Appendix for: A Unified Framework to BRIDGE Complete and Incomplete Deep Multi-View Clustering under Non-IID Missing Patterns

Xiaorui Jiang[1], Buyun He[1], Peng Yuan Zhou[3], Xinyue Chen[2], Jingcai Guo[4], Jie Xu[2,5,*], Yong Liao[1,*]
[1]University of Science and Technology of China; [2]University of Electronic Science and Technology of China; [3]Aarhus University; [4]The Hong Kong Polytechnic University; [5]Singapore University of Technology and Design

1. Related Work

1.1. Multi-View Clustering

In the field of multi-view clustering, traditional methods and deep learning-based methods dominate as two mainstream approaches. Specifically, traditional methods include graph-based [1, 19, 31], subspace-based [29, 30, 32], kernel-based [2, 9, 18], and matrix factorization-based [8, 33, 34]. Traditional methods have relatively strong interpretability and play an important role in certain specific fields. However, limited by their feature extraction capabilities, they are relatively weak in handling complex tasks. Deep learning-based methods are playing an increasingly important role, with the mainstream approach being the use of deep autoencoders [23, 26, 28]. Subsequently, the reconstructed encoder represents the data, followed by deeper representations achieved through self-supervised modules, fusion modules, clustering modules, and others [17]. Among the techniques used in deep MVC, the most representative include those based on deep divergence [16, 25], contrastive learning [5, 26], and mutual information [6, 7]. The core of them lies in fully exploring the complementary and consistent information of multi-view data to achieve the clearest clustering structure.

1.2. Incomplete Multi-View Clustering

In real life, multi-view datasets are likely to have missing instances in some data, which has promoted the development of incomplete multi-view clustering (IMVC) [6, 10, 27]. The characteristic of incomplete multi-view datasets is that some instances of certain samples are missing, rendering that portion of information unusable. In such scenarios, the core of IMVC research lies in maximizing the use of available knowledge and minimizing the negative impact of missing information, which corresponds to handling missing data at the methodological level. Imputation-based methods [4, 6, 7, 12, 15] and imputation-free methods [1, 22, 24] are the two main paradigms at present. The former can achieve the final clustering results as the missing data (features) are fully recovered, while the latter avoids imputation altogether, as predicting missing values inevitably introduces noise.

*Co-corresponding authors.

1.3. The Overlap Between Deep MVC and IMVC

Deep IMVC data consists of two components: the complete data and the incomplete data, where training on complete data essentially aligns with a typical deep MVC task. Many studies emphasize that before tackling incomplete data [3, 6, 7, 24], it is crucial to first establish a well-performing model on complete data. Theoretically, DCP [7] argues that cross-view consistency and data recovery are inherently interconnected—learning from complete data facilitates the recovery of incomplete data, and vice versa. DSIMVC [14] highlights that learning from both complete and incomplete data is no worse than learning solely from complete data, yet pretraining on complete data remains a fundamental step. DMVG [20] addresses IMVC by decomposing it into two independent sub-tasks: missing view generation and multiview learning. It first applies methods like VIGAN [12] and CRA [15] for data recovery before employing standard MVC techniques on the reconstructed dataset. However, even in DMVG, the view generation process still depends on complete data training. In summary, representations learned from complete data provide the basis for training on incomplete data, shaping our core motivation: when designing deep IMVC tasks, the primary focus should be on improving incomplete data processing, while minimizing efforts on complete data, as standard deep MVC methods can be directly leveraged for it.

2. Motivation

In Figure 1 in the main text, we illustrate the mainstream deep MVC and IMVC methods, and the following conclusions are drawn from our observations:

(1) **Deep MVC and IMVC share significant overlap in methodology**. Specifically, in IMVC, the training methods for the complete data portion are mostly similar to those used in MVC—this is unsurprising, as the two can almost be regarded as identical tasks. However, such repetitive

Table 1. Clustering performance gains (NMI%) compared with BRIDGE+MI (Part 1/2). $\mathcal{M}$ represents the missing rate, and the best results are highlighted in **bold**.

