Supplementary Material for A Survey on Incomplete Multi-view Clustering

Jie Wen, Zheng Zhang, *Senior Member*, *IEEE*, Lunke Fei, Bob Zhang, *Senior Member*, *IEEE*, Yong Xu, *Senior Member*, *IEEE*, Zhao Zhang, *Senior Member*, *IEEE*, Jinxing Li

TABLE I: The selected datasets for experiments.

Dataset	# Class	# View	# Samples	# Features
BUAA	10	2	90	100/100
BDGP	5	4	2500	79/1000/500/250
Caltech101	102	4	9144	254/1984/512/928
BBCSport	5	4	116	1991/2063/2113/2158
NUSWIDE	31	5	30000	64/144/73/128/225

I. Experimental results on the BBCSport, Caltech101, and NUSWIDE datasets

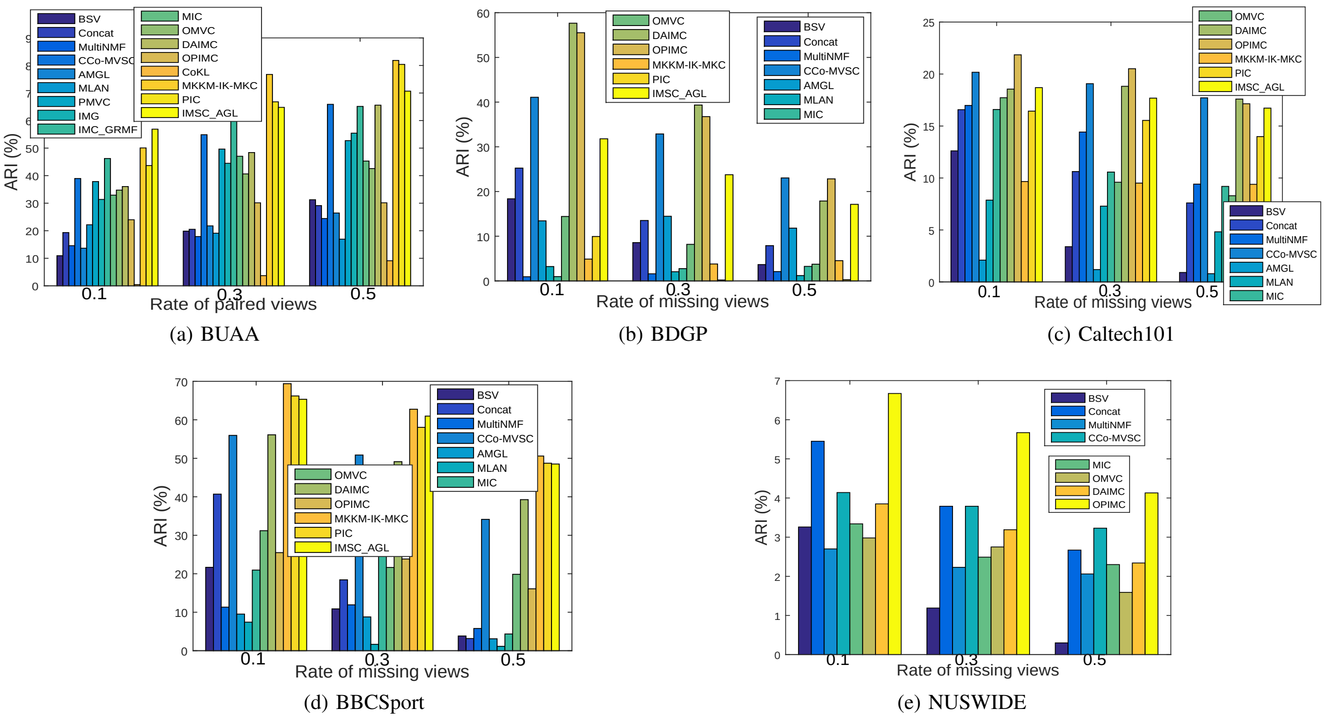
The experimental results *w.r.t.* ACC, NMI, and purity of different methods on the the BBCSport, Caltech101, and NUSWIDE datasets are listed in Table II. The experimental results in terms of adjusted rand index (ARI) of different methods on the five datasets mentioned in the main manuscript and Table I are shown in Fig.1. We can observe that the graph based IMC methods obtain much higher ARIs than the matrix factorization based methods on the BUAA and BBCSport datasets. However, on the BDGP and Caltech101 datasets, DAIMC and OPIMC obtain higher ARIs than PIC and IMSC_AGL. In addition, on the NUSWIDE dataset, the graph based method, *i.e.*, CCo-MVSC, obtains worse performance than Concat, OPIMC, and DAIMC. This indicates that it is difficult to capture the intrinsic similarity relationships from the large-scale datasets, which leads to a bad performance for the graph based methods.

