

Multi-View Incremental Learning with Structured Hebbian Plasticity for Enhanced Fusion Efficiency

Yuhong Chen^{1,2}, Ailin Song^{2,3}, Huifeng Yin⁴, Shuai Zhong², Fuhai Chen¹, Qi Xu⁵,
Shiping Wang^{1*}, Mingkun Xu^{2,4*}

¹College of Computer and Data Science, Fuzhou University, Fuzhou, China

²Guangdong Institute of Intelligence Science and Technology, Hengqin, Zhuhai, China

³University of Electronic Science and Technology of China, Chengdu, China

⁴Center for Brain Inspired Computing Research, Department of Precision Instrument, Tsinghua University, Beijing, China

⁵School of Computer Science and Technology, Dalian University of Technology, Dalian, China

yhchen2320@163.com, sal314159@163.com, yinhf22@mails.tsinghua.edu.cn, zhongshuai@gdiist.cn,
chenfuhai3c@163.com, xuqi@dlut.edu.cn, shipingwangphd@163.com, xumingkun@gdiist.cn

Abstract

The rapid evolution of multimedia technology has revolutionized human perception, paving the way for multi-view learning. However, traditional multi-view learning approaches are tailored for scenarios with fixed data views, falling short of emulating the intricate cognitive procedures of the human brain processing signals sequentially. Our cerebral architecture seamlessly integrates sequential data through intricate feed-forward and feedback mechanisms. In stark contrast, traditional methods struggle to generalize effectively when confronted with data spanning diverse domains, highlighting the need for innovative strategies that can mimic the brain's adaptability and dynamic integration capabilities. In this paper, we propose a bio-neurologically inspired multi-view incremental framework named MVIL aimed at emulating the brain's fine-grained fusion of sequentially arriving views. MVIL lies two fundamental modules: structured Hebbian plasticity and synaptic partition learning. The structured Hebbian plasticity reshapes the structure of weights to express the high correlation between view representations, facilitating a fine-grained fusion of view representations. Moreover, synaptic partition learning is efficient in alleviating drastic changes in weights and also retaining old knowledge by inhibiting partial synapses. These modules bionically play a central role in reinforcing crucial associations between newly acquired information and existing knowledge repositories, thereby enhancing the network's capacity for generalization. Experimental results on six benchmark datasets show MVIL's effectiveness over state-of-the-art methods.

Introduction

Swift advancements in multimedia technology have fundamentally altered human perception of the world and laid the foundation for advanced analytical methods like multi-view learning (Wen et al. 2023; Li et al. 2021; Wang et al. 2019; Wu et al. 2023; Zhuang et al. 2024; Lu et al. 2024a). This technology enhances information abundance by describing data in varied formats and views. In practical applications, multi-view data are collected and constructed dynamically,

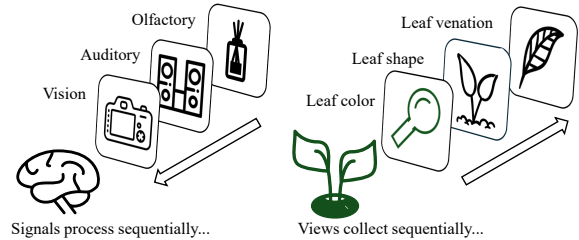


Figure 1: Sensory signals reach the brain sequentially, and in parallel with this, the view data increases dynamically.

especially in cutting-edge fields like brain-computer interfaces (Zhu et al. 2023; Kazi et al. 2022), which use multiple sensors, such as electrodes, to monitor neural signals and convert them into commands, quite similar to the complex way in which the human brain handles multiple data. As shown in Fig. 1, the human brain perceives and makes decisions by gathering information from diverse sensory channels, such as visual, olfactory, and auditory ones. Nevertheless, these sensory signals to the brain's processing center are not instantaneous (Pesnot Lerousseau et al. 2022). To compensate for this temporal disparity and seamlessly integrate the diverse streams of information, the brain deploys sophisticated mechanisms that intertwine feed-forward and feedback mechanisms (Hou et al. 2019; Zhang et al. 2022; Jiang et al. 2023; Wu et al. 2022). The brain continuously interacts with the external environment sequentially, constantly learning and optimizing its overall perception, thereby achieving a higher level of adaptability and intelligence (Chiel and Beer 1997; Ito et al. 2022).

However, traditional multi-view learning deals with scenarios where all data views are available in advance (Wan et al. 2024a; Lu et al. 2024b; Wan et al. 2023). Specifically, they prioritize learning new view data over the whole knowledge accumulation, thus lacking the adaptability and scalability to efficiently integrate sequentially acquired data from multiple views. In fact, they neglect the scenarios in which the human brain dynamically handles and conveys new view information, adapts, and optimizes overall cognition. This

*Corresponding Authors.

Copyright © 2025, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

contrasts sharply with the synaptic plasticity observed in the human brain. In neuroscience, synaptic strengthens or weakens itself based on excitation or inhibition between the pre- and post-neurons, which is known as Hebbian theory (Golkar et al. 2020; Triche, Maida, and Kumar 2022; Duan et al. 2023). As information streams in constantly, the feed-forward mechanism deftly extracts pertinent and consistent details, while leveraging the feedback mechanism to refine and rectify this information. This integrated approach fosters a profound interconnection between old and new knowledge, facilitating the comprehensive understanding and enhancing cognitive capabilities (Xiao et al. 2023; Shen et al. 2024; Zhang et al. 2019).

As mentioned above, in reality, views are generally added one by one, yet a view carries limited information, which makes it intractable to integrate view associations at a fine-grained scale to make the network stable when the domain changes. Several works have been done to tackle the dilemma in multi-view continual clustering (Wan et al. 2022; Yan et al. 2024; Wan et al. 2024b). Some methods learn a consensus weight matrix to preserve the old views' topology (Mitash, Bekris, and Boularias 2017; Cai et al. 2023). However, the views' inherent differences may cause the generated topology to be less comprehensive. And models are more inclined to have faith in past views which renders it harder to knowledge accumulation. Moreover, they focus on clustering and cannot tackle classification scenarios.

