Role Hypergraph Contrastive Learning for Multivariate Time-Series Analysis

Rundong Xue[1], Hao Hu[1], Zhitao Zeng[2], Xiangmin Han[2], Zhiqiang Tian[3], Shaoyi Du[1,*], and Yue Gao[2,*]

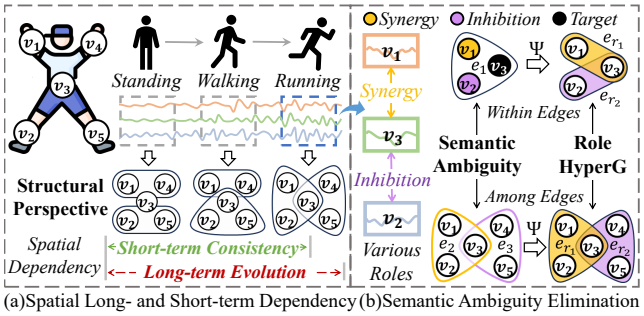
[1]State Key Laboratory of Human-Machine Hybrid Augmented Intelligence, National Engineering Research Center for Visual Information and Applications, and Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, China
[2]School of Software, Tsinghua University, China
[3]School of Software Engineering, Xi'an Jiaotong University, China
xuerundong@stu.xjtu.edu.cn, zengzhitao8@gmail.com, dushaoyi@xjtu.edu.cn, gaoyue@tsinghua.edu.cn

Abstract

Multivariate Time-Series (MTS) analysis is crucial across various domains. Considering the spatial and temporal consistency of MTS, existing methods leverage graph structures with temporal augmentation and contrastive learning to achieve robust learning of spatial dependencies and temporal patterns. Given the inherent high-order correlations in MTS, hypergraphs present a promising approach. However, two key challenges limit their further development: 1) Feature-based perspectives capture limited spatial information, while structural perspectives encode richer spatial consistency and evolution dependency; 2) Various semantic patterns (e.g., synergy, inhibition) entangle in sensor correlations, leading to semantic ambiguity. The underlying reason is that conventional hypergraph structures cannot distinguish specific semantic roles within or across hyperedges. Thus, we propose Role Hypergraph Contrastive Learning for MTS analysis. Specifically, we introduce the concept of role to generalize hypergraphs to Role Hypergraphs, enabling precise modeling of sensor correlations by assigning each vertex-hyperedge pair with a semantic role. Building on this structure, we design a role hypergraph contrastive learning paradigm to comprehensively capture the spatial and temporal dependencies: From a structural perspective, role hypergraph structural contrasting captures spatial short-term consistency and long-term evolution; from a feature perspective, alignment of complementary role information ensures sensor-level temporal consistency. Experiments on classification and forecasting tasks demonstrate the effectiveness and interpretability of our method.

Introduction

Multivariate Time-Series (MTS) data play an important role in diverse domains, including clinical diagnosis, transportation analysis, climate science, and energy management (Wang et al. 2023; Cirstea et al. 2022; Xue et al. 2025b). However, the inherent complexity and high dimensionality of MTS make label acquisition more challenging compared to other data, limiting its applicability in the real world (Pöppelbaum, Chadha, and Schwung 2022). Self-Supervised Learning (SSL) (Zhang et al. 2024) presents a promising solution by pre-training models to learn generalizable representations from unlabeled data, followed by fine-tuning on downstream tasks. Contrastive learning (Khosla et al. 2020), as a self-supervised paradigm, leverages inherent consistency in data by contrasting augmented views, thereby improving encoder robustness against perturbations.



(a)Spatial Long- and Short-term Dependency (b)Semantic Ambiguity Elimination

Figure 1: (a) Spatial short-term consistency and long-term evolution under a structural perspective. (b) Semantic ambiguity arising from the various entangled role patterns inherent in sensor correlations is eliminated via role hypergraphs.

Several advances in contrastive learning for time series focus primarily on temporal consistency (Luo et al. 2023; Liu and Chen 2024). These methods generate positive and negative pairs via temporal augmentations (e.g., warping, cropping) (Othman et al. 2020) to learn representations robust to temporal perturbations using contrastive learning. These approaches largely neglect spatial correlations among sensors within MTS (Wang et al. 2024d). Recent graph methods have incorporated spatial information into graph contrastive frameworks by modeling sensor correlations as graphs from a feature perspective (Wang et al. 2024c). However, relying solely on feature-level representations to capture the nuanced spatial structural variations in MTS evolution may be inadequate (Ye et al. 2022; Xue et al. 2025a). As illustrated in Fig. 1 (a), during activity transitions (e.g., from standing to walking), new edges may emerge between sensors (torso-foot connections), reflecting structural changes not evident at the feature level. Given that the spatial structure of MTS is stable in short time but changeable over long periods (Bao et al. 2024), necessitating approaches that characterize spatial dependency **from a structural perspective** to measure **short-term consistency** and **long-term evolution**.

Graph-based methods effectively model pairwise sensor

*Corresponding Authors

relationships (Chen and Eldardiry 2024) but fail to capture higher-order correlations inherent in MTS. Hypergraphs, as an extension of graphs, address this limitation by enabling hyperedges to connect multiple vertices, thus capturing beyond pairwise correlations (Feng et al. 2019). However, conventional hypergraphs suffer from **semantic ambiguity**: they cannot differentiate distinct semantic patterns among hyperedges within a hypergraph or vertices within a hyperedge (shown in Fig. 1 (b)). Specifically, existing methods for modeling sensor correlations into hypergraphs, e.g., kNN, sparse representation (SR), rely solely on the absolute correlation magnitude to construct hyperedges, ignoring the critical distinction between positive (synergy) and negative (inhibition) relationships. For instance, in SR, the resulting sparse solution includes both positive and negative coefficients, yet traditional methods group all non-zero coefficients indiscriminately into a single hyperedge (Chen et al. 2023a). Consequently, semantically contradictory vertices (synergy versus inhibition) may coexist within a single hyperedge, leading to semantic ambiguity. Even if some methods attempt to differentiate the semantic roles, conventional hypergraph convolution (HGNN) operations treat all hyperedges uniformly regardless of their roles (Ji et al. 2025). In summary, conventional hypergraph methods struggle to represent the diverse semantic relationships among sensors, limiting their capacity to model complex sensor interactions.

To address the above limitations, we propose Role Hypergraph Contrastive Learning (RHCL) for MTS analysis. Specifically, we introduce a *role* concept to generalize hypergraph, forming a novel **Role Hypergraph** structure. The core idea is to assign a specific semantic role (e.g., synergy or inhibition) to each vertex-hyperedge pair, enabling precise modeling of sensor interactions. Based on this structure, we design a Role HGNN to facilitate information propagation among sensors within and across roles. With updated sensor role features, we propose role hypergraph contrastive learning to achieve robust learning of spatial and temporal dependencies with the role view: **Role hypergraph structural contrasting** from a structural perspective to measure the short-term consistency and long-term evolution of the spatial role hypergraph; Considering sensor temporal evolution patterns in MTS, we **align the complementary role information** of each sensor across time windows from a feature perspective, ensuring sensor-level temporal consistency. The main contributions are summarized as follows:

1) We propose a generalized hypergraph structure that adapts and transfers the concept of semantic differentiation, enabling precise modeling of various roles by assigning each vertex-hyperedge pair with a semantic role.
2) To achieve robust learning of spatial and temporal dependence, we propose RHCL: From a structural perspective, we adopt role hypergraph structural contrasting to capture spatial short-term consistency and long-term evolution. From a feature perspective, alignment of complementary role information ensures temporal consistency.
3) Experimental results demonstrate the superiority of our methods. Moreover, RHCL offers enhanced interpretability through its ability to capture spatial dependencies and temporal consistency under the role perspective.