II. Statistical analysis on the experimental results

In fields of statistical learning, ‘t-test’ is a popular statistical analysis method for evaluating the significant difference of data [1-4]. In our work, we adopt it to analyze the significant difference of the clustering accuracies (ACC) obtained by different incomplete multi-view clustering (IMC) methods. Table III lists the p-value of ACC between MIC and the other IMC methods on the BUAA and BBCSport datasets with different rates of paired views or missing views, where MIC is selected as the baseline method and a smaller value of p-value indicates a larger significant difference between the two sets of data. From this table, we can clearly observe that the other matrix factorization based methods, such as PMVC, OMVC, and OPIMC, obtain comparative or even worse performance than MIC. However, the advanced kernel and graph learning based methods significantly outperform MIC on the two datasets, which demonstrates the effectiveness of transforming the feature space learning tasks into the graph or kernel space to address the incomplete clustering problem.

III. Running time analysis

In this section, we implement different methods on the five datasets mentioned in the main manuscript and Table I at the same computer platform to compare their computational efficiencies, where the missing-view or paired-view rate is commonly set as 0.3. For fair comparison, all methods except for MLAN and OPIMC, use the same kmeans code provided by Matlab2015a software and the computer platform is Inter i9-9900K CPU, 64GB RAM, and Win10 system. The real running times (seconds) and their computational complexities are shown in Table IV. From the experimental results, we can observe that 1) OPIMC performs much faster than the other methods. For instance, on the large-scale Caltech101 dataset, the efficiency of OPIMC is about 4000 times that of IMSC_AGL. This indicates that separating the whole samples into some batches and using the batch optimization strategy can greatly reduce the computational costs. 2) All IMC methods except for OPIMC have a very high computational cost on the large-scale Caltech101 and NUSWIDE datasets. Specifically, the graph based and kernel based IMC methods are not only very inefficient, but also require more storage space than the matrix factorization based methods. Generally speaking, for the data with n samples and l views, these methods need to calculate at least l large-scale graphs or kernels with the size of $n \times n$, which are very time-consuming and require a very large storage space. The high computational cost and storage requirements are the main reasons that limit the applications of the graph or kernel based methods in IMC tasks.



(a) BUAA (b) BDGP (c) Caltech101

(d) BBCSport (e) NUSWIDE

Fig. 1: ARIs (%) of different methods on the (a) BUAA, (b) BDGP, (c) Caltech101, (d) BBCSport, and (e) NUSWIDE datasets with different missing-view rates or paired-view rates.

TABLE II: Average and standard deviation values of ACC (%), NMI (%), purity (%) of different methods on the BBCSport, Caltech101, and NUSWIDE datasets.

