PRELIMINARY PREPRINT VERSION – DO NOT CITE
This is not the version of record. The final version will appear in the Proceedings of IJCAI-ECAI 2026.

McGCN: Learning Continuous Graph Dynamics for Multi-Channel Fusion

Na Song[1,2,3], **Zihan Fang**[4], **Weidong Zhang**[1,3,5*], **Zehua Jia**[1,3] and **Shiping Wang**[4]
[1]School of Information and Communication Engineering, Hainan University, Haikou, China
[2]College of Artificial Intelligence, Putian University, Putian, China
[3]Engineering Research Center of Marine Intelligent Systems, Ministry of Education, Haikou, China
[4]College of Computer and Data Science, Fuzhou University, Fuzhou, China
[5]School of Automation and Intelligent Sensing, Shanghai Jiao Tong University, Shanghai, China
nasong1010@163.com, fzihan11@163.com, wdzhang@sjtu.edu.cn, zhjia@hainanu.edu.cn, shipingwangphd@163.com

Abstract

Multi-channel data fusion is essential for capturing comprehensive representations in complex systems. While graph convolutional networks have demonstrated remarkable efficacy, existing fusion paradigms primarily rely on discrete architectures governed by first-order propagation. These models are typically confined to discrete message-passing mechanisms, which makes it difficult to characterize the continuous evolution of underlying system dynamics. To address these limitations, we propose a multi-channel Graph Continuous Network (mcGCN), a novel framework for multi-channel data fusion. By formulating the information propagation as a second-order partial differential equation on graphs, mcGCN transitions from discrete layer-wise updates to a continuous dynamical system. Specifically, mcGCN integrates a multi-channel encoding module with continuous feature dynamics to initialize and evolve the latent node representations over static graph topologies. This physical analogy enables more robust information flow and effectively mitigates the performance degradation typically associated with deep graph architectures. Extensive experiments on diverse benchmark datasets demonstrate that our method outperforms state-of-the-art baselines, validating its effectiveness and robustness.

1 Introduction

Multi-channel signal analysis is a fundamental component of various real-world applications, ranging from traditional industrial condition monitoring [Kou *et al.*, 2024] to emerging autonomous perception systems (e.g., unmanned surface and aerial vehicles) [Jia *et al.*, 2024b] [Sun *et al.*, 2024] [Li Can, 2025]. To model the spatial correlations among channels, Graph Neural Networks (GNNs) have been widely adopted [Wang *et al.*, 2026] [Kou *et al.*, 2025]. By representing channels as nodes and inter-channel dependencies as edges, GNNs enable structured information propagation over the sensor graph, allowing the extraction of high-level representations that jointly encode local channel properties and global spatial context [Wang *et al.*, 2020] [Song *et al.*, 2025].

Despite their promising performance, existing graph-based paradigms encounter two fundamental challenges. The primary bottleneck lies in the architectural mismatch between discrete model formulations and continuous physical realities [Sun *et al.*, 2025] [Kou *et al.*, 2026]. Most GNNs architectures propagate features through a discrete stacking of layers, which inherently contradicts the continuous-time dynamics of physical signals [Khemani *et al.*, 2024] [Yuan *et al.*, 2024]. In scenarios like smart ocean monitoring or low-altitude sensing, sensor measurements evolve via continuous physical processes governed by differential equations rather than fixed discrete steps [Jia *et al.*, 2024a] [Wu *et al.*, 2022]. This discrepancy introduces approximation errors, constraining the ability of GNNs to capture spatiotemporal variations [Song *et al.*, 2026] [He *et al.*, 2026]. Compounding this architectural limitation is the pervasive challenge of label scarcity [Yuan *et al.*, 2025] [Dong and Ansari, 2020]. Acquiring high-quality annotations for large-scale sensor data is prohibitively expensive and labor-intensive. Consequently, conventional fully supervised fusion frameworks struggle to generalize, failing to effectively exploit the intrinsic geometric structure present within the unlabeled multi-channel data.

To bridge these theoretical and practical gaps, we introduce mcGCN, a framework for learning continuous graph dynamics in multi-channel fusion. From a dynamics modeling perspective, we revisit the multi-channel fusion process through the lens of partial differential equations (PDEs). Instead of formulating graph convolution as a static stacking of discrete layers, we interpret message passing as the numerical discretization of a continuous diffusion process over the graph manifold. This perspective enables mcGCN to naturally capture the intrinsic spatiotemporal continuity of multi-channel signals, allowing feature representations to evolve smoothly in accordance with underlying physical dynamics. Meanwhile, to address label scarcity, we incorporate a label propagation mechanism that operates on the learned continuous graph structure. By diffusing supervisory signals from sparse annotations to unlabeled nodes, this diffusion process effectively enforces global consistency and substantially improves downstream fusion performance in low-resource engineering

*Corresponding author

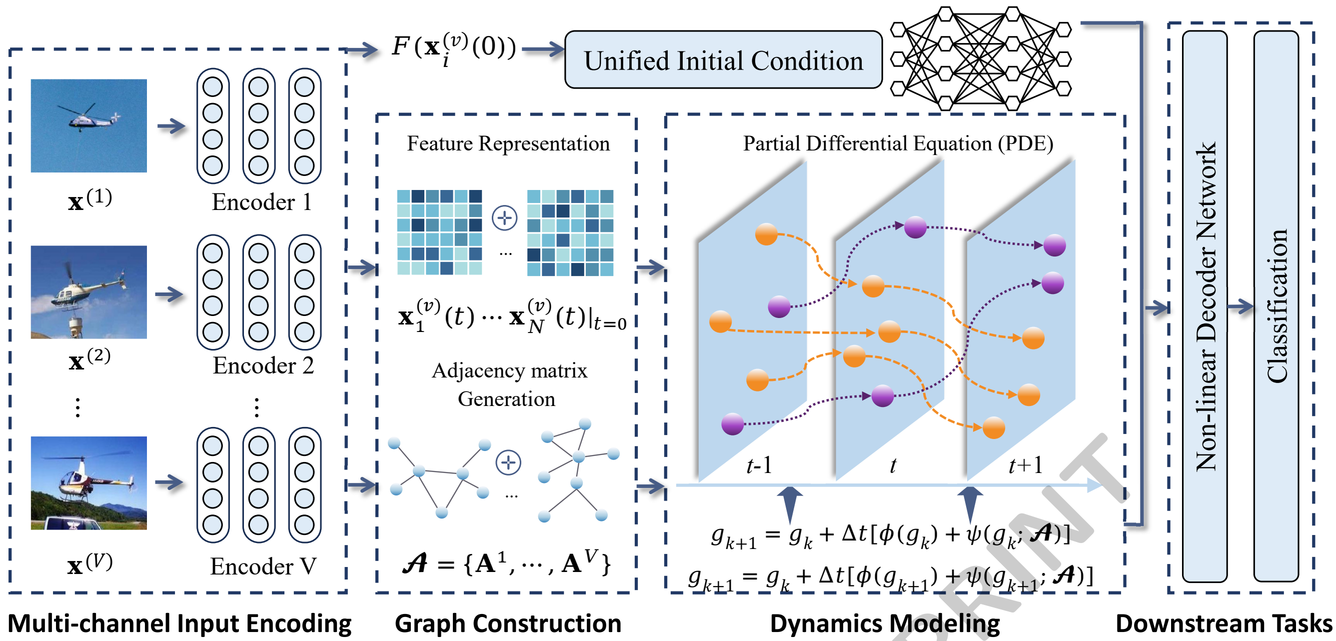


Figure 1: The architecture of the proposed mcGCN. The framework first encodes multi-channel inputs and constructs a graph representation, then propagates information through a second-order graph PDE to capture continuous dynamics for downstream tasks.