Method	BDGP			BBCSport			Hdigit			MNIST-USPS		
	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$
CDIMC-net	70.8-8.8	69.7-1.7	59.9-5.1	45.2+29.6	33.5+15.9	35.3+30.4	40.9-53.4	39.2-52.2	29.8-54.9	65.5-27.2	53.6-30.9	49.4-28.9
COMPLETER	54.1-25.5	27.7-43.7	29.8-35.2	4.2-11.4	2.6-15.0	4.1-0.8	89.1-5.2	93.4+2	89.2+4.5	93.1+0.4	88.7+4.2	81.5+3.2
DIMVC	64.1-15.5	62.8-8.6	42.3-22.7	4.1-11.5	1.3-16.3	3.5-1.4	92.7-1.6	90.7-0.7	88.9+4.2	80.3-12.4	79.6-4.9	73.6-4.7
DCP	40.9-38.7	50.4-21.0	26.4-38.6	8.0-7.6	8.7-8.9	1.2-3.7	91.5-2.8	**94.5**+3.1	90.7+6.0	82.3-10.4	84.6+0.1	71.1-7.2
DSIMVC	88.4+8.8	**87.0**+15.6	81.4+16.4	43.3+27.7	22.3+4.7	25.0+20.1	95.8+1.5	93.2+1.8	90.6+5.9	95.1+2.4	92.9+8.4	88.1+9.8
ProImp	85.3+5.7	69.4-2.0	68.5+3.5	28.7+13.1	25.6+8.0	14.8+9.9	94.7+0.4	93.8+2.4	89.7+5.0	86.9-5.8	80.9-3.6	78.1-0.2
APADC	69.3-10.3	63.4-8.0	58.5-6.5	6.8-8.8	5.8-11.8	5.9+1.0	44.6-49.7	41.3-50.1	62.8-21.9	93.0+0.3	88.9+4.4	83.5+5.2
CPSPAN	81.3+1.7	79.1+7.7	79.6+14.6	26.1+10.5	24.5+6.9	24.1+19.2	89.2-5.1	90.2-1.2	89.3+4.6	77.5-15.2	73.3-11.2	72.6-5.7
ICMVC	65.1-14.5	64.0-7.4	49.5-15.5	9.2-6.4	2.9-14.7	2.2-2.7	96.4+2.1	92.1+0.7	87.9+3.2	96.0+3.3	90.8+6.3	87.2+8.9
DIVIDE	81.8+2.2	74.6+3.2	50.8-14.2	1.4-14.2	2.1-15.5	1.8-3.1	93.8-0.5	90.3-1.1	84.6-0.1	89.0-3.7	83.7-0.8	78.8+0.5
BRIDGE+MI	79.6	71.4	65.0	15.6	17.6	4.9	94.3	91.4	84.7	92.7	84.5	78.3
BRIDGE+DD	86.6+7.0	81.6+10.2	75.0+10.0	**69.3** +53.7	**64.5**+46.9	**56.8**+51.9	86.2-8.1	84.3 -7.1	81.1-3.6	78.5 -14.2	72.9-11.6	67.8 -10.5
BRIDGE+CL	**91.6**+12.0	86.5+15.1	**81.8**+16.8	67.2+51.6	62.7+45.1	53.4+48.5	**97.0**+2.7	94.4 +3.0	**91.9**+7.2	**96.1**+3.4	**93.2** +8.7	**88.7** +10.4

Table 2. Clustering performance gains (ARI%) compared with BRIDGE+MI (Part 1/2). $\mathcal{M}$ represents the missing rate, and the best results are highlighted in **bold**.

Method	BDGP			BBCSport			Hdigit			MNIST-USPS		
	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$
CDIMC-net	69.9-5.4	67.3+1.7	60.7+4.3	25.5+15.8	14.3+7.0	19.3+18.3	22.2-73.1	26.1-66.4	8.8-73.9	45.4-47.2	31.2-46.2	32.9-37.0
COMPLETER	27.2-48.1	7.6-58.0	7.5-48.9	1.65-8.05	0.1-7.2	0.1-0.9	83.8-11.5	94.8+2.3	90.7+8.0	93.3+0.7	87.4+10.0	72.6+2.7
DIMVC	62.3-13.0	61.6-4.0	40.0-16.4	1.1-8.6	0.2-7.1	0.2-0.8	93.8-1.5	92.0-0.5	90.0+7.3	74.7-17.9	72.3-5.1	63.7-6.2
DCP	21.1-54.2	24.2-41.4	6.23-50.17	3.4-6.3	1.7-5.6	0.1-0.9	85.5-9.8	95.7+3.2	92.3+9.6	61.9-30.7	74.7-2.7	50.1-19.8
DSIMVC	90.5+15.2	89.2+23.6	83.5+27.1	37.7+28.0	17.4+10.1	20.3+19.3	96.8+1.5	94.5+2.0	92.0+9.3	95.9+3.3	93.6+16.2	88.9+19.0
ProImp	87.0+11.7	66.9+1.3	67.8+11.4	20.9+11.2	21.8+14.5	10.7+9.7	95.9+0.6	95.2+2.7	91.4+8.7	84.4-8.2	76.8-0.6	76.4+6.5
APADC	61.2-14.1	50.2-15.4	48.1-8.3	1.5-8.2	1.3-6.0	1.1+0.1	18.5-76.8	21.9-70.6	45.6-37.1	93.8+1.2	89.9+12.5	83.2+13.3
CPSPAN	**97.6**+22.3	84.0+18.4	82.6+26.2	11.1+1.4	12.7+5.4	11.3+10.3	89.8-5.5	91.4-1.1	90.5 +7.8	69.7-22.9	65.6-11.8	64.4-5.5
ICMVC	58.6-16.7	59.4-6.2	42.0-14.4	5.6-4.1	1.9-5.4	1.0	97.3+2.0	93.5+1.0	89.5+6.8	96.8+4.2	91.8+14.4	87.7+17.8
DIVIDE	82.8+7.5	75.4+9.8	48.0-8.4	0.1-9.6	0.6-6.7	0.8-0.2	95.1-0.2	91.9-0.6	86.3+3.6	89.7-2.9	81.4+4.0	78.5+8.6
BRIDGE+MI	75.3	65.6	56.4	9.7	7.3	1.0	95.3	92.5	82.7	92.6	77.4	69.9
BRIDGE+DD	89.7+14.4	84.8+19.2	77.5+21.1	**69.9**+60.2	**63.4**+56.1	**53.3**+52.3	87.0-8.3	85.3 -7.2	81.3-1.4	74.1 -18.5	65.8-11.6	59.0 -10.9
BRIDGE+CL	93.7+18.4	**89.4**+23.8	**84.4**+28.0	62.1+52.4	59.2+51.9	53.0+52.0	**97.7**+2.4	**95.7**+3.2	**93.4**+10.7	**96.9** +4.3	**94.4**+17.0	**89.8**+19.9

