

Article

Not peer-reviewed version

Less for Better: A View Filter-Driven Graph Representation Fusion Network

[Yue Wang](#), [Xibei Yang](#), [Keyu Liu](#), [Qihang Guo](#), [Xun Wang](#)*

Posted Date: 1 December 2025

doi: 10.20944/preprints202512.0016.v1

Keywords: multi-view learning; graph representation fusion; graph neural networks; graph entropy



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Less for Better: A View Filter-Driven Graph Representation Fusion Network

Yue Wang ¹ , Xibei Yang ¹, Keyu Liu ¹, Qihang Guo ² and Xun Wang ^{1,*}

¹ School of Computer, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu, 212003, China

² School of Economics and Management, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu, 212100, China

* Correspondence: wangxun@just.edu.cn

Abstract

Multi-view learning has recently gained considerable attention in graph representation learning, as it enables the fusion of complementary information from multiple views to enhance representation quality. However, most existing studies neglect that irrelevant views may introduce noise and negatively affect representation quality. To address the issue, we propose a novel multi-view representation learning framework called a View Filter-driven graph representation fusion network named ViFi. Following the "less for better" principle, the framework focuses on filtering informative views while discarding irrelevant ones. Specifically, entropy-based adaptive view filter is designed to dynamically filter the most informative views by evaluating their feature-topology entropy characteristics, aiming to not only reduce irrelevance among views but also enhance their complementarity. In addition, to promote more effective fusion of informative views, we propose an optimized fusion mechanism that leverages the filtered views to identify the optimal integration strategy using a novel information gain function. Through extensive experiments on classification and clustering tasks, ViFi demonstrates clear performance advantages over existing state-of-the-art approaches.

Keywords: multi-view learning; graph representation fusion; graph neural networks; graph entropy

1. Introduction

Graph-structured data increasingly exhibit multi-view characteristics with the development of graph representation learning. For instance, social networks encompass multiple views, including user profiles, interaction histories, and content preferences. Multi-view learning is crucial as it integrates diverse perspectives, enhancing model performance on downstream tasks and enabling more reliable analysis of complex data [1,2]. Multi-view learning has witnessed significant advances across diverse domains, including computer vision [3], natural language processing [4], and bioinformatics [5].

Recently, Owing to the powerful capability of Graph Neural Networks (GNNs) in capturing complex structural relationships and semantic information, the collaboration between multi-view learning and GNNs facilitates the extraction of view-specific representations and the discovery of inter-view correlations [6,7]. This collaboration enables multi-view GNNs to fully exploit complementary information distributed across views, leading to their promising performance in various real-world applications [8,9]. For example, in recommendation systems, multi-view GNNs fuse heterogeneous user-item interactions, achieving excellent recommendation accuracy.

Currently, to enhance the effectiveness of multi-view fusion in representation learning, existing approaches typically employ attention [10], gating [11], alignment [12], or fusion modules [13] to jointly integrate information from all views. These methods assign weights or scores to each view and aggregate views accordingly, allowing models to emphasize informative views while suppressing irrelevant ones. However, these approaches aggregate all views based on the assumption that all views contribute positively to the learning process. We argue that this assumption is not always valid, thereby we conduct a preliminary experiment to examine its validity. To assess the contribution of

each view, we measured its information content using information entropy [14]. A greedy strategy was applied, where the raw view (Raw) was sequentially integrated with other views (V1, V2, V3) in descending order of entropy. As illustrated in Figure 1, the classification accuracy initially increases but subsequently declines as more views are added, indicating that irrelevant views may interfere with model performance. Therefore, it is vital to design an effective mechanism to identify and retain only the views that contribute positively to representation, while filtering out irrelevant views that contribute negatively.

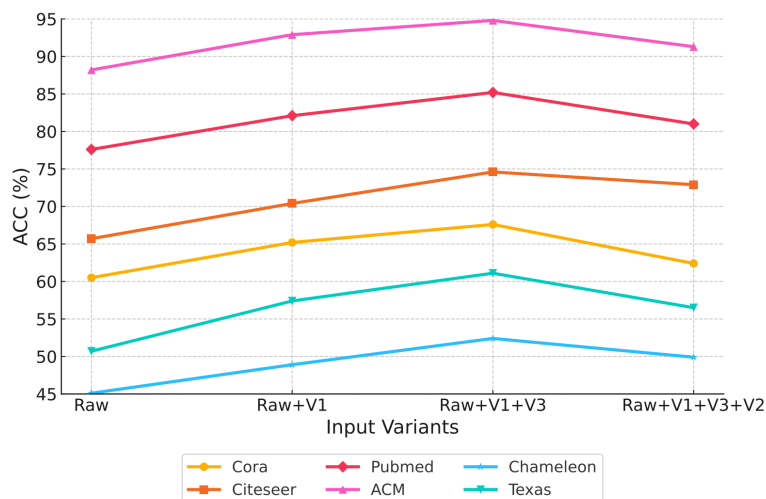


Figure 1. Validating the necessity of view filtering. As the number of views increases, the classification accuracy (ACC%) reaches a peak and subsequently declines, indicating that not all views contribute positively to model performance.

To address the above issue, we propose a novel multi-view representation learning framework called a View filter-driven graph representation fusion network named ViFi. We advocate a "less for better" principle, aiming to obtain superior representations by leveraging fewer yet more informative views.

Initially, entropy-based adaptive view filter is designed to identify and retain only the most informative views in a multi-view learning system, filtering out irrelevant views. The filter quantifies each view's information content through feature-topology entropy characteristics [15], which effectively reflect the uncertainty and diversity of feature and topology distributions, serving as an indicator of a view's information richness. By maximizing an entropy-based objective function, the module adaptively determines the optimal number of views to retain, filtering most informative views while discarding irrelevant views. The objective is formulated to reward views with substantial feature-topology entropy while penalizing irrelevance, thereby promoting a compact subset that preserves maximal informational diversity. The filter further stabilizes the subsequent fusion process by providing a compact and informative set of views.

In addition, to promote more effective fusion of informative views, we propose an optimized fusion mechanism that is introduced after filtering to adaptively determine the optimal integration strategy which yields the most informative representations. A novel information gain function is proposed to evaluate candidate view groupings based on entropy balance and structural complementarity. Entropy balance is achieved by normalizing the entropy of each view matrix and computing an equilibrium degree, which encourages uniform information contribution and avoids interference in fusion. Structural complementarity is enforced by normalizing structural differences and defining an information-gain term that favors the grouping of view matrices exhibiting significant topological diversity. By selecting the subset that maximizes overall group gain as optimal integration strategy, this module achieves balanced and complementary fusion, thereby facilitating effective collaboration among views.

To conclude, the main contributions of this study are presented as follows:

(1) A novel multi-view representation learning framework is proposed, which systemically combines view filtering with optimized fusion to produce compact, informative multi-view graph representations.

(2) An entropy-based adaptive view filter is developed, which evaluates view contribution through feature–topology entropy and dynamically filters the most informative views to reduce irrelevance and enhance complementarity.

(3) A novel information gain function is designed to evaluate the contribution of different view integration strategies and guide the selection of an optimal integration strategy that achieves entropy balance and structural complementarity, thereby strengthening inter-view collaboration.

(4) Based on comprehensive experiments on classification and clustering tasks, the proposed method consistently achieves superior performance over existing state-of-the-art approaches.

2. Related work

2.1. Multi-View Learning

Multi-view learning exploits complementary information from different data perspectives, whose integration enables more comprehensive representations. Traditional multi-view learning methods can be divided into three major categories [16]: co-training style algorithms [17,18], co-regularization style algorithms [19] and margin-consistency style algorithms [20,21]. For instance, Qiao et al. [22] proposed Deep Co-Training, which trains multiple networks as distinct views under adversarial perturbations to maintain complementary and diverse view representations. Wang et al. [23] introduced Deep Canonically Correlated Autoencoders (DCCA), combining CCA with autoencoder reconstruction to co-regularize representations across views. The hybrid objective offers a flexible way to align and denoise multi-view features. Mao et al. [24] proposed Soft-Margin-Consistency Multi-View MED (SMVMED), relaxing hard margin-equality constraints into a soft consistency principle to improve scalability while preserving discriminative power. However, these studies generally integrate all available views, overlooking the potential interference of irrelevant views in graph representation learning. Inspired by these works, we attempt to design a view filtering mechanism that filtering most informative views while discarding irrelevant views.

2.2. Multi-View Representation Fusion

As a critical component of multi-view learning, multi-view representation fusion is not a novel topic, and it has been extensively explored in a range of existing studies. Based on the time of multi-view fusion during the learning process, traditional works can be divided into two primary types [25]. (1) Early fusion integrates features from multiple views at the feature level before model training through typical methods such as feature concatenation, pooling-based integration and convolutional fusion. For example, Kachole et al. [26] concatenated RGB and event-stream features at the feature level to construct a unified multi-view representation, which effectively enhanced view complementarity and improved segmentation accuracy. Wei et al. [27] aggregates per-view CNN descriptors by adaptive view pooling over a learned view-graph to obtain a unified global shape descriptor. Liang et al. [28] employs 1×1 convolutions for channel-wise fusion of LiDAR and camera features before joint detection, significantly improving multi-view 3D detection accuracy. (2) Late fusion aggregates the predictions from different views after training separate models for each view through methods such as score averaging, weighted voting, and meta-classifier stacking. For instance, Simonyan et al. [29] averaged softmax scores from spatial and temporal streams to obtain a single decision, which consistently outperformed either stream alone on multi-view action recognition. Su et al. [30] used weighted voting to aggregate the prediction scores from multiple views, enabling a more robust fusion that improved the reliability of biomarker identification. Wang et al. [31] employed a meta-classifier to integrate the cross-validated outputs of view-specific CNN classifiers, enabling stacking-based fusion that better exploited heterogeneous cues and enhanced overall prediction accuracy.

Building upon the above fusion strategies, LoGo-GNN [32] can be regarded as a hybrid fusion framework that integrates both early and late fusion principles through a local-to-global architecture. In this framework, the local module performs fusion by integrating multiple view-specific representations through fixed aggregation pathways, where neighborhood information from different local views is jointly embedded to capture interrelated patterns within each view. The global module further aligns these aggregated representations across views to ensure overall consistency. However, its architecture relies on fixed local view strategy, limiting its flexibility in adapting to heterogeneous graph structures and diverse view relationships. In our study, we inspired from LoGo-GNN and extends its local-global architecture by designing an optimised fusion mechanism that dynamically selects optimal integration strategy.

3. ViFi

3.1. Notation

We consider a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with adjacency matrix \mathbf{A} and node-feature matrix \mathbf{X} . The i -th graph view is denoted by M_i , and S denotes the selected subset of views with $k = |S|$ the number of selected high-quality views; T_3 is the minimum number of views required. For each view M_i , $H_f(M_i)$ and $H_s(M_i)$ denote its feature entropy and topology entropy, respectively; their combination yields the view information score $I(M_i)$. A subset S is evaluated by the subset score $F(S)$. To characterize complementarity between views, we use the entropy balance B_{ij} and the normalized structural difference Δ_{ij} , whose product defines the pairwise gain $G(M_i, M_j)$; aggregating pairwise gains over S gives the group gain $G(S)$. $g(\cdot)$ denotes a GNN encoder. Table 1 summarizes the notations used in this paper.