scenarios, as illustrated in Figure 1. The main contributions of this work are summarized as follows:

- **Unified Continuous Dynamics**. Propose mcGCN, which harmonizes representation learning and label propagation within a single continuous dynamic system, thereby capturing the diffusive nature of signal flow to enable robust fusion in label-scarce scenarios.
- **A PDE-Based Perspective**. Formulate graph convolution via PDEs by interpreting network depth as continuous evolution time. This approach resolves the mismatch between discrete layers and continuous signal dynamics, capturing fine-grained spatiotemporal variations that discrete GNNs struggle to resolve.
- **Unified Theoretical Perspective and Robustness**. Theoretically prove that several representative GNNs are specific discretizations of our continuous framework. This unified perspective translates into state-of-the-art performance and robustness in label-scarce scenarios.

2 Background and Preliminaries

In this section, we revisit the standard graph convolution formulations that underpin our framework.

2.1 Revisiting Single-Channel Graph Convolution

We denote an attributed graph as an ordered quadruple $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{A}, \mathbf{X})$, where $\mathcal{V}$ is the node set with $|\mathcal{V}| = N$, $\mathbf{A} \in \{0,1\}^{N\times N}$ is the adjacency matrix, and $\mathbf{X} \in \mathbb{R}^{N\times D}$ represents the node feature matrix with feature dimension D. Given a single-channel graph $\mathcal{G}$, graph convolution aims to learn an operator $g_\theta : \mathbb{R}^D \to \mathbb{R}^d$ such that for any graph signal $\mathbf{x} \in \mathbb{R}^D$ defined on the nodes, the output $g_\theta * \mathbf{x}$ preserves the graph domain. Then we revisit graph convolution formulations on a single attributed graph. Given an attributed graph $\mathcal{G} = (\mathcal{V}, \mathbf{A}, \mathbf{X})$, the spectral graph convolution of a signal $\mathbf{x}$ is formulated as

$$g_\theta * \mathbf{x} = \mathbf{U} g_\theta \mathbf{U}^\top \mathbf{x}, \tag{1}$$

where $\mathbf{U}$ is the eigenvectors of the normalized graph Laplacian $\mathbf{L} = \mathbf{I} - \mathbf{D}^{-\frac{1}{2}}\mathbf{A}\mathbf{D}^{-\frac{1}{2}}$, and $g_\theta = \mathrm{diag}(\theta)$ denotes a spectral filter parameterized by θ. From a graph signal processing perspective, g_θ serves as a filter in the graph Fourier domain, expressible as a function of the Laplacian eigenvalues $\mathbf{\Lambda}$.

To avoid the computational burden of eigendecomposition, Hammond et al. [Hammond *et al.*, 2011] proposed approximating the spectral filter via a truncated Chebyshev polynomial expansion

$$g_{\theta'}(\mathbf{\Lambda}) \approx \sum_{k=0}^{K} \theta'_k T_k(\mathbf{\Lambda}'), \tag{2}$$

where $\mathbf{\Lambda}' = \frac{2}{\lambda_{\max}}\mathbf{\Lambda} - \mathbf{I}$ is the rescaled eigenvalue matrix, $\lambda_{\max}$ denotes the largest eigenvalue of $\mathbf{L}$, and $\theta' \in \mathbb{R}^K$ represents the Chebyshev coefficients. Building upon this formulation, [Kipf and Welling, 2017] simplifies the convolution by restricting the polynomial order to $K = 1$, deriving a first-order approximation as

$$g_{\theta'} * \mathbf{x} \approx \theta'_0\mathbf{x} + \theta'_1(\mathbf{L} - \mathbf{I})\mathbf{x} = \theta'\tilde{\mathbf{A}}\mathbf{x}, \tag{3}$$

where $\tilde{\mathbf{A}} = \mathbf{I} + \mathbf{D}^{-\frac{1}{2}}\mathbf{A}\mathbf{D}^{-\frac{1}{2}}$ denotes the graph filter derived by enforcing the parameter constraint $\theta' = \theta'_0 = -\theta'_1$.

Generalizing to high-dimensional node features $\mathbf{X} \in \mathbb{R}^{N\times D}$, the graph convolution operation is formulated as $g_{\theta'} * \mathbf{X} \approx \theta'_0\mathbf{X} + \theta'_1(\mathbf{L} - \mathbf{I})\mathbf{X} = \widehat{\mathbf{A}}\mathbf{X}\mathbf{\Theta}$, where $\mathbf{\Theta} \in \mathbb{R}^{D\times d}$ denotes the trainable weight matrix, and $\widehat{\mathbf{A}} = \tilde{\mathbf{D}}^{-\frac{1}{2}}\mathbf{A}\tilde{\mathbf{D}}^{-\frac{1}{2}}$ is the renormalized adjacency matrix. The output $g_{\theta'} * \mathbf{X}$ yields the updated node representations following a single message-passing iteration.

2.2 Continuous Graph Dynamics

Conventionally, GNNs formulate node representation learning as a discrete state evolution process. By stacking a finite sequence of propagation layers, the feature update at the discrete step l is typically modeled as a recurrence relation:

$$\mathbf{H}^{(l+1)} = \sigma\big(\mathcal{P}(\mathbf{A})\mathbf{H}^{(l)}\mathbf{W}^{(l)}\big), \tag{4}$$

where $\mathbf{H}^{(l)}$ represents the node states at layer l, $\mathcal{P}(\mathbf{A})$ denotes the normalized structural operator, $\mathbf{W}^{(l)}$ governs the parametric transformation, and $\sigma(\cdot)$ is an element-wise nonlinearity. This formulation implicitly treats representation learning on graphs as a sequence of discrete transformations indexed by the layer depth.

From a dynamical perspective, this layer-wise propagation can be interpreted as the explicit Euler discretization of a continuous evolution process. By reformulating the update rule as a difference quotient, we have:

$$\frac{\mathbf{H}^{(l+1)} - \mathbf{H}^{(l)}}{\Delta t} \approx f(\mathbf{H}^{(l)}, \mathbf{L}), \tag{5}$$

where Δt represents the discretization step size. In the continuous limit where $\Delta t \to 0$, this discrete sequence converges to a Graph Ordinary Differential Equation (Graph ODE):

$$\frac{d\mathbf{H}(t)}{dt} = f(\mathbf{H}(t), \mathbf{L}), \tag{6}$$

where $\mathbf{H}(t)$ represents the node states at continuous time t. This continuous formulation provides a principled connection between graph convolution and diffusion processes on graphs, where information propagates smoothly according to the graph topology. It also reveals inherent limitations of purely discrete architectures: coarse discretization with a small number of layers may fail to accurately approximate the underlying continuous dynamics, leading to suboptimal modeling of fine-grained spatiotemporal variations. These observations motivate the development of graph learning frameworks that explicitly model continuous graph dynamics, offering a more faithful representation of physical signal evolution and structured data propagation.