Related Work

Contrastive Learning for MTS Contrastive learning, as a self-supervised paradigm, aims to learn robust representations by minimizing the distance between positive pairs while maximizing it for negative pairs (Khosla et al. 2020; Chen et al. 2020). This approach offers a distinct alternative to pretext-task-dependent methods for robust learning from unlabeled data. Recent work has adapted contrastive frameworks to MTS, primarily focusing on achieving temporal consistency (Zhang et al. 2022; Hao et al. 2023; Yue et al. 2022). These methods create view pairs by temporal augmentations (e.g., warping, flipping, and cropping) to perturb time series, and then employ contrastive learning to learn robust representations under temporal perturbations.

While the above works advance temporal consistency through contrastive learning, they overlook spatial consistency in MTS data. Some studies have incorporated spatial correlations into contrastive approaches. For instance, TS-GAC (Wang et al. 2024d) proposes a graph contrast learning method that aligns spatial features derived from strong and weak augmentation views. In summary, existing methods primarily achieve spatial consistency by minimizing spatial feature distance from the perspective of features. However, relying solely on feature-level consistency to represent spatial consistency may be insufficient. A promising direction is to characterize spatial information from a structural perspective (e.g., graph structure) to measure graph structural consistency. Building on this, considering the nature of MTS data (the evolution of spatial structures), incorporating short-term spatial consistency in this long-term evolutionary process offers potential for achieving robust learning on MTS.

Graph and Hypergraph for MTS Graphs are increasingly popular in time series analysis (Chen et al. 2023b; Wang et al. 2024b), representing variables as vertices and correlations as edges. While effective for pairwise relationships, this graph-based approach faces limitations in representing higher-order, more complex correlations inherent in MTS. Hypergraphs (Feng et al. 2019), an extension of graphs, address this by allowing hyperedges to connect multiple vertices, thereby capturing high-order correlations beyond pairwise interactions. This capability has made HGNN a versatile tool adopted in diverse research fields (Han et al. 2024; Feng et al. 2024). However, hypergraph-based methods remain nascent in time series analysis (Shang et al. 2024; Han et al. 2025). A critical limitation of conventional hypergraphs is their inability to represent differences in semantic roles. In a hypergraph, there are varying semantic roles among hyperedges within a hypergraph and vertices within a hyperedge (e.g., cooperative or inhibitory patterns in sensor correlations). However, these differences in roles cannot be adequately captured by standard hypergraphs.

Methodology

Problem Formulation Given the input MTS samples $\mathcal{X} = \{\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n\}$, where each $\boldsymbol{X}_i \in \mathbb{R}^{N \times T}$ is collected from N sensors over T timestamps. The objective

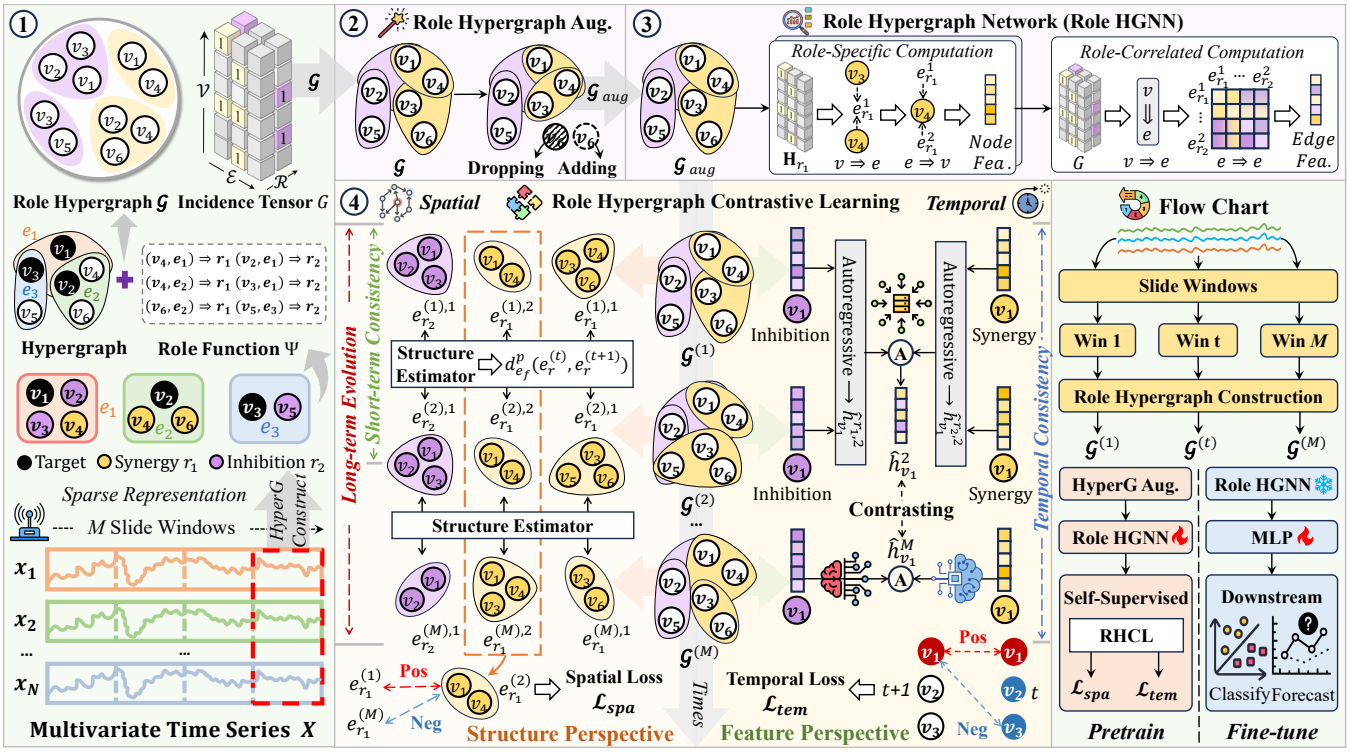


Figure 2: Overview of RHCL. The central section is Role Hypergraph Contrastive Learning, integrating role hypergraph structural contrasting (left) and role complementary information alignment (right). Flowchart is depicted in the lower right corner.

is to develop a hypergraph contrastive learning scheme that learns general representations for $\mathcal{X}$, enabling the training of an encoder in an unsupervised manner during the pre-training phase. During the fine-tuning phase, the pre-trained encoder parameters remain frozen to preserve learned representations, while an MLP is employed for downstream tasks.

Hypergraph Hypergraphs model high-order correlations through hyperedges. A hypergraph $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ consists of a vertex set $\mathcal{V}$ and a hyperedge set $\mathcal{E}$. The relationships between $\mathcal{V}$ and $\mathcal{E}$ can be represented by the incidence matrix $\mathbf{H}$, where $\mathbf{H}(v, e) = 1$ if $v \in e$, and 0 otherwise.

Due to the dynamic properties of MTS, the correlation patterns within MTS evolve over time. To capture this temporal evolution, we segment the time series $\boldsymbol{X}$ with length L into $M = T/L$ windows, i.e., $\boldsymbol{X} = \{\mathbf{X}^{(t)}\}_{t=1}^{M}$. The hypergraph is formalized as a sequence $\{\mathcal{H}^{(t)} = (\mathcal{V}, \mathcal{E}^{(t)})\}_{t=1}^{M}$.