Table 1. Notation description.

Notation	Description
\mathcal{G}	A graph
\mathcal{V}	Node set
\mathcal{E}	Edge set
\mathbf{A}	Adjacency matrix
\mathbf{X}	Node feature matrix
M_i	i -th graph view
S	Selected subset of views
$H_f(M_i)$	Feature entropy of view M_i
$H_s(M_i)$	Topology entropy of view M_i
$I(M_i)$	View information score
k	Order of optimal integration strategy
$F(S)$	Subset score
T_3	Minimum number of views
B_{ij}	Entropy balance
Δ_{ij}	Normalized structural difference
$G(M_i, M_j)$	Pairwise gain
$G(S)$	Group gain
$g(\cdot)$	GNN encoder

3.2. Framework Overview

The overall structure of ViFi is illustrated in Figure 2. It mainly includes two modules. (1) Entropy-based adaptive view filter evaluates each view based on feature and topological entropy characteristics to adaptively filter those with the highest informational utility, thereby establishing a compact and diverse set of informative views. (2) Optimized fusion mechanism employs a novel information gain function, which is constructed from entropy balance and normalized structural difference across views. It effectively fuses the representations by dynamically selecting optimal view integration strategy that maximize the group gain.

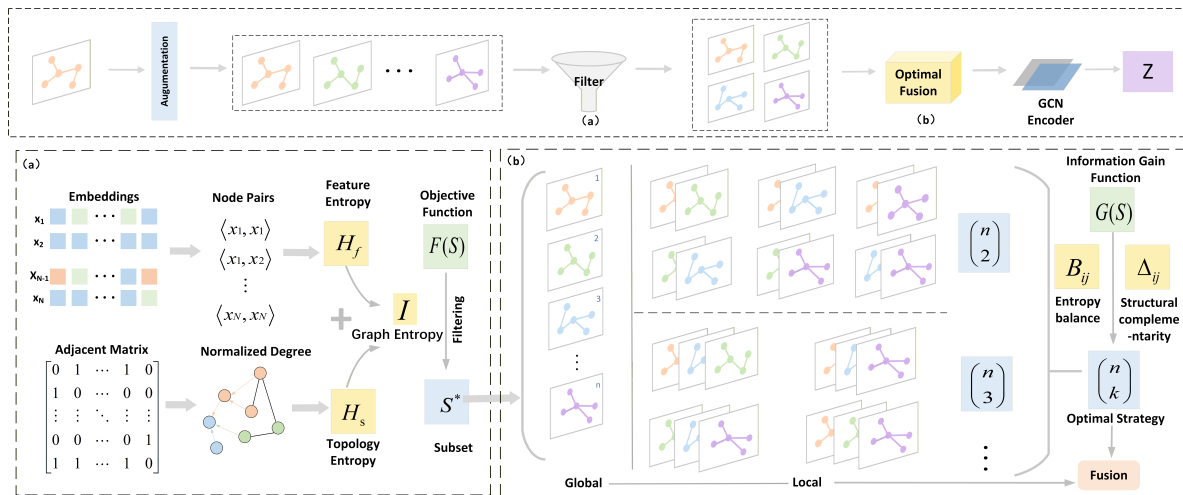


Figure 2. The framework of ViFi, where modules (a) and (b) represent the core components.

3.3. Entropy-Based Adaptive View Filter

Multi-view learning often contains irrelevant views. For example, in recommendation graphs, views derived from clicks, purchases, and wishlists may overlap in signal, and sparsity varies across them. In molecular property prediction, 2D topology, 3D conformers, and substructure fingerprints provide complementary but uneven information. Passing all views downstream raises computation and risks over-fitting to noisy or irrelevant signals. Therefore, entropy-based adaptive view filter is designed to address this issue by ranking and filtering only informative views and by adapting the number of retained views to the data.

We quantify the information content in each view by feature–topology entropy characteristics [15]: feature entropy and topology entropy. Feature entropy reflects uncertainty in node attributes under locality assumptions. Topology entropy captures higher-order topology using normalized structural statistics. Entropy is an appropriate metric here because it summarizes distributional uncertainty in a model-agnostic way and requires no task labels. Let a view $M_i = (A_i, X_i)$. We define:

$$H_f(M_i) = -\sum_i p_i^{(f)} \log p_i^{(f)}, \quad H_s(M_i) = -\sum_i p_i^{(s)} \log p_i^{(s)}, \quad (1)$$

where $p^{(f)}$ and $p^{(s)}$ are normalized distributions induced by features and structure following Luo et al [15]. We combine them as:

$$I(M_i) = \alpha H_f(M_i) + (1 - \alpha) H_s(M_i), \quad (2)$$

where $\alpha \in [0, 1]$, and use $I(M_i)$ only to drive the filter. Higher scores indicate that the view contains richer information and greater complementarity.

We now use entropy to filter a subset of views and to adaptively decide how many to keep. The key idea is to retain views with high information content while preventing the inclusion of an excessively large set that would reintroduce redundancy and increase computational cost. We score any candidate subset S by:

$$F(S) = \frac{1}{|S|} \sum_{M_i \in S} I(M_i) - \mu g(|S|), \quad |S| \geq T_3. \quad (3)$$

Here S denotes the filtered set of candidate views. $I(M_i)$ is the per-view information score defined above. The parameter $\mu > 0$ controls the trade-off between information and complexity. The function $g(\cdot)$ is an increasing size penalty, with common choices $g(n) = \max(n - T_3, 0)$ or $g(n) = n$. The constant $T_3 = 3$ imposes a minimum of three views to preserve basic multi-view complementarity.

Maximizing $F(S)$ yields a filter that keeps high-information views and discards redundant ones while automatically determining how many views to retain:

$$S^* = \arg \max_S F(S). \quad (4)$$

This design makes the filter adaptive, reduces computation, and stabilizes downstream fusion by avoiding noisy or overlapping inputs.

Entropy-based adaptive filter uses feature–topology entropy characteristics to quantify view information, then maximizes $F(S)$ to select an adaptive-size subset. The module filters out irrelevant views. It preserves complementary signal and yields a compact, informative set with less redundancy. This not only improves efficiency and robustness but also reduces the burden on the subsequent fusion stage.

Algorithm 1: Entropy-based adaptive view filter

Input: Graph views $\{M_1, M_2, \dots, M_n\}$, where each $M_i = (A_i, X_i)$; trade-off parameter $\alpha \in [0, 1]$; penalty coefficient $\mu > 0$; minimum view count T_3 ; size penalty function $g(\cdot)$.

Output: Filtered subset of views S^* .

```

1:   for each view  $M_i$  do
2:     Compute feature entropy  $H_f(M_i)$  and topology entropy  $H_s(M_i)$  via Eq.1;
3:     Calculate feature–topology entropy as view content information via Eq.2;
4:   end for
5:   Initialize: candidate subset  $S$  with  $|S| \geq T_3$ ;
6:   Compute the subset score  $F(S)$  by combining the average information score of views in  $S$  with
   a penalty term based on the subset size, as defined in Eq.3;
7:   Select the most informative subset  $S^*$  by maximizing the subset score  $F(S)$  over all candidate
   subsets, following the process in Eq.4;
8:   return  $S^*$ .

```

3.4. Optimized Fusion Mechanism

Furthermore, to enable more effective fusion of informative views, we introduce a fusion mechanism based on a local-to-global framework, which promotes hierarchical integration of view-specific information and enhances overall representation consistency. To further investigate how different integration strategies influence representation quality, we conduct a preliminary experiment that evaluates the effect of different integration strategies on node classification accuracy across multiple datasets. The results shown in Figure 3, indicate that varying integration strategies lead to significant differences in classification performance. This observation motivates the need to adaptively identify the most effective view integration strategy for fusion mechanism.

After obtaining a set of informative views, we need to decide which subset of the filtered views should be integrated so that their fused representation is maximally informative. The fusion objective is to balance the information contribution of each view and to exploit structural complementarity among them. To achieve this, we introduce a novel information gain function that scores candidate groups and selects the subset with the highest gain.

The goal of this module is to select an optimal subset S of filtered view matrices for joint fusion. Let $\{M_1, \dots, M_n\}$ represent the candidate view matrices obtained from Module 1. We group these matrices into sets of size $k \geq 2$, since fusion requires at least two views. The objective is to find the grouping that maximizes the overall fusion score. Each group should maintain balanced information entropy across views and preserve structural complementarity. In this way, the fusion process avoids dominance by a single view and reduces irrelevance among similar structures.

To quantify these ideas, we start by defining a *pairwise gain* $G(M_i, M_j)$ between any two view matrices M_i and M_j . First, let $H(M_i)$ denote the Shannon entropy of matrix M_i , interpreted as the amount of information it contains. We define two factors:

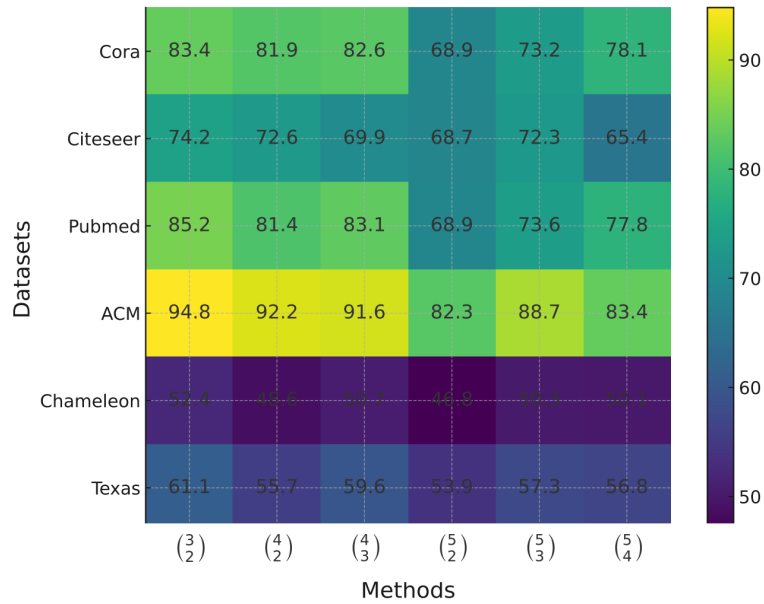


Figure 3. Heatmap of classification accuracy (ACC%) across six datasets and six view integration strategies. Brighter cells indicate higher accuracy. These patterns indicate that varying integration strategies lead to significant differences in classification performance.