3 The Proposed Method

Leveraging the continuous dynamics perspective established in Subsection 2.2, we present mcGCN, a unified framework for multi-channel signal fusion on graphs. Fundamentally, mcGCN reformulates the feature propagation task not as discrete steps, but as a continuous diffusion process evolving over the graph topology. We first derive a continuous graph convolution extending message passing to the time domain. Subsequently, we detail the architecture instantiated for label-scarce scenarios. Ultimately, we prove that mcGCN unifies classical models as special cases.

3.1 Notations and Problem Setup

In this work, we address the challenge of representation learning on multi-channel attributed graphs, particularly in label-scarce scenarios. To model multi-channel signals, we define a multi-channel attributed graph set as $\mathbb{G} = \{\mathcal{G}^v\}_{v=1}^V$, where each channel is denoted by $\mathcal{G}^v = (\mathcal{V}, \mathbf{A}^v, \mathbf{X}^v)$. In practical multi-sensor engineering systems, the node set $\mathcal{V}$ is typically shared across channels, representing the same set of physical entities. Conversely, each channel is characterized by a specific adjacency matrix $\mathbf{A}^v \in \mathbb{R}^{N\times N}$ and a feature matrix $\mathbf{X}^v \in \mathbb{R}^{N\times D_v}$, reflecting heterogeneous sensing modalities and distinct topological interactions.

Formally, the goal of multi-channel graph convolution is to synthesize the set of channel-wise graph signals $\{\mathbf{X}^v\}_{v=1}^V$ into a unified node representation. This process defines a fusion operator: $m_\theta : \prod_{v=1}^V \mathbb{R}^{N\times D_v} \to \mathbb{R}^{N\times D}$, which integrates heterogeneous information from multiple views while preserving the underlying structural consistency of the shared graph domain. Assuming there are C classes, the model follows the geometric properties of a shared graph structure, aiming to efficiently propagate information from a few labeled samples to the majority of unlabeled nodes.

3.2 Unified Multi-Channel Graph Convolution

In this subsection, we derive a unified multi-channel graph convolution operator grounded in continuous-time dynamics. Formally, we model the information propagation on each specific channel v as a continuous dynamical system governed by the following ODE:

$$\begin{cases} \frac{\partial \mathbf{h}_i^v(t)}{\partial t} = \underbrace{F(\mathbf{h}_i^v(t))}_{\text{Self-evolution}} + \underbrace{\sum_{j\in\mathcal{V}} \mathbf{A}_{ij}^v G(\mathbf{h}_i^v(t), \mathbf{h}_j^v(t))}_{\text{Neighbor Interaction}}, & t > 0, \\ \mathbf{h}_i^v(0) = f_{\Theta_E^v}(\mathbf{x}_i^v), & t = 0. \end{cases} \tag{7}$$

Here, $\mathbf{h}_i^v(t) \in \mathbb{R}^D$ denotes the continuous latent state of node i in channel v at time t. The initial state $\mathbf{h}_i^v(0)$ is projected from the raw observation $\mathbf{x}_i^v \in \mathbb{R}^{D_v}$ via a channel-specific encoder $f_{\Theta_E^v}$. In Eq. (7), the function $F : \mathbb{R}^D \to \mathbb{R}^D$ characterizes the autonomous self-dynamics of the node, while $G : \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}^D$ models the pairwise interaction dynamics between connected neighbors.

To provide a unified theoretical justification for multi-channel coupling, we hypothesize the existence of a global latent potential $\chi(t, \mu)$, where μ represents a continuous coordinate in the channel space. Crucially, while real-world observations consist of a finite number of discrete channels, we conceptualize them as discrete samples drawn from this underlying continuous manifold. Under this formulation, the specific observation from channel v corresponds to the gradient of this potential with respect to the view coordinate

μ_v. Consequently, the multi-view learning process can be modeled as the evolution of this potential field. By lifting the first-order dynamics in Eq. (7) to a higher-order formulation, we derive the following second-order partial differential equation:

$$\begin{cases} \frac{\partial^2 \chi(t,\mu_v)}{\partial t \partial \mu_v} = \phi\left(\frac{\partial \chi(t,\mu_v)}{\partial \mu_v}\right) + \psi\left(\frac{\partial \chi(t,\mu_v)}{\partial \mu_v}; \mathbf{A}^v\right), & t > 0, \\ \frac{\partial \chi(t,\mu_v)}{\partial \mu_v}\Big|_{t=0} = f_{\Theta_E^v}(\mathbf{X}^v), & t = 0, \end{cases} \tag{8}$$

where $\phi(\cdot)$ and $\psi(\cdot)$ retain their roles as the self-dynamics and interaction operators, respectively. Furthermore, the mixed derivative term $\frac{\partial^2 \chi(t,\mu_v)}{\partial t \partial \mu_v}$ provides an intuitive physical justification for heterogeneous fusion. It explicitly models the cross-channel interaction dynamics, capturing how the spatial diffusion of information within a specific modality is synergistically driven by the temporal evolution of the global potential field. This perspective interprets the observed node features not as static states, but as the tangent vectors of an underlying potential field, thereby unifying heterogeneous channels as directional derivatives on a single latent manifold.

However, Eq. (8) fundamentally presents a differential perspective, describing only the local variations along each channel's direction. This decoupled formulation poses a significant challenge for fusion, as the global latent state remains implicit. To resolve this, we aggregate these directional derivatives into a unified gradient vector field, rather than attempting an intractable mathematical inversion to reconstruct the scalar field χ explicitly.

Definition 1. *(Additivity Under Concatenation, AUC) Let $\mathcal{V} = \bigcup_{k=0}^{\infty} \mathbb{R}^k$ denote the space of all finite-dimensional vectors. An operator $f : \mathcal{V} \to \mathcal{S}$, which maps vectors to an additive group $(\mathcal{S}, \uplus)$, satisfies additivity under concatenation if, for all $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, the following holds:*

$$f(x \oplus y) = f(x) \uplus f(y), \tag{9}$$

where $x \oplus y \in \mathbb{R}^{n+m}$ denotes the concatenation of x and y.