Table 3. Clustering performance gains (PUR%) compared with BRIDGE+MI (Part 1/2). $\mathcal{M}$ represents the missing rate, and the best results are highlighted in **bold**.

Method	BDGP			BBCSport			Hdigit			MNIST-USPS		
	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$
CDIMC-net	86.4-1.7	84.6+1.3	82.2+3.5	51.0-2.6	41.1-6.5	42.6-8.8	45.9-52.0	44.4-52.2	32.1-60.2	51.4-45.1	45.9-44.8	41.4-45.2
COMPLETER	55.5-32.6	42.8-40.5	44.4-34.3	36.2-17.4	36.6-11.0	37.0-14.4	87.1-10.8	97.6+1.0	95.7+3.4	96.9+0.4	93.8+3.1	83.3-3.3
DIMVC	80.3-7.8	81.7-1.6	65.3-13.4	35.5-18.1	35.8-11.8	36.1-15.3	97.2-0.7	96.3-0.3	95.4+3.1	82.6-13.9	83.0-7.7	77.8-8.8
DCP	51.8-36.3	58.3-25.0	36.2-42.5	41.0-12.6	39.7-7.9	35.7-15.7	88.5-9.4	**98.1**+1.5	96.5+4.2	68.8-27.7	77.8-12.9	57.6-29.0
DSIMVC	96.1+8.0	95.5+12.2	93.0+14.3	68.6+15.0	52.6+5.0	56.3+4.9	98.5+0.6	97.5+0.9	96.3+4.0	98.1+1.6	97.1+6.4	94.8+8.2
ProImp	94.6+6.5	84.2+0.9	84.8+6.1	60.5+6.9	56.8+9.2	49.6-1.8	98.1+0.2	97.8+1.2	96.0+3.7	92.9-3.6	89.2-1.5	88.7+2.1
APADC	79.7-8.4	75.3-8.0	68.4-10.3	40.4-13.2	37.7-9.9	37.7-13.7	44.8-53.1	41.5-55.1	65.1-27.2	97.2+0.7	95.5+4.8	92.1+5.5
CPSPAN	93.3+5.2	92.7+9.4	92.6+13.9	53.5-0.1	53.9+6.3	50.7-0.7	95.3-2.6	96.0-0.6	95.6+3.3	81.7-14.8	79.3-11.4	78.0-8.6
ICMVC	73.6-14.5	75.6-7.7	65.8-12.9	41.0-12.6	39.9-7.7	38.2-13.2	98.8+0.9	97.0+0.4	95.1+2.8	98.6+2.1	96.2+5.5	94.3+7.7
DIVIDE	92.6+4.5	89.2+5.9	74.6-4.1	30.6-23.0	31.5-16.1	33.1-18.3	97.7-0.2	96.3-0.3	93.6+1.3	95.2-1.3	91.7+1.0	89.6+3.0
BRIDGE+MI	88.1	83.3	78.7	53.6	47.6	51.4	97.9	96.6	92.3	96.5	90.7	86.6
BRIDGE+DD	95.7+7.6	93.6 +10.3	90.2+11.5	**86.6**+33.0	**83.1**+35.5	**76.7**+25.3	93.9-4.0	93.1-3.5	91.1-1.2	86.5-10.0	79.5-11.2	75.2-11.4
BRIDGE+CL	**97.4**+9.3	**95.6**+12.3	**93.4**+14.7	83.5+29.9	79.6+32.0	76.0+24.6	**99.0**+1.1	98.0 +1.4	**97.0**+4.7	**98.6**+2.1	**97.4**+6.7	**95.3**+8.7

designs still take up considerable space and effort, making it difficult to focus on the processing of incomplete data. In other words, in IMVC, the module for processing complete data can be replaced, which would significantly reduce the cost of developing a new IMVC method. This motivates us to shift our focus to the handling of incomplete data.

(2) **Imputation-based and imputation-free methods represent the two main paradigms in IMVC**. As discussed, these approaches are suitable for different scenarios and are not inherently superior or inferior. The former provides a complete clustering structure, while the latter avoids introducing noise and is computationally simpler. Given these considerations, a potential direction is to integrate both paradigms, incorporating imputation and imputation-free modules into a unified framework.