		ACC			NMI			Purity		
Data	Method	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5
Caltch101	BSV	22.57±0.25	19.36±0.37	16.55±0.33	42.94±0.12	33.57±0.11	23.57±0.18	40.98±0.21	33.31±0.15	25.77±0.29
	Concat	22.51±0.66	17.59±0.46	14.12±0.32	43.38±0.22	37.13±0.14	31.38±0.31	42.14±0.39	35.95±0.13	30.85±0.29
	MultiNMF	21.42±0.62	19.28±0.95	14.98±0.86	44.71±0.41	39.70±0.47	31.92±0.66	44.06±0.53	39.13±0.18	31.43±0.71
	CCo-MVSC	26.26±0.46	24.92±0.25	22.53±0.25	46.77±0.31	45.86±0.16	40.97±0.29	46.94±0.42	45.18±0.38	42.39±0.47
	AMGL	19.02±0.75	17.26±1.06	15.44±0.61	37.03±0.62	32.43±0.97	28.13±0.73	38.42±0.56	34.45±0.96	30.34±0.43
	MLAN	21.10±1.72	16.58±0.75	15.90±0.69	28.76±1.69	22.30±0.99	21.00±0.67	35.04±2.06	29.07±0.84	27.60±0.61
	MIC	22.82±0.57	20.12±0.75	18.44±0.36	44.55±0.79	38.12±0.94	31.78±0.56	43.11±0.31	35.77±0.59	32.15±0.54
	OMVC	23.45±0.77	20.59±0.34	19.75±0.89	45.05±0.53	38.52±0.51	33.81±1.11	44.25±0.40	36.35±0.25	34.76±0.52
	DAIMC	24.93±1.41	25.15±0.31	23.21±1.12	47.81±0.50	46.81±0.25	40.97±1.29	47.03±0.45	46.13±0.52	40.49±1.26
	OPIMC	26.78±1.24	25.04±1.47	21.63±1.17	31.39±0.57	25.63±1.37	33.58±0.18	33.01±0.95	29.64±1.04	28.24±0.37
	MKKM-IK-MKC	15.11±0.28	14.35±0.21	14.83±0.50	32.72±0.14	32.88±0.28	33.58±0.29	32.34±0.30	32.17±0.28	33.14±0.31
	PIC	25.33±2.38	24.53±1.02	22.24±1.11	46.46±3.24	45.40±1.32	41.94±1.10	47.10±2.84	46.49±1.41	43.31±1.26
	IMSC_AGL	**27.77±0.55**	**26.18±0.89**	**23.57±1.25**	**48.55±0.77**	**47.45±0.83**	**42.19±1.05**	**48.53±0.77**	**47.77±1.06**	**44.39±0.98**
BBCSport	BSV	58.62±3.94	51.31±5.33	44.03±3.78	43.73±7.43	31.03±2.08	21.40±2.61	65.79±5.52	55.07±1.51	47.59±2.28
	Concat	70.62±3.76	58.72±5.42	33.21±2.19	61.69±6.72	38.92±7.87	18.61±1.44	80.59±4.59	63.24±5.82	37.00±1.54
	MultiNMF	48.58±4.25	42.75±7.33	40.34±8.07	23.48±3.15	18.25±7.74	14.79±8.36	47.17±5.12	44.66±6.29	43.10±7.78
	CCo-MVSC	72.76±4.13	70.06±4.50	61.38±6.17	62.79±6.52	57.69±3.54	39.55±8.00	81.49±4.81	78.62±2.71	66.72±5.81
	AMGL	44.48±2.51	44.14±5.78	38.53±4.02	26.03±2.85	24.86±6.93	16.07±2.83	51.12±2.88	50.26±5.53	43.45±2.94
	MLAN	44.14±3.72	34.66±2.84	31.41±2.44	23.53±4.28	9.38±2.14	8.72±2.04	49.22±4.17	37.76±2.36	38.53±2.78
	MIC	51.21±4.21	46.21±4.71	46.03±5.19	29.90±6.25	25.84±3.24	24.01±5.39	55.00±4.15	51.72±4.27	52.41±6.23
	OMVC	53.33±3.21	51.38±3.06	48.79±3.10	30.64±2.00	41.57±2.79	40.63±2.45	56.49±2.81	59.20±2.12	57.47±2.80
	DAIMC	68.62±4.59	63.45±10.97	56.89±5.59	56.62±4.60	50.17±9.91	37.89±6.22	76.90±5.89	71.72±10.76	61.03±5.08
	OPIMC	54.14±4.78	52.93±4.93	45.69±6.00	35.66±4.71	31.56±6.10	21.75±6.44	58.28±4.82	56.72±5.76	50.86±6.87
	MKKM-IK-MKC	**77.55±2.01**	**75.66±3.01**	67.07±3.51	**72.91±3.29**	64.42±4.69	53.52±4.74	**88.76±2.01**	84.03±3.22	77.00±3.55
	PIC	75.52±1.57	74.48±3.32	**69.48±6.02**	70.94±2.22	64.18±2.74	53.91±6.22	87.41±1.44	82.41±2.48	77.14±6.04
	IMSC_AGL	76.41±3.07	74.48±5.80	69.31±3.48	70.46±3.83	**66.11±5.17**	**54.57±4.87**	87.41±2.25	**85.00±5.12**	**78.10±3.08**
NUSWIDE	BSV	11.92±0.69	11.53±0.49	11.50±0.26	10.37±1.06	8.86±0.89	6.51±0.59	23.48±1.90	20.93±1.55	18.39±0.97
	Concat	14.76±0.28	12.44±0.30	10.45±0.18	**13.93±0.06**	**10.98±0.13**	**8.34±0.09**	**26.33±0.12**	**23.38±0.33**	20.61±0.10
	MultiNMF	12.23±0.57	12.17±0.69	12.19±0.81	10.51±0.36	8.53±0.39	7.74±0.31	21.48±0.44	19.79±0.51	19.19±0.43
	CCo-MVSC	13.20±0.24	12.59±0.23	11.50±0.46	10.74±0.23	10.27±0.30	9.37±0.37	22.95±0.37	22.65±0.29	**22.20±0.34**
	MIC	11.91±0.40	11.31±0.16	11.50±0.46	11.06±0.19	9.30±0.44	8.05±0.43	22.81±0.22	20.98±0.34	20.34±0.39
	OMVC	11.52±0.46	11.16±0.89	9.05±0.22	8.94±0.25	7.90±0.13	5.56±0.24	21.25±0.76	20.74±0.18	18.48±0.44
	DAIMC	13.19±0.50	12.02±0.54	10.51±0.63	10.59±0.25	9.04±0.26	7.08±0.28	22.40±0.32	20.87±0.40	19.40±0.69
	OPIMC	**19.90±0.79**	**19.03±1.17**	**16.67±1.00**	10.66±0.41	9.00±0.27	7.19±0.50	22.80±0.89	21.34±0.45	19.58±0.86