Overview

Fig. 2 presents an overview of RHCL. Specifically, due to the local dynamic nature of MTS, each sample is segmented into multiple windows, within which a hypergraph is constructed. Next, we introduce the function Ψ that assigns a role r to each (v, e), forming a role hypergraph $\mathcal{G}$ with its incidence tensor G. Following augmentation and Role HGNN encoding, role hypergraph contrastive learning is employed to explore spatial and temporal dependencies. Structurally, it captures spatial short-term consistency and long-term evolution by measuring role hypergraph structural variations across adjacent and distant windows. Feature-wise, it summarizes and aligns complementary role information across windows to preserve sensor-level temporal consistency.

Role Hypergraph Representation

Here, we introduce a role hypergraph representation method for MTS to fine-grained semantics within high-order relationships that account for the various role correlations (e.g., synergy, inhibition) that exist among sensors.

Hypergraph Construction for MTS For MTS, the N sensors are defined as the vertex set $\mathcal{V} = \{v_1, \ldots, v_N\}$, with hyperedges encoding high-order correlations. Given the hypergraph sequence $\{\mathcal{H}^{(1)}, \cdots, \mathcal{H}^{(t)}, \cdots, \mathcal{H}^{(M)}\}$, the t-th hypergraph is constructed by sparse representation.

Specifically, $\mathbf{X}^{(t)} = [x_1, \ldots, x_i, \ldots, x_N]^\top$, each sensor time series x_i is regarded as a sparse linear combination of other N-1 sensor time series (Chen et al. 2015):

$$x_i = A_i \boldsymbol{\alpha}_i + \boldsymbol{\tau}_i, \quad i = 1, 2, \ldots, N, \tag{1}$$

where $A_i = [x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_N]^\top$ contains all sensor time series except the i-th one. $\boldsymbol{\alpha}_i$ quantifies interaction weights, and $\boldsymbol{\tau}_i$ is a noise term. The resulting sparse coefficients $\boldsymbol{\alpha}_i$ can be approximately recovered by solving a standard l_1-norm regularized optimization problem:

$$\min_{\boldsymbol{\alpha}_i} \| x_i - A_i \boldsymbol{\alpha}_i \|_2 + \lambda \| \boldsymbol{\alpha}_i \|_1, \tag{2}$$

where λ is a regularization parameter controlling the sparsity of the solution. The resulting sparse coefficients $\boldsymbol{\alpha}_i$ induce hyperedge construction through non-zero coefficients.

Role Hypergraph Representation We observe that $\boldsymbol{\alpha}_i$, which encodes the high-order correlations among sensors within a hyperedge, contains both positive and negative values. This indicates distinct, potentially conflicting semantic correlation patterns among sensors (e.g., synergy *vs.* inhibition) within the same hyperedge. Such semantic ambiguity reflects diverse semantic relationships among vertices within a hyperedge or hyperedges within a hypergraph, which conventional GNN/HGNNs fail to differentiate. Thus, we introduce the notion of $role$ to generalize hypergraphs.

Definition 1 *The role hypergraph is formally defined as $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{R}, \Psi)$, where $\mathcal{V}$ and $\mathcal{E}$ denote the sets of vertex and hyperedge, and $\mathcal{R} = \{r_1, \ldots, r_R\}$ represents a set of semantic roles (e.g., $\{$synergy, inhibition$\}$). The key innovation of this structure is the role assignment function $\Psi : \mathcal{V} \times \mathcal{E} \to \mathcal{R}$, which assigns a specific semantic role $\Psi(e, v) \in \mathcal{R}$ to each pair of vertex and hyperedge (e, v). This approach allows for a more detailed modeling of roles, rather than applying a fixed pattern to an entire hyperedge.*

To represent this structure computationally, we define a role incidence tensor $G \in \mathbb{R}^{|\mathcal{R}|\times|\mathcal{V}|\times|\mathcal{E}|}$, where:

$$G(r, v, e) = \begin{cases} 1 & \text{if } v \in e \text{ and } \Psi(e, v) = r, \\ 0 & \text{otherwise}. \end{cases} \tag{3}$$

In this paper, we focus exclusively on the two roles of MTS: synergy and inhibition. Following Definition 1, we derive the role hypergraph $\mathcal{G}^{(t)}$ with $\boldsymbol{G}^{(t)}$ for each window.

Role Hypergraph Augmentation Conventional hypergraph augmentation typically randomly replaces or drops vertices within hyperedges. However, in the case of role hypergraphs $\mathcal{G}$, such indiscriminate operations can disrupt the semantic roles assigned to vertex-hyperedge pairs, distorting sensor correlations expressed through synergy and inhibition roles. Thus, leveraging the role-specific insights, we introduce a two-stage role hypergraph augmentation strategy that considers sensor specificity and sensor correlations.

Vertex Dropping Strategy: In the first stage, to learn robust sensor specificity while maintaining MTS semantic consistency, we employ a vertex-dropping strategy to randomly drop vertices from each role hyperedge $e_r \in \mathcal{E}_r$ (defined in Definition 2) with a probability p:

$$\mathcal{E}'_r = \{e_r - \mathcal{D}(e_r; p) \mid e_r \in \mathcal{E}_r\}, \tag{4}$$

where $\mathcal{D}(e_r; p)$ is the dropped subset of vertices from e_r.

Vertex Adding Strategy: In the second stage, to address potential hyperedge redundancy caused by vertex dropping, we introduce a constrained vertex-adding strategy. Unlike naive vertex adding, to avoid semantic conflicts, a candidate vertex v is added to e_r only if it is compatible with the hyperedge's existing role r (i.e., $\Psi(e_r, v) = r$).

Definition 2 ***Projection from Role Hypergraph to Role-Specific Hypergraph****: Given a role hypergraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{R}, \Psi)$ as defined in Definition 1, we can project it onto conventional hypergraphs specific to each semantic role $r \in \mathcal{R}$. For a given hyperedge $e \in \mathcal{E}$ and specific role $r \in \mathcal{R}$, a role-specific hyperedge e_r is defined as:*

$$e_r = \{v \in \mathcal{V} \mid \Psi(e, v) = r\}. \tag{5}$$

This represents the vertices in e assigned to role r. For role r, the set of role hyperedges can be defined as:

$$\mathcal{E}_r = \{e_r \mid e \in \mathcal{E}\}. \tag{6}$$

Consequently, the $\mathcal{H}_r = (\mathcal{V}, \mathcal{E}_r)$ constitutes a role r-specific conventional hypergraph, where each role hyperedge is the projection of the original structure under role r.

Role Hypergraph Network With the augmented role hypergraphs, we redefine the message passing mechanism on the role hypergraph $\mathcal{G}$ to update sensor correlations.