To tackle the relentless expansion and increasing complexity of data, multi-view continual learning has emerged (Sun et al. 2020; Wang et al. 2021). However, it primarily focuses on scenarios where new classes of data emerge, with each task encompassing more than one non-overlapping classes. It emphasizes the continual learning effects in different tasks, with the aim that decision-making for old tasks is not interfered with by new knowledge. However, it overlooks the scenarios of the growth of views to describe data. Thus we introduce multi-view incremental learning, which aims to gradually merge the view information to achieve a holistic knowledge of the data by utilizing the consistency and complementarity of multi-view data. Each task in this framework contains all classes of data, ensuring that no new classes emerge. Critically, each task presents the data from a unique view, so an ongoing fusion of complementary information is urgently needed. Specifically, we design a biologically plausible module named structured Hebbian plasticity, which facilitates the seamless integration of new incoming view data with existing view representation by reshaping the structure of weights. Besides, we devise a synaptic partition learning that depresses some of the synapses to minimize drastic changes in weights. The former reinforces synapses, while the latter suppresses them. The interaction between both mechanisms enables the model to effectively regulate weight changes, thereby mitigating the oblivion of knowledge. Our method offers an efficient solution for real-world applications. We conducted extensive experimental evaluations to demonstrate the superior performance of our method for node classification. The main contributions are as below:

- We discern the constraints inherent in traditional multi-view learning when confronted with the reality of increasing view data, i.e., the lack of ability to generalize to new data.
- We propose a bio-inspired multi-view incremental learning framework, which designs a structured Hebbian plasticity to reinforce views' associations for efficient knowledge integration of sequentially arriving data.
- The proposed method is employed in multi-view semi-supervised classification tasks, showcasing superior performance compared with other state-of-the-art graph-based algorithms.

Related Work

In this section, we present a concise introduction of the most pertinent work pertaining to our study in this paper, including multi-view learning and multi-view continual learning.

Multi-view Learning

Multi-view learning (Wang et al. 2015; Song et al. 2024; Pi et al. 2024) involves integrating and encoding information from multiple sources to derive a low-dimensional representation that captures both consistency and complementarity across views. It is worth noting that techniques from many emerging fields like graph learning (Chen et al. 2024a, 2023b; Wu et al. 2024; Wu, Zhang, and Fan 2023), have prompted the fast development of multi-view learning methods. By leveraging the ability to construct a comprehensive understanding of objects, multi-view learning provides versatile solutions for complex data analysis tasks in different domains, such as bioinformatics (Huang et al. 2023; Luo et al. 2018; Thammasorn et al. 2021; Yu et al. 2024), social network analysis (Lan et al. 2017; Chen et al. 2023c; Ma et al. 2017; Chen et al. 2024b), recommender systems (Zou et al. 2022; Wang et al. 2020; Cui et al. 2018). However, conventional multi-view learning methods presuppose that all views are constantly available. This poses a significant challenge: as data is continuously updated or augmented, the multi-view learning model necessitates retraining, significantly impeding efficiency. Therefore, to effectively accommodate the continually expanding number of view data, multi-view learning must possess the ability for continual knowledge retention and generalization.

Multi-view Continual Learning

Current multi-view continual learning has two dominant classifications: task-incremental and class-incremental types (Van de Ven, Tuytelaars, and Tolias 2022). Multi-view task-incremental learning (Li, Chandrasekaran, and Huan 2017; Sun et al. 2018) establishes connections between tasks and extracts consistent information across multiple views to complement each task learning, thus reducing the forgetting of old tasks as new ones arrive. Notably, task-incremental learning requires receiving prompts for new tasks, where one task contains multiple views; whereas our view incremental learning does not receive prompts for new tasks, and only one view is included in a task. However, in practice,

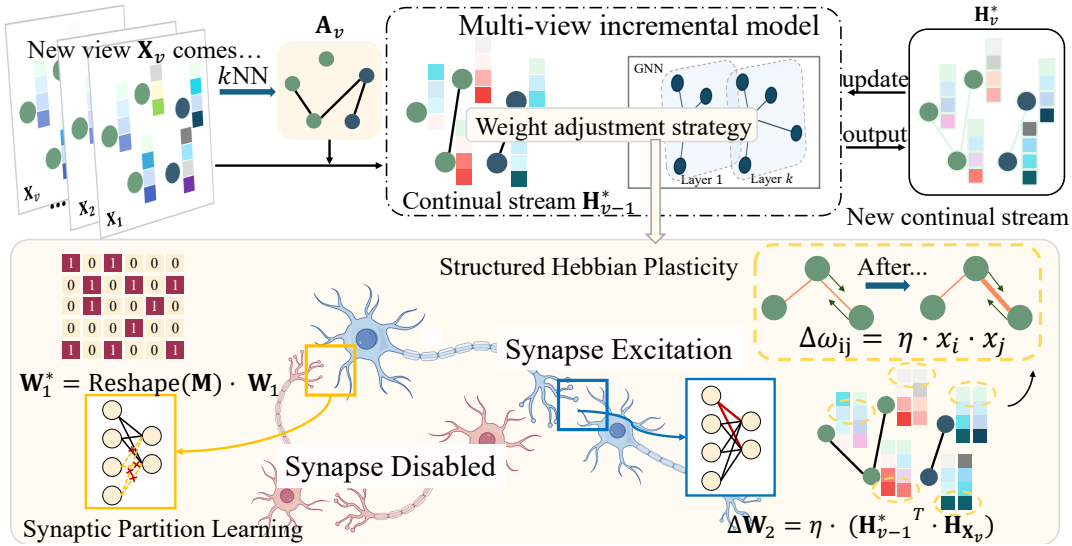


Figure 2: An overview of the proposed framework MVIL, which is drawn by Figdraw. MVIL introduces a weight adjustment strategy that incorporates structured Hebbian plasticity and synaptic partition learning, aiming to allow the model to learn with higher biological fidelity. The structured Hebbian plasticity reshapes the weight structure by reinforcing the consistent information between the old and new views; while the latter slows down weight changes by suppressing some of the synapses.

the prompts for new tasks are hard to distinguish. Consequently, detecting the task boundary becomes a formidable task. Thus, the multi-view class-incremental learning (Yang et al. 2022; Li et al. 2024) has been proposed, where new class data is seamlessly introduced without explicit prompting to the model, requiring the model to autonomously discern the onset of a new task. These methods focus on regularizing the update of weights, thus alleviating the forgetting of old tasks. Class-incremental learning assumes that data growth leads to new unknown classes, where a task contains more than one classes of data with multiple views, while our view incremental learning assumes that data growth leads to new descriptions, and a task includes all the classes from one view, with no class growth. We primarily emphasize overall cognitive capture, whereas continual learning exclusively centers on discrete task-specific decisions.

Preliminary

In this section, we start with the introduction of some basic notations used. Denote given multi-view feature matrices as

Notations	Descriptions
$\mathbf{X}_v \in \mathbb{R}^{n \times d_v}$	The v -th given multi-view feature matrix.
$\mathbf{A}_v \in \mathbb{R}^{n \times n}$	The similarity matrix evaluated from $\mathbf{X}^{(v)}$ by k NN.
$\mathbf{H}_v \in \mathbb{R}^{n \times c}$	The v th view initial representation.
$\mathbf{H}_v^* \in \mathbb{R}^{n \times c}$	The v th view representation after Hebb reinforcement.
$\mathbf{M} \in \mathbb{R}^{d_v \times d}$	The binary matrix of synaptic partition learning.
$\mathbf{W}_1 \in \mathbb{R}^{d_v \times d}$	The first layer weight.
$\mathbf{W}_1^* \in \mathbb{R}^{d_v \times d}$	The first layer weight after the partition mask.
$\mathbf{W}_2 \in \mathbb{R}^{d \times c}$	The second layer weight.
$\mathbf{W}_2^* \in \mathbb{R}^{d \times c}$	The second layer weight after Hebb reinforcement.