TABLE III: p-value *w.r.t.* ACC between MIC and the other IMC methods on the BUAA and BBCSport datasets with different rates of paired views or missing views. Note: 1) MIC is selected as the baseline method. 2) Significance level is set as 0.05 and '$p < 0.05$' means a large significant difference for the corresponding two sets of data. 3) The value marked by '•' denotes that the ACC obtained by the corresponding method increases significantly with respect to MIC. 4) The value marked by 'blue color' means the ACC obtained by the corresponding method decreases significantly with respect to MIC.

Data	Rate	PMVC	IMG	IMC_GRMF	OMVC	DAIMC	OPIMC	CoKL	MKKM-IK-MKC	PIC	IMSC_AGL
BUAA	0.1	0.621	0.622	0.049•	0.587	0.739	0.059	5.239×10^{-5}	0.005•	0.004•	0.005•
	0.3	0.746	0.986	0.007•	0.192	0.817	0.001	1.584×10^{-5}	0.002•	0.001•	0.002•
	0.5	0.933	0.591	0.075•	0.156	0.382	0.002	0.001	0.001•	0.004•	0.017•
BBCSport	0.1	–	–	–	0.245	4.826×10^{-4}•	0.890	–	4.508×10^{-4}•	1.55×10^{-4}•	1.339×10^{-4}•
	0.3	–	–	–	0.948	0.166	0.755	–	0.002•	5.103×10^{-4}•	0.010•
	0.5	–	–	–	0.071	0.004•	0.057	–	4.602×10^{-5}•	8.914×10^{-4}•	1.812×10^{-4}•

TABLE IV: Running time (seconds) of different methods on the five datasets mentioned in the main manuscript and Table I with the missing-view rate or paired-view rate of 30%. 'T' denotes the iterations of inner loop. 'm_{max}' denotes the largest dimensionality of views. 't' denotes the iterations of k-means clustering. 'd' denotes the cluster number. The other variables are defined in Table I. '–' indicates that the corresponding method cannot handle the data with such an incomplete case. '~*~' means that the corresponding method cannot process the data on such a computer with 64GB RAM.