Role-Specific Hypergraph Computation enables message propagation within specific roles defined by the incidence tensor G. Following Definition 2, the incidence tensor G can be mapped to the sequence of hypergraph incidence matrix $\{\mathbf{H}_r\}_{r=1}^{R}$, where $\mathbf{H}_r(v, e) = G(r, v, e)$. For each role r, we perform hypergraph convolution using its specific incidence structure with vertex degree matrix $\mathbf{D}_{v,r}$ and hyperedge degree matrix $\mathbf{D}_{e,r}$ at layer l:

$$\mathbf{X}_r^{(l+1)} = \sigma(\mathbf{D}_{v,r}^{-1/2}\mathbf{H}_r\mathbf{D}_{e,r}^{-1}\mathbf{H}_r^\top\mathbf{D}_{v,r}^{-1/2}\mathbf{X}_r^{(l)}\Theta^{(l)}), \tag{7}$$

Role-Correlated Hypergraph Computation: Since sensor correlations often span multiple semantic roles, we introduce the role hyperedge latent representations h_{e_r} to manifest the characteristics of hyperedges assigned with roles. To model these role-correlated dependencies and facilitate information propagation across various roles, we adopt a role-aware hyperedge attention mechanism that operates on h_{e_r} to learn high-order sensor correlations across roles:

$$\|_{r=1}^{|\mathcal{R}|} h_{e_r}^{(l+1)} = \text{softmax}(\tfrac{QK^\top}{\sqrt{D_K}})(\|_{r=1}^{|\mathcal{R}|} \mathbf{D}_{e,r}^{-1}\mathbf{H}_r^\top\mathbf{X}_r^{(l+1)}\Theta^{(l)}) \tag{8}$$

Role Hypergraph Contrastive Learning

The role hypergraph, with its semantic partitioning of sensor correlations, offers a potential advantage for modeling spatial dependency and temporal consistency: sensors that share the semantic role across time windows often exhibit more similar interaction patterns. Motivated by this, we propose role hypergraph contrastive learning, a framework that learns generalizable representations by leveraging the properties of role hypergraphs. It employs role hypergraph structural contrasting to capture short-term spatial consistency and long-term structural evolution, while ensuring temporal consistency via role complementary information alignment.

Role Hypergraph Structural Contrasting Existing studies for spatial consistency often rely on feature-level contrastive learning with weak and strong augmented views, which fail to comprehensively characterize spatial dependencies (Wang et al. 2024d,a). Meanwhile, these vertical contrast mechanisms overlook the horizontal structural evolution over time. Given that spatial structure is stable in short time but changeable over long periods for MTS (Bao et al. 2024), we propose a role hypergraph structural contrasting module that captures the spatial short-term consistency and long-term evolution from a structural perspective.

Role Hypergraph Structure Estimator: Due to the complexity of the role hypergraph structure, we cannot directly measure the structural variations between role hypergraph

pairs. However, we note that role hypergraph structure evolution manifests as changes to specific role hyperedges. Intuitively, we can assess structural evolution by measuring the distances between corresponding role hyperedge pairs across time windows, ensuring validity while restricting the distance computation to the relevant role hyperedges, thereby significantly reducing computational complexity.

Inspired by (Chowdhury and Mémoli 2019; Chowdhury et al. 2024), we employ optimal transport (OT) to quantify the structural variations between corresponding role hyperedges. To generalize the hypergraph structure estimator to role hypergraphs, we propose the following approach based on Definition 1 and Definition 2: Given the $\mathcal{G}^{(t)}$ and $\mathcal{G}^{(t+1)}$ in two windows, project them onto role-specific hypergraphs $\{(\mathcal{H}_r^{(t)}, \mathcal{H}_r^{(t+1)})\}_{r\in\mathcal{R}}$. For each role $r \in \mathcal{R}$, the $\mathcal{H}_r^{(t)}$ and $\mathcal{H}_r^{(t+1)}$ are transformed into corresponding measure hypergraphs $H_r^{(t)} = (\mathcal{V}, \mu_r^{(t)}, \mathcal{E}_r^{(t)}, \eta_r^{(t)}, \kappa_r^{(t)})$ and $H_r^{(t+1)} = (\mathcal{V}, \mu_r^{(t+1)}, \mathcal{E}_r^{(t+1)}, \eta_r^{(t+1)}, \kappa_r^{(t+1)})$ via OT, where μ and η denote Borel probability measures on $\mathcal{V}$ and $\mathcal{E}$. $\kappa : \mathcal{V} \times \mathcal{E} \to \mathbb{R}$ is a measurable, bounded hypergraph function. The structural variations within the corresponding role hyperedges pairs $(e_r^{(t)}, e_r^{(t+1)})$ can be formulated as:

$$d_r^p(e_r^{(t)}, e_r^{(t+1)}) = \inf_{\pi\in\Pi(\mu_r^{(t)}, \mu_r^{(t+1)})} \Big(\int_{\mathcal{V}\times\mathcal{V}} |\kappa(v, e_r^{(t)}) - \kappa'(v, e_r^{(t+1)})|^p \pi(dv \times dv)\Big)^{1/p}, \tag{9}$$

where π are couplings $(\mu_r^{(t)}, \mu_r^{(t+1)})$. Following this, the structural variations for role-r hypergraph pairs is defined as $d_r^p(\mathcal{H}_r^{(t)}, \mathcal{H}_r^{(t+1)})$. Finally, the variations between role hypergraph pairs $(\mathcal{G}^{(t)}, \mathcal{G}^{(t+1)})$ via aggregation is defined as:

$$d^p(\mathcal{G}^{(t)}, \mathcal{G}^{(t+1)}) = \sum_{r\in\mathcal{R}} d_r^p(\mathcal{H}_r^{(t)}, \mathcal{H}_r^{(t+1)}). \tag{10}$$

Spatial Structure Contrastive Learning: Built on the variations under a structural perspective, we employ a role hypergraph structural contrasting scheme to capture spatial short-term consistency and long-term evolution. Specifically, we quantify the distance between role hypergraph pairs as the variation of structural information with parameter λ:

$$\mathcal{I}_r^{(t,t+1)} = e^{-\lambda \cdot d^p(\mathcal{H}_r^{(t)}, \mathcal{H}_r^{(t+1)})}. \tag{11}$$

Within $\mathcal{E}_r^{(t)}$, each pair of corresponding role hyperedges in adjacent windows $(e_r^{(t)}, e_r^{(t+1)})$ forms a positive sample pair. This reflects the short-term spatial consistency of the sensor correlation patterns for role r. Conversely, for a given $e_r^{(t)}$, its corresponding role hyperedges from subsequent windows $\hat{\mathcal{E}}_r^{(t)} = \{e_r^{(t+k)} | k > 2\}$ are treated as negative samples. This strategy enables the model to track the long-term evolution of the spatial structure. The role hypergraph spatial contrastive loss for role hyperedge $e_r^{(t)}$ is formulated as:

$$l_{e_r}^{(t)} = -\log \frac{\exp(h_{e_r}^{(t)} \cdot h_{e_r}^{(t+1)}/\tau)(-\mathcal{I}_r^{(t,t+1)})}{\sum_{\hat{e}_r^{(t+k)}\in\hat{\mathcal{E}}_r^{(t)}} \exp(h_{e_r}^{(t)} \cdot h_{\hat{e}_r}^{(t+k)}/\tau)(-\mathcal{I}_r^{(t,t+k)})}, \tag{12}$$

where $h_{e_r}^{(t)}$ and $h_{e_r}^{(t+1)}$ are the embeddings of hyperedges extracted from the encoder (Role HGNN), and τ denotes the temperature parameter. The spatial contrastive loss $\mathcal{L}(\mathcal{G}^{(t)})$ for the t-th window role hypergraph $\mathcal{G}^{(t)}$ can be defined as:

$$\mathcal{L}(\mathcal{G}^{(t)}) = \sum_{r\in\mathcal{R}} \mathcal{L}(\mathcal{H}_r^{(t)}) = \sum_{r\in\mathcal{R}} \Big(\frac{1}{|\mathcal{E}_r^{(t)}|} \sum_{e_r\in\mathcal{E}_r^{(t)}} l_{e_r}^{(t)}\Big). \tag{13}$$

The overall spatial loss $\mathcal{L}_{spa}$ can be formulated as:

$$\mathcal{L}_{spa} = \sum_{t=1}^{M-1} \mathcal{L}(\mathcal{G}^{(t)}). \tag{14}$$

Role Complementary Information Alignment Role hypergraph structural contrasting captures comprehensive spatial dependency, while sensors also exhibit temporal consistency across multiple windows. Within each window, the role hypergraph encodes semantic information through complementary roles for each sensor. It implies that the complementary role information of each sensor shows temporal dependence across windows. Based on this, we introduce a complementary information alignment of the role hypergraph to achieve sensor-level temporal consistency.