1) Entropy balance:

$$B_{ij} = \frac{\min(H(M_i), H(M_j))}{\max(H(M_i), H(M_j))}, \quad (5)$$

where lies in $[0, 1]$. When $H(M_i) \approx H(M_j)$, B_{ij} is close to 1, indicating balanced information; if one entropy is much smaller than the other, B_{ij} approaches 0, reflecting imbalance.

2) Normalized structural difference:

$$\Delta_{ij} = \frac{\|M_i - M_j\|_F}{\|M_i\|_F + \|M_j\|_F}, \quad (6)$$

where $\|\cdot\|_F$ is the Frobenius norm. This measures how different the two matrices are, normalized by their magnitudes. The value $\Delta_{ij} \in [0, 1]$ is 0 if $M_i = M_j$ (no structural difference) and approaches 1 as the difference grows large relative to their norms. This normalization makes Δ_{ij} comparable across different data scales.

The pairwise gain is then defined as the product of these two factors:

$$G(M_i, M_j) = B_{ij} \times \Delta_{ij} = \frac{\min(H(M_i), H(M_j))}{\max(H(M_i), H(M_j))} \cdot \frac{\|M_i - M_j\|_F}{\|M_i\|_F + \|M_j\|_F}. \quad (7)$$

By construction, $G(M_i, M_j)$ is high only when M_i and M_j have comparable entropies (B_{ij} is near 1) and significant structural difference (Δ_{ij} is large). If either condition fails (one view dominates in entropy or the two views are nearly identical), the product ensures $G(M_i, M_j)$ is small. This multiplicative form naturally enforces a "both-or-none" threshold: both entropy balance and structural complementarity must be present for a large gain. It is worthy to note that no additional parameters are needed in this metric.

Next, we extend the pairwise gain to an entire group (subset) of views. For any subset $S \subseteq \{M_1, \dots, M_n\}$ of size $|S| = k$, we define its group gain $G(S)$ as the average of all pairwise gains within the group:

$$G(S) = \frac{1}{\binom{k}{2}} \sum_{\substack{M_i, M_j \in S \\ i < j}} G(M_i, M_j). \quad (8)$$

This average pairwise gain is the group quality score. The mean over $\binom{k}{2}$ pairs normalizes for group size. High $G(S)$ indicates complementary structure and balanced information. A weak member lowers many pair scores and reduces the average. Hence $G(S)$ is a reliable indicator of fusion value. By maximizing the overall group gain $G(S)$ to adaptively identify the view subset corresponding to the optimal integration strategy, this module achieves a balanced and complementary fusion, thereby facilitating effective collaboration among views.

Algorithm 2: Optimized fusion mechanism

Input: Filtered view set $\{M_1, M_2, \dots, M_m\}$; GNN encoder $g(\cdot)$; group size constraint $k \geq 2$;

Output: Fused representation \tilde{Y} .

```

1:   for each pair of views  $(M_i, M_j)$  do
2:     Compute entropy balance  $B_{ij}$  and normalized structural difference  $\Delta_{ij}$  via Eq.5, 6;
3:     Determine the pairwise gain  $G(M_i, M_j)$  by multiplying the entropy balance and structural
4:     difference metrics to assess complementarity via Eq.7;
5:   end for
6:   for each candidate subset  $S \subseteq \{M_1, \dots, M_m\}$  with  $|S| = k$  do
7:     Compute the group gain by averaging all pairwise gains within the subset to evaluate
8:     overall fusion potential via Eq.8 ;
9:   end for;
10:  Select the subset as optimal view integration strategy by identifying the candidate with the
11:  highest group gain value;
12:  Fuse the subset using  $g(\cdot)$  and attention mechanism to integrate their representations;
13:  Return  $\tilde{Y}$ .

```

3.5. Model Details

Our framework is compatible with a wide range of graph neural network models, imposing no architectural limitations. In this study, to keep the design concise, we adopt a standard one-layer Graph Convolutional Network [33] as the basic module used to merge the representations.

$$g(\cdot) : \mathbb{R}^{N \times F} \times \mathbb{R}^{N \times N} \rightarrow \mathbb{R}^{N \times F_h}. \quad (9)$$

Its propagation rule is:

$$g(\mathbf{X}, \mathbf{A}) = \sigma(\tilde{\mathbf{A}} \mathbf{X} \Theta), \quad (10)$$

where $\sigma(\cdot)$ is a standard nonlinearity (e.g., ReLU, softmax, sigmoid, or tanh). The symmetrically normalized adjacency is:

$$\tilde{\mathbf{A}} = \hat{\mathbf{D}}^{-\frac{1}{2}} \hat{\mathbf{A}} \hat{\mathbf{D}}^{-\frac{1}{2}}, \quad \hat{\mathbf{A}} = \mathbf{A} + \mathbf{I}, \quad \hat{\mathbf{D}} = \mathbf{D} + \mathbf{I}. \quad (11)$$

$\mathbf{X} \in \mathbb{R}^{N \times F}$ contains node features and $\mathbf{A} \in \mathbb{R}^{N \times N}$ encodes graph connectivity; adding \mathbf{I} introduces self-loops. The matrix $\Theta \in \mathbb{R}^{F \times F_h}$ maps F -dimensional inputs to F_h -dimensional fused features. Symmetric normalization scales messages by node degrees for stable training. This single layer mixes information over one hop, and any other GNN backbone can replace it without altering the fusion pipeline.

Local and global view encoders. Each view encoder is composed of a GCN-based encoder together with its corresponding set of input views. These inputs may consist of different combinations of views, selected according to the chosen integration strategy. Rather than operating on a single view as in a typical GCN, both the local and global encoders receive multiple views simultaneously and generate their representations by applying mean pooling to the aggregated features:

$$g_{\Theta} = \frac{1}{k} \sum_{\mathcal{G}_j \in \{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_k\}} \text{ReLU}(\tilde{\mathbf{A}} \mathbf{X}_j \Theta). \quad (12)$$

Attention mechanism. To merge embeddings from the local and global encoders, we use an attention module to obtain a richer semantic representation.

Given m embeddings $\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_m$, attention scores their contributions:

$$(\alpha_1, \alpha_2, \dots, \alpha_m) = \text{att}(\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_m). \quad (13)$$

Here $\alpha_i \in \mathbb{R}^{N \times 1}$ gives node-wise weights for \mathbf{H}_i . The coefficients follow $\alpha_{ij} = \text{softmax}(w_{ij})$, with $w_{ij} = \mathbf{q}_i^T \cdot \tanh(\mathbf{\Theta}_{att} \cdot (\mathbf{h}_j)^T + \mathbf{b}_i)$, where $\mathbf{q}_i \in \mathbb{R}^{F' \times 1}$ is a shared attention vector, $\mathbf{\Theta}_{att} \in \mathbb{R}^{F' \times F'}$ is the weight matrix, and $\mathbf{b}_i \in \mathbb{R}^{F' \times 1}$ is the bias.

The fused embedding is the attention-weighted sum:

$$\mathbf{H} = \alpha_1 \cdot \mathbf{H}_1 + \alpha_2 \cdot \mathbf{H}_2 + \dots + \alpha_m \cdot \mathbf{H}_m. \quad (14)$$

3.5.1. Semi-Supervised Learning Architecture

In settings where only partial labels are available, typified by node classification, the framework concludes its representation extraction stage with a single-layer GCN that functions as the terminal encoder. The fused embedding \mathbf{H} is subsequently processed by a single-layer Graph Convolutional Network, which functions as the final stage of encoding and generates the model's representations $\tilde{\mathbf{Y}}$:

$$\tilde{\mathbf{Y}} = \text{softmax}(\tilde{\mathbf{A}}\mathbf{H}\mathbf{\Theta}_{final}). \quad (15)$$

3.5.2. Unsupervised Learning Architecture

In unsupervised settings, incorporating a decoder enables the model to exploit self-supervised signals from the input graphs [34]. ViFi comprises a suite of encoders that capture complementary perspectives, whose outputs are aggregated into a unified embedding. Reconstruction is then achieved by pairing this fused representation with a set of decoders, each tasked with reconstituting a specific input view. Beyond the direct correspondence between input and output, an auxiliary decoder is incorporated to reconstruct the structural information of the original view. This reconstruction module is built from multiple two-layer components, with each layer crafted to serve as an approximate inverse of the encoder it mirrors.

The decoding stage produces two reconstructions: the node-feature matrix $\tilde{\mathbf{X}}_i$ and the recovered view topology $\tilde{\mathbf{A}}_{re}$. For the decoder associated with the i -th input view, the reconstructed node attributes are computed as:

$$\tilde{\mathbf{X}}_i^{re} = \text{softmax}(\tilde{\mathbf{A}} \text{ReLU}(\tilde{\mathbf{A}}\mathbf{H}\mathbf{\Theta}^{(0)}\theta'_i)\mathbf{\Theta}^{(1)}\theta'_i), \quad (16)$$

where $\mathbf{\Theta}^{(l)}\theta'_i$ represents the learnable weights associated with the i -th view-specific reconstruction decoder at layer l .

The output of decoder is:

$$\tilde{\mathbf{H}} = \tilde{\mathbf{A}} \text{ReLU}(\tilde{\mathbf{A}}\mathbf{H}\mathbf{\Theta}_{\theta'}^{(0)})\mathbf{\Theta}_{\theta'}^{(1)}, \quad \tilde{\mathbf{A}}^{re} = \text{sigmoid}(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^T). \quad (17)$$

3.6. Training

The training objective of ViFi integrates two core terms: a complementary-learning objective and object loss [32]. The former, denoted as (\mathcal{L}_c) , aims to merge information drawn from multiple encoder viewpoints, whereas the latter, (\mathcal{L}_o) , drives the model to perform well on the specific downstream application.

We introduce a global loss function denoted as \mathcal{L} to governs the entire training process:

$$\mathcal{L} = \eta \cdot \mathcal{L}_c + (1 - \eta) \cdot \mathcal{L}_o. \quad (18)$$

Discriminator and complementary loss. To mitigate the issue that the global encoder alone may struggle to learn an optimal representation, a local encoder is incorporated to provide additional corrective signals. This design encourages the two encoding pathways to capture mutually informative cues, and the interaction between them is strengthened through a graph-based contrastive learning scheme. On this basis, the complementary objective (\mathcal{L}_c) [32] is defined as follows:

$$\mathcal{L}_c = \exp\left(-\frac{1}{|\mathcal{V}|} \sum_{l \in \mathcal{V}} \sum_{i=1}^{m-1} \sum_{j \neq i}^m \left\| \mathbf{Z}_{l,i} - \mathbf{Z}_{l,j} \right\|_2\right). \quad (19)$$

where $|\mathcal{V}|$ denotes the total number of nodes, and $\mathbf{Z}_{l,i}$ ($1 < i \leq m$) represents the feature embedding of node l produced by the i -th view encoder.