	Running time (seconds)					
Methods	BUAA	BDGP	Caltech101	BBCSport	NUSWIDE	Computational complexity
BSV	0.183	43.213	1.622×10^3	0.788	92.431	$O\left(t\sum_{v=1}^{l} m_v^2 n\right)$ [5]
Concat	0.114	0.092	2.364×10^3	0.895	81.690	$O\left(tn\left(\sum_{k=1}^{v} m_k\right)^2\right)$ [5]
MultiNMF	2.464	33.831	2.507×10^3	18.448	1.789×10^3	$O\left(\tau T\sum_{v=1}^{l} m_v dn + tnd^2\right)$ [6]
CCo-MVSC	0.371	7.231	436.347	0.359	4.277×10^3	$O\left(\tau\left(ln^3+n^3\right)+tnd^2\right)$
AMGL	0.196	24.986	416.389	0.182	~*~	$O\left(\tau n^3 + tnd^2\right)$ [7]
MLAN	0.107	18.183	203.386	0.104	~*~	$O\left(\tau n^3\right)$ [8]
PMVC	0.269	–	–	–	–	$O\left(\tau\sum_{v=1}^{l} m_v n_v d + tnd^2\right)$ [9, 10]
IMG	1.092	–	–	–	–	$O\left(\tau\left(2n^3 + T\sum_{v=1}^{l} m_v dn\right) + tnd^2\right)$ [11]
IMC_GRMF	0.330	–	–	–	–	$O\left(\tau\sum_{v=1}^{l} m_v d^2 + tnd^2\right)$ [12]
MIC	1.040	124.340	1.407×10^4	3.843	9.234×10^4	$O\left(\tau T\sum_{v=1}^{l} m_v dn + tnd^2\right)$ [6]
OMVC	0.830	166.328	1.372×10^4	4.341	8.932×10^4	$O\left(\tau T\sum_{v=1}^{l} m_v dn + tnd^2\right)$ [6]
DAIMC	0.396	24.932	1.861×10^3	148.501	4.172×10^3	$O\left(\tau\left(Tndm_{\max} + lm_{\max}^3\right) + tnd^2\right)$ [13]
OPIMC	0.075	0.746	6.642	0.089	1.261	$O\left(\tau ldnm_{\max}\right)$ [14]
CoKL	0.360	–	–	–	–	$O\left(6n^3 + tnd^2\right)$
MKKM-IK-MKC	0.522	77.846	1.263×10^3	0.454	~*~	$O\left(\tau\left(n^3 + ln^3 + l^3\right) + tnd^2\right)$ [15]
PIC	0.389	33.305	1.611×10^3	0.424	~*~	$O\left(l^3n^2d + l^3 + n^3 + tnd^2\right)$ [16]
IMSC_AGL	1.862	579.101	2.437×10^4	4.504	~*~	$O\left(\tau\left(ln^3 + n^3 + \sum_{v=1}^{l} n_v^3 + tnd^2\right)\right)$ [17]

TABLE V: Notations and definitions used in the main document.