(3) **Existing methods rarely consider the Non-IID miss-**

Table 4. Clustering performance gains (NMI%) compared with BRIDGE+MI (Part 2/2). $\mathcal{M}$ represents the missing rate, OOM denotes out of memory, and the best results are highlighted in **bold**. Methods supporting ≥3 views are compared.

Method	Cifar10			Cifar100			Reuters			NUSWIDE		
	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$
CDIMC-net	OOM	OOM	OOM	OOM	OOM	OOM	1.4-15.7	0.9-9.0	0.2-10.0	6.9-4.9	5.2-5.0	3.0-4.3
COMPLETER	90.8-0.9	87.3+1.4	84.0+1.3	71.7+8.9	74.3-1.1	73.6+8.3	10.8-6.3	19.0+9.1	12.3+2.1	8.6-3.2	5.0-5.2	5.8-1.5
DIMVC	90.0-1.7	86.0+0.1	82.7	98.9+36.1	97.9+22.5	98.0+32.7	27.1+10.0	28.4+18.5	27.9+17.7	16.7+4.9	13.7+3.5	10.3+3.0
DCP	90.8-0.9	87.3+1.4	**84.1**+1.4	81.4+18.6	78.1+2.7	80.5+15.2	13.5-3.6	16.2+6.3	11.0+0.8	6.1-5.7	8.2-2.0	6.4-0.9
DSIMVC	89.3-2.4	84.6-1.3	80.2-2.5	50.9-11.9	43.5-31.9	53.9-11.4	28.5+11.4	28.5+18.6	25.9+15.7	16.1+4.3	21.2+11.0	17.8 +10.5
CPSPAN	53.1-38.6	58.7-27.2	63.6-19.1	87.9+25.1	88.8+13.4	87.0+21.7	21.6+4.5	22.1+12.2	19.9+9.7	11.0-0.8	14.0+3.8	13.4+6.1
DIVIDE	81.2-10.5	75.7-10.2	69.1-13.6	94.5+31.7	91.5+16.1	87.6+22.3	10.1-7.0	7.4-2.5	5.5-4.7	11.5-0.3	6.7-3.5	3.6-3.7
BRIDGE+MI	**91.7**	85.9	82.7	62.8	75.4	65.3	17.1	9.9	10.2	11.8	10.2	7.3
BRIDGE+DD	89.8-1.9	86.0+0.1	80.3-2.4	97.2+34.4	96.1+20.7	94.8+29.5	**36.3**+19.2	**35.6**+25.7	**30.6**+20.4	20.8+9.0	19.0+8.8	16.6+9.3
BRIDGE+CL	91.4-0.3	**87.8**+1.9	84.3+1.6	**99.7**+36.9	**98.9**+23.5	**99.6**+34.3	31.3+14.2	30.4+20.5	22.3+12.1	**30.2**+18.4	**25.9**+15.7	**22.6**+15.3

Table 5. Clustering performance gains (ARI%) compared with BRIDGE+MI (Part 2/2). $\mathcal{M}$ represents the missing rate, OOM denotes out of memory, and the best results are highlighted in **bold**. Methods supporting ≥3 views are compared.

Method	Cifar10			Cifar100			Reuters			NUSWIDE		
	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$
CDIMC-net	OOM	OOM	OOM	OOM	OOM	OOM	**45.7**+35.4	**44.4**+38.0	**31.3**+25.5	**51.4**+40.0	**45.9**+40.6	**40.6**+34.6
COMPLETER	91.9-1.1	88.2+2.0	84.9+1.3	19.7-1.4	20.6-13.0	23.1+4.3	1.6-8.7	3.3-3.1	3.0-2.8	2.4-9.0	0.5-4.8	0.5-5.5
DIMVC	91.1-1.9	87.0+0.8	83.7+0.1	94.3+73.2	89.3+55.7	90.7+71.9	22.1+11.8	22.2+16.2	22.0+14.5	22.1+10.7	22.2+16.9	22.0+16.0
DCP	92.0-1.0	88.4+2.2	85.0+1.4	35.6+14.5	24.3-9.3	30.5+11.7	2.8-7.5	4.0-2.4	1.4-4.4	1.5-9.9	4.1-1.2	0.8-5.2
DSIMVC	90.1-2.9	84.5-1.7	78.2-5.4	14.9-6.2	9.6-24.0	17.4-1.4	22.3+12.0	21.3+14.9	19.3+13.5	12.0+0.6	19.2+13.9	15.6 +9.6
CPSPAN	41.3-51.7	36.0-50.2	51.2-32.4	60.8+39.7	67.9+34.3	61.7+42.9	13.9+3.6	15.0+8.6	13.2+7.4	6.6-4.8	8.4+3.1	8.1+2.1
DIVIDE	74.7-18.3	65.4-20.8	56.9-26.7	79.5+58.4	59.7+26.1	55.0+36.2	5.3-5.0	5.2-1.2	3.7-2.1	11.1-0.3	6.2+0.9	2.9-3.1
BRIDGE+MI	**93.0**	86.2	83.6	21.1	33.6	18.8	10.3	6.4	5.8	11.4	5.3	6.0
BRIDGE+DD	91.1-1.9	87.4+1.2	81.1-2.5	91.9+70.8	91.1+57.5	89.9+71.1	30.4+20.1	29.0+22.6	25.6+19.8	18.4+7.0	17.2+11.9	14.9+8.9
BRIDGE+CL	92.7-0.3	**89.1**+2.9	**85.6**+2.0	**98.7**+77.6	**95.1**+61.5	**98.5**+79.7	24.5+14.2	25.3+18.9	18.4+12.6	29.1+17.7	24.8+19.5	21.4+15.4