Specifically, for each sensor v_i, we employ a transformer to summarize the role-specific embeddings $h_{v_i,r}^{(t)}$ (obtained by Role HGNN) in past d windows following (Eldele et al. 2021): $h_{v_i}^{r,d} = \mathcal{F}(h_{v_i,r}^{(1)}, \cdots, h_{v_i,r}^{(d)})$. Then, we aggregate its role complementary embeddings $h_{v_i}^{r,d}$ across all roles $r \in \mathcal{R}$ to form a comprehensive sensor representation: $\hat{h}_{v_i}^d = \sum_{r\in\mathcal{R}} h_{v_i}^{r,d}$. To achieve temporal consistency, we select its adjacent window s of d to form a positive pair $(\hat{h}_{v_i}^d, \hat{h}_{v_i}^s)$. The comprehensive representations within the same sensor should exhibit higher similarity compared to other sensors $\hat{h}_{v_j}^d$. Formally, we obtain the temporal consistency loss:

$$\mathcal{L}_{tem} = -\frac{1}{N} \sum_{i}^{N} \log \frac{\exp(f_{sim}(\hat{h}_{v_i}^d, \hat{h}_{v_i}^s)/\tau)}{\sum_{v_j\in\hat{\mathcal{V}}} \exp(f_{sim}(\hat{h}_{v_i}^d, \hat{h}_{v_j}^d)/\tau)} \tag{15}$$

where $f_{sim}(\cdot)$ is used to measure the cosine similarity of vectors, and $\hat{\mathcal{V}}$ is the set of sensors excluding the sensor i.

Experiments

Experimental Setting

Datasets The proposed RHCL is evaluated on two downstream tasks: Classification and Long-term Forecasting, including *eight MTS classification datasets*: Human Activity Recognition (HAR) dataset (Anguita et al. 2012), the ISRUC dataset (Khalighi et al. 2016), and six datasets from the UEA archive: Articulary Word Recognition (AWR), Finger Movements (FM), Face Detection (FD), Insect Wingbeat (IW), Spoken Arabic Digits Equivalent (SAD), and Self Regulation SCP1 (SRSCP1) and *four long-term MTS forecasting datasets*: ETT (Zhou et al. 2021), Traffic (Lai et al. 2018), Weather (Zeng et al. 2023), and Electricity (Wu et al. 2020).

Baselines We compare our method with several methods for Classification and Forecasting tasks: ***Classification***: SVP-T (Zuo et al. 2023), TodyNet (Liu et al. 2024), TS-GAC (Wang et al. 2024d), ShapeFormer (Le et al. 2024),

Datasets	Metrics	SVP-T *AAAI'23*	TodyNet *Info.'24*	TS-GAC *AAAI'24*	ShapeFormer *KDD'24*	ST *NN'25*	MPTSNet *AAAI'25*	SAGoG *TKDE'25*	**RHCL (Our)**
HAR	ACC(%)	$87.74_{\pm 0.84}$	$91.75_{\pm 0.43}$	$92.07_{\pm 0.12}$	$91.57_{\pm 0.35}$	$\underline{92.94}_{\pm 0.24}$	$92.45_{\pm 0.27}$	$90.42_{\pm 0.42}$	$\mathbf{95.14}_{\pm \mathbf{0.09}}$
	MF1(%)	$86.82_{\pm 0.88}$	$91.10_{\pm 0.50}$	$92.01_{\pm 0.15}$	$90.83_{\pm 0.35}$	$\underline{92.47}_{\pm 0.28}$	$91.97_{\pm 0.29}$	$90.02_{\pm 0.47}$	$\mathbf{94.85}_{\pm \mathbf{0.12}}$
ISRUC	ACC(%)	$78.45_{\pm 1.24}$	$76.37_{\pm 2.14}$	$\underline{84.43}_{\pm 0.89}$	$79.45_{\pm 0.82}$	$83.27_{\pm 0.57}$	$84.02_{\pm 0.24}$	$81.78_{\pm 1.04}$	$\mathbf{85.96}_{\pm \mathbf{0.15}}$
	MF1(%)	$77.83_{\pm 1.67}$	$75.86_{\pm 2.35}$	$\underline{83.88}_{\pm 0.96}$	$78.98_{\pm 0.93}$	$82.76_{\pm 0.61}$	$83.34_{\pm 0.30}$	$81.29_{\pm 1.22}$	$\mathbf{85.16}_{\pm \mathbf{0.19}}$
AWR	ACC(%)	$92.13_{\pm 2.66}$	$96.11_{\pm 1.04}$	$97.52_{\pm 1.02}$	$\underline{98.44}_{\pm 0.54}$	$97.80_{\pm 0.23}$	$97.73_{\pm 0.82}$	$94.09_{\pm 3.26}$	$\mathbf{98.85}_{\pm \mathbf{0.11}}$
	MF1(%)	$92.00_{\pm 3.74}$	$95.97_{\pm 2.18}$	$97.68_{\pm 1.45}$	$\underline{98.44}_{\pm 0.54}$	$97.63_{\pm 0.45}$	$97.46_{\pm 1.14}$	$94.01_{\pm 3.81}$	$\mathbf{98.78}_{\pm \mathbf{0.13}}$
FM	ACC(%)	$57.31_{\pm 3.45}$	$58.79_{\pm 2.64}$	$54.46_{\pm 4.75}$	$59.82_{\pm 3.54}$	$62.38_{\pm 1.87}$	$\underline{63.11}_{\pm 3.23}$	$61.59_{\pm 2.41}$	$\mathbf{66.10}_{\pm \mathbf{2.21}}$
	MF1(%)	$52.76_{\pm 5.18}$	$55.75_{\pm 3.13}$	$48.72_{\pm 5.26}$	$\underline{57.58}_{\pm 4.39}$	$54.94_{\pm 2.58}$	$57.43_{\pm 4.72}$	$53.64_{\pm 2.87}$	$\mathbf{61.98}_{\pm \mathbf{3.45}}$
FD	ACC(%)	$56.71_{\pm 2.64}$	$\underline{66.53}_{\pm 2.74}$	$60.20_{\pm 1.39}$	$65.72_{\pm 0.87}$	$59.76_{\pm 1.41}$	$61.82_{\pm 1.40}$	$62.47_{\pm 0.63}$	$\mathbf{67.63}_{\pm \mathbf{0.42}}$
	MF1(%)	$56.47_{\pm 3.17}$	$\underline{66.27}_{\pm 4.34}$	$60.05_{\pm 2.01}$	$65.46_{\pm 1.14}$	$59.51_{\pm 2.28}$	$61.60_{\pm 2.51}$	$62.24_{\pm 0.98}$	$\mathbf{67.17}_{\pm \mathbf{0.37}}$
IW	ACC(%)	$29.63_{\pm 5.51}$	$30.65_{\pm 4.81}$	$\underline{60.41}_{\pm 2.27}$	$32.64_{\pm 2.45}$	$50.15_{\pm 3.37}$	$51.05_{\pm 1.07}$	$59.59_{\pm 1.79}$	$\mathbf{64.71}_{\pm \mathbf{1.51}}$
	MF1(%)	$29.36_{\pm 6.87}$	$30.24_{\pm 5.14}$	$\underline{60.06}_{\pm 3.02}$	$32.22_{\pm 3.39}$	$50.04_{\pm 4.64}$	$54.52_{\pm 1.93}$	$59.08_{\pm 2.56}$	$\mathbf{64.44}_{\pm \mathbf{1.28}}$
SAD	ACC(%)	$95.66_{\pm 1.12}$	$96.83_{\pm 0.36}$	$98.03_{\pm 0.85}$	$97.53_{\pm 0.12}$	$98.40_{\pm 0.17}$	$\mathbf{98.70}_{\pm \mathbf{1.32}}$	$97.19_{\pm 0.53}$	$\underline{98.45}_{\pm 0.07}$
	MF1(%)	$95.47_{\pm 2.43}$	$96.70_{\pm 0.52}$	$97.90_{\pm 1.27}$	$97.53_{\pm 0.12}$	$98.36_{\pm 0.22}$	$\mathbf{98.51}_{\pm \mathbf{1.64}}$	$97.04_{\pm 0.89}$	$\underline{98.40}_{\pm 0.09}$
SRSCP1	ACC(%)	$78.45_{\pm 0.76}$	$82.83_{\pm 2.35}$	$81.58_{\pm 2.73}$	$84.43_{\pm 1.54}$	$80.27_{\pm 0.76}$	$\underline{88.18}_{\pm 1.08}$	$\mathbf{88.24}_{\pm \mathbf{1.35}}$	$87.97_{\pm 0.87}$
	MF1(%)	$78.18_{\pm 0.83}$	$82.66_{\pm 2.63}$	$81.31_{\pm 3.15}$	$83.88_{\pm 1.75}$	$80.04_{\pm 0.83}$	$\mathbf{88.04}_{\pm \mathbf{1.42}}$	$87.60_{\pm 1.88}$	$\underline{87.89}_{\pm 1.01}$