Table 1: The symbolic notations and their descriptions.

$\mathcal{X} = \{\mathbf{X}_v\}_{v=1}^V$, where $\mathbf{X}_v = [\mathbf{x}_1; \dots; \mathbf{x}_n] \in \mathbb{R}^{n \times d_v}$ is the data from the v th view for any $v \in \{1, \dots, V\}$. n represents the quantity of samples, and V signifies the count of views. d_v indicates the distinct dimensions of the feature matrices, and c denotes the number of classes. Furthermore, to better understand the mathematical notation usages which are primarily used, we list the notation explanations in Table 1.

Methodology

As previously mentioned, the major existing methods overlook the scenarios where data observations of new views accumulate gradually over time. Consequently, we have devised a multi-view incremental framework MVIL specifically tailored for scenarios where continuously expanding new view data. This framework mimics the human brain feed-forward and feedback mechanisms to empower the model to adapt to the ever-growing amount of view data, thereby achieving knowledge accumulation. Fig. 2 provides a comprehensive depiction of our proposed method. Notably, we have incorporated a biologically inspired weight adjustment strategy that integrates structured Hebbian plasticity and synaptic partition learning.

Streaming View Representation Learning

Firstly, we design a streaming graph learning model wherein upon the arrival of new view data \mathbf{X}_v , and construct the corresponding adjacency \mathbf{A}_v by applying k -Nearest Neighbor (k NN). When the first view data arrives, we utilize a two-layer GCN with shared weights, which can be represented as,

$$\mathbf{H}_1^* = \hat{\mathbf{A}}_1 \sigma(\hat{\mathbf{A}}_1 \mathbf{X}_1 \mathbf{W}_1^*) \mathbf{W}_2. \quad (1)$$

$\hat{\mathbf{A}}_1 = \tilde{\mathbf{D}}_1^{-\frac{1}{2}} \tilde{\mathbf{A}}_1 \tilde{\mathbf{D}}_1^{-\frac{1}{2}}$ and $\tilde{\mathbf{A}}_1 = \mathbf{A}_1 + \mathbf{I}$. $\mathbf{A}_1 \in \mathbb{R}^{n \times n}$ corresponds to an adjacency matrix constructed by \mathbf{X}_1 . Herein,

$\mathbf{W}_1^* \in \mathbb{R}^{d_v \times d}$ and $\mathbf{W}_2 \in \mathbb{R}^{d \times c}$ denote the learnable shared weight matrices, which would be shared by all the view. At this stage, the initialization of the streaming model has been completed. Streaming models continually retain the currently deposited representations without storing the entered data, thus effectively enhancing the privacy of the model.

When the new view data \mathbf{X}_v comes, we introduce \mathbf{X}_v into the continual stream model and initially integrate the previous knowledge. As the view v increases, the representation \mathbf{H}_v can be expressed as,

$$\mathbf{H}_v = \underbrace{\hat{\mathbf{A}}_v \sigma(\hat{\mathbf{A}}_v \mathbf{X}_v \mathbf{W}_1^*)}_{\text{new data stream}} \mathbf{W}_2 + \underbrace{\alpha \mathbf{H}_{v-1}^*}_{\text{incremental stream}}, \quad (2)$$

where $v \in \{2, \dots, V\}$. $\hat{\mathbf{A}}_v = \tilde{\mathbf{D}}_v^{-\frac{1}{2}} \tilde{\mathbf{A}}_v \tilde{\mathbf{D}}_v^{-\frac{1}{2}}$ and $\tilde{\mathbf{A}}_v = \mathbf{A}_v + \mathbf{I}$. $\mathbf{A}_v \in \mathbb{R}^{n \times n}$ corresponds to the v -th adjacency matrix constructed by particular \mathbf{X}_v . $\mathbf{H}_{v-1}^* \in \mathbb{R}^{n \times c}$ denotes the representation of the past views when the view v has not been increased. α refers to a learnable parameter to balance the weight of the new view data with the past knowledge. As new data continues to arrive, the streaming model continuously assimilates new knowledge, integrating representation learning. However, simplistic fusion approaches fail to capture the intricate interactions among continual views, and the challenge of knowledge oblivion caused by the incremental addition of views remains unresolved.

Structured Hebbian Plasticity

Inspired by biological neurology, we firstly introduce the Hebbian learning (Golkar et al. 2020; Triche, Maida, and Kumar 2022) into multi-view graph representation learning. The core of Hebbian’s theory is that continuous repetitive stimulation of pre-synaptic and post-synaptic neurons results in an increase in synaptic transmission efficacy. Thus, we design the structured Hebbian plasticity, which facilitates the selective strengthening or weakening of connections based on structured patterns or relations in the data, enhancing the domain’s generalization ability.

Specifically, the Hebbian learning rule updates the connection weights between two neurons based on the correlation between them. If neuron i and neuron j are active at the same moment, then the connection weight between them will increase. Therefore, the rule can be expressed as:

$$\Delta w_{ij} = \eta \cdot x_i \cdot x_j, \quad (3)$$

where Δw_{ij} indicates the amount of change in weights, η is the learning rate, x_i and x_j are the outputs of neuron i and neuron j .

Instead of uniformly treating all weights, structured Hebbian plasticity selectively identifies biologically plausible node representations based on the similarity between representations across different views, thereby optimizing specific weights, enhancing learning efficiency, and improving generalization performance. Simultaneous activation of the front and back neurons strengthens the weights connecting this pair of neurons, which coincides with a plausible fusion of the past and new representations. This happens to be available for enhanced fine-grained representation; if the

past and new representations stimulate similar node representations at the same time, we enhance this weight to emphasize the consistent information. Therefore, we apply it to the weight \mathbf{W}_2 .

$$\Delta \mathbf{W}_2 = \eta \cdot \mathbf{H}_{new}^T \mathbf{H}_{past}, \quad (4)$$

where $\mathbf{H}_{new}^T \in \mathbb{R}^{n \times d}$ denotes the new view data and $\mathbf{H}_{past} \in \mathbb{R}^{n \times c}$ refers to the past representation. Here, we develop Eq. (4) in more detail, and update \mathbf{W}_2^* by using the Hebbian’s rule:

$$\mathbf{W}_2^* = \mathbf{W}_2 + \varepsilon (\hat{\mathbf{A}}_v \mathbf{X}_v \mathbf{W}_1)^T \mathbf{H}_{v-1}, \quad (5)$$

where ε refers to the learning rate. The latter item is a detailed expansion of Eq. (4). \mathbf{W}_2^* emphasizes the consistent information between old and new representations. Herein, \mathbf{W}_2^* reconstruct the weight through Hebbian’s rule, which is more suitable for the intermingling of old and new knowledge. Then, it acts on forward propagation to acquire new representation and can be denoted as,

$$\mathbf{H}_v^* = \mathbf{H}_v + \varepsilon (\hat{\mathbf{A}}_v \mathbf{X}_v \mathbf{W}_1) \mathbf{W}_2^*, \quad (6)$$

where $\mathbf{H}_v^* \in \mathbb{R}^{n \times c}$ is the final fused representation after structured reinforcement. By mining the inherent structure in the multi-view data, structured Hebbian plasticity enables more efficient learning, and improves domain generalization.