Table 6. Clustering performance gains (PUR%) compared with BRIDGE+MI (Part 2/2). $\mathcal{M}$ represents the missing rate, OOM denotes out of memory, and the best results are highlighted in **bold**. Methods supporting ≥3 views are compared.

Method	Cifar10			Cifar100			Reuters			NUSWIDE		
	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$	$\mathcal{M}=0.3$	$\mathcal{M}=0.5$	$\mathcal{M}=0.7$
CDIMC-net	OOM	OOM	OOM	OOM	OOM	OOM	45.7+2.6	44.4+0.7	31.3-9.8	51.4	45.9-3.3	40.6+0.3
COMPLETER	96.3-0.5	94.5+0.9	92.8+0.6	33.5-2.4	34.0-19.9	28.1-9.0	26.8-16.3	33.4-10.3	27.8-13.3	31.2-20.2	25.6-23.6	25.4-14.9
DIMVC	95.8-1.0	93.9+0.3	92.2	95.0+59.1	91.7+37.8	92.6+55.5	49.0+5.9	50.8+7.1	48.3+7.2	42.9-8.5	39.6-9.6	36.4-3.9
DCP	96.3-0.5	94.5+0.9	92.9+0.7	49.5+13.6	46.1-7.8	49.7+12.6	29.0-14.1	32.1-11.6	25.7-15.4	28.3-23.1	31.7-17.5	26.8-13.5
DSIMVC	95.4-1.4	92.4-1.2	87.3-4.9	22.8-13.1	17.5-36.4	24.4-12.7	49.7+6.6	46.1+2.4	46.3+5.2	39.7-11.7	49.0-0.2	43.9 +3.6
CPSPAN	57.7-39.1	46.8-46.8	66.2-26.0	62.3+26.4	65.9+12.0	62.1+25.0	44.2+1.1	43.8+0.1	41.3+0.2	36.5-14.9	38.7-10.5	38.5-1.8
DIVIDE	88.7-8.1	83.3-10.3	79.6-12.6	94.2+58.3	90.4+36.5	86.3+49.2	43.8+0.7	36.6-7.1	29.7-11.4	49.0-2.4	42.5-6.7	35.9-4.4
BRIDGE+MI	**96.8**	93.6	92.2	35.9	53.9	37.1	43.1	43.7	41.1	51.4	49.2	40.3
BRIDGE+DD	95.8-1.0	94.1+0.5	90.9-1.3	96.7+60.8	95.6+41.7	95.5+58.4	**60.1**+17.0	**58.6**+14.9	**55.9**+14.8	51.0-0.4	49.9+0.7	48.3+8.0
BRIDGE+CL	96.6-0.2	**94.9**+1.3	**93.1**+0.9	**99.6**+63.7	**98.1**+44.2	**99.4** +62.3	52.7+9.6	52.2+8.5	47.3+6.2	**59.2**+7.8	**56.1**+6.9	**53.8**+13.5

ing patterns. In practice, it is highly likely that data within the same view may have inconsistent distributions. For example, high temperatures might cause a sensor to collect biased data, leading to distributions that differ from those collected under normal temperatures. Based on this consideration, we aim to align the distributions within the same view, enabling the incomplete data to leverage well-represented complete data for more accurate predictions.

(4) **There is currently no comprehensive framework for deep IMVC**. A unified deep IMVC framework would facilitate the development and evaluation of new methods, providing a fairer platform for comparison within the field. With this in mind, we seek to establish a comprehensive and unified deep IMVC framework to drive a transformation in the research priorities of this field.

3. Experimental Details

This section provides additional details and results about the experiments.

3.1. Datasets

Our experiments use eight datasets, with the number of views, samples, and classes summarized in Table 7.

3.2. Comparison Methods

We choose 10 classic and state-of-the-art deep IMVC methods for an in-depth comparison, including CDIMC-net [21], COMPLETER [6], DIMVC [22], DCP [7], DSIMVC [13],

Table 7. Statistics of the related datasets.