Table 1: Quantitative Comparison of MTS Classification

Method	**Metrics**	**Weather**	**Electric**	**ETTh1**	**Traffic**
MSGNet *AAAI'24*	MSE	0.250	0.191	0.454	0.430
	MAE	0.279	0.283	0.458	0.282
iTransformer *ICLR'24*	MSE	0.254	0.175	0.446	0.428
	MAE	0.274	0.266	0.439	0.279
Ada-Hyper *NeurIPS'24*	MSE	0.233	0.167	<u>0.418</u>	0.419
	MAE	<u>0.259</u>	0.259	<u>0.426</u>	0.267
HyperMixer *AAAI'25*	MSE	<u>0.228</u>	<u>0.165</u>	0.424	**0.405**
	MAE	0.260	<u>0.257</u>	0.428	**0.258**
RHCL (Our)	MSE	**0.217**	**0.152**	**0.404**	<u>0.408</u>
	MAE	**0.246**	**0.246**	**0.413**	<u>0.260</u>

Table 2: Comparison of MTS Forecasting (Length 192)

ST (Du et al. 2025), MPTSNet (Mu, Shahzad, and Zhu 2025), SAGoG (Wang et al. 2025). ***Forecasting***: MSGNet (Cai et al. 2024), iTransformer (Liu et al. 2023), Ada-Hyper (Shang et al. 2024), HyperMixer (Tian et al. 2025).

Implementation This study focuses exclusively on two roles of MTS: synergy and inhibition. For pre-training, the Role HGNN is utilized as the encoder. During fine-tuning, the classifier or predictor is employed in downstream tasks. The classification performance of the models is evaluated using the Accuracy (ACC) and Macro-F1 score (MF1). Mean Squared Error (MSE) and Mean Absolute Error (MAE) are used for predictive performance.

Experimental Results

Classification Table 1 presents the quantitative classification results. The RHCL achieves the best performance on six out of eight datasets. Notably, RHCL demonstrates substantial improvements on FM and IW, surpassing the best prior results by 2.99% and 4.30% in accuracy, respectively. In the remaining cases (SAD and SRSCP1), RHCL attains competitive results. The superior performance of RHCL stems from its ability to partition semantic roles when constructing role hypergraphs, along with Role HGNN, enabling detailed modeling of high-order sensor correlations.

Forecasting Table 2 reveals that RHCL outperforms all compared baselines across four forecasting datasets. Besides the inherent advantages of role hypergraphs, the superiority can be attributed to the spatial structure variations captured through role hypergraph structural contrasting. Moreover, it benefits from the alignment of role complementary information, capturing temporal continuity and consistency in MTS.

Ablation Study

Three categories of variants of the RHCL are created to evaluate core components: 1) Type: Replacing the Role HGNN with a standard HGNN; 2) Role: Disabling either the synergy or inhibition roles within RHCL; and 3) Contrast: Removing either the temporal or spatial contrast module.

The ablation results in Table 3 validate the effectiveness of our design. Specifically, replacing the Role HGNN with a standard HGNN results in a notable performance drop, demonstrating the necessity of incorporating role differentiation in sensor correlations. For roles, while the joint modeling of both achieves the best performance (confirming their complementary nature), inhibition-only outperforms synergy-only, suggesting antagonistic relationships may encode more discriminative features. Regarding the contrast mechanisms, the removal of the spatial or temporal contrast module results in performance deterioration, emphasizing the significance of capturing the spatial dependencies and achieving temporal consistency based on the role structure.

Type		Role		Contrast		HAR		ISRUC		ETTh1	
HGNN	*RHGNN*	*Synergy*	*Inhibit*	*Spatial*	*Temporal*	ACC (↑)	MF1 (↑)	ACC (↑)	MF1 (↑)	MSE (↓)	MAE (↓)
✓				✓	✓	$90.68_{\pm 0.75}$	$90.13_{\pm 0.78}$	$80.25_{\pm 0.30}$	$79.76_{\pm 0.36}$	0.458	0.462
	✓	✓		✓	✓	$91.80_{\pm 0.34}$	$91.08_{\pm 0.36}$	$82.86_{\pm 0.15}$	$82.16_{\pm 0.23}$	0.429	0.437
	✓		✓	✓	✓	$92.45_{\pm 0.40}$	$91.97_{\pm 0.54}$	$83.29_{\pm 0.36}$	$82.85_{\pm 0.40}$	0.432	0.436
	✓	✓	✓	✓		$94.14_{\pm 0.17}$	$93.92_{\pm 0.24}$	$85.04_{\pm 0.24}$	$85.00_{\pm 0.31}$	0.409	0.419
	✓	✓	✓		✓	$93.59_{\pm 0.13}$	$92.89_{\pm 0.21}$	$84.55_{\pm 0.31}$	$84.39_{\pm 0.34}$	0.420	0.429
	✓	✓	✓	✓	✓	$\mathbf{95.14}_{\pm 0.09}$	$\mathbf{94.85}_{\pm 0.12}$	$\mathbf{85.96}_{\pm 0.15}$	$\mathbf{85.16}_{\pm 0.19}$	**0.404**	**0.413**

Table 3: Ablation Studies on proposed RHCL

Datasets	HAR ACC (↑)	ISRUC ACC (↑)	Weather MSE (↓)	Traffic MSE (↓)
TodyNet	91.75	76.37	/	/
TodyNet+Role	92.34	77.03	/	/
Ada-Hyper	/	/	0.233	0.419
Ada-Hyper+Role	/	/	0.220	0.410
RHCL w/o Role	90.68	80.25	0.262	0.435
RHCL	95.14	85.96	0.217	0.408

Table 4: Impact of Role Information on Conventional Model

Discussion

Can Role Information Improve Conventional Methods? To test if role information can improve other (hyper)graph methods, we incorporate the role information into baselines (TodyNet and Ada-Hyper) by separately modeling synergy and inhibition roles across four datasets. Results in Table 4 indicate that introducing role leads to modest improvements over the original versions. However, a notable performance gap remains compared to RHCL. This highlights that naively modeling roles as separate structures fails to capture overall role characteristics, underscoring the necessity of modeling roles within a unified framework as RHCL. Notably, while both graph and hypergraph methods benefit from role information, the hypergraph-based baseline (Ada-Hyper) achieves higher gains than the graph-based one. This divergence aligns with the nature of role interactions, which exhibit high-order cluster patterns among vertices. Thus, hypergraphs prove more compatible with role structures.