Object loss for semi-supervised and unsupervised learning. For the semi-supervised setting, node classification is optimized by applying a cross-entropy objective to the embeddings corresponding to the labeled nodes:

$$\mathcal{L}_o = - \sum_{v_k \in \mathcal{V}_L} \sum_j^C Y_{kj}^L \log(\tilde{Y}_{kj}). \quad (20)$$

In the unsupervised setting, the model relies on a combination of reconstruction loss and contrastive loss to realize self-supervised training. Accordingly, the task-specific loss for this scenario is formulated as follows:

$$\begin{aligned} \mathcal{L}_o = & \sum_{j \in \{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_k\}} \left\| \tilde{\mathbf{X}}_j - \tilde{\mathbf{X}}_j^{re} \right\|_2 \\ & + \lambda_1 \left\| \tilde{\mathbf{A}} - \tilde{\mathbf{A}}^{re} \right\|_2 + \lambda_2 \sum_i^m \sum_{j \in \mathcal{G}_{set_i}} \sum_{k \in \mathcal{G}_{set_i}, j \neq k} \mathcal{I}_{j,k}^i. \end{aligned} \quad (21)$$

where λ_1 and λ_2 are hyperparameters. \mathcal{G}_{set_i} is the input set for the i -th view encoder.

4. Experiments

In this section, we conduct extensive experiments to validate the superiority of ViFi and answer the following research questions.

Q1: What is the performance of ViFi in node classification and graph classification tasks?

Q2: How can we verify the advantage of the ViFi framework?

Q3: What roles are fulfilled by the individual components within the proposed framework?

Q4: In what ways is the robustness of ViFi demonstrated?

Q5: How effective is the approach when it is utilized for node-clustering applications?

Q6: To what extent do changes in hyperparameter settings influence the behavior of ViFi?

Datasets. We conduct our evaluation on six benchmark datasets, whose main statistics are listed in Table 2. Cora, Citeseer, Pubmed, and DBLP are citation networks, whereas ACM and Chameleon come from academic and Wikipedia sources respectively.

ViFi. In the experimental analysis, **OURS** denotes the semi-supervised variant of ViFi, whereas **OURS-UN** refers to its unsupervised counterpart. For the unsupervised setting, the learned representations are subsequently fed into linear classifiers, allowing the model to produce node-classification results.

Baselines. We compare ViFi with state-of-the-art methods. 1) Base encoder: GCN [33]. 2) Attention-based encoder: GAT [35], MAGCN [36], DGCN [37] and PA-GCN [38]. 3) Multi-view information fusion-based encoder for node classification: MixHop [39], N-GCN [40], MOGCN [41], MAGCN [36], DGCN [37], PA-GCN [38], LoGo-GNN[32], StrucGCN[42] and ND-GCN[43]. 4) Multi-view information fusion-based encoder for graph classification: Co-GCN [44], LGCN-FF [45], SLFNet [46], HGCM-MVSC [47], MGCN-DNS [48]. 5) Contrastive learning-based encoder: NCLA [49], PA-GCN

[38], GraphCL [50], IGCL [51], and GCA [52]. 5) Unsupervised learning model: K-means, Deepwalk [53], GAE [34], and VGAE [34].

Parameter Setting. A complete list of hyperparameter settings for each dataset is provided in Table 3.

Implementation Details. Our training procedure adopts a full-batch strategy for each epoch. The method is implemented in Pytorch, and parameter updates are carried out using the Adam [54] algorithm. For the standard graph benchmarks, we randomly sample different numbers of labeled nodes per class for training, while keeping 1,000 nodes fixed for testing. In superpixel datasets, evaluation is conducted on a set of 10,000 images. All classification accuracies (ACC) are averaged over 10 independent runs using the data splits described above. The hyperparameter η is explored across $\{0.05, 0.1, 0.15, \dots, 0.95\}$, while λ_1 and λ_2 are adjusted within $\{0, 0.1, 0.2, \dots, 1\}$. Additionally, the cosine-threshold is tuned in $\{0.1, 0.15, 0.2, \dots, 0.5\}$, and k is varied over $\{5, 10, 15, \dots, 30\}$. The final results for each dataset are reported using the hyperparameter combination and iteration count that yield optimal performance.

Table 2. Data details.

Dataset	Nodes	Edges	Features	Classes	Training	Test
Cora	2708	5429	1433	7	140	1000
Citeseer	3327	4732	3703	6	120	1000
ACM	3025	13128	1870	3	60	1000
Chameleon	2277	36101	2325	4	80	1000
DBLP	17716	105734	1639	4	80	1000
Pubmed	19717	44338	500	3	60	1000
MNIST	—	—	—	10	55000	10000
ALOI	—	—	—	100	800	200
NUS-WIDE	—	—	—	81	161789	107859
MSRC-v1	—	—	—	7	105	200

Table 3. Hyperparameter settings.

Dataset	Learning rate	Weight decay	Training epochs	Dropout
Cora	0.0015	5.00E-04	200	0.3
Citeseer	0.0015	5.00E-04	200	0.3
ACM	0.0015	5.00E-04	200	0.3
Chameleon	0.0015	5.00E-04	200	0.3
DBLP	0.0015	5.00E-04	200	0.3
Pubmed	0.0015	5.00E-04	200	0.3
MNIST	0.001	1.00E-05	5000	0.5
ALOI	0.001	1.00E-05	5000	0.5
NUS-WIDE	0.001	1.00E-05	5000	0.5
MSRC-v1	0.001	1.00E-05	5000	0.5

Evaluation Metrics. Following established practices in node and graph classification, we evaluate the performance of both baseline methods and ViFi on node classification tasks using classification accuracy (ACC). For each dataset, ACC is computed across all test samples. In addition, to assess clustering effectiveness, we employ normalized mutual information (NMI) [55] and adjusted rand index (ARI) [56], providing complementary measures of how well ViFi and competing approaches capture underlying cluster structures.

Table 4. Notation description.

Notation	Description
Raw	The raw input view.
RawC	The convolution augmentation view via Equation (22).
PR	The augmented view based on raw view relations.
PC	The augmented view based on cosine similarity view relations.
PK	The augmented view based on k NN view relations.
Global	Learned representation of global views ({RAW, PC, PK}).
Local1	Learned representation of local views ({RAW, PC}).
Local2	Learned representation of local views ({RAW, PK}).
Local3	Learned representation of local views ({PC, PK}).
Local-Global	Fused representation from the local-to-global views.
No-Opt (manual)	Fusion without optimization; a fixed manual strategy is used to integrate filtered views.
Random-Select (avg)	Randomly selects view subsets of the same size for fusion and averages the results.
Only-Entropy	Employs only the entropy-balance term in the gain function.
Only-Structure	Employs only the normalized structural-difference term in the gain function.
Greedy-by-Entropy	Sequentially adds views based on their individual information scores until the gain ceases to improve.
Full (Ours)	Complete optimized fusion mechanism using both entropy balance and structural complementarity to adaptively select the optimal view integration.

4.1. Performance on Node and Graph Classification(Q1)

4.1.1. Performance on Node Classification

This section reports the average classification accuracy (ACC) along with its standard deviation across 10 independent trials. For reference, the results of DeepWalk [53], NCLA [49], LGCN-FF [45], and SLFNet [46] are adopted from their respective original studies. The outcomes of semi-supervised node classification are compiled in Table 5, with the key insights summarized as follows:

1. Compared with the baseline models, ViFi demonstrates consistently superior performance across most datasets. Notably, the semi-supervised version of ViFi (**OURS**) achieves better results than its unsupervised counterpart (**OURS-UN**). This improvement can be explained by the fact that the semi-supervised framework of ViFi adopts an end-to-end fusion mechanism, in which the available label information effectively guides and refines the embedding fusion process during model training.
2. The semi-supervised version of ViFi (**OURS**) consistently outperforms other models that incorporate multi-topology or multi-view information fusion, such as MAGCN, MOGCN, and PA-GCN. Moreover, the unsupervised version of ViFi (**OURS-UN**) also achieves better performance than most graph contrastive learning methods, including GCA and IGCL. This advantage primarily stems from ViFi's ability to identify and retain only the views that contribute positively to representation, while filtering out irrelevant views that contribute negatively, thereby enhancing representation quality and improving the model's robustness against noisy or irrelevant information.
3. Building on the failure cases listed in Table 5, we select the Pubmed dataset as a representative example for a more detailed analysis. In this dataset, the performance of ViFi is inferior to that of the NCLA model. This phenomenon can be attributed to a key characteristic of NCLA: it employs an augmentation-based learning strategy, which facilitates effective extraction of self-supervised signals between the original graph and its augmented versions, provided that the underlying relationships in the raw graph are trustworthy. A comparable pattern emerges when examining the Citeseer dataset. In this case, ViFi also trails behind the LA-GCN model. This performance gap

can be attributed to the design of LA-GCN, which incorporates a trainable local augmentation mechanism grounded in the structural relations of the original graph. Meanwhile, it places greater emphasis on enhancing the informative value of data from the perspective of feature engineering, which further explains its performance advantage over ViFi in this dataset. However, overall, ViFi continues to exhibit better performance compared with these competing approaches. The superior performance of ViFi can be mainly attributed to its entropy-driven adaptive view filter and optimized fusion mechanism. The adaptive view filter evaluates the similarity and connectivity among nodes to characterize the concentration of information distribution. Based on this, it filters the views that contain richer and more informative content for subsequent fusion. In the optimized fusion mechanism, a novel information gain function is proposed to evaluate candidate view groupings based on entropy balance and structural complementarity. It is further employed to determine the most effective integration strategy, thereby enabling a more efficient and complementary fusion of multi-view representations. The superiority of ViFi becomes particularly pronounced in the presence of noise. The following experimental section examines and quantifies the robustness of the model.

Table 5. The classification accuracy (ACC%) for node level tasks is reported, with Y representing the corresponding label information.(Bold: best).