Datasets	#Samples	#Views	#Classes
MNIST-USPS	5,000	2	10
BDGP	2,500	2	5
BBCSport	544	2	5
Hdigit	10,000	2	10
Reuters	1,200	5	6
NUSWIDE	5,000	5	5
Cifar10	50,000	3	10
Cifar100	50,000	3	100

ProImp [4], APADC [24], CPSPAN [3], ICMVC [1], and DIVIDE [11]. The specific descriptions of these methods are as follows:

- CDIMC-net introduces the Cognitive Deep Incomplete Multi-view Clustering Network, which captures high-level features and local structures of each view by integrating view-specific deep encoders and graph embedding strategies into its framework.
- COMPLETER is a framework that integrates representation learning and data recovery from an information-theoretic perspective.
- DIMVC is a novel deep IMVC framework without imputation or fusion to address issues such as inaccurate input or imputation of missing data and low-quality views.
- DCP is a unified framework designed to address the following two challenging issues in incomplete multi-view representation learning: i) how to learn a consistent representation that unifies different views, and ii) how to recover missing views.
- DSIMVC is a novel framework designed to mitigate the risk of clustering performance degradation caused by semantically inconsistent input views.
- ProImp is a new dual-stream model that employs dual attention layers and dual contrastive learning losses to learn view-specific prototypes and model the sample-prototype relationships.
- APADC proposes a hypothesis-free deep IMVC method, which considers distribution alignment in feature learning.
- CPSPAN proposes a cross-view partial sample and prototype alignment network for deep incomplete multi-view clustering.
- ICMVC is a novel high-confidence-guided incomplete contrastive multi-view clustering method.
- DIVIDE is a novel decoupled contrastive multi-view clustering method with higher-order random walks, which progressively identifies data pairs globally rather than locally using random walks.

Table 8. Accuracy results of incomplete data after adding noise (%)

σ	BDGP	BBCSport	Reuters	NUSWIDE	MNIST-USPS
0.0	93.48	77.98	46.17	53.90	94.98
0.1	93.44	78.53	46.08	53.82	95.10
0.2	93.00	78.35	46.58	53.30	94.98
0.3	92.40	77.61	45.75	52.68	94.76
0.4	91.92	76.88	45.75	52.22	94.48
0.5	91.08	76.33	45.42	51.58	94.14
0.6	90.16	75.60	45.08	50.70	93.82
0.7	88.92	75.60	44.75	49.76	93.46
0.8	88.04	75.41	44.75	48.80	92.86
0.9	87.40	72.84	44.83	48.22	92.06

3.3. Metrics

In our work, we introduce a new metric, the Final-to-Complete Accuracy Ratio (FCR), which is defined as the ratio of the model's final accuracy to its accuracy on the complete data:

$$\mathrm{FCR} = \frac{\mathrm{ACC}}{\mathrm{ACC_{cp}}}. \tag{1}$$

It can be seen that the value of FCR is constrained by both the complete and incomplete data. Specifically, if the model achieves higher ACC on the complete data but performs poorly when transferred to incomplete data, resulting in a lower ratio of overall ACC to complete ACC (i.e., $\mathrm{ACC_{cp}}$), FCR will decrease, reflecting the model's poor handling of incomplete data.

Notably: Since the ACC on all data is generally lower, the range of FCR is typically 0-1, but it is also possible for the model to represent incomplete data better, resulting in a value greater than 1. This phenomenon (FCR $>$ 1) frequently appears in simple datasets (e.g., BBCSport), where the complete portion has very few samples, and noise during transfer stage results in better representation of incomplete data by the model. For most datasets, when the sample size is not extremely small, FCR is less than 1, which aligns with intuition.

4. Additional Experiments

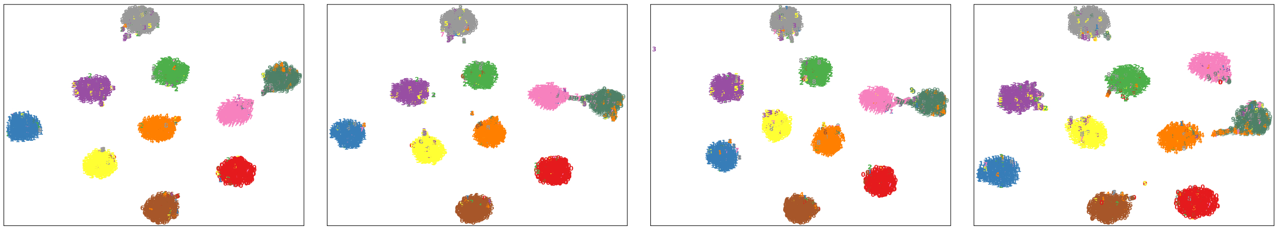
In this section, we provide additional experimental results and conduct further experiments.

4.1. Results under more metrics

We supplement the experimental results of other metrics corresponding to the main text (Table 1 and Table 2 in the main text), including NMI, ARI, and PUR, as shown in Tables (1)-(6). On these metrics, BRIDGE performs similarly to the results in the main text, excelling in most cases.

4.2. Results under varying levels of uniform noise intensity

To evaluate the robustness of our proposed method, we introduce random noise to each view, with the noise intensity



(a) $\sigma = 0.3$ (b) $\sigma = 0.5$ (c) $\sigma = 0.7$ (d) $\sigma = 0.9$

Figure 1. Visualization of different noise intensities on the MNIST-USPS dataset with a missing rate of 0.7.

Table 9. Verification of the effectiveness of the distribution alignment module under different missing rates with added noise. ACC+ indicates the addition of the distribution alignment module (i.e., the complete BRIDGE), while ACC- indicates the removal of the module. The same applies to other metrics (%).