Interpretability of Spatial Dependencies To intuitively present the spatial dependencies captured by RHCL, we visualize the sensor correlations from a structural perspective in the HAR dataset. As shown in Figure 3, the correlation structure exhibits short-term stability across adjacent windows, which is critical for robust learning (Jia et al. 2020). As time progresses, the figure also reveals gradual structural variations corresponding to state transitions. Specifically, the patterns of synergistic and inhibitory sensor relationships among sensors progressively adapt in alignment with underlying state transitions. Consequently, capturing such structural variations is essential for identifying state transitions in tasks such as MTS classification. RHCL adapts to stable and evolving interaction patterns, allowing it to remain sensitive to nuanced contextual shifts while preserving robustness.

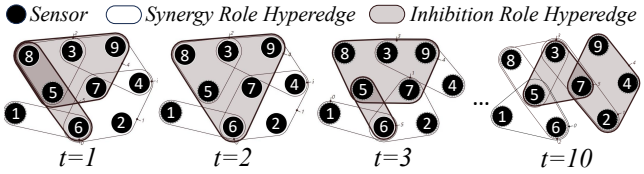


Figure 3: Spatial Structural Dependency across Windows

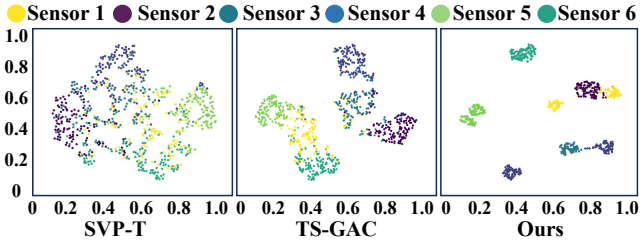


Figure 4: Comparison of Sensor-level Temporal Consistency

Interpretability of Temporal Consistency To investigate sensor-level temporal dependency, we visualize the complementary role information for each sensor across windows on SRSCP1 from a feature perspective. Figure 4 shows that RHCL exhibits a more compact feature distribution for the same sensor compared to baselines, indicating effective capture of sensor-level temporal consistency by aligning role information. Interpretively, the compactness of feature distribution signifies that RHCL preserves the semantic stability of each sensor across temporal contexts, whereas baselines produce scattered or drifting representations that undermine temporal integrity. This property is critical in MTS analysis, as the overall robustness of representations fundamentally relies on the stability of each sensor (Queen et al. 2023; Li et al. 2023). Ensuring sensor-level temporal consistency is essential for robust learning in downstream tasks.

Conclusion

In this paper, we introduce a generalized hypergraph structure, enabling precise modeling of sensor correlations. Building on this structure, we propose RHCL for MTS analysis. This approach captures spatial short-term consistency and long-term evolution through role hypergraph structural contrasting. Moreover, aligning complementary role information ensures temporal consistency. Experiments demonstrate the effectiveness and interpretability of RHCL.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant Nos. U24A20252 and 62088102, the Key Research and Development Program of Shaanxi Province of China under Grant Nos. 2024PT-ZCK-66 and 2024CY2-GJHX-48.

References

Anguita, D.; Ghio, A.; Oneto, L.; Parra, X.; and Reyes-Ortiz, J. L. 2012. Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. In *International workshop on ambient assisted living*, 216–223. Springer.

Bao, P.; Li, J.; Yan, R.; and Liu, Z. 2024. Dynamic graph contrastive learning via maximize temporal consistency. *Pattern Recognition*, 148: 110144.

Cai, W.; Liang, Y.; Liu, X.; Feng, J.; and Wu, Y. 2024. Msgnet: Learning multi-scale inter-series correlations for multivariate time series forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, 11141–11149.

Chen, H.; and Eldardiry, H. 2024. Graph time-series modeling in deep learning: a survey. *ACM Transactions on Knowledge Discovery from Data*, 18(5): 1–35.

Chen, H.; Rossi, R. A.; Mahadik, K.; Kim, S.; and Eldardiry, H. 2023a. Hypergraph neural networks for time-series forecasting. In *2023 IEEE International Conference on Big Data (BigData)*, 1076–1080. IEEE.

Chen, L.; Chen, D.; Shang, Z.; Wu, B.; Zheng, C.; Wen, B.; and Zhang, W. 2023b. Multi-scale adaptive graph neural network for multivariate time series forecasting. *IEEE Transactions on Knowledge and Data Engineering*, 35(10): 10748–10761.

Chen, T.; Kornblith, S.; Norouzi, M.; and Hinton, G. 2020. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, 1597–1607. PmLR.

Chen, Z.; Zuo, W.; Hu, Q.; and Lin, L. 2015. Kernel sparse representation for time series classification. *Information Sciences*, 292: 15–26.

Chowdhury, S.; and Mémoli, F. 2019. The Gromov–Wasserstein distance between networks and stable network invariants. *Information and Inference: A Journal of the IMA*, 8(4): 757–787.

Chowdhury, S.; Needham, T.; Semrad, E.; Wang, B.; and Zhou, Y. 2024. Hypergraph co-optimal transport: Metric and categorical properties. *Journal of Applied and Computational Topology*, 8(5): 1171–1230.

Cirstea, R.-G.; Yang, B.; Guo, C.; Kieu, T.; and Pan, S. 2022. Towards spatio-temporal aware traffic time series forecasting. In *2022 IEEE 38th International Conference on Data Engineering (ICDE)*, 2900–2913. IEEE.

Du, M.; Wei, Y.; Tang, Y.; Zheng, X.; Wei, S.; and Ji, C. 2025. ST-Tree with interpretability for multivariate time series classification. *Neural Networks*, 183: 106951.

Eldele, E.; Ragab, M.; Chen, Z.; Wu, M.; Kwoh, C. K.; Li, X.; and Guan, C. 2021. Time-series representation learning via temporal and contextual contrasting. *arXiv preprint arXiv:2106.14112*.

Feng, Y.; Huang, J.; Du, S.; Ying, S.; Yong, J.-H.; Li, Y.; Ding, G.; Ji, R.; and Gao, Y. 2024. Hyper-yolo: When visual object detection meets hypergraph computation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Feng, Y.; You, H.; Zhang, Z.; Ji, R.; and Gao, Y. 2019. Hypergraph neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, 3558–3565.

Han, X.; Xue, R.; Du, S.; and Gao, Y. 2024. Inter-intra high-order brain network for ASD diagnosis via functional MRIs. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 216–226. Springer.

Han, X.; Xue, R.; Feng, J.; Feng, Y.; Du, S.; Shi, J.; and Gao, Y. 2025. Hypergraph foundation model for brain disease diagnosis. *IEEE Transactions on Neural Networks and Learning Systems*.

Hao, S.; Wang, Z.; Alexander, A. D.; Yuan, J.; and Zhang, W. 2023. MICOS: Mixed supervised contrastive learning for multivariate time series classification. *Knowledge-Based Systems*, 260: 110158.

Ji, S.; Feng, Y.; Di, D.; Ying, S.; and Gao, Y. 2025. Mode Hypergraph Neural Network. *IEEE Transactions on Neural Networks and Learning Systems*.