Synaptic Partition Learning

In the process of biological neural networks, emphasizing only relevant synaptic connections without suppressing irrelevant synaptic connections may result in the model being interfered with by a large amount of noise and redundant information. Thus, we design a synaptic partition learning mechanism, which focuses on partitioning synapses into distinct groups or partitions and independently updating the synaptic weights within each partition. We adopt it to enable more efficient and targeted learning by selectively updating subsets of synapses.

Specifically, we first define a one-dimensional matrix of the same size as the weight \mathbf{W}_1 and mask it randomly, ensuring that the number of places 1 satisfies a given ratio θ , as shown below,

$$\mathbf{M} = \mathbf{M} \cdot \mathbf{S}, \quad (7)$$

where $\mathbf{M} \in \mathbb{R}^{nd \times 1}$ is an all-1 matrix and $\mathbf{S} \in \mathbb{R}^{nd \times 1}$ a randomly initialized matrix for each epoch. Here, the number of elements in \mathbf{S} that have 1 value is no more than $\lfloor \theta \cdot \mathbf{M.size}() \rfloor$. We then partially disable the neurons with randomly initialized \mathbf{M} . As we can see, \mathbf{W}_1 can be represented as,

$$\mathbf{W}_1^* = \text{Reshape}(\mathbf{M}) \cdot \mathbf{W}_1. \quad (8)$$

$\text{Reshape}(\mathbf{M}) \in \mathbb{R}^{d_v \times d}$ is a function to adjust the dimension of the weight \mathbf{W}_1 . In each round of training, we set $\theta \ll 1/V$ and use θ to randomize the setup of \mathbf{S} . Instead of globally adjusting all synapses simultaneously, synaptic partition learning allows for local adjustments within specific subsets of synapses, potentially leading to faster convergence, improved generalization, and reduced susceptibility to oblivion of knowledge.

In summary, the synaptic partitioning learning module aims to eliminate irrelevant connections by inhibiting some synapses, thus reducing the interference of redundant information. While the structured Hebbian plasticity module focuses on activating highly correlated synapses to facilitate meaningful information transfer. The synergistic effect of these two modules can improve the generalization ability of the model.

Loss Function

Moreover, our model attempts to regulate the entire activity level of neurons to keep it within a certain range. Thus, we adopt a loss function to penalize shifting weights and thus mitigate large shifts in weights. If the activity of certain neurons begins to change excessively, this loss works to restore balance, i.e.,

$$\mathcal{L}_{RE} = \frac{1}{2} \sum_{i=1}^2 \|\mathbf{W}_i - (\mathbf{W}_i)_{old}\|_F^2, \quad (9)$$

where $(\mathbf{W}_1)_{old}$ and $(\mathbf{W}_2)_{old}$ are the weights of the previous representation. We expect the two shared weights \mathbf{W}_1 and \mathbf{W}_2 not to fluctuate unduly, thus maintaining the holistic memory of the old knowledge.

Besides, for semi-supervised node classification, we calculate CrossEntropy loss and update parameters in our method, as follows

$$\mathcal{L}_{CE} = - \sum_{i \in \Omega} \sum_{j=1}^c y_{ij} \ln \hat{y}_{ij}, \quad (10)$$

where Ω refers to the index set of labeled samples, $\hat{y} = \text{softmax}(\mathbf{H}_v^*)$ and c is the amount of classes. The final loss can be expressed as

$$\mathcal{L}_{all} = \mathcal{L}_{CE} + \beta \mathcal{L}_{RE}. \quad (11)$$

\mathcal{L}_{all} is the final loss for our method, which consists of two components. β is a parameter that resizes the loss \mathcal{L}_{RE} . We conclude the training procedure of MVIL in Algorithm 1.

Experiments

To further validate the effectiveness of the proposed method, we have designed the experimental section intended to answer the following key evaluation questions (EQs):

- **EQ1** Does MVIL achieve superior performance compared to its competitors for the semi-supervised node classification task?
- **EQ2** Does MVIL capture the tight association between old and new views when new ones add up, and does it promote learning about the whole knowledge?

MVIL has performed ablation experiments, parameter sensitivity analysis, and t-Sne demonstration in the Appendix.

Datasets & Compared Methods

Datasets We adopt six multi-view graph datasets widely used in different domains to evaluate the performance of MVIL compared to state-of-the-art baselines. Table 2 illustrates the details of all six datasets.

Algorithm 1: MVIL

Require: Multi-view feature matrices $\{\mathbf{X}\}_{v=1}^V$, label matrix $\mathbf{Y} \in \mathbb{R}^{|\Omega| \times c}$, hyperparameter k in k NN.
Ensure: Predictive output \mathbf{H}_v^* .
1: **for** new view data \mathbf{X}_v comes... **do**
2: **while** not convergent **do**
3: Construct the adjacency \mathbf{A}_v by utilizing k NN;
4: Calculate the weight \mathbf{W}_1^* by Eq. (8);
5: Calculate \mathbf{H}_v by Eq. (2);
6: Update the \mathbf{W}_2^* by Eq. (5);
7: Obtain the final representation \mathbf{H}_v^* by Eq. (6);
8: Compute the loss \mathcal{L}_{RE} and \mathcal{L}_{CE} by Eqs. (9) and (10);
9: Optimize the trainable parameters from network by backward propagation;
10: **end while**
11: **end for**
12: Obtain the predictive output \mathbf{H}_v^* ;
13: **return** Predictive output \mathbf{H}_v^* .

Datasets	Views	Samples	Features	Classes	Data Types
100leaves	3	1,600	64	100	plant species
Animals	2	10,158	4,096	50	images
Flower17	7	1,360	7	17	images
NGs	3	500	2,000	5	documents
NoisyMNIST_15000	2	15,000	784	10	images
YaleB_Extended	5	2,424	1,024	38	images

Table 2: A brief description of the test multi-view datasets.

Compared Methods We compare the proposed method with classical and state-of-the-art baseline models, including GAT (Veličković et al. 2018), SI (Zenke, Poole, and Ganguli 2017), MAS (Aljundi et al. 2018), LGCNFF (Chen et al. 2023a), IHGCN (Zou et al. 2024), TMC (Han et al. 2020) and MVCIL (Li et al. 2024).

Experimental Setting

Experimental rounds of the proposed method are 600 per view under 10% randomly labeled data and the average value is taken by repeating the run 3 times. Four well-known classification evaluation metrics are applied to our experiments, including Accuracy (ACC), Precision (P), Recall (R), and Macro-averaged F1 Score (MAF1). ACC is defined as the number of correctly classified samples divided by the number of samples. P reflects how accurate the model is in predicting the positive class while R is the ability to recognize samples in the positive class. MAF1 is the arithmetic mean of the F1 values over all classes. The values of these metrics fall within the range of [0, 1], where higher values indicate better performance.