Metric	$\mathcal{M} = 0.3$	$\mathcal{M} = 0.5$	$\mathcal{M} = 0.7$	$\mathcal{M} = 0.9$
ACC-	93.12	90.96	85.20	76.08
ACC+	93.16	91.08	87.40	76.48
NMI-	80.15	75.39	63.92	50.47
NMI+	80.26	75.60	67.46	51.78
ARI-	83.70	78.92	66.62	50.01
ARI+	83.81	79.16	71.28	51.48
PUR-	93.12	90.96	85.20	76.08
PUR+	93.16	91.08	87.40	76.48

controlled by the parameter σ—a larger σ implies stronger perturbation.

Results. Table 8 shows the experimental results of our BRIDGE+CL on different datasets. It can be seen that the introduction of noise negatively impacts BRIDGE, but it still demonstrates a considerable degree of robustness. Additionally, Table 9 validates the performance of BRIDGE on the BDGP dataset at different missing rates, where $\mathcal{M}$ represents the missing rate. It can be seen that, in the presence of noise, the distribution alignment module in BRIDGE effectively aligns the consistency between the complete and incomplete data, leading to good and accurate predictions during the imputation phase.

Visualization. Additionally, we present the visualization on the MNIST-USPS dataset with a missing rate of 0.7 in Figure 1, where the noise intensity σ takes values of 0.3, 0.5, 0.7, and 0.9. It can be observed that the visualization results under different noise intensities are very similar, demonstrating the robustness of BRIDGE to noise.

References

[1] Guoqing Chao, Yi Jiang, and Dianhui Chu. Incomplete contrastive multi-view clustering with high-confidence guiding. In Proceedings of the AAAI Conference on Artificial Intelligence, pages 11221–11229, 2024. 1, 4

[2] Lynn Houthuys, Rocco Langone, and Johan AK Suykens. Multi-view kernel spectral clustering. Information Fusion, 44:46–56, 2018. 1

[3] Jiaqi Jin, Siwei Wang, Zhibin Dong, Xinwang Liu, and En Zhu. Deep incomplete multi-view clustering with cross-view partial sample and prototype alignment. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 11600–11609, 2023. 1, 4

[4] Haobin Li, Yunfan Li, Mouxing Yang, Peng Hu, Dezhong Peng, and Xi Peng. Incomplete multi-view clustering via prototype-based imputation. arXiv preprint arXiv:2301.11045, 2023. 1, 4

[5] Yunfan Li, Peng Hu, Zitao Liu, Dezhong Peng, Joey Tianyi Zhou, and Xi Peng. Contrastive clustering. In Proceedings of the AAAI conference on artificial intelligence, pages 8547–8555, 2021. 1

[6] Yijie Lin, Yuanbiao Gou, Zitao Liu, Boyun Li, Jiancheng Lv, and Xi Peng. Completer: Incomplete multi-view clustering via contrastive prediction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 11174–11183, 2021. 1, 3

[7] Yijie Lin, Yuanbiao Gou, Xiaotian Liu, Jinfeng Bai, Jiancheng Lv, and Xi Peng. Dual contrastive prediction for incomplete multi-view representation learning. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(4):4447–4461, 2022. 1, 3

[8] Jialu Liu, Chi Wang, Jing Gao, and Jiawei Han. Multi-view clustering via joint nonnegative matrix factorization. In Proceedings of the 2013 SIAM international conference on data mining, pages 252–260. SIAM, 2013. 1

[9] Jing Liu, Fuyuan Cao, Xiao-Zhi Gao, Liqin Yu, and Jiye Liang. A cluster-weighted kernel k-means method for multi-view clustering. In Proceedings of the Aaai conference on artificial intelligence, pages 4860–4867, 2020. 1

[10] Xinwang Liu, Xinzhong Zhu, Miaomiao Li, Lei Wang, Chang Tang, Jianping Yin, Dinggang Shen, Huaimin Wang, and Wen Gao. Late fusion incomplete multi-view clustering. IEEE transactions on pattern analysis and machine intelligence, 41 (10):2410–2423, 2018. 1

[11] Yiding Lu, Yijie Lin, Mouxing Yang, Dezhong Peng, Peng Hu, and Xi Peng. Decoupled contrastive multi-view clustering with high-order random walks. In Proceedings of the AAAI Conference on Artificial Intelligence, pages 14193–14201, 2024. 4

[12] Chao Shang, Aaron Palmer, Jiangwen Sun, Ko-Shin Chen, Jin Lu, and Jinbo Bi. Vigan: Missing view imputation with generative adversarial networks. In 2017 IEEE International conference on big data (Big Data), pages 766–775. IEEE, 2017. 1

[13] Huayi Tang and Yong Liu. Deep safe multi-view clustering: Reducing the risk of clustering performance degradation caused by view increase. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 202–211, 2022. 3

[14] Huayi Tang and Yong Liu. Deep safe incomplete multi-view clustering: Theorem and algorithm. In International Conference on Machine Learning, pages 21090–21110. PMLR, 2022. 1