Remark. Standard linear operators (e.g., matrix multiplication) inherently satisfy the AUC condition. Leveraging the AUC property, we can now materialize the integration principle proposed in the previous subsection. Assuming that the operators $\phi(\cdot)$ and $\psi(\cdot)$ satisfy AUC, Problem (8) reduces to the following second-order PDE:

$$\begin{cases} \frac{\partial^2 \chi(t)}{\partial \mu, t} = \phi(\frac{\partial \chi(t)}{\partial \mu}) + \psi(\frac{\partial \chi(t)}{\partial \mu}; \mathcal{A}), & t > 0, \\ \frac{\partial \chi(t)}{\partial \mu}|_{t=0} = [f_{\Theta_E^1}(\mathbf{X}^1), \cdots, f_{\Theta_E^V}(\mathbf{X}^V)], & t = 0, \end{cases} \tag{10}$$

where $\mathcal{A} = \{\mathbf{A}^1, \cdots, \mathbf{A}^V\}$. Here, the initial condition of the PDE is specified by a first-order derivative. By aggregating the partial derivatives along the channel coordinate μ, we define the global state variable $g(t)$ as the concatenation of channel-specific states. Consequently, the second-order problem in Eq. (8) can be reduced to a unified first-order PDE

$$\begin{cases} \frac{\mathrm{d}g(t)}{\mathrm{d}t} = \phi(g(t)) + \psi(g(t); \mathcal{A}), & t > 0, \\ g(0) = [f_{\Theta_E^1}(\mathbf{X}^1), \cdots, f_{\Theta_E^V}(\mathbf{X}^V)], & t = 0. \end{cases} \tag{11}$$

It is worth noting that the term $\phi(g(0))$ can be theoretically absorbed into $\psi(g(t); \mathcal{A})$. Crucially, while the solution is computed via this first-order ODE, its derivation from the second-order potential field validates that the simple concatenation of channels is not an arbitrary heuristic, but mathematically grounded in the unified evolution of an underlying continuous manifold. However, we retain it explicitly to emphasize the distinct roles of node-intrinsic evolution versus structural diffusion, preserving the physical interpretability of the dynamics.

The solution to the differential equation in (11) can be expressed in its integral form. For any $t \geq 0$, we have

$$g(t) = g(0) + \int_0^t [\phi(g(s)) + \psi(g(s); \mathcal{A})] \mathrm{d}s. \tag{12}$$

Since the nonlinear operators $\phi(\cdot)$ and $\psi(\cdot)$ generally render an analytical solution intractable, we resort to numerical discretization. Given a step size $\Delta t > 0$, the continuous evolution can be approximated via the explicit and implicit Euler

$$\frac{g(t + \Delta t) - g(t)}{\Delta t} \approx \phi(g(t)) + \psi(g(t); \mathcal{A}), \tag{13}$$

$$\frac{g(t + \Delta t) - g(t)}{\Delta t} \approx \phi(g(t + \Delta t)) + \psi(g(t + \Delta t); \mathcal{A}), \tag{14}$$

which yields the following iterative formulation:

$$g_{k+1} = g_k + \Delta t\big[\phi(g_k) + \psi(g_k; \mathcal{A})\big], \tag{15}$$

$$g_{k+1} = g_k + \Delta t\big[\phi(g_{k+1}) + \psi(g_{k+1}; \mathcal{A})\big], \tag{16}$$

where $g_k \approx g(k\Delta t)$. The index k ranges over $\{0, 1, \ldots, K\}$ where $K = T/\Delta t$, with the initial condition $g_0 = g(0)$.

To obtain the final prediction, we introduce a nonlinear projector $h_{\Theta_P} : \mathbb{R}^D \to \mathbb{R}^C$ to map the evolved latent representation to the label space, yielding $\widehat{\mathbf{Y}} = h_{\Theta_P}(g(T))$.

3.3 Network Architecture and Model Training

The proposed network architecture for multi-channel signal fusion is organized into three core modules: feature alignment encoder, multi-channel convolution module, and label propagation decoder. Specifically, the feature alignment encoder projects multi-channel signals into a unified and aligned embedding space. The multi-channel convolution module then fuses heterogeneous signals over graphs in a time-continuous manner. Finally, the label propagation decoder aligns the refined feature representations with the ground-truth labels. Formally, the multi-channel fusion process is formulated as the following optimization problem

$$\min_{\Theta} \mathcal{L}(\mathbf{Y}, \widehat{\mathbf{Y}}), \text{ s.t. } \widehat{\mathbf{Y}} = h_{\Theta_P} \circ g \circ f_{\Theta_E}(\mathcal{X}), \tag{17}$$

where $\Theta_E = \{\Theta_E^v\}_{v=1}^V$ refers to the channel-specific parameters within the encoder, and Θ_P denotes the parameters of the classifier. Here, $\mathbf{Y}$ and $\widehat{\mathbf{Y}}$ denote the ground-truth and predicted label matrices, respectively. The loss function $\mathcal{L}(\cdot, \cdot)$ is instantiated as the cross-entropy loss, given by:

$$\mathcal{L}(\mathbf{Y}, \widehat{\mathbf{Y}}) = -\sum_{i \in \Omega} \sum_{j=1}^{C} \mathbf{Y}_{ij} \log(\widehat{\mathbf{Y}}_{ij}), \tag{18}$$

Algorithm 1 The proposed mcGCN

Input: Multi-channel features $\{\mathbf{X}^v\}_{v=1}^V$; ground-truth labels $\mathbf{Y}$; training index set Ω and testing index set $\overline{\Omega}$; maximum iteration number I_{max}.
Output: Predicted labels for testing samples $\{\widehat{\mathbf{y}}_i\}_{i\in\overline{\Omega}}$.

1: **Initialization**: Initial model parameters $\Theta = \{\Theta_E, \Theta_P\}$;
2: // MODEL TRAINING //
3: **while** not converged **and** iteration $i < I_{max}$ **do**
4: Extract feature representations $\{f_{\Theta_E^v}(\mathbf{X}^v)\}_{v=1}^V$ via V encoder networks; ▷ `Feature Encoding`
5: Compute the evoluted feature representation $g(T)$ by solving PDE in Eq. (11); ▷ `Multi-Channel Graph Convolution`
6: Generate label predictions $\widehat{\mathbf{Y}} = h_{\Theta_P}(g(T))$ via the projector; ▷ `Label Alignment`
7: Update parameters Θ by minimizing Eq. (18) via gradient descent; ▷ `Backward Propagation`
8: **end while**
9: **return** Prediction $\{\widehat{y}_i\}_{i\in\overline{\Omega}}$ with $\widehat{y}_i = \arg\max_{j\in[C]} \widehat{\mathbf{Y}}_{ij}$.

where Ω denotes the set of indices for training samples, and C represents the number of classes. The complete training scheme is outlined in Algorithm 1.

3.4 Connections to Existing Methods

In this subsection, we theoretically elucidate the connections between the proposed framework and a broad spectrum of representative GNNs, demonstrating that our continuous dynamics unify these distinct models as special cases.