Method	Training Data	Datasets					
		Cora	Citeseer	ACM	Chameleon	DBLP	Pubmed
DeepWalk [53]	A	67.2	43.2	-	-	-	65.3
NCLA [49]	X,A	82.2	71.7	-	-	-	82.0
GCA [52]	X,A	82.7 ± 0.6	71.8 ± 0.7	89.4 ± 0.6	35.6 ± 0.2	70.6 ± 0.5	76.8 ± 0.5
IGCL [51]	X,A	83.5 ± 0.5	72.1 ± 0.6	89.2 ± 0.7	47.6 ± 0.4	73.6 ± 0.3	80.8 ± 0.6
GraphCL [50]	X,A	83.1 ± 0.6	72.1 ± 0.6	90.2 ± 0.6	48.6 ± 0.5	72.9 ± 0.5	80.6 ± 0.6
OURS-UN	X,A	83.5 ± 0.4	73.0 ± 0.3	91.6 ± 0.7	51.5 ± 0.6	73.8 ± 0.3	80.9 ± 0.8
GCN [33]	X,A,Y	81.5 ± 0.2	70.4 ± 0.4	87.8 ± 0.2	47.6 ± 0.4	70.2 ± 0.5	79.0 ± 0.6
GAT [35]	X,A,Y	83.2 ± 0.7	72.6 ± 0.7	87.4 ± 0.3	47.9 ± 0.4	71.0 ± 0.3	79.0 ± 0.6
PA-GCN [38]	X,A,Y	83.6 ± 0.2	70.4 ± 0.3	90.9 ± 0.3	49.0 ± 0.3	72.0 ± 0.4	79.3 ± 0.3
MOGCN [41]	X,A,Y	82.4 ± 1.2	72.4 ± 0.8	90.1 ± 1.4	46.9 ± 0.4	70.9 ± 0.7	79.2 ± 0.3
N-GCN [40]	X,A,Y	83.0 ± 0.5	72.2 ± 0.5	88.0 ± 0.3	48.9 ± 0.4	71.3 ± 0.2	79.5 ± 0.4
MixHop [39]	X,A,Y	81.9 ± 0.4	71.4 ± 0.6	87.9 ± 0.7	40.6 ± 0.6	70.9 ± 0.3	80.8 ± 0.2
DGCN [37]	X,A,Y	84.1 ± 0.3	73.3 ± 0.1	90.2 ± 0.2	48.9 ± 0.4	72.3 ± 0.2	80.2 ± 0.3
MAGCN [36]	X,A,Y	84.5 ± 0.5	73.3 ± 0.3	90.6 ± 0.3	50.4 ± 0.5	72.5 ± 0.3	80.6 ± 0.8
LA-GCN [57]	X,A,Y	84.6 ± 0.7	74.7 ± 0.5	89.9 ± 0.4	48.3 ± 0.6	72.0 ± 0.5	81.7 ± 0.7
LoGo-GNN [32]	X,A,Y	83.9 ± 0.4	73.4 ± 0.5	91.8 ± 0.5	51.7 ± 0.3	73.9 ± 0.4	81.6 ± 0.7
StrucGCN [42]	X,A,Y	82.8 ± 0.6	72.3 ± 0.4	90.7 ± 0.3	50.6 ± 0.5	73.1 ± 0.3	81.4 ± 0.4
ND-GCN [43]	X,A,Y	83.2 ± 0.4	73.6 ± 0.3	89.9 ± 0.6	51.2 ± 0.3	73.5 ± 0.4	80.8 ± 0.5
OURS	X,A,Y	84.2 ± 0.4	73.8 ± 0.5	94.8 ± 0.5	52.7 ± 0.3	74.9 ± 0.5	83.6 ± 0.7

4.1.2. Performance on Graph Classification

We use the multi view datasets for graph classification and take GCN-fusion [33], Co-GCN [44], LGCN-FF [45], SLFNet [46], HGCN-MVSC [47] and MGCN-DNS [48] as baselines. The reported classification accuracy (ACC) represents the average over 10 independent runs. Graph classification outcomes are summarized in Table 6, demonstrating that ViFi maintains strong competitiveness in graph-level classification tasks.

Table 6. The classification accuracy (ACC%) for graph level tasks.(Bold: best).

Method	Datasets			
	ALOI	MNIST	NUS-WIDE	MSRC-v1
GCN-fusion [33]	89.4	86.6	36.7	58.5
Co-GCN [44]	90.2	90.8	26.5	67.2
LGCN-FF [45]	93.4	90.2	38.3	83.4
SLFNet [46]	91.7	95.6	45.7	76.1
HGCN-MVSC [47]	92.3	91.1	44.3	81.5
MGCN-DNS [48]	92.2	94.7	45.2	82.1
OURS-UN	93.4	95.2	47.3	85.2
OURS	94.8	96.3	48.8	86.7

4.2. ViFi Architecture Study (Q2)

To simplify the analysis, we employ only the best-performing variant of ViFi (OURS) built upon a semi-supervised learning framework. The model is applied to node classification and semantic similarity tasks on the Cora dataset, providing additional validation for the effectiveness of the proposed entropy-based adaptive view filtering and the optimized fusion strategy.

- Multi views-GCN+MLP: A GCN framework combining local and global encoder views via MLP is utilized to perform node classification tasks;
- Multi views-GAT+GAT: A graph attention network (GAT) architecture that merges local and global encoder views using a single-layer GAT is employed for node level classification tasks;
- Global-GCN+GCN: A GCN architecture adopting a global encoder perspective, implemented with a single-layer GCN for node classification tasks;
- **OURS-UN-w/o**: ViFi implemented under an unsupervised learning framework, operating without the complementary loss term \mathcal{L}_c ;
- **OURS-w/o**: ViFi implemented under an semi-supervised learning framework, operating without the complementary loss term \mathcal{L}_c ;
- **OURS-UN**: ViFi configured within an unsupervised learning framework;
- **OURS**: ViFi configured within an semi-supervised learning framework;

These results lead to several noteworthy observations, as discussed below:

1. To further assess the efficacy of the ViFi framework, we utilize a standard MLP or a Graph Attention Network as the terminal classifier. The experimental results indicate that incorporating graph attention networks (GAT) yields superior performance compared with both the MLP-based and our proposed configurations. This improvement suggests that GAT provides a more efficient mechanism for capturing the topological dependencies within the fused node representations. Moreover, the performance of our model surpasses that of the Multi-view GCN+MLP, implying that the fusion process preserves the essential relational patterns among nodes.
2. Compared with the Global-GCN+GCN variant, ViFi delivers superior results, underscoring the necessity of incorporating localized structural cues when global information is sparse or partially missing, as is often the case under limited label supervision. Furthermore, architectures that jointly exploit global and local views tend to exhibit more stable and effective behavior than approaches relying solely on a global-view encoder.
3. Acquiring complementary information from different views plays a vital role. As shown in Table 7, the ViFi (**OURS-UN** and **OURS**) trained with the global loss function achieves better performance than the ViFi trained solely with the object loss \mathcal{L}_o . This observation further validates our rationale for introducing the global loss function.

Table 7. Classification accuracy (ACC%) for ViFi and its variant models (Bold: best).

Method	Cora	Citeseer	ACM	Chameleon	DBLP	Pubmed
Multi views-GCN+MLP	84.2	72.4	90.1	49.5	72.1	80.3
Multi views-GAT+GAT	84.7	73.3	92.0	52.9	75.2	83.6
Global-GCN+GCN	84.5	73.2	91.6	51.7	74.5	81.2
OURS-UN-w/o	82.5	71.5	90.4	51.2	72.8	80.5
OURS-w/o	84.3	73.3	91.5	51.3	73.1	80.9
OURS-UN	83.5	71.7	91.6	51.5	73.8	80.9
OURS	84.2	73.8	94.8	52.7	74.9	83.6

4.3. Ablation Study (Q3)

The effectiveness of the proposed ViFi is confirmed through the comparative experiments discussed above. In addition, to further validate the contribution of each individual component within ViFi, we conduct a series of ablation studies in this subsection.

4.3.1. Entropy-Based Adaptive View Filter

First, we apply various aggregation strategies to generate augmented graphs, each serving as a distinct view. For comparison, we incorporate the GCN aggregation method [33] into the graph augmentation process (Equation 22). Table 4 summarizes the descriptions of the various views. To visualize the representational capacity of each view (PR, PC, RawC, Raw), a broken-line chart is employed, tracking changes after each iteration. The alignment between each augmented view and the original input view (Raw) is subsequently quantified by examining their semantic similarity. Both the augmented and original views are encoded using a shared GCN encoder under identical parameter settings, and their learned graph representations are subsequently compared to assess semantic alignment.

The results in Table 8 indicate that the performance of augmented views varies notably across datasets, reflecting the distinct structural and semantic characteristics captured by each view. The PC view attains the highest accuracy on the Cora and ACM datasets, suggesting that cosine similarity preserves meaningful feature correlations beneficial for classification, while the PK view performs slightly better on Pubmed, implying that neighborhood-based relations are more informative in this dataset. Such inconsistency among single views highlights the necessity of filtering and integrating multiple complementary views. When multiple views are combined, classification accuracy improves substantially compared with individual views. Among them, Raw+PC+PK consistently achieves the best results across all datasets, indicating that combining feature-similarity and structural-proximity views enhances the representational capacity of the model and leads to more effective multi-view learning.

The analysis of semantic similarity and representation ability of Figure 4 provides a key explanation for this phenomenon. The figure shows that the semantic similarity and representation ability between different views show different evolution trajectories with the change of the number of iterations. The PC view can obtain higher representation ability with fewer iterations, and maintain a moderate semantic distance from the original view, which can provide effective complementary information, but said the ability to improve the limited. This inconsistency in performance and properties implies that directly fusing all views without any filtering mechanism may introduce interfering or task-irrelevant views, thereby degrading model performance. Moreover, the model would fail to adaptively identify the most informative views for the given task, ultimately resulting in suboptimal fusion outcomes. Therefore, a filtering mechanism that evaluates and adaptively filters views is essential to ensure that the fusion process focuses on the most informative and complementary views.

$$\tilde{x}_i = \hat{A}_{i,i}x_i + \sum_{x_j \in \mathcal{N}(x_i)} \hat{A}_{i,j}x_j. \quad (22)$$

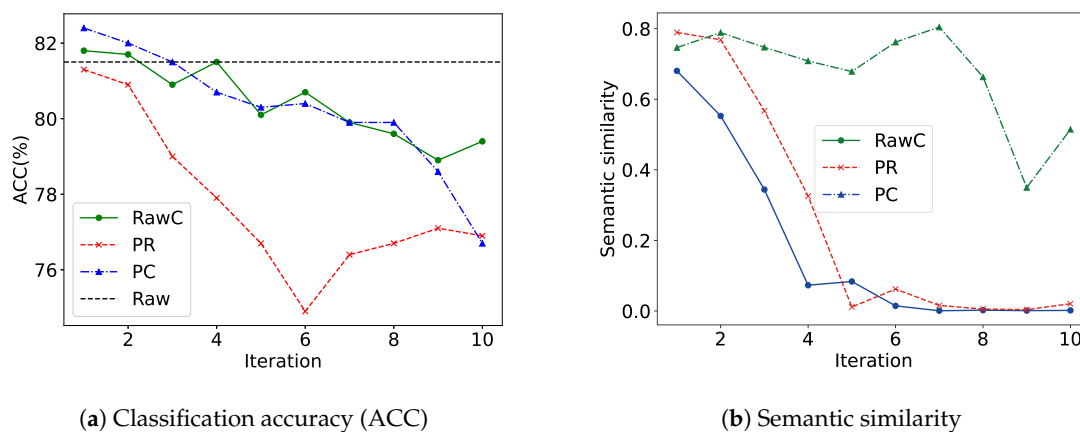


Figure 4. Comparison of classification accuracy and semantic similarity across iterations for RawC, PR, and PC.

Table 8. The accuracy (%) achieved by the GCN encoder when operating on various individual views or integration of views. (Bold: best).