[15] Luan Tran, Xiaoming Liu, Jiayu Zhou, and Rong Jin. Missing modalities imputation via cascaded residual autoencoder. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1405–1414, 2017. 1

[16] Daniel J Trosten, Sigurd Lokse, Robert Jenssen, and Michael Kampffmeyer. Reconsidering representation alignment for multi-view clustering. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 1255–1265, 2021. 1

[17] Daniel J Trosten, Sigurd Løkse, Robert Jenssen, and Michael C Kampffmeyer. On the effects of self-supervision and contrastive alignment in deep multi-view clustering. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 23976–23985, 2023. 1

[18] Grigorios Tzortzis and Aristidis Likas. Kernel-based weighted multi-view clustering. In 2012 IEEE 12th international conference on data mining, pages 675–684. IEEE, 2012. 1

[19] Hao Wang, Yan Yang, Bing Liu, and Hamido Fujita. A study of graph-based system for multi-view clustering. Knowledge-Based Systems, 163:1009–1019, 2019. 1

[20] Jie Wen, Shijie Deng, Waikeung Wong, Guoqing Chao, Chao Huang, Lunke Fei, and Yong Xu. Diffusion-based missing-view generation with the application on incomplete multi-view clustering. In Forty-first International Conference on Machine Learning. 1

[21] Jie Wen, Zheng Zhang, Yong Xu, Bob Zhang, Lunke Fei, and Guo-Sen Xie. Cdimc-net: Cognitive deep incomplete multi-view clustering network. In Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20, pages 3230–3236. International Joint Conferences on Artificial Intelligence Organization, 2020. Main track. 3

[22] Jie Xu, Chao Li, Yazhou Ren, Liang Peng, Yujie Mo, Xiaoshuang Shi, and Xiaofeng Zhu. Deep incomplete multi-view clustering via mining cluster complementarity. In Proceedings of the AAAI conference on artificial intelligence, pages 8761–8769, 2022. 1, 3

[23] Jie Xu, Huayi Tang, Yazhou Ren, Liang Peng, Xiaofeng Zhu, and Lifang He. Multi-level feature learning for contrastive multi-view clustering. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 16051–16060, 2022. 1

[24] Jie Xu, Chao Li, Liang Peng, Yazhou Ren, Xiaoshuang Shi, Heng Tao Shen, and Xiaofeng Zhu. Adaptive feature projection with distribution alignment for deep incomplete multi-view clustering. IEEE Transactions on Image Processing, 32: 1354–1366, 2023. 1, 4

[25] Jie Xu, Yazhou Ren, Huayi Tang, Zhimeng Yang, Lili Pan, Yang Yang, Xiaorong Pu, Philip S. Yu, and Lifang He. Self-supervised discriminative feature learning for deep multi-view clustering. IEEE Transactions on Knowledge and Data Engineering, 35(7):7470–7482, 2023. 1

[26] Weiqing Yan, Yuanyang Zhang, Chenlei Lv, Chang Tang, Guanghui Yue, Liang Liao, and Weisi Lin. Gcfagg: Global and cross-view feature aggregation for multi-view clustering. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 19863–19872, 2023. 1

[27] Mouxing Yang, Yunfan Li, Peng Hu, Jinfeng Bai, Jiancheng Lv, and Xi Peng. Robust multi-view clustering with incomplete information. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(1):1055–1069, 2022. 1

[28] Xihong Yang, Jin Jiaqi, Siwei Wang, Ke Liang, Yue Liu, Yi Wen, Suyuan Liu, Sihang Zhou, Xinwang Liu, and En Zhu. Dealmvc: Dual contrastive calibration for multi-view clustering. In Proceedings of the 31st ACM International Conference on Multimedia, pages 337–346, 2023. 1

[29] Qiyue Yin, Shu Wu, Ran He, and Liang Wang. Multi-view clustering via pairwise sparse subspace representation. Neurocomputing, 156:12–21, 2015. 1

[30] Qiyue Yin, Shu Wu, and Liang Wang. Incomplete multi-view clustering via subspace learning. In Proceedings of the 24th ACM international on conference on information and knowledge management, pages 383–392, 2015. 1

[31] Kun Zhan, Changqing Zhang, Junpeng Guan, and Junsheng Wang. Graph learning for multiview clustering. IEEE transactions on cybernetics, 48(10):2887–2895, 2017. 1

[32] Changqing Zhang, Qinghua Hu, Huazhu Fu, Pengfei Zhu, and Xiaochun Cao. Latent multi-view subspace clustering. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 4279–4287, 2017. 1

[33] Chen Zhang, Siwei Wang, Jiyuan Liu, Sihang Zhou, Pei Zhang, Xinwang Liu, En Zhu, and Changwang Zhang. Multi-view clustering via deep matrix factorization and partition alignment. In Proceedings of the 29th ACM international conference on multimedia, pages 4156–4164, 2021. 1

[34] Linlin Zong, Xianchao Zhang, Long Zhao, Hong Yu, and Qianli Zhao. Multi-view clustering via multi-manifold regularized non-negative matrix factorization. Neural Networks, 88:74–89, 2017. 1