Connection to Graph Convolutional Networks. GCNs [Kipf and Welling, 2017] adopt the following updating rules:

$$\mathbf{H}^{(l+1)} = \sigma\left(\widehat{\mathbf{A}}\mathbf{H}^{(l)}\mathbf{\Theta}^{(l+1)}\right), \tag{19}$$

where $\{\mathbf{H}^{(l)}\}_{l=1}^L$ denotes the layer-wise feature representations, and $\{\mathbf{\Theta}^{(l)}\}_{l=1}^L$ represents the set of trainable weights across a network of depth L. Simple Graph Convolution (SGC) [Wu *et al.*, 2019] reduces the layer-wise update to a single linear transformation

$$\mathbf{H}^{(L)} = \sigma\left(\widehat{\mathbf{A}}^L\mathbf{H}^{(0)}\mathbf{\Theta}\right). \tag{20}$$

By constraining $\phi(\cdot) \equiv 0$ and specifying the dynamics as $\psi(g(t);\mathcal{A}) = -\mathbf{L}g(t)$, our framework recovers the GCNs with a time step of $\Delta t = 1$, while a step size of $\Delta t = \frac{1}{L}$ yields the SGC, consistent with the theoretical analysis in [Wang *et al.*, 2021].

Connection to Personalized PageRank (PPNP). Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, PPNP [Klicpera *et al.*, 2019] defines the node representations as the steady state of a personalized propagation process:

$$\mathbf{H} = \alpha(\mathbf{I} - (1-\alpha)\widetilde{\mathbf{A}})^{-1}\mathbf{H}^{(0)}. \tag{21}$$

As a scalable approximation of PPNP, APPNP [Klicpera *et al.*, 2019] replaces the matrix inversion with L-hop iterative propagation, formulated as:

$$\mathbf{H}^{(t+1)} = (1-\alpha)\widehat{\mathbf{A}}\mathbf{H}^{(t)} + \alpha\mathbf{H}^{(0)}, \tag{22}$$

where $\alpha \in (0,1)$ is the teleport probability. Theoretically, these two models can be unified under our framework through different discretization schemes. Specifically, by setting $\phi(g(0)) = (\alpha-1)g(0)$ and $\psi(g(t);\mathcal{A}) = (1-\alpha)\widetilde{\mathbf{A}}g(t)$ with a step size $\Delta t = 1$, PPNP corresponds to a single-step integration using the backward Euler method (Eq. 16). In contrast, configuring $\phi(g(0)) = \alpha g(0)$ and $\psi(g(t);\mathcal{A}) = [(1-\alpha)\widehat{\mathbf{A}} - \mathbf{I}]g(t)$ with $\Delta t = 1$ demonstrates that the forward Euler scheme (Eq. (15)) exactly recovers the APPNP updating rule in Eq. (22).

Connection to Spectral GCNs. While GCNs and PPNP operate in the spatial domain, Spectral GCNs fundamentally define convolutions in the Fourier domain.

$$\mathbf{H}^{(L)} = \sigma\left(\sum_{l=0}^{L}\alpha_l\mathbf{L}^l\mathbf{H}^{(0)}\mathbf{\Theta}\right), \tag{23}$$

where $\{\alpha_l\}_{l=0}^L$ denotes the coefficients of the L-th order polynomial approximation. By adopting different polynomial bases, various spectral GNN architectures can be instantiated [He *et al.*, 2021; Wang and Zhang, 2022]. Crucially, if we initialize $\phi(g(0)) = \alpha_0 g(0)$, define the dynamics as $\psi(g(t);\mathcal{A}) = -\mathbf{L}g(t) = (\widehat{\mathbf{A}} - \mathbf{I})g(t)$, and employ variable step sizes $\{\Delta t_n\}_{n=1}^L$, the spectral graph convolution in Eq. (23) is analytically equivalent to performing L steps of the forward Euler integration defined in Eq. (15).

Connection to Graph Isomorphism Network. To achieve maximum discriminative capability among graph structures, GIN [Xu *et al.*, 2019] defines the layer-wise update rule for node v as:

$$\mathbf{h}_v^{(l+1)} = \mathrm{MLP}^{(l)}\left((1+\epsilon^{(l)})\mathbf{h}_v^{(l)} + \sum_{u\in\mathcal{N}(v)}\mathbf{h}_u^{(l)}\right), \tag{24}$$

where $\epsilon^{(l)}$ is a learnable parameter or a fixed scalar, and $\mathcal{N}(v)$ denotes the set of neighbors of node $v \in \mathcal{V}$. By focusing on the linear propagation component (i.e., omitting the MLP-based transformation), the GIN aggregation rule simplifies to:

$$\mathbf{H}^{(l+1)} = \left[(1+\epsilon^{(l)})\mathbf{I} + \mathbf{A}\right]\mathbf{H}^{(l)}, \tag{25}$$

which is analytically equivalent to the forward Euler discretization in Eq. (15) of our dynamics, obtained by setting $\phi(g(0)) = 0$, $\psi(g(t);\mathcal{A}) = (\epsilon^{(l)}\mathbf{I} - \mathbf{L})g(t)$, and $\Delta t = 1$.

Remark. Synthesizing the above analysis, the layer-wise propagation in GNNs can be fundamentally interpreted as the numerical discretization of a PDE on the graph manifold. In this paradigm, the graph topology constitutes the discrete spatial domain, while the network depth corresponds to the integration steps of the continuous temporal evolution. Crucially, this formulation empowers the model to intrinsically capture the diffusive nature of signal flow, ensuring that feature information propagates smoothly across coupled channels in adherence to underlying physical laws.

4 Experiment

In this section, we evaluate the proposed method on a variety of real-world datasets through classification tasks.