Notation	Definition	Notation	Definition
$T^{(v)} \in R^{m_v \times n}$	Original incomplete multi-view data whose missing instances are denoted by 'NaN'.	m_v	Feature dimension of instances of the vth view.
l	Number of views.	c	Number of clusters.
n	Number of data samples.	n_v	Number of available instances of the vth view.
n_c	Number of samples that have all views.	$\mathit{1}$	Column vector with all elements as 1.
$X^{(v)} = \left[x_1^{(v)}, \ldots, x_{n_v}^{(v)}\right]$	Set of the available instances of the vth view.	I	Identity matrix.
$x_i^{(v)}$	The ith instance of the vth view.	$X_{i,j}^{(v)}$	The (i,j)th element of matrix $X^{(v)}$.
$X_c^{(v)}$	Set of paired instances in the vth view.	$\widehat{X}^{(v)}$	Instance set of samples that are only observed in the vth view.
$P^{(v)}$	Representation of the available samples in the vth view.	P_c	Common Representation shared by the paired-samples that are observed in all views.
$\bar{P}^{(v)}$	Full representation of the available instances and missing instances in the vth view.	$P_s^{(v)}$	Representation of the samples that are only observed in the vth view.
P	Common representation of all samples shared by all views.	$U^{(v)}$	Basis matrix of the vth view in the MF model.
$Z^{(v)}$	Similarity graph of the vth view.	Z^*	Consensus similarity graph shared by all views.
$K^{(v)}$	Kernel matrix of the vth view.	$L_{Z^{(v)}}$	Laplacian matrix of the corresponding similarity graph in the vth view.
λ_1, λ_2, λ_3, γ_v, β_v, $\alpha^{(v)}$	Hyper-penalty parameter.	$L^{(v)}$	Laplacian kernel matrix of the vth view.
$Y^{(v)} \in R^{m_v \times n}$	Set of available instances and missing instances of the vth view, where the missing instances are set as 0 or the average vector of the available instances.	$W^{(v)}$	A diagonal matrix whose element $W_{i,i}^{(v)} = 1$ if the ith sample has the vth view, otherwise $W_{i,i}^{(v)} = 0$.
$G^{(v)} \in \{0,1\}^{n \times n_v}$	$G_{i,j}^{(v)} = 1$ if the jth available instance of the vth view is from the ith sample, otherwise $G_{i,j}^{(v)} = 0$.	ϕ and φ	Boundary constraint and regularization constraint of the corresponding variables, respectively.

TABLE VI: Comparison of the representative IMC methods presented in the main document.