Comparison to SOTA (EQ1)

We next evaluate the effectiveness of our method on the node classification task compared with several classical and state-of-the-art methods in Table 3, where the best performance is highlighted in bold and the second-best results are un-

Classification		Static Multi-view Learning			Multi-view Incremental Learning				
Datasets	Metric	TMC	LGCNFF	IHGCN	GAT	SI	MAS	MVCIL	MVIL
100leaves	ACC	79.20 (0.08)	87.41 (0.62)	91.03 (0.70)	43.52 (0.94)	43.25 (0.25)	43.46 (0.32)	69.50 (0.43)	91.50 (0.47)
	P	79.20 (0.08)	88.76 (0.83)	92.29 (0.53)	42.36 (1.64)	45.16 (1.18)	45.73 (1.30)	70.31 (0.06)	92.55 (0.46)
	R	80.56 (0.07)	87.41 (0.62)	91.03 (0.70)	43.52 (0.94)	43.25 (0.25)	43.46 (0.32)	69.50 (0.43)	91.50 (0.47)
	MAF1	78.50 (0.09)	86.58 (0.56)	90.76 (0.80)	40.27 (0.96)	40.58 (0.24)	40.89 (0.24)	67.87 (0.30)	91.33 (0.48)
Animals	ACC	81.35 (0.53)	74.15 (1.32)	83.13 (0.04)	39.20 (0.36)	45.86 (0.14)	46.29 (0.16)	67.03 (0.03)	84.42 (0.10)
	P	74.23 (0.65)	69.07 (1.71)	80.46 (0.35)	34.18 (1.33)	39.92 (0.18)	39.89 (0.28)	64.86 (0.10)	81.59 (0.31)
	R	79.19 (1.01)	65.49 (1.46)	76.17 (0.04)	32.45 (0.44)	38.15 (0.04)	38.59 (0.18)	60.53 (0.03)	78.10 (0.19)
	MAF1	74.65 (0.76)	65.20 (1.59)	76.82 (0.03)	31.53 (0.69)	37.03 (0.13)	37.19 (0.02)	61.05 (0.00)	78.32 (0.14)
Flower17	ACC	52.44 (0.63)	45.35 (4.20)	58.94 (0.17)	17.78 (1.56)	44.81 (0.20)	47.10 (0.20)	40.93 (1.63)	60.54 (0.72)
	P	52.44 (0.63)	55.44 (6.80)	62.43 (0.27)	10.53 (3.75)	43.26 (0.22)	45.82 (0.33)	42.01 (1.62)	61.98 (1.01)
	R	54.06 (0.48)	45.32 (4.20)	58.94 (0.17)	17.78 (1.56)	44.81 (0.20)	47.10 (0.20)	40.93 (1.63)	60.54 (0.72)
	MAF1	50.16 (0.74)	39.75 (5.54)	59.14 (0.20)	10.19 (2.77)	42.86 (0.22)	45.61 (0.31)	38.53 (2.18)	60.22 (0.70)
NGs	ACC	93.11 (0.01)	90.65 (1.27)	96.89 (0.14)	72.52 (4.74)	79.11 (1.33)	83.56 (0.67)	58.89 (7.11)	95.04 (0.64)
	P	93.11 (0.01)	92.02 (0.66)	96.89 (0.14)	75.00 (4.66)	82.56 (0.07)	84.13 (0.62)	60.38 (8.56)	95.16 (0.60)
	R	93.14 (0.01)	90.67 (1.27)	96.89 (0.14)	72.52 (4.74)	79.11 (1.33)	83.56 (0.67)	58.89 (7.11)	95.04 (0.64)
	MAF1	93.03 (0.01)	90.57 (1.23)	96.87 (0.14)	72.50 (5.12)	79.31 (1.30)	83.60 (0.64)	58.81 (7.67)	95.05 (0.64)
NoisyMNIST_15000	ACC	87.01 (0.35)	89.78 (0.48)	93.64 (0.46)	73.43 (0.61)	82.74 (2.10)	90.02 (0.29)	90.79 (0.47)	92.43 (0.64)
	P	87.03 (0.26)	89.75 (0.47)	93.61 (0.47)	74.02 (0.87)	85.44 (0.97)	90.20 (0.04)	90.64 (0.46)	92.55 (0.49)
	R	86.59 (0.37)	89.58 (0.49)	93.51 (0.48)	72.89 (0.63)	82.45 (2.07)	89.80 (0.30)	90.60 (0.47)	92.29 (0.62)
	MAF1	86.52 (0.40)	89.58 (0.49)	93.53 (0.47)	72.43 (0.71)	82.25 (1.96)	89.83 (0.28)	90.59 (0.48)	92.30 (0.62)
YaleB_Extended	ACC	50.59 (1.10)	34.01 (1.43)	47.69 (0.36)	31.19 (0.82)	32.76 (0.25)	32.86 (0.16)	64.16 (0.82)	66.99 (0.28)
	P	81.62 (0.84)	48.32 (6.29)	75.85 (0.17)	60.14 (4.80)	36.12 (0.33)	36.98 (0.45)	67.45 (0.97)	71.06 (0.23)
	R	50.61 (1.10)	33.99 (1.40)	47.69 (0.36)	31.21 (0.86)	32.83 (0.25)	32.92 (0.15)	64.24 (0.82)	67.07 (0.28)
	MAF1	58.13 (0.77)	33.24 (3.38)	53.78 (0.26)	34.01 (1.87)	33.51 (0.23)	33.85 (0.00)	64.86 (0.81)	67.70 (0.27)

Table 3: Node classification performance with various algorithms.

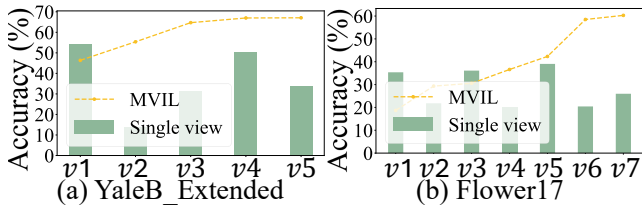


Figure 3: A comparison of MVIL streaming input and single-view learning performance.

derlined. We divide the compared methods into two classes: static multi-view learning and view incremental learning.

Static Multi-view Learning We compare it with the leading multi-view methods, where all view information is collected completely and then trained on the left side of Table 3. The three compared methods, TMC, LGCNFF, and IHGCN simultaneously obtain all the views and can be well mined for correlations between the views. TMC skillfully blends credible evidence of view judgment by determining the confidence level of view. In contrast, LGCNFF and IHGCN achieve very impressive performance by capturing consistent and complementary information of multi-view data. In particular, static multi-view learning obtains leading performance in datasets with a small number of views and excellent view quality; in the face of datasets with a large number of views, it is very challenging to extract the overall cognition and determine the view quality.