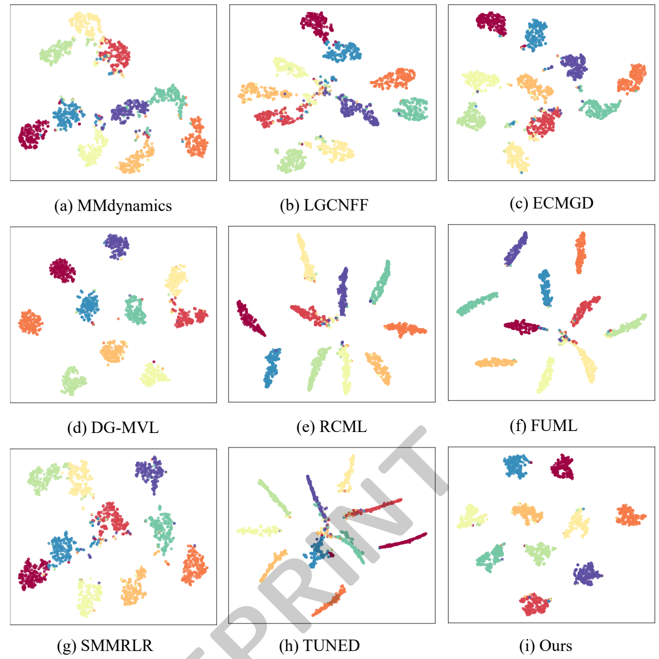
	Text-attributed graph			
Method	**Citeseer**	**Cora**	**Cornell**	**Texas**
MMdynamics	59.50 (0.00)	60.64 (0.46)	55.68 (0.00)	56.55 (0.00)
LGCNFF	37.66 (2.22)	55.94 (2.88)	49.55 (3.22)	62.50 (2.10)
ECMGD	66.96 (0.34)	72.68 (0.58)	**67.05 (2.30)**	68.10 (1.23)
DG-MVL	65.38 (0.81)	72.67 (1.19)	43.31 (13.10)	69.67 (2.97)
RCML	57.46 (0.04)	70.75 (0.10)	64.43 (0.99)	60.95 (0.61)
FUML	62.45 (1.74)	69.32 (3.93)	48.07 (6.69)	61.18 (8.07)
SMMRL	54.81 (7.50)	71.12 (1.10)	58.52 (2.67)	63.69 (4.58)
TUNED	63.13 (0.58)	67.08 (0.33)	50.57 (1.02)	56.21 (0.00)
mcGCN-*F*	**70.95 (0.50)**	**77.82 (0.38)**	65.23 (2.26)	**73.57 (0.71)**
	Multi-view text dataset			
Method	**20Newsgroups**	**3Sources**	**BBCNews**	**BBCSports**
MMdynamics	94.44 (0.00)	63.82 (0.00)	76.34 (0.00)	83.14 (0.24)
LGCNFF	78.18 (5.04)	83.03 (0.97)	76.95 (5.07)	87.45 (3.09)
ECMGD	83.11 (4.42)	63.16 (0.72)	75.98 (2.97)	80.94 (2.27)
DG-MVL	57.75 (17.86)	55.91 (21.40)	72.71 (10.87)	80.09 (12.22)
RCML	81.47 (0.64)	63.29 (0.26)	71.18 (0.06)	78.98 (0.20)
FUML	59.91 (7.40)	49.15 (10.96)	74.72 (4.06)	72.90 (9.54)
SMMRL	88.36 (2.17)	62.24 (1.53)	77.63 (1.69)	86.31 (2.33)
TUNED	42.18 (0.43)	67.32 (0.00)	79.19 (1.39)	73.14 (0.10)
mcGCN-*F*	**96.18 (0.17)**	**83.55 (0.42)**	**89.72 (0.08)**	**98.37 (0.00)**
	Multi-feature image datasets			
Method	**Animals**	**Caltech7**	**Handwritten**	**SUNRGBD**
MMdynamics	72.96 (0.15)	87.33 (0.15)	92.36 (0.07)	35.28 (0.07)
LGCNFF	64.30 (3.41)	91.73 (0.70)	91.80 (0.23)	28.06 (3.90)
ECMGD	80.91 (0.17)	95.23 (0.17)	94.43 (0.24)	**50.81 (0.33)**
DG-MVL	81.74 (0.43)	92.42 (1.81)	96.86 (0.30)	45.87 (0.62)
RCML	81.92 (0.04)	89.62 (1.15)	95.99 (0.15)	33.78 (0.11)
FUML	76.30 (1.58)	89.83 (1.93)	95.87 (0.33)	45.97 (0.23)
SMMRL	78.03 (1.16)	86.76 (2.44)	92.42 (0.52)	19.35 (3.14)
TUNED	81.57 (0.04)	92.28 (0.38)	91.54 (0.62)	48.78 (0.08)
mcGCN-*F*	**83.72 (1.42)**	**95.35 (0.22)**	**97.27 (0.19)**	49.79 (0.82)

Table 1: ACC (%) performance on all datasets, where the best and the second-best results are highlighted in **boldface** and underlined respectively.

4.1 Experimental Setup

Datasets and Compared Methods. To comprehensively evaluate the effectiveness and robustness of the proposed method, we conduct experiments on a diverse collection of benchmark datasets, covering text-attributed graph, multi-view text datasets, and visual multi-feature image datasets. The training set and the test set are split in a ratio of 10%/90%. And we conduct comparisons with state of the art methods, including MMdynamics [Han *et al.*, 2022], LGCNFF [Chen *et al.*, 2023], ECMGD [Lu *et al.*, 2024a], DG-MVL [Lu *et al.*, 2024b], RCML [Xu *et al.*, 2024], FUML [Duan *et al.*, 2025], SMMRL [Fang *et al.*, 2025], TUNED [Huang *et al.*, 2025].

Implementation Details. Following the proposed mcGCN, we model multi-view feature propagation via Laplacian-based diffusion and discretize the continuous dynamics using Euler schemes. Specifically, for each view v, a k-nearest-neighbor (kNN) graph is constructed from the input features to obtain the adjacency matrix $\mathbf{A}^v$, from which the normalized graph Laplacian $\mathbf{L}^v$ is derived. In all experiments, we adopt the Forward Euler scheme (mcGCN-F), which results in the following iterative propagation rule: $\mathbf{H}^{(k+1)} = (\mathbf{I} - \alpha\mathbf{L})\mathbf{H}^{(k)}, \quad k = 0, \ldots, K-1$, where α denotes the diffusion strength controlling the intensity of neighborhood propagation, and K specifies the propagation depth. Unless otherwise specified, the hyperparameters are set to $k = 10$, $K = 3$, and $\alpha = 3$.



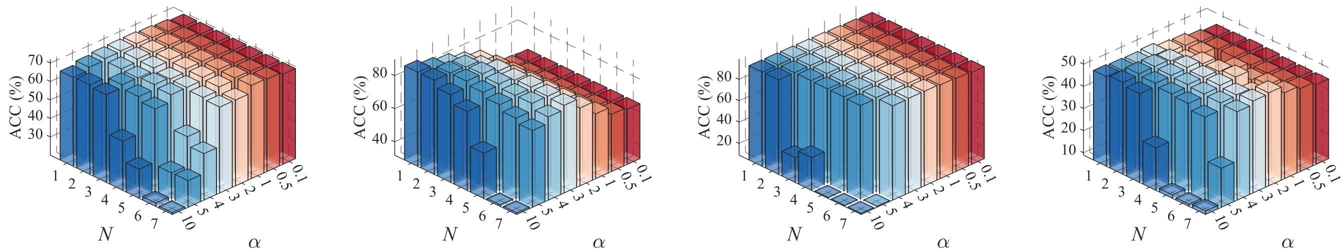
(a) MMdynamics (b) LGCNFF (c) ECMGD (d) DG-MVL (e) RCML (f) FUML (g) SMMRLR (h) TUNED (i) Ours

Figure 2: t-SNE visualization of the comparative models and the proposed method on the Handwritten dataset.

4.2 Comparison with State-of-the-Art Methods

Quantative Results. Table 1 reports the ACC results on all multi-view datasets. On citation and web graph datasets, which are typically sparse and noisy, performance critically depends on stable neighborhood propagation without excessive smoothing. Methods such as LGCNFF and DG-MVL either rely on learnable fusion or feature disentanglement, and thus exhibit weaker or unstable behavior when structural information dominates. For multi-view text and image datasets, where different channels are complementary but often corrupted by noise, uncertainty-aware methods (*e.g.*, MMdynamics, RCML, FUML) mainly operate at the decision level and do not explicitly couple multi-channel interactions with graph-based propagation. In contrast, by interpreting feature propagation as the numerical solution of a continuous-time dynamical system, our method provides explicit control over propagation strength and depth, resulting in more stable representations.

