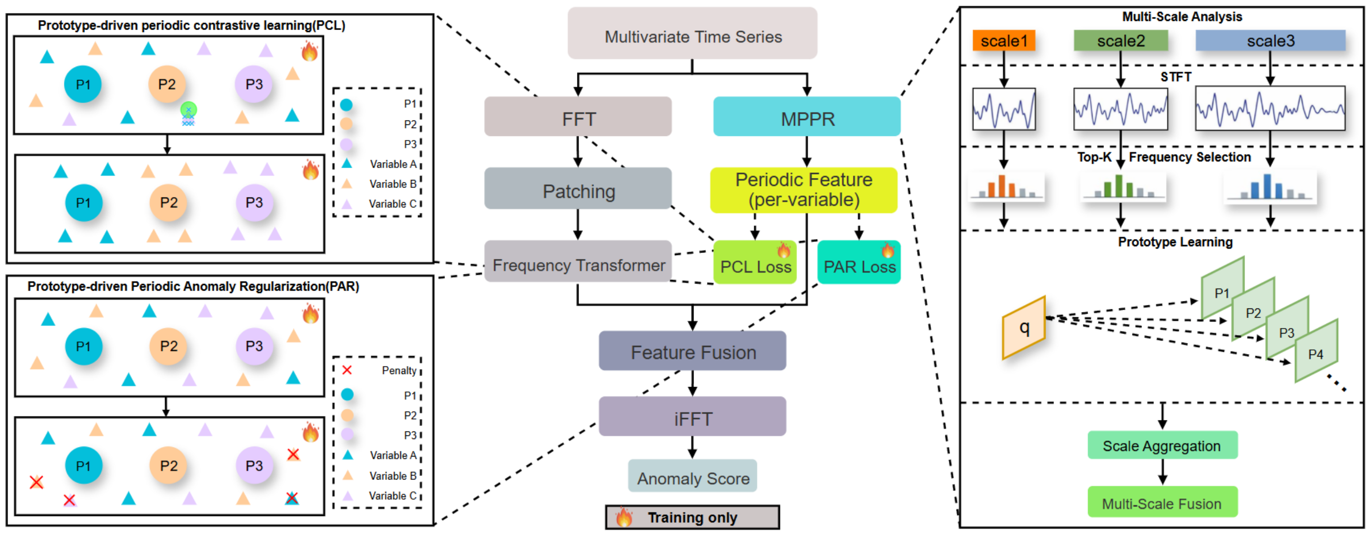

Fig. 2. Overall architecture of PPRL. The time–frequency reconstruction backbone reconstructs the input via FFT-based patching and frequency Transformer encoding, while MPPR extracts periodic features using multi-scale STFT, top-k peak mining, and prototype matching, and injects them into the backbone through lightweight fusion. During training, PCL aligns representations with consistent periodic patterns to shared prototypes, and PAR penalizes deviations from assigned prototypes to enlarge anomaly-aware margins; during inference, anomaly scoring relies solely on reconstruction residuals.

E. Periodic Contrastive Learning

To encourage consistent periodic representations, we introduce a PCL objective. Given the periodic representation $\mathbf{p}_d^{(i)}$ of variable d in window i produced by MPPR, we treat representations associated with the same periodic prototype as positives, and those associated with different prototypes or variables as negatives.

Specifically, for an anchor $\mathbf{p}_d^{(i)}$, we adopt an InfoNCE-style objective:

$$\mathcal{L}_{\text{pcl}} = -\log \frac{\sum_{j \in \mathcal{P}(i,d)} \exp\left(\text{sim}\left(\mathbf{p}_d^{(i)}, \mathbf{p}_d^{(j)}\right)/\tau\right)}{\sum_{k \in \mathcal{P}(i,d) \cup \mathcal{N}(i,d)} \exp\left(\text{sim}\left(\mathbf{p}_d^{(i)}, \mathbf{p}_d^{(k)}\right)/\tau\right)}, \quad (8)$$

where $\mathcal{P}(i,d)$ and $\mathcal{N}(i,d)$ denote positive and negative sets defined by prototype assignment, $\text{sim}(\cdot,\cdot)$ is cosine similarity, and τ is a temperature parameter.