Multi-view Incremental Learning First, we introduce GAT into the view incremental framework and we find that GAT is explicitly failing to memorize the old view informa-

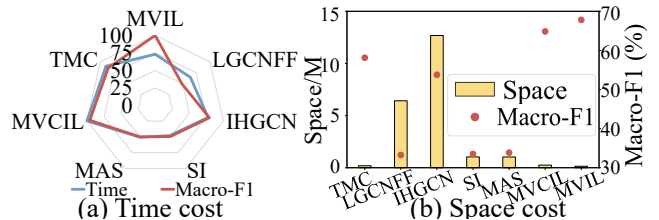


Figure 4: The time-space cost of all compared algorithms.

tion. Their performance is intricately linked to the quality of the most recent view data rather than the integration of the overall information. We then compare this to the continual learning approaches associated with synaptic plasticity. SI and MAS express the role of synaptic plasticity in capturing overall information, but it is not easy to reconcile the poor quality of views and the large number of categories. Moreover, we compare it with relevant methods for multi-view class-incremental learning MVCIL. MVCIL shows excellent performance in learning datasets with multiple view counts due to its ability to memorize old knowledge well.

From Table 3, it is evident that the proposed method has achieved a significant leading position. More specifically, MVIL improves over the suboptimal method by **1.6%** on the Flower17 dataset and **2.83%** on the YaleB_Extended dataset. This illustrates that even with non-static multi-view data, we are able to capture a comprehensive representation of them as the number of views increases. Incidentally, it reaffirms the ability of our model to extract maximum information about the whole view and fuse the ever-growing view data in a comprehensive and efficient manner.

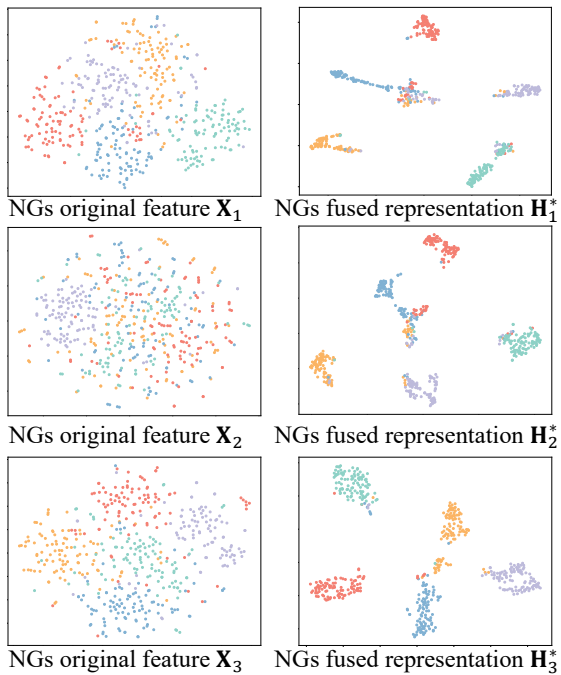


Figure 5: The initial features and learned fused representations for each view of the NGs.

Time and Space Cost To validate the low energy consumption and high efficiency of MVIL, we plotted the time-space cost shown in Fig. 4. Herein, Fig. 4(a) illustrates all methods’ time (as a percentage of the longest time) and performance (as a percentage of the highest performance), while Fig. 4(b) depicts space cost and performance. Our MVIL utilizes very few parameters and incurs minimal time costs while delivering leading performance.

Knowledge Accumulation (EQ2-1)

We plot the change in overall model performance as the views continue to increase, and add the results of training with only one view as a comparison, in Fig. 3. It is evident that the performance of our MVIL improves as the number of views increases, surpassing the performance achieved with single-view learning. However, partial single-view learning sometimes outperforms MVIL. Our synaptic partitioning learning module restricts neuronal activity, which helps retain old knowledge but can also limit performance. This not only confirms the effectiveness of the multi-view incremental strategy but also clearly indicates its ability to outperform traditional single-view learning paradigms. Specifically, by fusing the complementary information embedded in multiple views, the model is able to understand the data features more comprehensively, which in turn leads to more accurate and robust prediction or classification. This process not only validates our model’s superior ability to accumulate and integrate knowledge but also demonstrates its ability to retain old knowledge while continuously absorbing and utilizing new knowledge from additional views.

Knowledge Leads to Learning (EQ2-2)

To further validate the ability of our MVIL to integrate information from all views, we plot the initial features and the learned fused representation for each new view data, as shown in Fig. 5. In the original feature \mathbf{X}_1 , the inter-class and intra-class distances of the initial features are almost equal, yet it is fortunate that the intra-class nodes are clustered together. Thus \mathbf{H}_1^* is well differentiated. When view data \mathbf{X}_2 comes in, the features are muddled and the same class of nodes is scattered, because of the favorable guiding effect of \mathbf{H}_1^* , the inter-class distance of \mathbf{H}_2^* is still large, but the intra-class nodes become slightly decentralized. When view data \mathbf{X}_3 enters, the features are sharply differentiated, and with the guiding effect of \mathbf{H}_2^* , the nodes in the middle of the two classes are pulled back into the class.

Due to the consideration of fine-grained correlations between nodes, the old representations can still guide the representation learning well even if the new view data is scattered, and the learned representations adaptively fuse the merits of the two. This adequately illustrates the importance of structured Hebbian plasticity. When the new incoming view data is complementary to the old knowledge, it learns a well orthogonal representation to accommodate both; when the new data is not well differentiated, the module takes full advantage of the guidance of the old view in order to learn serviceable representations from it. We capitalize on the association between the old and new views, thus mitigating the effects of poor view quality on a single view. It not only enhances the accuracy in understanding the overall view but continuously improves MVIL’s perceptual effectiveness.

Conclusion

Traditional multi-view learning paradigms are constrained by static multi-view data, impeding their ability to effectively accommodate the dynamic and ever-expanding amount of view data encountered in real-world settings. In this study, drawing inspiration from biological neurology, we proposed a multi-view incremental learning framework named MVIL. Therein, the underlying problem structure remains the same, but the input distribution changes. Different from continual learning, multi-view incremental learning prioritizes overall cognition when interacting with streaming data and each task includes only one view. Our method first applied Hebbian learning principles to multi-view graph representation learning. We devised a weight adjustment strategy that includes structured Hebbian plasticity and synaptic partitioning learning, which fine-grain fused representations thereby better mimicking the brain’s cognitive decision-making. Experimental evaluations demonstrate the ability of the proposed MVIL to effectively fine-grained fusion of ever-increasing view representations.

Acknowledgments

This work is in part supported by the National Key Research and Development Plan of China under Grant 2021YFB3600503 and National Natural Science Foundation of China under Grants No.U21A20472, No.62276065, No.62204140, No.62476035 and No.62206037.