Qualitative Results. As shown in Figure 2, the proposed method produces representation with clearly separated clusters and compact intra-class distributions, indicating strong discriminative capability. In particular, the clusters produced by mcGCN show clearer margins and less boundary mixing, whereas several baselines exhibit either partial overlap among neighboring classes (e.g., MMdynamics, LGCNFF, and SMMRLR) or elongated structures with larger intra-class spread (e.g., RCML and TUNED). Overall, the proposed method consistently achieves desired performance across almost all benchmarks, demonstrating the effectiveness and robustness of the unified multi-channel graph convolution



(a) Citeseer (b) 20Newsgroups (c) Handwritten (d) SUNRGBD

Figure 3: Sensitivity analysis of α and N in the proposed method on four multi-view datasets.

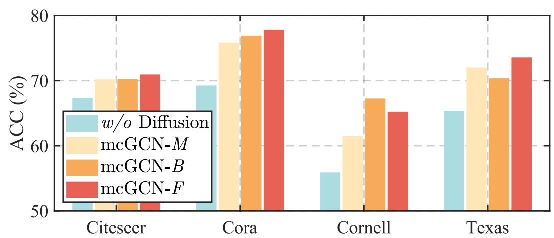
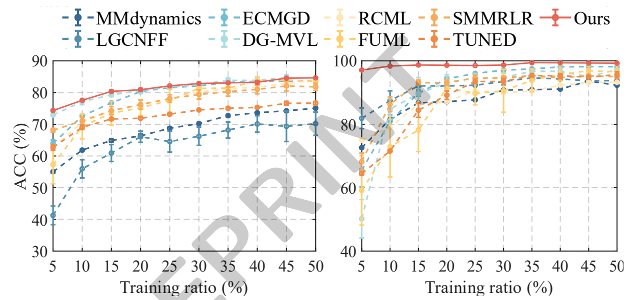


Figure 4: The performance of the proposed method on text-attributed graph under varied Euler schemes.



(a) Cora (b) BBCSports

Figure 5: The performance of the proposed method under varying training ratios on Cora and BBCSports datasets.

framework.

4.3 Model Analysis

Effect of Diffusion Solver. Figure 4 reports the ablation results on graph datasets. *w/o* Diffusion removes PDE-driven propagation and only keeps the view-wise encoders and a fusion head. The consistent performance gap to our full model indicates that the gain mainly comes from the proposed diffusion-based multi-channel propagation rather than the encoder capacity. Among the diffusion strategies, the Forward Euler variant (mcGCN-*F*) consistently yields the best performance, while Backward Euler (mcGCN-*B*) and Modified Euler (mcGCN-*M*) show slightly inferior but still competitive results.

Effect of Diffusion Strength and Propagation Depth. Figure 3 studies the influence of diffusion strength α and propagation depth K. A larger α enhances neighborhood aggregation within each step and accelerates information propagation when K is small. On high-dimensional multi-view text datasets, moderate to relatively large α with shallow or moderate depths consistently yields strong performance, indicating effective cross-view information exchange. Graph datasets are more sensitive to these parameters: large α or deep propagation quickly leads to over-smoothing due to sparse connectivity. For multi-feature image datasets, moderate α balances noise suppression and discriminative feature preservation, while excessive diffusion introduces cross-class interference. Overall, a moderate diffusion strength with shallow propagation provides a robust setting across datasets.

Effect of Training Ratio. Figure 5 illustrates the performance under varying training ratios on two datasets. On the citation network Cora, our method consistently achieves higher accuracy across all training ratios, with particularly clear advantages in low-label regimes. On BBCSports, where views are semantically well-aligned and labels are relatively sufficient, all methods benefit from increased training data, yet our approach consistently remains on the upper performance envelope. This demonstrates that the unified multi-channel evolution not only improves robustness under scarce supervision but also preserves discriminative representations when sufficient labels are available.

5 Conclusion and Future Work

This paper presented a novel perspective on multi-channel signal fusion by formulating it as a continuous evolution process governed by partial differential equations. We theoretically demonstrated that our proposed mcGCN serves as a generalizable parent framework, from which various discrete GNN paradigms can be derived via specific discretization schemes. Empirical evaluations on real-world datasets validated that the theoretical unification translates into superior performance, particularly in distinguishing complex topological structures under limited supervision. Currently, our framework primarily focuses on diffusive dynamics. In future work, we aim to incorporate adaptive feedback control mechanisms into the continuous evolution process. This integration seeks to impart intrinsic robustness against heterogeneous sensor noise and adversarial perturbations, while enhancing decision-making transparency through physically interpretable stability guarantees in safety-critical applications.

Acknowledgements

This paper is partly supported by the National Natural Science Foundation of China (U24A20260), Hainan Province Science and Technology Special Fund (ZDYF2024GXJS003), the Fujian Provincial Natural Science Foundation of China under Grant 2025J011036.

Contribution Statement

Na Song and Zihan Fang contribute equally to this work. Weidong Zhang is the Corresponding author (wdzhang@sjtu.edu.cn).

References

[Chen *et al.*, 2023] Zhaoliang Chen, Lele Fu, Jie Yao, Wenzhong Guo, Claudia Plant, and Shiping Wang. Learnable graph convolutional network and feature fusion for multi-view learning. *Information Fusion*, 95:109–119, 2023.

[Dong and Ansari, 2020] Mianxiong Dong and Nirwan Ansari. Guest editorial: special section on cyber-physical social systems—integrating human into computing. *IEEE Transactions on Emerging Topics in Computing*, 8:4–5, 2020.

[Duan *et al.*, 2025] Siyuan Duan, Yuan Sun, Dezhong Peng, Guiduo Duan, Xi Peng, and Peng Hu. Deep fuzzy multi-view learning for reliable classification. In *Proceedings of the Forty-second International Conference on Machine Learning*, pages 1–11, 2025.

[Fang *et al.*, 2025] Zihan Fang, Ying Zou, Shiyang Lan, Shide Du, Yanchao Tan, and Shiping Wang. Scalable multi-modal representation learning networks. *Artificial Intelligence Review*, 58(7):209, 2025.

[Hammond *et al.*, 2011] David K. Hammond, Pierre Vandergheynst, and Rémi Gribonval. Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, 30(2):129–150, 2011.

[Han *et al.*, 2022] Zongbo Han, Fan Yang, Junzhou Huang, Changqing Zhang, and Jianhua Yao. Multimodal dynamics: Dynamical fusion for trustworthy multimodal classification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20707–20717, 2022.

[He *et al.*, 2021] Mingguo He, Zhewei Wei, Zengfeng Huang, and Hongteng Xu. Bernnet: Learning arbitrary graph spectral filters via bernstein approximation. In *Proceedings of the Advances in Neural Information Processing Systems*, pages 14239–14251, 2021.

[He *et al.*, 2026] Jin He, Wei Wang, Zhijian Gou, Xiaozhao Jin, Gexiang Zhang, and Mianxiong Dong. A dual-branch fusion network based on convolutional neural network and mamba for outdoor fire detection. *Engineering Applications of Artificial Intelligence*, 167:113838, 2026.