Jia, Z.; Lin, Y.; Wang, J.; Zhou, R.; Ning, X.; He, Y.; and Zhao, Y. 2020. GraphSleepNet: Adaptive Spatial-Temporal Graph Convolutional Networks for Sleep Stage Classification. In *IJCAI*, 1324–1330.

Khalighi, S.; Sousa, T.; Santos, J. M.; and Nunes, U. 2016. ISRUC-Sleep: A comprehensive public dataset for sleep researchers. *Computer methods and programs in biomedicine*, 124: 180–192.

Khosla, P.; Teterwak, P.; Wang, C.; Sarna, A.; Tian, Y.; Isola, P.; Maschinot, A.; Liu, C.; and Krishnan, D. 2020. Supervised contrastive learning. *Advances in neural information processing systems*, 33: 18661–18673.

Lai, G.; Chang, W.-C.; Yang, Y.; and Liu, H. 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.

Le, X.-M.; Luo, L.; Aickelin, U.; and Tran, M.-T. 2024. Shapeformer: Shapelet transformer for multivariate time series classification. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 1484–1494.

Li, Z. L.; Zhang, G. W.; Yu, J.; and Xu, L. Y. 2023. Dynamic graph structure learning for multivariate time series forecasting. *Pattern Recognition*, 138: 109423.

Liu, H.; Yang, D.; Liu, X.; Chen, X.; Liang, Z.; Wang, H.; Cui, Y.; and Gu, J. 2024. Todynet: temporal dynamic graph neural network for multivariate time series classification. *Information Sciences*, 677: 120914.

Liu, J.; and Chen, S. 2024. Timesurl: Self-supervised contrastive learning for universal time series representation

learning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 38, 13918–13926.

Liu, Y.; Hu, T.; Zhang, H.; Wu, H.; Wang, S.; Ma, L.; and Long, M. 2023. itransformer: Inverted transformers are effective for time series forecasting. *arXiv preprint arXiv:2310.06625*.

Luo, D.; Cheng, W.; Wang, Y.; Xu, D.; Ni, J.; Yu, W.; Zhang, X.; Liu, Y.; Chen, Y.; Chen, H.; et al. 2023. Time series contrastive learning with information-aware augmentations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, 4534–4542.

Mu, Y.; Shahzad, M.; and Zhu, X. X. 2025. MPTSNet: Integrating multiscale periodic local patterns and global dependencies for multivariate time series classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 19572–19580.

Othman, S.; Cohn, J. E.; Burdett, J.; Daggumati, S.; and Bloom, J. D. 2020. Temporal augmentation: a systematic review. *Facial Plastic Surgery*, 36(03): 217–225.

Pöppelbaum, J.; Chadha, G. S.; and Schwung, A. 2022. Contrastive learning based self-supervised time-series analysis. *Applied Soft Computing*, 117: 108397.

Queen, O.; Hartvigsen, T.; Koker, T.; He, H.; Tsiligkaridis, T.; and Zitnik, M. 2023. Encoding time-series explanations through self-supervised model behavior consistency. *Advances in Neural Information Processing Systems*, 36: 32129–32159.

Shang, Z.; Chen, L.; Wu, B.; and Cui, D. 2024. Ada-MSHyper: adaptive multi-scale hypergraph transformer for time series forecasting. *Advances in Neural Information Processing Systems*, 37: 33310–33337.

Tian, C.; Lu, Z.; Zhang, Z.; Yang, H.; Cao, W.; Guo, Z.; Sun, X.; and Jin, L. 2025. HyperMixer: Specializable Hypergraph Channel Mixing for Long-term Multivariate Time Series Forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 20885–20893.

Wang, B.; Wang, P.; Zhang, Y.; Wang, X.; Zhou, Z.; Bai, L.; and Wang, Y. 2024a. Towards dynamic spatial-temporal graph learning: A decoupled perspective. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, 9089–9097.

Wang, S.; Zhang, Y.; Lin, X.; Hu, Y.; Huang, Q.; and Yin, B. 2025. SAGoG: Similarity-Aware Graph of Graphs Neural Networks for Multivariate Time Series Classification. *IEEE Transactions on Knowledge and Data Engineering*.

Wang, Y.; Han, Y.; Wang, H.; and Zhang, X. 2023. Contrast everything: A hierarchical contrastive framework for medical time-series. *Advances in Neural Information Processing Systems*, 36: 55694–55717.

Wang, Y.; Long, H.; Zheng, L.; and Shang, J. 2024b. Graphformer: Adaptive graph correlation transformer for multivariate long sequence time series forecasting. *Knowledge-Based Systems*, 285: 111321.

Wang, Y.; Xu, Y.; Yang, J.; Wu, M.; Li, X.; Xie, L.; and Chen, Z. 2024c. Fully-connected spatial-temporal graph for multivariate time-series data. In *Proceedings of the AAAI conference on artificial intelligence*, volume 38, 15715–15724.

Wang, Y.; Xu, Y.; Yang, J.; Wu, M.; Li, X.; Xie, L.; and Chen, Z. 2024d. Graph-aware contrasting for multivariate time-series classification. In *Proceedings of the AAAI conference on artificial intelligence*, volume 38, 15725–15734.

Wu, Z.; Pan, S.; Long, G.; Jiang, J.; Chang, X.; and Zhang, C. 2020. Connecting the dots: Multivariate time series forecasting with graph neural networks. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, 753–763.

Xue, R.; Han, X.; Hu, H.; Zhang, Z.; Du, S.; and Gao, Y. 2025a. Adaptive Embedding for Long-Range High-Order Dependencies via Time-Varying Transformer on fMRI. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 46–55. Springer.

Xue, R.; Hu, H.; Zhang, Z.; Han, X.; Wang, J.; Gao, Y.; and Du, S. 2025b. DHGFormer: Dynamic Hierarchical Graph Transformer for Disorder Brain Disease Diagnosis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 268–278. Springer.

Ye, J.; Liu, Z.; Du, B.; Sun, L.; Li, W.; Fu, Y.; and Xiong, H. 2022. Learning the evolutionary and multi-scale graph structure for multivariate time series forecasting. In *Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining*, 2296–2306.

Yue, Z.; Wang, Y.; Duan, J.; Yang, T.; Huang, C.; Tong, Y.; and Xu, B. 2022. Ts2vec: Towards universal representation of time series. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, 8980–8987.

Zeng, A.; Chen, M.; Zhang, L.; and Xu, Q. 2023. Are transformers effective for time series forecasting? In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, 11121–11128.

Zhang, K.; Wen, Q.; Zhang, C.; Cai, R.; Jin, M.; Liu, Y.; Zhang, J. Y.; Liang, Y.; Pang, G.; Song, D.; et al. 2024. Self-supervised learning for time series analysis: Taxonomy, progress, and prospects. *IEEE transactions on pattern analysis and machine intelligence*.

Zhang, X.; Zhao, Z.; Tsiligkaridis, T.; and Zitnik, M. 2022. Self-supervised contrastive pre-training for time series via time-frequency consistency. *Advances in Neural Information Processing Systems*, 35: 3988–4003.

Zhou, H.; Zhang, S.; Peng, J.; Zhang, S.; Li, J.; Xiong, H.; and Zhang, W. 2021. Informer: Beyond efficient transformer for long sequence time-series forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, 11106–11115.

Zuo, R.; Li, G.; Choi, B.; Bhowmick, S. S.; Mah, D. N.-y.; and Wong, G. L. 2023. SVP-T: A shape-level variable-position transformer for multivariate time series classification. In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, 11497–11505.