This objective pulls representations with consistent periodic patterns toward shared prototypes, while pushing apart mismatched or cross-variable periodic representations.

F. Prototype-driven Periodic Anomaly Regularization

While PCL enforces consistency among normal periodic patterns, it does not explicitly regulate abnormal deviations. To this end, we introduce a PAR loss, which penalizes large deviations from the assigned periodic prototype.

For a periodic representation $\mathbf{p}_d^{(i)}$ and its matched prototype $\mathbf{c}_d$, we define:

$$\mathcal{L}_{\text{par}} = \max\left(0, \left\|\mathbf{p}_d^{(i)} - \mathbf{c}_d\right\|_2 - m\right), \quad (9)$$

where m is a margin hyperparameter that controls tolerance to normal periodic variation.

This regularization enlarges the representation margin between normal periodic patterns and anomalous deviations, encouraging abnormal periodic behaviors to remain unreconstructed rather than being absorbed into normal prototypes.

G. Prototype-driven Periodic Contrastive Learning (Training-only)

PCL is applied only during training to stabilize the geometry of periodic features. Let $\mathbf{P} \in \mathbb{R}^{B \times D \times d_p}$ be periodic features from a mini-batch and denote the flattened embeddings as $\{\mathbf{e}_{b,d}\}$ with variable index d.

IV. Experiments

A. Experimental Setup

We evaluate PPRL on four widely used multivariate time series anomaly detection benchmarks: PSM, MSL, SMAP, and SMD. All datasets are processed following standard protocols, where models are trained using normal data only and evaluated on the official test splits. Performance is assessed using two commonly adopted measures: Average A-R (AUC-ROC) and Aff-F (Affiliated-F1), which jointly evaluate detection accuracy and the temporal coherence of detected anomaly segments. All experiments are conducted on a single NVIDIA RTX 3090 GPU using PyTorch 2.1.2 with CUDA 11.8 and Python 3.8.20. For fair comparison, all methods adopt the same sliding-window strategy and training schedule. Hyperparameters related to periodic modeling are treated as tunable and fixed across datasets once selected.

B. Comparison with State-of-the-Art Methods

We compare PPRL with several representative reconstruction-based and frequency-aware methods, including CATCH, DC, DualTF, and ATrans. The quantitative results under AUC-ROC and Aff-F are summarized in **TABLE I**.

TABLE I
PERFORMANCE COMPARISON ON FOUR BENCHMARK DATASETS UNDER THE AVERAGE A-R (AUC-ROC) AND AFF-F (AFFILIATED-F1) ACCURACY MEASURES.

Dataset	Metric	PPRL	CATCH	DC	DualTF	ATrans
MSL	Aff-F	**0.757**	0.740	0.694	0.588	0.692
	A-R	**0.674**	0.664	0.507	0.576	0.508
PSM	Aff-F	**0.847**	0.829	0.682	0.725	0.710
	A-R	**0.653**	0.648	0.501	0.600	0.514
SMAP	Aff-F	**0.717**	0.699	0.701	0.674	0.703
	A-R	0.452	0.504	**0.522**	0.478	0.511
SMD	Aff-F	**0.856**	0.847	0.675	0.679	0.724
	A-R	0.802	**0.811**	0.502	0.631	0.508

Overall, PPRL delivers strong and stable performance across all benchmarks. Under the Aff-F metric, it outperforms all compared methods on PSM, MSL, SMAP, and SMD, which indicates that it is particularly effective at detecting temporally coherent anomaly segments rather than producing fragmented alarms. This advantage is especially clear on datasets with pronounced periodic structure, such as PSM and MSL, where prototype-driven periodic representation learning provides a more reliable notion of normal periodicity and makes periodic violations easier to surface.