References

- Aljundi, R.; Babiloni, F.; Elhoseiny, M.; Rohrbach, M.; and Tuytelaars, T. 2018. Memory aware synapses: Learning what (not) to forget. In *ECCV*, 139–154.
- Cai, H.; Tan, Y.; Huang, S.; and Lv, J. 2023. Lifelong multi-view spectral clustering. In *IJCAI*, 3488–3496.
- Chen, Y.; Chen, F.; Wu, Z.; Chen, Z.; Cai, Z.; Tan, Y.; and Wang, S. 2024a. Heterogeneous Graph Embedding with Dual Edge Differentiation. *Neural Networks*, 106965.
- Chen, Z.; Fu, L.; Yao, J.; Guo, W.; Plant, C.; and Wang, S. 2023a. Learnable graph convolutional network and feature fusion for multi-view learning. *Information Fusion*, 95: 109–119.
- Chen, Z.; Wu, Z.; Lin, Z.; Wang, S.; Plant, C.; and Guo, W. 2023b. AGNN: Alternating graph-regularized neural networks to alleviate over-smoothing. *IEEE Transactions on Neural Networks and Learning Systems*, 35(10): 13764–13776.
- Chen, Z.; Wu, Z.; Wang, S.; and Guo, W. 2023c. Dual low-rank graph autoencoder for semantic and topological networks. In *AAAI*, volume 37, 4191–4198.
- Chen, Z.; Wu, Z.; Zhong, L.; Plant, C.; Wang, S.; and Guo, W. 2024b. Attributed Multi-order Graph Convolutional Network for Heterogeneous Graphs. *Neural Networks*, 174: 106225.
- Chiel, H. J.; and Beer, R. D. 1997. The brain has a body: adaptive behavior emerges from interactions of nervous system, body and environment. *Trends in neurosciences*, 20(12): 553–557.
- Cui, Q.; Wu, S.; Liu, Q.; Zhong, W.; and Wang, L. 2018. MV-RNN: A multi-view recurrent neural network for sequential recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 32(2): 317–331.
- Duan, Y.; Jia, Z.; Li, Q.; Zhong, Y.; and Ma, K. 2023. Hebbian and gradient-based plasticity enables robust memory and rapid learning in rnns. In *ICLR*.
- Golkar, S.; Lipshutz, D.; Bahroun, Y.; Sengupta, A.; and Chklovskii, D. 2020. A simple normative network approximates local non-Hebbian learning in the cortex. In *NeurIPS*, volume 33, 7283–7295.
- Han, Z.; Zhang, C.; Fu, H.; and Zhou, J. T. 2020. Trusted multi-view classification. In *ICLR*.
- Hou, H.; Zheng, Q.; Zhao, Y.; Pouget, A.; and Gu, Y. 2019. Neural correlates of optimal multisensory decision making under time-varying reliabilities with an invariant linear probabilistic population code. *Neuron*, 104(5): 1010–1021.
- Huang, S.; Xiao, S.; Liu, W.; Lu, J.; Wu, Z.; Wang, S.; and Rajapakse, J. C. 2023. Multi-level Knowledge Integration with Graph Convolutional Network for Cancer Molecular Subtype Classification. In *BIBM*, 1983–1988.
- Ito, T.; Yang, G. R.; Laurent, P.; Schultz, D. H.; and Cole, M. W. 2022. Constructing neural network models from brain data reveals representational transformations linked to adaptive behavior. *Nature Communications*, 13(1): 673.
- Jiang, T.; Xu, Q.; Ran, X.; Shen, J.; Lv, P.; Zhang, Q.; and Pan, G. 2023. Adaptive deep spiking neural network with global-local learning via balanced excitatory and inhibitory mechanism. In *ICLR*.
- Kazi, A.; Markova, V.; Kondamadugula, P. R.; Liu, B.; Adly, A.; Faghihroohi, S.; and Navab, N. 2022. DG-GRU: dynamic graph based gated recurrent unit for age and gender prediction using brain imaging. In *Medical Imaging 2022: Computer-Aided Diagnosis*, volume 12033, 291–295.
- Lan, C.; Yang, Y.; Li, X.; Luo, B.; and Huan, J. 2017. Learning social circles in ego-networks based on multi-view network structure. *IEEE Transactions on Knowledge and Data Engineering*, 29(8): 1681–1694.
- Li, D.; Wang, T.; Chen, J.; Kawaguchi, K.; Lian, C.; and Zeng, Z. 2024. Multi-view class incremental learning. *Information Fusion*.
- Li, X.; Chandrasekaran, S. N.; and Huan, J. 2017. Lifelong multi-task multi-view learning using latent spaces. In *IEEE Big Data*, 37–46.
- Li, Z.; Tang, C.; Liu, X.; Zheng, X.; Zhang, W.; and Zhu, E. 2021. Consensus graph learning for multi-view clustering. *IEEE Transactions on Multimedia*, 24: 2461–2472.
- Lu, J.; Wu, Z.; Chen, Z.; Cai, Z.; and Wang, S. 2024a. Towards Multi-view Consistent Graph Diffusion. In *ACM Multimedia*, 186–195.
- Lu, J.; Wu, Z.; Zhong, L.; Chen, Z.; Zhao, H.; and Wang, S. 2024b. Generative Essential Graph Convolutional Network for Multi-View Semi-Supervised Classification. *IEEE Transactions on Multimedia*, 26: 7987–7999.
- Luo, Y.; Song, G.; Li, P.; and Qi, Z. 2018. Multi-task medical concept normalization using multi-view convolutional neural network. In *AAAI*, volume 32.
- Ma, G.; Lu, C.-T.; He, L.; Philip, S. Y.; and Ragin, A. B. 2017. Multi-view graph embedding with hub detection for brain network analysis. In *ICDM*, 967–972.
- Mitash, C.; Bekris, K. E.; and Boularias, A. 2017. A self-supervised learning system for object detection using physics simulation and multi-view pose estimation. In *IROS*, 545–551.
- Pesnot Lerousseau, J.; Parise, C. V.; Ernst, M. O.; and van Wassenhove, V. 2022. Multisensory correlation computations in the human brain identified by a time-resolved encoding model. *Nature Communications*, 13(1): 2489.
- Pi, Y.; Wu, Y.; Huang, Y.; Shi, Y.; and Wang, S. 2024. Inhomogeneous Diffusion-Induced Network for Multiview Semi-Supervised Classification. *IEEE Transactions on Neural Networks and Learning Systems*.
- Shen, J.; Ni, W.; Xu, Q.; and Tang, H. 2024. Efficient spiking neural networks with sparse selective activation for continual learning. In *AAAI*, volume 38, 611–619.
- Song, N.; Du, S.; Wu, Z.; Zhong, L.; Yang, L. T.; Yang, J.; and Wang, S. 2024. GAF-Net: Graph attention fusion network for multi-view semi-supervised classification. *Expert Systems with Applications*, 238: 122151.