Input	Cora	Citeseer	ACM	Chameleon	DBLP	Pubmed
RAW	81.2	70.6	87.5	47.9	70.4	78.8
PR	81.0	70.3	88.0	47.1	71.0	78.1
PC	82.3	71.2	88.5	49.4	71.7	80.0
PK	82.1	70.8	87.6	48.1	70.2	80.4
Raw+PR+PC	83.6	71.3	90.9	50.4	72.7	80.3
Raw+PR+PK	83.1	71.0	91.1	50.1	71.8	80.9
Raw+PC+PK	84.2	73.8	94.8	52.7	74.9	83.6

4.3.2. Optimized Fusion Mechanism

Furthermore, we design a series of comparative experiments, to investigate the role of the optimized fusion mechanism. Specifically, we implement several fusion variants, including manual fusion, random subset selection, and simplified versions that retain only the entropy-balance or structural-difference term. Each variant employs the same set of filtered views obtained from the first module, ensuring that the observed differences arise solely from the fusion mechanism. The descriptions of these variants are summarized in Table 4. Subsequently, we compare their classification accuracy across multiple datasets to evaluate the impact of different fusion mechanisms.

The experimental results presented in Table 9 reveal several important observations. The manual fusion setting, yields the lowest accuracy across all datasets, indicating that fixed integration schemes fail to capture the complementary information among views effectively. The random selection strategy shows only marginal improvement, suggesting that arbitrary integrations of views can not guarantee consistent information gain. The variants using only the entropy-balance or only the structural-difference term both lead to moderate performance gains, implying that each component contributes partially to the overall fusion objective. The greedy-by-entropy approach achieves slightly higher accuracy, demonstrating that prioritizing informative views provides a limited but noticeable benefit. The full optimized fusion mechanism achieves the highest accuracy across datasets, confirming that jointly considering entropy balance and structural complementarity enables more adaptive and synergistic view integration. Overall, the results validate the necessity of automatic strategy selection in promoting balanced and effective multi-view fusion.

Table 9. Ablation study on the optimized fusion mechanism (ACC%). (Bold: best).

Method	Cora	Citeseer	ACM	Chameleon	DBLP	Pubmed
No-Opt (manual)	83.7	71.5	91.0	50.6	72.9	80.4
Random-Select (avg)	83.9	71.9	91.2	51.0	73.2	80.7
Only-Entropy	84.0	72.1	91.4	51.4	73.6	81.0
Only-Structure	84.1	72.3	91.5	52.0	73.8	81.2
Greedy-by-Entropy	84.2	72.6	91.6	52.1	74.0	81.3
Full (Ours)	84.2	73.8	94.8	52.7	74.9	83.6

4.4. Robustness Analysis (Q4)

As both the unsupervised (**OURS-UN**) and semi-supervised (**OURS**) variants of VF-GRFN employ an identical fusion framework, only the semi-supervised model (**OURS**), which demonstrates superior performance, is utilized in the subsequent experiments to streamline the analysis.

First, we evaluate the expressive capability of each view through visualization and graph reconstruction techniques. To accomplish the reconstruction of the latent representations derived from each view-specific encoder on the Cora dataset, we utilize a Variational Graph Auto-Encoder (VGAE) [34]. The reconstruction quality is quantified using the AUC (area under the ROC curve) metric, and the corresponding results are illustrated in Figure 5. Furthermore, to intuitively demonstrate the distribution of the learned representations, we visualize the embeddings of each view encoder using t-SNE [58], as shown in Figure 6.

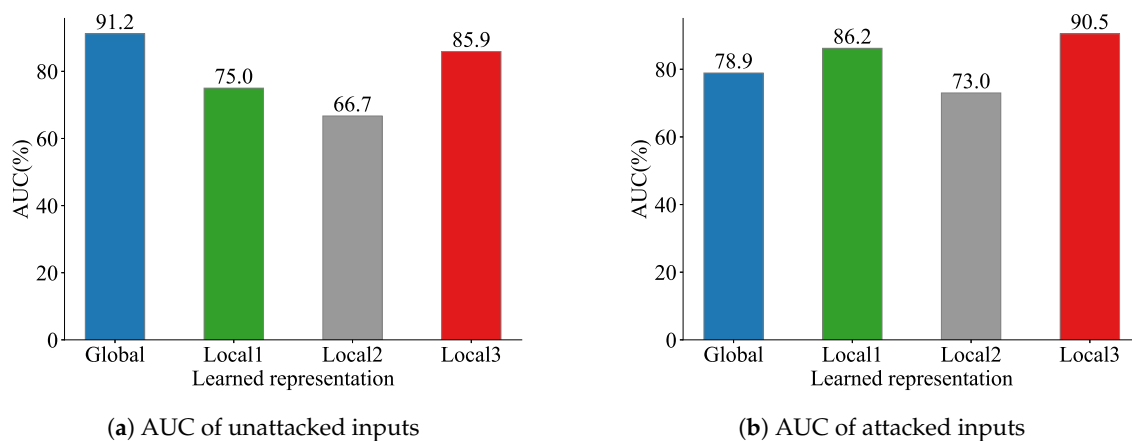


Figure 5. VGAE graph reconstruction performance on Cora across views. (a) AUC for unattacked inputs (learned from original views). (b) AUC for attacked inputs (learned from views with 80% edges randomly deleted).

On the Cora benchmark, we monitor the evolution of semantic similarity across embeddings from the global and local views throughout the learning process. This is presented through a line-based visualization computed at selected training epochs {1; 5; 10; 20; 40; 100}. The embedding vectors for all views are obtained through the same GCN encoder. Their semantic similarity is subsequently quantified by analyzing the resulting feature representations, as depicted in Figure 7. Table 4 shows various views' description. When the input graphs are perturbed, the expressive capability of each view encoder declines as both its visual distribution and its reconstruction quality deteriorate. In contrast, the fused representation that integrates local and global views preserves stronger expressiveness under noisy conditions, indicating that ViFi is more robust. The global-only representation performs less effectively, while local views exhibit higher discriminability and better adaptation to structural noise. Moreover, the semantic discrepancy between local and global representations grows with the proposed contrastive loss \mathcal{L}_c , confirming the significance of introducing local views. Our objective is enabled to enhance diversity among encoders from different views while maintaining the strong learning ability of key encoders. However, this divergence does not continually increase with training epochs, as the optimization is jointly constrained by contrastive and objective losses.

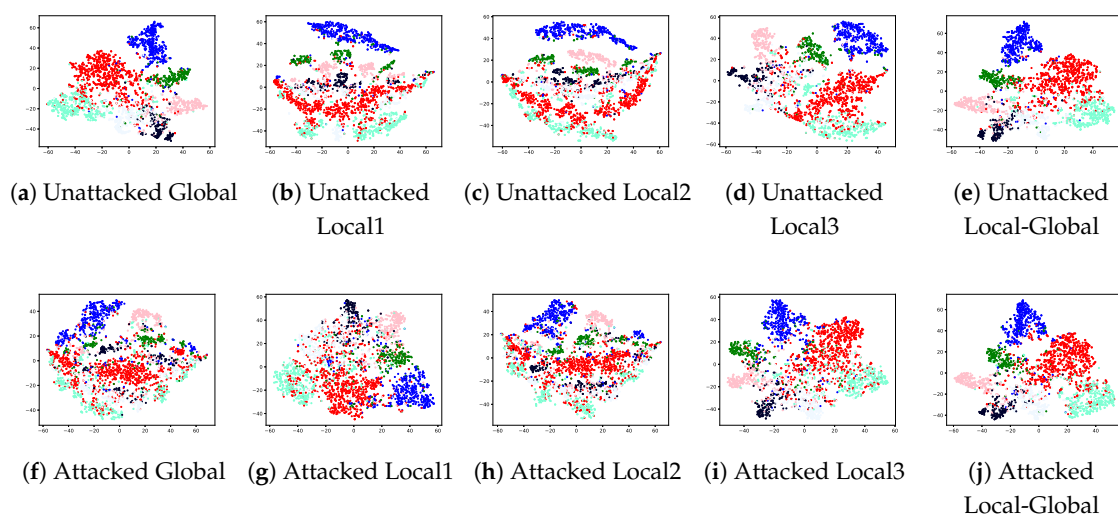


Figure 6. Visualization of learned embeddings on Cora across views. Unattacked (a-e): from original views. Attacked (f-j): from views with 80% edges randomly deleted.

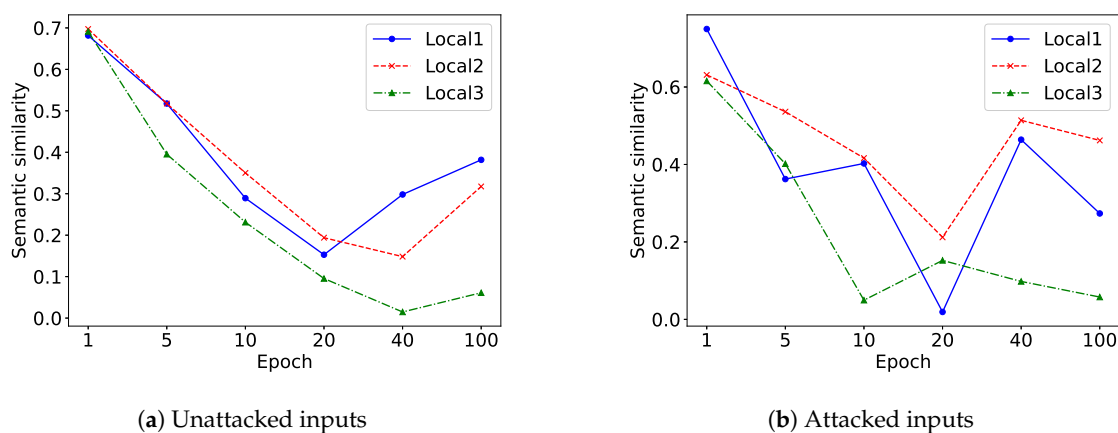


Figure 7. Semantic similarity between local and global view embeddings on Cora. (a) Unattacked inputs (from original views). (b) Attacked inputs (from views with 80% edges randomly removed).

To assess the robustness of ViFi, we further subject ViFi, its variants Global-GNN, and GCN [33] to a set of four uncertainty scenarios. These conditions are used to examine how reliably each model maintains classification performance when exposed to varying views. These uncertainties may introduce disturbances to the graph structure or node semantics, thereby affecting classification accuracy. All experiments are conducted on the Cora dataset. The number of labeled nodes is set to {14; 21; 28; 54}. The attack and mask ratios are {0.2; 0.4; 0.6; 0.8}, and the noise level is adjusted to {0.001; 0.01; 0.1; 1}. For topology and feature masking, we randomly remove a portion of edges or node features according to the specified ratio. The modified graphs are then used to test classification performance on the corrupted data. In the case of feature corruption caused by node noise, we perturb the input attributes by injecting Gaussian disturbances, where the parameter noise level controls the strength of the added variation. This process intentionally distorts the original feature distribution to evaluate model stability under corrupted inputs. The corresponding outcomes are presented in Figure 8. Drawing from these observations, several key findings can be outlined as follows.