On SMD, PPRL achieves the highest Aff-F score, suggesting improved robustness when the data exhibit complex, heterogeneous periodic patterns across variables and operating modes. On SMAP, while PPRL does not obtain the top A-R score, it still achieves the best Aff-F performance. This gap is informative: PPRL tends to emphasize precise and structurally consistent localization of anomalous segments, instead of aggressively expanding anomaly coverage to maximize recall. In our setting, this behavior is desirable because it is consistent with the core goal of mitigating over-reconstruction, where overly broad detections can be a side effect of unstable residual responses.

C. Qualitative Case Study (PSM)

To complement the quantitative results in Table I, we visualize the anomaly score response profiles on a representative subsequence of PSM. Specifically, we select a contiguous segment from the test stream with global indices `[40500, 43000)`, which is re-indexed from zero for clarity in Fig. 3. The normalized anomaly scores produced by PPRL and CATCH are plotted under the same sliding-window setting.

Since this dataset release does not provide official segment-level labels for visualization, this figure is not used to claim correctness but to examine score dynamics. As shown in Fig. 3, CATCH exhibits an isolated and extremely high spike, indicating that its response can be dominated by rare, window-specific error amplification. In contrast, PPRL avoids such isolated score explosions and produces more distributed and moderate responses across the segment, with salient regions emerging through sustained elevations rather than a single extreme peak. This behavior is consistent with the design of PPRL, where periodic prototypes regularize frequency-domain representations and reduce sensitivity to accidental amplification while remaining responsive to meaningful deviations from periodic structure.

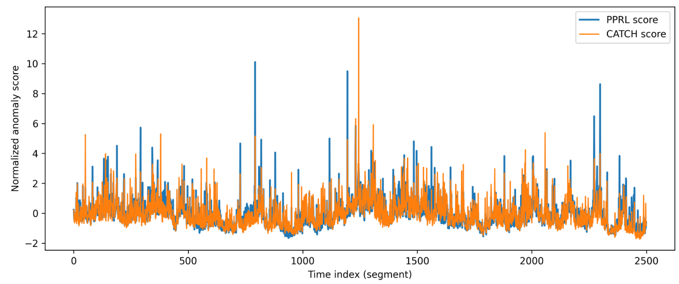

Fig. 3. PSM case study on a selected test subsequence with global indices `[40500, 43000)`. The x-axis re-indexes the selected segment from 0 for clarity. Since this dataset release does not provide official labels for visualization, the figure is used to compare score response profiles of PPRL and CATCH rather than claiming correctness.

D. Score Dynamics Analysis

To complement Fig. 3, we quantify score stability on the same PSM subsequence using z-scored anomaly scores. We report total variation (TV) to measure overall jitter, the maximum score to characterize extreme responses, and the Top1/Top5 ratio to assess isolated-spike dominance. As shown in Table II, PPRL yields lower TV (1670.53 vs. 1797.68), a lower maximum score (10.32 vs. 13.48), and a lower Top1/Top5 ratio (1.23 vs. 1.85) than CATCH, indicating more stable and less spike-dominated score dynamics. This behavior is consistent with Fig. 3, where CATCH exhibits an isolated extreme spike while PPRL produces a more regularized response profile.

TABLE II
SCORE DYNAMICS ANALYSIS ON A PSM CASE SEGMENT USING Z-SCORED ANOMALY SCORES. LOWER VALUES INDICATE MORE STABLE AND LESS SPIKE-DOMINATED RESPONSES.

Dataset	TV_P	TV_C	Max_P	Max_C	$\mathrm{Top1/Top5}_\mathrm{P}$	$\mathrm{Top1/Top5}_\mathrm{C}$
PSM	1670.53	1797.68	10.32	13.48	1.23	1.85

E. Ablation Study

We conduct an ablation study to analyze the contribution of different components within the PPRL framework. Starting from a minimal configuration that includes only the MPPR, we progressively enable the proposed training objectives. The results under Aff-F and Average A-R are reported in **TABLE III** and **TABLE IV**, respectively.