- Sun, G.; Cong, Y.; Li, J.; and Fu, Y. 2018. Robust lifelong multi-task multi-view representation learning. In *ICBK*, 91–98.
- Sun, G.; Cong, Y.; Zhang, Y.; Zhao, G.; and Fu, Y. 2020. Continual multiview task learning via deep matrix factorization. *IEEE Transactions on Neural Networks and Learning Systems*, 32(1): 139–150.
- Thammasorn, P.; Chaovalitwongse, W. A.; Hippe, D. S.; Wootton, L. S.; Ford, E. C.; Spraker, M. B.; Combs, S. E.; Peeken, J. C.; and Nyflot, M. J. 2021. Nearest neighbor-based strategy to optimize multi-view triplet network for classification of small-sample medical imaging data. *IEEE Transactions on Neural Networks and Learning Systems*, 34(2): 586–600.
- Triche, A.; Maida, A. S.; and Kumar, A. 2022. Exploration in neo-Hebbian reinforcement learning: Computational approaches to the exploration–exploitation balance with bio-inspired neural networks. *Neural Networks*, 151: 16–33.
- Van de Ven, G. M.; Tuytelaars, T.; and Tolias, A. S. 2022. Three types of incremental learning. *Nature Machine Intelligence*, 4(12): 1185–1197.
- Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Liò, P.; and Bengio, Y. 2018. Graph Attention Networks. In *ICLR*.
- Wan, X.; Liu, J.; Liang, W.; Liu, X.; Wen, Y.; and Zhu, E. 2022. Continual multi-view clustering. In *ACM Multimedia*, 3676–3684.
- Wan, X.; Liu, J.; Liu, X.; Wen, Y.; Yu, H.; Wang, S.; Yu, S.; Wan, T.; Wang, J.; and Zhu, E. 2024a. Decouple then Classify: A Dynamic Multi-view Labeling Strategy with Shared and Specific Information. In *ICML*.
- Wan, X.; Liu, X.; Liu, J.; Wang, S.; Wen, Y.; Liang, W.; Zhu, E.; Liu, Z.; and Zhou, L. 2023. Auto-weighted multi-view clustering for large-scale data. In *AAAI*, 10078–10086.
- Wan, X.; Xiao, B.; Liu, X.; Liu, J.; Liang, W.; and Zhu, E. 2024b. Fast continual multi-view clustering with incomplete views. *IEEE Transactions on Image Processing*.
- Wang, M.; Lin, Y.; Lin, G.; Yang, K.; and Wu, X.-m. 2020. M2GRL: A multi-task multi-view graph representation learning framework for web-scale recommender systems. In *SIGKDD*, 2349–2358.
- Wang, Q.; Sun, G.; Dong, J.; Wang, Q.; and Ding, Z. 2021. Continuous Multi-View Human Action Recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(6): 3603–3614.
- Wang, R.; Nie, F.; Wang, Z.; Hu, H.; and Li, X. 2019. Parameter-free weighted multi-view projected clustering with structured graph learning. *IEEE Transactions on Knowledge and Data Engineering*, 32(10): 2014–2025.
- Wang, W.; Arora, R.; Livescu, K.; and Bilmes, J. 2015. On deep multi-view representation learning. In *ICML*, 1083–1092.
- Wen, Y.; Liu, S.; Wan, X.; Wang, S.; Liang, K.; Liu, X.; Yang, X.; and Zhang, P. 2023. Efficient Multi-View Graph Clustering with Local and Global Structure Preservation. In *ACM Multimedia*, 3021–3030.
- Wu, Y.; Zhao, R.; Zhu, J.; Chen, F.; Xu, M.; Li, G.; Song, S.; Deng, L.; Wang, G.; Zheng, H.; et al. 2022. Brain-inspired global-local learning incorporated with neuromorphic computing. *Nature Communications*, 13(1): 65.
- Wu, Z.; Chen, Z.; Du, S.; Huang, S.; and Wang, S. 2024. Graph Convolutional Network with elastic topology. *Pattern Recognition*, 151: 110364.
- Wu, Z.; Lin, X.; Lin, Z.; Chen, Z.; Bai, Y.; and Wang, S. 2023. Interpretable graph convolutional network for multi-view semi-supervised learning. *IEEE Transactions on Multimedia*, 25: 8593–8606.
- Wu, Z.; Zhang, Z.; and Fan, J. 2023. Graph Convolutional Kernel Machine versus Graph Convolutional Networks. In *NeurIPS*, volume 36, 19650–19672.
- Xiao, M.; Meng, Q.; Zhang, Z.; He, D.; and Lin, Z. 2023. Hebbian Learning based Orthogonal Projection for Continual Learning of Spiking Neural Networks. In *ICLR*.
- Yan, X.; Gan, Y.; Mao, Y.; Ye, Y.; and Yu, H. 2024. Live and Learn: Continual Action Clustering with Incremental Views. In *AAAI*, volume 38, 16264–16271.
- Yang, D.; Zhou, Y.; Zhang, A.; Sun, X.; Wu, D.; Wang, W.; and Ye, Q. 2022. Multi-view correlation distillation for incremental object detection. *Pattern Recognition*, 131: 108863.
- Yu, J.; Wu, Z.; Cai, J.; Jia, A. L.; and Fan, J. 2024. Kernel Readout for Graph Neural Networks. In *IJCAI*, 2505–2514.
- Zenke, F.; Poole, B.; and Ganguli, S. 2017. Continual learning through synaptic intelligence. In *ICML*, 3987–3995.
- Zhang, A.; Zhou, H.; Li, X.; and Zhu, W. 2019. Fast and robust learning in spiking feed-forward neural networks based on intrinsic plasticity mechanism. *Neurocomputing*, 365: 102–112.
- Zhang, J.; Gu, Y.; Chen, A.; and Yu, Y. 2022. Unveiling Dynamic System Strategies for Multisensory Processing: From Neuronal Fixed-Criterion Integration to Population Bayesian Inference. *Research*.
- Zhu, S.; Hosni, S. I.; Huang, X.; Wan, M.; Borgheai, S. B.; McLinden, J.; Shahriari, Y.; and Ostadabbas, S. 2023. A dynamical graph-based feature extraction approach to enhance mental task classification in brain–computer interfaces. *Computers in Biology and Medicine*, 153: 106498.
- Zhuang, S.; Huang, S.; Huang, W.; Chen, Y.; Wu, Z.; and Liu, X. 2024. Enhancing Multi-view Graph Neural Network with Cross-view Confluent Message Passing. In *ACM Multimedia*, 10065–10074.
- Zou, D.; Wei, W.; Mao, X.-L.; Wang, Z.; Qiu, M.; Zhu, F.; and Cao, X. 2022. Multi-level cross-view contrastive learning for knowledge-aware recommender system. In *SIGIR*, 1358–1368.
- Zou, Y.; Fang, Z.; Wu, Z.; Zheng, C.; and Wang, S. 2024. Revisiting multi-view learning: A perspective of implicitly heterogeneous Graph Convolutional Network. *Neural Networks*, 169: 496–505.