- (1) As illustrated in Figure 8(a), the baseline model experiences a sharp decline in performance as the label rate decreases. In contrast, OURS maintains superior accuracy even under extremely low label availability. This observation indicates that the introduction of a view filter and an optimized fusion mechanism enhances the model's ability to capture essential features under limited supervision. Notably, OURS substantially outperforms Global-GNN in such low-label scenarios, further validating the effectiveness of the proposed design.
- (2) As illustrated in Figure 8(b), **OURS** consistently achieves superior performance compared to the baseline models under higher attack ratios. This advantage arises from its ability to effectively capture information from entropy-based view filter. In general, all methods exhibit a sharp decline in performance as the random attack ratio increases.
- (3) In alignment with earlier findings, the results presented in Figures 8(c)–(d) indicate that **OURS** maintains a clear advantage over both GCN and Global-GNN across varying noise and masking conditions. Overall, as the proportion of masked features or the strength of the injected noise increases, every method exhibits a substantial reduction in predictive accuracy.

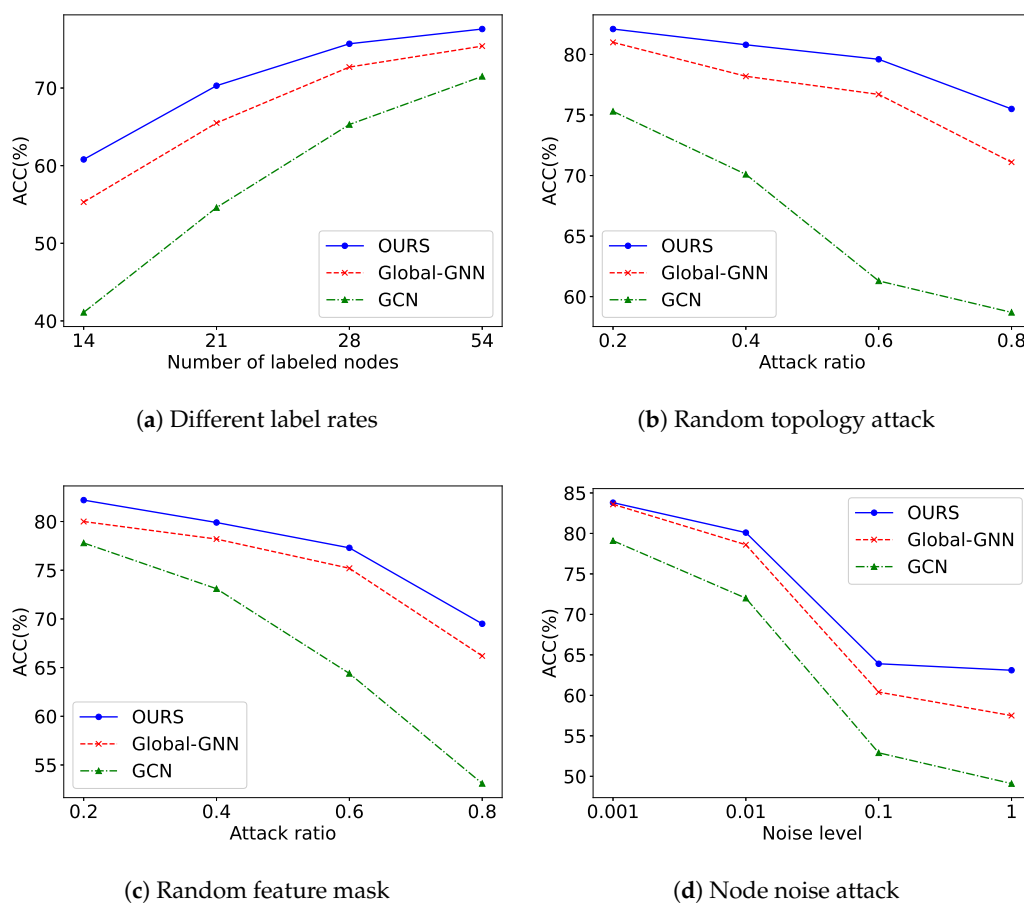


Figure 8. Performance Comparison of GCN, Global-GNN, and **OURS** on Cora. (a) Different label rates. (b) Random topology attack. (c) Random feature mask. (d) Node noise attack.

4.5. Performance on Node Clustering (Q5)

We further examine the capability of the framework in unsupervised graph representation learning. Once the model is trained, the fused node representations $H^{(1)}$ are extracted and subsequently grouped using the K-means clustering procedure on three benchmark datasets—Cora, Citeseer, and Pubmed. The evaluation is based on two metrics, normalized mutual information (NMI) and adjusted rand index (ARI), and their mean values along with standard errors are reported in Table 10. To establish competitive baselines, we incorporate a range of widely used unsupervised and contrastive graph learning methods, such as K-means, DeepWalk [53], GraphCL [50], IGCL [51], GCA [52], GAE [34], and VGAE [34]. To further illustrate the flexibility of the ViFi framework, the objective \mathcal{L}_o is equipped with alternative contrastive components by substituting its \mathcal{I} term with the loss formulations of GCA [52], IGCL [51], and GraphCL [50]. These variants are referred to as **OURS**, **OURS(GraphCL)**, and **OURS(IGCL)**, respectively. Experimental results indicate that ViFi achieves performance comparable to other state-of-the-art methods across most datasets. These results confirm that ViFi exhibits strong competitiveness in unsupervised clustering tasks. Furthermore, the model’s performance is highly sensitive to the choice of \mathcal{I} loss, underscoring its inherent flexibility.

Table 10. The clustering performance of various approaches under different input settings. (Bold: best).

Method	Training Data	Cora		Citeseer		Pubmed	
		NMI	ARI	NMI	ARI	NMI	ARI
K-mean	X	32.1	23.0	30.5	27.9	0.1	0.2
DeepWalk	X	32.7	24.3	8.8	9.2	27.9	29.9
GAE	X,A	42.9	34.7	17.6	12.4	27.7	27.9
VGAE	X,A	23.9	17.5	15.6	9.3	22.9	21.3
GRAGE	X,A	54.4 ± 1.2	43.4 ± 3.1	35.2 ± 1.8	34.1 ± 1.7	30.0 ± 3.4	29.5 ± 2.1
OURS	X,A	55.6 ± 2.3	47.2 ± 2.5	39.2 ± 2.3	35.1 ± 2.3	32.5 ± 2.6	33.7 ± 2.0
GraphCL	X,A	55.3 ± 3.2	54.5 ± 2.7	42.1 ± 3.9	44.1 ± 1.8	34.1 ± 3.3	34.2 ± 2.8
OURS(GraphCL)	X,A	56.9 ± 2.2	55.0 ± 2.3	45.0 ± 2.0	44.3 ± 2.1	35.6 ± 2.3	33.8 ± 1.5
IGCL	X,A	56.6 ± 3.0	53.9 ± 2.0	43.5 ± 1.7	45.2 ± 1.9	33.4 ± 2.6	32.9 ± 1.1
OURS(IGCL)	X,A	57.9 ± 1.3	55.2 ± 2.2	45.2 ± 2.3	44.8 ± 2.6	37.5 ± 2.4	35.7 ± 1.8

* Results reported as mean ± standard deviation.

4.6. Parameter Sensitivity (Q6)

Sensitivity studies are conducted on the key hyperparameters to examine their influence on model performance. ViFi is trained with α values ranging from 0.05 to 0.95, with an interval of 0.05. Experimental observations indicate that α values between 0.3 and 0.7 produce the most favorable outcomes, suggesting that maintaining a balanced emphasis between feature entropy and structural entropy enables the model to effectively capture both semantic richness and topological diversity, thereby achieving more comprehensive and discriminative representations. ViFi is trained with μ values ranging from 0.01 to 0.5, with an interval of 0.01. Experimental results show that moderate μ values, typically between 0.1 and 0.3, yield the best performance, suggesting that an appropriate regularization strength helps balance the trade-off between maximizing information gain and controlling the size of selected view subsets, thereby preventing overfitting to irrelevant views while maintaining informative diversity. We train ViFi with η values from 0.05 to 0.95 and observe that 0.4–0.8 yields the best performance, indicating that appropriately weighting cross-encoder representation relationships improves the model.

Furthermore, we investigate the effects of the hyperparameters λ_1 and λ_2 on the model's performance. In this experiment, other hyperparameters fixed while λ_1 and λ_2 vary from 0 to 0.9. The classification results of all nodes are visualized in a 3D bar chart, as shown in Figure 9. The findings reveal that higher values of λ_1 and λ_2 generally lead to improved self-supervised learning performance.

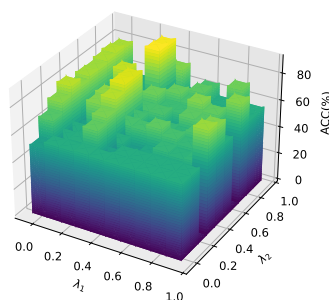


Figure 9. Impact of hyperparameters λ_1 and λ_2 on **OURS-UN**'s node classification accuracy (ACC%).

5. Conclusion

In this paper, we proposed a novel multi-view representation learning framework called the View Filter-driven Graph Representation Fusion Network (ViFi). Following the "less for better" principle, ViFi aims to obtain more effective graph representations by leveraging fewer but more informative views. The framework first evaluates the feature–topology entropy of each view to measure its information quality, adaptively filtering those that provide complementary signals. It then integrates the filtered views through an information gain function that balances information diversity and structural consistency, ensuring collaborative and complementary fusion. Extensive

experiments on classification and clustering tasks verified the effectiveness and superior performance of the proposed ViFi.

Despite the promising results, ViFi still has several limitations. The current entropy estimation relies on predefined statistical formulations, which may not fully capture higher-order correlations or latent semantic dependencies among views. Moreover, the fusion process is optimized based on pairwise information gain, potentially overlooking more complex multi-view interactions that could further enhance representation quality. Future work will explore learnable entropy modeling and more flexible fusion mechanism to better capture inter-view dependencies and improve the generalization of the framework.