Methods	All kinds of incomplete cases	Property
PMVC [9]	×	Merit: Simple and efficient. Drawbacks: 1) Can only handle the incomplete case that only has paired-samples and single-view samples. 2) It ignores the locality structure of the data. 3) The input data should be non-negative.
IMG [11]	×	Merit: Captures the local distance structure of the data. Drawbacks: 1) It is only suitable to the data with two views. 2) High computational complexity.
PMSC [18]	×	Merit: Captures the representation structure of the data. Drawbacks: 1) It is only suitable to data with two views. 2) High computational complexity.
IMC_GRMF [12]	×	Merit: It can handle the data with more than two views directly. Drawbacks: 1) It is only suitable to the incomplete data with paired-samples and single-view-samples. 2) It requires the feature dimension c of the new representation to be less than the minimum feature dimension of all views, *i.e.*, $\min(m_1, \ldots, m_v)$.
MIC [19]	✓	Merit: Can handle all kinds of incomplete cases. Drawbacks: 1) Ignores the geometric structure of the data. 2) The input data should be non-negative. 3) It is not suitable for data with a large missing rate.
OMVC [6]	✓	Merits: 1) Can handle all kinds of incomplete cases. 2) Provides a chunk by chunk training approach, which can greatly reduce the memory cost and improve the efficiency. Drawbacks: 1) Ignores the geometric structure of the data. 2) The input data should be non-negative. 3) It is not suitable to data with a large missing rate.
GPMVC [20]	✓	Merits: 1) Can handle all kinds of incomplete cases. 2) Captures the locality structure. Drawbacks:1) Many tunable hyper-parameters. 2) The input data should be non-negative.
DAIMC [13]	✓	Merit: Can handle all kinds of incomplete cases. Drawback: Ignores the local structure of the data.
OPIMC [14]	✓	Merits: 1) Can handle all kinds of incomplete cases. 2) High efficiency and low memory cost. Drawbacks: 1) Ignores the local structure of the data. 2) It is sensitive to the block size.
UEAF [21]	✓	Merits: 1) Can handle all kinds of incomplete cases. 2) Has the potential to recover the missing instances and explore their hidden information. Drawback: High computational complexity.
MIKC_OCK [22]	×	Merit: Can recover the rows and columns of the kernel matrix associated with the missing instances. Drawbacks: 1) Can only handle the incomplete data with two views where one view is complete. 2) Cannot obtain the optimal representation. 3) It is sensitive to the quality of the preconstructed kernels.
CoKL [23]	×	Merit: Can recover the rows and columns of the kernel matrices associated with the missing instances. Drawbacks: 1) Can only handle the incomplete data with two views. 2) Cannot obtain the optimal representation. 3) It is sensitive to the quality of the preconstructed kernels.
IMKKC [24]	✓	Merits: 1) Can handle all kinds of incomplete cases. 2) It is a penalty-parameter-free method. Drawbacks: 1) Ignores the local distribution of the data. 2) High computational complexity. 3) It is sensitive to the quality of the preconstructed kernels.
MKKM-IK-MKC [15]	✓	Merits: 1) Can handle all kinds of incomplete cases. 2) Obtain more reasonable kernel matrices. Drawbacks: 1) Ignores the local distribution of the data. 2) High computational complexity. 3) May be trapped into a local minimum. 4) It is sensitive to the quality of the preconstructed kernels.
LIMKKC [25]	✓	Merits: 1) Can handle all kinds of incomplete cases. 2) Obtains more reasonable kernel matrices. Drawbacks: 1) Ignores the local distribution of the data. 2) High computational complexity. 3) It may be trapped into a local minimum. 4) It is sensitive to the quality of the preconstructed kernels.
EE-IMVC [26]	✓	Merits: 1) Can handle all kinds of incomplete cases. 2) It is a penalty-parameter-free method. 3) Relatively low storage cost and high efficiency. Drawback: 1) Ignores the local distribution of the data. 2) It is sensitive to the quality of the preconstructed kernels.
SCIMC [27]	✓	Merits Can handle all kinds of incomplete cases. Drawbacks: 1) Cannot learn the optimal consensus representation. 2) The initialization of similarity graphs is unreasonable.
PIC [16]	✓	Merit: Can handle all kinds of incomplete cases. Drawbacks: 1) Cannot obtain the optimal consensus graph. 2) It is sensitive to the quality of the pre-constructed graphs.
CGL_IMC [28]	✓	Merit: Can handle all kinds of incomplete cases. Drawback: It is sensitive to the quality of the pre-constructed graphs.
IMSC_AGL [17]	✓	Merits: 1) Can handle all kinds of incomplete cases. 2) Provides a joint learning framework to adaptively obtain the optimal graphs and consensus representation. Drawback: High computational complexity.
IMC_DSM [29]	×	Merit: Has the potential to extract high-level features to improve the performance. Drawback: 1) It is only suitable to the case with paired samples and single-view samples. 2) It requires that the latent representation must be non-negative.
PMVC_CGAN [30]	×	Merit: 1) Has the potential to extract high-level features to improve the performance. 2) Can recover the missing views. 3) Can be applied to large-scale datasets. Drawback: It is only suitable to incomplete multi-view data with two views.
AIMC [31]	×	Merits: 1) Has the potential to extract high-level features to improve the performance. 2) Can recover the missing views. 3) Can be applied to large-scale datasets. Drawback: It is only suitable to incomplete multi-view data without paired views.

REFERENCES

[1] E. Kreyszig, *Introductory Mathematical Statistics*. John Wiley, 1970.

[2] J. Wen, X. Fang, Y. Xu, C. Tian, and L. Fei, "Low-rank representation with adaptive graph regularization," *Neural Networks*, vol. 108, pp. 83–96, 2018.

[3] S. Li and Y. Fu, "Learning robust and discriminative subspace with low-rank constraints," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 11, pp. 2160–2173, 2015.

[4] Z. Zhang, Z. Lai, Y. Xu, L. Shao, J. Wu, and G.-S. Xie, "Discriminative elastic-net regularized linear regression," *IEEE Transactions on Image Processing*, vol. 26, no. 3, pp. 1466–1481, 2017.