TABLE III
ABLATION STUDY OF PPRL UNDER THE AFF-F (AFFILIATED-F1) METRIC.

Method	**PSM**	**MSL**	**SMAP**	**SMD**
PPRL (MPPR only)	0.837	0.751	0.696	0.853
PPRL + PAR	0.840	0.755	0.701	0.853
PPRL + PCL	0.842	0.754	0.705	0.855
PPRL (Full)	**0.847**	**0.757**	**0.717**	**0.856**

TABLE IV
ABLATION STUDY OF PPRL UNDER THE AVERAGE A-R (AUC-ROC) METRIC.

Method	PSM	MSL	SMAP	SMD
PPRL (MPPR only)	0.650	0.667	0.449	0.801
PPRL + PAR	0.651	0.668	0.451	0.801
PPRL + PCL	0.652	0.667	0.448	0.801
PPRL (Full)	**0.653**	**0.673**	**0.452**	**0.802**

Starting from MPPR only, we already obtain strong and consistent performance across all datasets, indicating that anchoring recurring periodic patterns in a shared prototype space provides an effective reference for normal periodic structure and alleviates over-reconstruction. Adding PAR further improves results (especially Aff-F) by penalizing large deviations from the matched periodic prototypes during training, which helps preserve abnormal periodic behaviors instead of smoothing them into clean cycles. PCL brings complementary gains by regularizing the prototype space, promoting more consistent periodic representations while maintaining separability across different periodic regimes. With both objectives enabled, the full PPRL achieves the best or near-best performance, confirming that MPPR, PAR, and PCL address different aspects of periodic anomaly detection and work synergistically.

V. Conclusion

In this paper, we proposed PPRL, a self-supervised framework for multivariate time series anomaly detection that models periodic structure with multi-scale periodic prototypes. PPRL mitigates over-reconstruction through a learnable prototype space and prototype-driven objectives while retaining residual-based inference. Results on multiple benchmarks show consistent gains, particularly on affiliation-based metrics, without extra inference-time cost. Future work will study adaptive prototype management and extensions to other time series tasks.

ACKNOWLEDGMENT

This work was supported by the National Key Research and Development Program of China (2023YFB4004505) and the Key R&D Program of Shandong Province, China (2025CXGC020102, 2024CXGC010501).

References

[1] H. Ren, B. Xu, Y. Wang, C. Yi, C. Huang, X. Kou, T. Xing, M. Yang, J. Tong, and Q. Zhang, "Time-series anomaly detection service at microsoft," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)*, 2019, pp. 3009–3017.

[2] K. Hundman, V. Constantinou, C. Laporte, I. Colwell, and T. Söderström, "Detecting spacecraft anomalies using LSTMs and nonparametric dynamic thresholding," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)*, 2018, pp. 387–395.

[3] Z. Z. Darban, G. I. Webb, S. Pan, C. C. Aggarwal, and M. Salehi, "Deep learning for time series anomaly detection: A survey," *ACM Computing Surveys*, vol. 57, no. 1, 2024.

[4] S. Schmidl, P. Wenig, and T. Papenbrock, "Anomaly detection in time series: A comprehensive evaluation," *Proceedings of the VLDB Endowment*, vol. 15, no. 9, pp. 1779–1797, 2022.

[5] P. Malhotra, A. Ramakrishnan, G. Anand, L. Vig, P. Agarwal, and G. Shroff, "LSTM-based encoder-decoder for multi-sensor anomaly detection," *arXiv preprint arXiv:1607.00148*, 2016.

[6] C. Zhang, D. Song, Y. Chen, X. Feng, C. Lumezanu, W. Cheng, J. Ni, B. Zong, H. Chen, and N. V. Chawla, "A deep neural network for unsupervised anomaly detection and diagnosis in multivariate time series data," in *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2019, pp. 1409–1416.