Author Contributions: Conceptualization, Y.W.; methodology, Y.W.; software, Y.W. and X.Y; validation, Y.W., X.Y. and X.W.; formal analysis, Y.W.; investigation, Y.W. and Q.G.; resources, K.L. and X.W; data curation, Y.W. and Q.G.; writing—original draft preparation, Y.W.; writing—review and editing, X.Y. and X.W. and K.L.; visualization, Y.W. and Q.G.; supervision, X.Y. and X.W. and K.L.; project administration, X.Y. and X.W; funding acquisition, X.W. and K.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the National Natural Science Foundation of China (nos. 62506145)

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: During the preparation of this study, the authors used Chatgpt for the purposes of language polishing. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Xiao, S.; Li, J.; Lu, J.; et al. Graph neural networks for multi-view learning: a taxonomic review. *Artificial Intelligence Review* **2024**, *57*, 341.
2. Bi, J.; Dornaika, F. Sample-weighted fused graph-based semi-supervised learning on multi-view data. *Information Fusion* **2024**, *104*, 102175.
3. Hong, Y.; Lin, C.; Du, Y.; et al. 3d concept learning and reasoning from multi-view images. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 9202–9212.
4. Ma, Z.; Luo, M.; Guo, H.; et al. Event-radar: Event-driven multi-view learning for multimodal fake news detection. In Proceedings of the Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics, 2024, Vol. 1, pp. 5809–5821.
5. Tian, X.; Deng, Z.; Ying, W.; et al. Deep multi-view feature learning for EEG-based epileptic seizure detection. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **2019**, *27*, 1962–1972.
6. Brynte, L.; Iglesias, J.P.; Olsson, C.; et al. Learning structure-from-motion with graph attention networks. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 4808–4817.
7. Cong, H.; Yang, X.; Liu, K.; et al. Feature-topology cascade perturbation for graph neural network. *Engineering Applications of Artificial Intelligence* **2025**, *152*, 110657.
8. Dai, S.; Wang, J.; Huang, C.; et al. Dynamic multi-view graph neural networks for citywide traffic inference. *ACM Transactions on Knowledge Discovery from Data* **2023**, *17*, 1–22.
9. Wang, D.; Zhang, X.; Yin, Y.; et al. Multi-view enhanced graph attention network for session-based music recommendation. *ACM Transactions on Information Systems* **2023**, *42*, 1–30.
10. Song, N.; Du, S.; Wu, Z.; et al. GAF-Net: Graph attention fusion network for multi-view semi-supervised classification. *Expert Systems with Applications* **2024**, *238*, 122151.
11. Xiong, L.; Yuan, X.; Hu, Z.; et al. Gated fusion adaptive graph neural network for urban road traffic flow prediction. *Neural Processing Letters* **2024**, *56*, 9.
12. Li, C.; Zhu, X.; Yan, Y.; et al. Mhgnn: Multi-view fusion based heterogeneous graph neural network. *Applied Intelligence* **2024**, *54*, 8073–8091.

13. Zhang, S.; Liu, J.; Jiao, Y.; et al. A Multimodal Semantic Fusion Network with Cross-Modal Alignment for Multimodal Sentiment Analysis. *ACM Transactions on Multimedia Computing, Communications and Applications* **2025**.
14. Tsai, D.Y.; Lee, Y.; Matsuyama, E. Information entropy measure for evaluation of image quality. *Journal of Digital Imaging* **2008**, *21*, 338–347.
15. Luo, G.; Li, J.; Peng, H.; et al. Graph Entropy Guided Node Embedding Dimension Selection for Graph Neural Networks. In Proceedings of the Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, 2021, pp. 2767–2774.
16. Zhao, J.; Xie, X.; Xu, X.; Sun, S. Multi-view learning overview: Recent progress and new challenges. *Information Fusion* **2017**, *38*, 43–54.
17. Ma, F.; Meng, D.; Dong, X.; et al. Self-paced multi-view co-training. *Journal of Machine Learning Research* **2020**, *21*, 1–38.
18. Chen, X.; Yuan, Y.; Zeng, G.; et al. Semi-supervised semantic segmentation with cross pseudo supervision. In Proceedings of the Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition, 2021, pp. 2613–2622.
19. Hwang, H.; Kim, G.H.; Hong, S.; et al. Multi-view representation learning via total correlation objective. *Advances in Neural Information Processing Systems* **2021**, *34*, 12194–12207.
20. Niu, Y.; Shang, Y.; Tian, Y. Multi-view SVM classification with feature selection. *Procedia Computer Science* **2019**, *162*, 405–412.
21. Xu, R.; Wang, H. Multi-view learning with privileged weighted twin support vector machine. *Expert Systems with Applications* **2022**, *206*, 117787.
22. Qiao, S.; Shen, W.; Zhang, Z.; et al. Deep co-training for semi-supervised image recognition. In Proceedings of the Proceedings of the European Conference on Computer Vision, 2018, pp. 135–152.
23. Wang, W.; Arora, R.; Livescu, K.; et al. On deep multi-view representation learning. In Proceedings of the International Conference on Machine Learning, 2015, pp. 1083–1092.
24. Mao, L.; Sun, S. Soft margin consistency based scalable multi-view maximum entropy discrimination. In Proceedings of the International Joint Conference on Artificial Intelligence, 2016, pp. 1839–1845.
25. Yu, Z.; Dong, Z.; Yu, C.; et al. A review on multi-view learning. *Frontiers of Computer Science* **2025**, *19*, 197334.
26. Kachole, S.; Huang, X.; Naeini, F.B.; et al. Bimodal SegNet: Fused instance segmentation using events and RGB frames. *Pattern Recognition* **2024**, *149*, 110215.
27. Wei, X.; Yu, R.; Sun, J. View-GCN: View-based graph convolutional network for 3D shape analysis. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 1850–1859.
28. Liang, T.; Xie, H.; Yu, K.; et al. BEVFusion: A Simple and Robust LiDAR-Camera Fusion Framework. In Proceedings of the Advances in Neural Information Processing Systems, 2022, Vol. 35, pp. 10421–10434.
29. Simonyan, K.; Zisserman, A. Two-stream convolutional networks for action recognition in videos. *Advances in Neural Information Processing Systems* **2014**, *27*.
30. Su, B.; Wang, W.; Lin, X.; Liu, S.; Huang, X. Identifying the potential miRNA biomarkers based on multi-view networks and reinforcement learning for diseases. *Briefings in Bioinformatics* **2024**, *25*, bbad427.
31. Wang, S.; Du, Y.; Zhao, S.; et al. Research on the construction of weaponry indicator system and intelligent evaluation methods. *Scientific Reports* **2023**, *13*, 19370.
32. Guo, Q.; Yang, X.; Li, M.; et al. Collaborative graph neural networks for augmented graphs: A local-to-global perspective. *Pattern Recognition* **2025**, *158*, 111020.
33. Kipf, T.N.; Welling, M. Semi-Supervised Classification with Graph Convolutional Networks. In Proceedings of the Proceedings of 5th International Conference on Learning Representations, 2017.
34. Kipf, T.N.; Welling, M. Variational Graph Auto-Encoders. *Computing Research Repository* **2016**. <https://doi.org/abs/1611.07308>.
35. Velickovic, P.; Cucurull, G.; Casanova, A.; et al. Graph Attention Networks. In Proceedings of the Proceedings of 6th International Conference on Learning Representations, 2018.
36. Yao, K.; Liang, J.; Liang, J.; et al. Multi-view graph convolutional networks with attention mechanism. *Artificial Intelligence* **2022**, *307*, 103708.
37. Jin, T.; Dai, H.; Cao, L.; et al. Deepwalk-aware graph convolutional networks. *Science China Information Sciences* **2022**, *65*, 1–15.

38. Guo, Q.; Yang, X.; Zhang, F.; et al. Perturbation-augmented Graph Convolutional Networks: A Graph Contrastive Learning architecture for effective node classification tasks. *Engineering Applications of Artificial Intelligence* **2024**, *129*, 107616.
39. Abu-El-Haija, S.; Perozzi, B.; Kapoor, A.; et al. MixHop: Higher-Order Graph Convolutional Architectures via Sparsified Neighborhood Mixing. In Proceedings of the Proceedings of the 36th International Conference on Machine Learning, 2019.
40. Abu-El-Haija, S.; Kapoor, A.; Perozzi, B.; et al. N-GCN: Multi-scale Graph Convolution for Semi-supervised Node Classification. In Proceedings of the Proceedings of the 35th Conference on Uncertainty in Artificial Intelligence, 2019.
41. Wang, J.; Liang, J.; Cui, J.; et al. Semi-supervised learning with mixed-order graph convolutional networks. *Information Sciences* **2021**, *573*, 171–181.
42. Zhang, J.; Li, M.; Xu, Y.; et al. StrucGCN: Structural enhanced graph convolutional networks for graph embedding. *Information Fusion* **2025**, *117*, 102893.
43. Liu, S.; He, D.; Yu, Z.; et al. Beyond homophily: Neighborhood distribution-guided graph convolutional networks. *Expert Systems with Applications* **2025**, *259*, 125274.
44. Li, S.; Li, W.T.; Wang, W. Co-GCN for multi-view semi-supervised learning. In Proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence, 2020, Vol. 34, pp. 4691–4698.
45. Chen, Z.; Fu, L.; Yao, J.; et al. Learnable graph convolutional network and feature fusion for multi-view learning. *Information Fusion* **2023**, *95*, 109–119.
46. Teng, Q.; Yang, X.; Sun, Q.; et al. Sequential attention layer-wise fusion network for multi-view classification. *International Journal of Machine Learning and Cybernetics* **2024**, *15*, 5549–5561.
47. Wang, S.; Huang, S.; Wu, Z.; et al. Heterogeneous graph convolutional network for multi-view semi-supervised classification. *Neural Networks* **2024**, *178*, 106438.
48. Chen, Z.; Fu, L.; Xiao, S.; et al. Multi-view graph convolutional networks with differentiable node selection. *ACM Transactions on Knowledge Discovery from Data* **2023**, *18*, 1–21.
49. Shen, X.; Sun, D.; Pan, S.; et al. Neighbor Contrastive Learning on Learnable Graph Augmentation. In Proceedings of the Proceedings of 37th AAAI Conference on Artificial Intelligence, 2023, pp. 9782–9791.
50. You, Y.; Chen, T.; Sui, Y.; et al. Graph Contrastive Learning with Augmentations. In Proceedings of the Proceedings of 33rd Annual Conference on Neural Information Processing Systems, 2020.
51. Liang, H.; Du, X.; Zhu, B.; et al. Graph contrastive learning with implicit augmentations. *Neural Networks* **2023**, *163*, 156–164.
52. Zhu, Y.; Xu, Y.; Yu, F.; et al. Graph Contrastive Learning with Adaptive Augmentation. In Proceedings of the Proceedings of The Web Conference, 2021.
53. Perozzi, B.; Al-Rfou, R.; Skiena, S. DeepWalk: online learning of social representations. In Proceedings of the Proceedings of 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2022.
54. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. In Proceedings of the Proceedings of 3rd International Conference on Learning Representations, 2015.
55. Strehl, A.; Ghosh, J. Cluster ensembles—a knowledge reuse framework for combining multiple partitions. *Journal of Machine Learning Research* **2002**, *3*, 583–617.
56. Hubert, L.; Arabie, P. Comparing partitions. *Journal of Classification* **1985**, *2*, 193–218.
57. Liu, S.; Ying, R.; Dong, H.; et al. Local Augmentation for Graph Neural Networks. In Proceedings of the Proceedings of International Conference on Machine Learning, 2022.
58. Wang, X.; Zhu, M.; Bo, D.; et al. AM-GCN: Adaptive Multi-channel Graph Convolutional Networks. In Proceedings of the Proceedings of 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2020.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.