[Huang *et al.*, 2025] Haojian Huang, Chuanyu Qin, Zhe Liu, Kaijing Ma, Jin Chen, Han Fang, Chao Ban, Hao Sun, and Zhongjiang He. Trusted unified feature-neighborhood dynamics for multi-view classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 17413–17421, 2025.

[Jia *et al.*, 2024a] Zehua Jia, Haibo Lu, Hongtian Chen, and Weidong Zhang. Robust distributed cooperative rendezvous control for heterogeneous marine vehicles using model predictive control. *IEEE Transactions on Vehicular Technology*, 73:11002–11013, 2024.

[Jia *et al.*, 2024b] Zehua Jia, Kunwu Zhang, Yang Shi, and Weidong Zhang. Safety-preserving lyapunov-based model predictive rendezvous control for heterogeneous marine vehicles subject to external disturbances. *IEEE Transactions on Cybernetics*, 54:5244–5256, 2024.

[Khemani *et al.*, 2024] Bharti Khemani, Shruti Patil, Ketan Kotecha, and Sudeep Tanwar. A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions. *Journal of Big Data*, 11:18, 2024.

[Kipf and Welling, 2017] Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *Proceedings of the 5th International Conference on Learning Representations*, pages 1–14, 2017.

[Klicpera *et al.*, 2019] Johannes Klicpera, Aleksandar Bojchevski, and Stephan Günnemann. Predict then propagate: Graph neural networks meet personalized pagerank. In *Proceedings of the 7th International Conference on Learning Representations*, pages 1–15, 2019.

[Kou *et al.*, 2024] Zhiqiang Kou, Jing Wang, Jiawei Tang, Yuheng Jia, Boyu Shi, and Xin Geng. Exploiting multi-label correlation in label distribution learning. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, pages 4326–4334, 2024.

[Kou *et al.*, 2025] Zhiqiang Kou, Si Qin, Hailin Wang, Jing Wang, Mingkun Xie, Shuo Chen, Yuheng Jia, Tongliang Liu, Masashi Sugiyama, and Xin Geng. Label distribution learning with biased annotations assisted by multi-label learning. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, pages 16–22, 2025.

[Kou *et al.*, 2026] Zhiqiang Kou, Junxiang Wu, Wenke Huang, Wenwen He, Ming-Kun Xie, Changwei Wang, Yuheng Jia, Di Jiang, Yang Liu, Xin Geng, and Qiang Yang. Fedharmony: Harmonizing heterogeneous label correlations in federated multi-label learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2026.

[Li Can, 2025] Zhang Xiaoxuan Pan Quan Li Can, Wang Zengfu. Land-sea clutter classification method based on multi-channel graph convolutional networks. *Journal of Radars*, 14:322–337, 2025.

[Lu *et al.*, 2024a] Jielong Lu, Zhihao Wu, Zhaoliang Chen, Zhiling Cai, and Shiping Wang. Towards multi-view consistent graph diffusion. In *Proceedings of the 32nd ACM International Conference on Multimedia*, pages 186–195, 2024.

[Lu *et al.*, 2024b] Mingfei Lu, Qi Zhang, and Badong Chen. Divergence-guided disentanglement of view-common and view-unique representations for multi-view data. *Information Fusion*, 114:102661, 2024.

[Song *et al.*, 2025] Na Song, Laurence T. Yang, Jing Yang, Xiangli Yang, Jieming Yang, Shundong Yang, and Shiping Wang. Enhancing interpretability of graph convolutional networks for multi-view learning. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 2025.

[Song *et al.*, 2026] Na Song, Zhenheng Lin, Shiyang Lan, Zhongyue Lei, and Shiping Wang. High-order relation driven multi-view representation learning. *Knowledge-Based Systems*, page 116028, 2026.

[Sun *et al.*, 2024] Yuan Sun, Yang Qin, Yongxiang Li, Dezhong Peng, Xi Peng, and Peng Hu. Robust multi-view clustering with noisy correspondence. *IEEE Transactions on Knowledge and Data Engineering*, 36:9150–9162, 2024.

[Sun *et al.*, 2025] Yuan Sun, Yongxiang Li, Zhenwen Ren, Guiduo Duan, Dezhong Peng, and Peng Hu. Roll: Robust noisy pseudo-label learning for multi-view clustering with noisy correspondence. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 30732–30741, 2025.

[Wang and Zhang, 2022] Xiyuan Wang and Muhan Zhang. How powerful are spectral graph neural networks. In *Proceedings of the International Conference on Machine Learning*, pages 23341–23362, 2022.

[Wang *et al.*, 2020] Xiao Wang, Meiqi Zhu, Deyu Bo, Peng Cui, Chuan Shi, and Jian Pei. Am-gcn: Adaptive multi-channel graph convolutional networks. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1243–1253, 2020.

[Wang *et al.*, 2021] Yifei Wang, Yisen Wang, Jiansheng Yang, and Zhouchen Lin. Dissecting the diffusion process in linear graph convolutional networks. In *Advances in Neural Information Processing Systems*, pages 5758–5769, 2021.

[Wang *et al.*, 2026] Jianghua Wang, Zhicheng Wei, Zhibin Shi, Na Song, and Shiping Wang. Robustness controlled adversarial graph convolutional network for multi-view semi-supervised classification. *Knowledge-Based Systems*, page 115789, 2026.

[Wu *et al.*, 2019] Felix Wu, Amauri H. Souza Jr., Tianyi Zhang, Christopher Fifty, Tao Yu, and Kilian Q. Weinberger. Simplifying graph convolutional networks. In *Proceedings of the 36th International Conference on Machine Learning*, pages 6861–6871, 2019.

[Wu *et al.*, 2022] Wentao Wu, Zhouhua Peng, Lu Liu, and Dan Wang. A general safety-certified cooperative control architecture for interconnected intelligent surface vehicles with applications to vessel train. *IEEE Transactions on Intelligent Vehicles*, 7:627–637, 2022.

[Xu *et al.*, 2019] Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks? In *Proceedings of the 7th International Conference on Learning Representations*, pages 1–13, 2019.

[Xu *et al.*, 2024] Cai Xu, Jiajun Si, Ziyu Guan, Wei Zhao, Yue Wu, and Xiyue Gao. Reliable conflictive multi-view learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 16129–16137, 2024.

[Yuan *et al.*, 2024] Honglin Yuan, Shiyun Lai, Xingfeng Li, Jian Dai, Yuan Sun, and Zhenwen Ren. Robust prototype completion for incomplete multi-view clustering. In *Proceedings of the 32nd ACM international conference on multimedia*, pages 10402–10411, 2024.

[Yuan *et al.*, 2025] Honglin Yuan, Yuan Sun, Fei Zhou, Jing Wen, Shihua Yuan, Xiaojian You, and Zhenwen Ren. Prototype matching learning for incomplete multi-view clustering. *IEEE Transactions on Image Processing*, 34:828–841, 2025.