[7] Y. Su, Y. Zhao, C. Niu, R. Liu, W. Sun, and D. Pei, "Robust anomaly detection for multivariate time series through stochastic recurrent neural network," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)*, 2019, pp. 2828–2837.

[8] J. Audibert, P. Michiardi, F. Guyard, S. Marti, and M. A. Zuluaga, "USAD: Unsupervised anomaly detection on multivariate time series," in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)*, 2020, pp. 3395–3404.

[9] S. Tuli, G. Casale, and N. R. Jennings, "TranAD: Deep transformer networks for anomaly detection in multivariate time series data," *Proceedings of the VLDB Endowment*, vol. 15, no. 6, pp. 1201–1214, 2022.

[10] J. Xu, H. Wu, J. Wang, and M. Long, "Anomaly transformer: Time series anomaly detection with association discrepancy," in *International Conference on Learning Representations (ICLR)*, 2022.

[11] Y. Nam, S. Yoon, Y. Shin, M. Bae, H. Song, J.-G. Lee, and B. S. Lee, "Breaking the time-frequency granularity discrepancy in time-series anomaly detection," in *Proceedings of The Web Conference (WWW)*, 2024, pp. 4204–4215.

[12] X. Wu, X. Qiu, Z. Li, Y. Wang, J. Hu, C. Guo, H. Xiong, and B. Yang, "CATCH: Channel-aware multivariate time series anomaly detection via frequency patching," in *International Conference on Learning Representations (ICLR)*, 2025.

[13] A. Huet, J. M. Navarro, and D. Rossi, "Local evaluation of time series anomaly detection algorithms," in *Proceedings of the 28th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)*, 2022, pp. 635–645.

[14] H. Xu, W. Chen, N. Zhao, Z. Li, J. Bu, Z. Li, Y. Liu, Y. Zhao, D. Pei, Y. Feng *et al.*, "Unsupervised anomaly detection via variational auto-encoder for seasonal KPIs in web applications," in *Proceedings of the 2018 World Wide Web Conference (WWW '18)*, 2018, pp. 187–196.

[15] M. Munir, S. A. Siddiqui, A. Dengel, and S. Ahmed, "DeepAnT: A deep learning approach for unsupervised anomaly detection in time series," *IEEE Access*, vol. 7, pp. 1991–2005, 2019.

[16] D. Li, D. Chen, B. Jin, L. Shi, J. Goh, and S. Ng, "MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks," in *Artificial Neural Networks and Machine Learning – ICANN 2019: Text and Time Series*, ser. Lecture Notes in Computer Science, vol. 11730. Springer, 2019, pp. 703–716.

[17] A. Geiger, D. Liu, S. Alnegheimish, A. Cuesta-Infante, and K. Veeramachaneni, "TadGAN: Time series anomaly detection using generative adversarial networks," in *2020 IEEE International Conference on Big Data (Big Data)*, 2020, pp. 33–43.

[18] H. Zhao, J. Wang, D. Yuan, Q. Hu, M. Huang, Y. Duan, Y. Liu, and Q. Liao, "Multivariate time-series anomaly detection via graph attention network," in *2020 IEEE International Conference on Data Mining (ICDM)*, 2020, pp. 841–850.

[19] A. Deng and B. Hooi, "Graph neural network-based anomaly detection in multivariate time series," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 5, pp. 4027–4035, 2021.

[20] Z. Li, Y. Zhao, J. Han, Y. Su, R. Jiao, X. Wen, and D. Pei, "Multivariate time series anomaly detection and interpretation using hierarchical inter-metric and temporal embedding," in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD '21)*, 2021, pp. 3220–3230.

[21] S. Tonekaboni, D. Eytan, and A. Goldenberg, "Unsupervised representation learning for time series with temporal neighborhood coding," in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2021.

[22] Z. Yue, Y. Wang, J. Duan, T. Yang, C. Huang, Y. Tong, and B. Xu, "TS2Vec: Towards universal representation of time series," in *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, vol. 36, no. 8, 2022, pp. 8980–8987.