[5] J. Guo and J. Ye, "Anchors bring ease: An embarrassingly simple approach to partial multi-view clustering," in *AAAI Conference on Artificial Intelligence*, vol. 33, 2019, pp. 118–125.

[6] W. Shao, L. He, C.-t. Lu, and S. Y. Philip, "Online multi-view clustering with incomplete views," in *IEEE International Conference on Big Data*. IEEE, 2016, pp. 1012–1017.

[7] F. Nie, J. Li, and X. Li, "Parameter-free auto-weighted multiple graph learning: a framework for multiview clustering and semi-supervised classification," in *International Joint Conference on Artificial Intelligence*, 2016, pp. 1881–1887.

[8] F. Nie, G. Cai, and X. Li, "Multi-view clustering and semi-supervised classification with adaptive neighbours," in *AAAI Conference on Artificial Intelligence*, 2017, pp. 2408–2414.

[9] S.-Y. Li, Y. Jiang, and Z.-H. Zhou, "Partial multi-view clustering," in *AAAI Conference on Artificial Intelligence*, 2014, pp. 1969–1974.

[10] C.-J. Hsieh and I. S. Dhillon, "Fast coordinate descent methods with variable selection for non-negative matrix factorization," in *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2011, pp. 1064–1072.

[11] H. Zhao, H. Liu, and Y. Fu, "Incomplete multi-modal visual data grouping." in *International Joint Conferences on Artificial Intelligence*, 2016, pp. 2392–2398.

[12] J. Wen, Z. Zhang, Y. Xu, and Z. Zhong, "Incomplete multi-view clustering via graph regularized matrix factorization," in *European Conference on Computer Vision Workshops*, 2019, pp. 593–608.

[13] M. Hu and S. Chen, "Doubly aligned incomplete multi-view clustering," in *International Joint Conferences on Artificial Intelligence*, 2019, pp. 2262–2268.

[14] ——, "One-pass incomplete multi-view clustering," in *AAAI Conference on Artificial Intelligence*, 2019, pp. 3838–3845.

[15] X. Liu, X. Zhu, M. Li, L. Wang, E. Zhu, T. Liu, M. Kloft, D. Shen, J. Yin, and W. Gao, "Multiple kernel k-means with incomplete kernels," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

[16] H. Wang, L. Zong, B. Liu, Y. Yang, and W. Zhou, "Spectral perturbation meets incomplete multi-view data," in *Proceedings of the 28th International Joint Conference on Artificial Intelligence*. AAAI Press, 2019, pp. 3677–3683.

[17] J. Wen, Y. Xu, and H. Liu, "Incomplete multiview spectral clustering with adaptive graph learning," *IEEE Transactions on Cybernetics*, vol. 50, no. 4, pp. 1418–1429, 2020.

[18] N. Xu, Y. Guo, X. Zheng, Q. Wang, and X. Luo, "Partial multi-view subspace clustering," in *ACM Multimedia Conference on Multimedia*. ACM, 2018, pp. 1794–1801.

[19] W. Shao, L. He, and S. Y. Philip, "Multiple incomplete views clustering via weighted nonnegative matrix factorization with $l\_2, 1$ regularization," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 2015, pp. 318–334.

[20] N. Rai, S. Negi, S. Chaudhury, and O. Deshmukh, "Partial multi-view clustering using graph regularized nmf," in *International Conference on Pattern Recognition (ICPR)*. IEEE, 2016, pp. 2192–2197.

[21] J. Wen, Z. Zhang, Y. Xu, B. Zhang, L. Fei, and H. Liu, "Unified embedding alignment with missing views inferring for incomplete multi-view clustering," in *AAAI Conference on Artificial Intelligence*, 2019, pp. 5395–5400.

[22] A. Trivedi, P. Rai, H. Daumé III, and S. L. DuVall, "Multiview clustering with incomplete views," in *Annual Conference on Neural Information Processing Systems Workshop*, vol. 224, 2010.

[23] W. Shao, X. Shi, and S. Y. Philip, "Clustering on multiple incomplete datasets via collective kernel learning," in *IEEE International Conference on Data Mining*. IEEE, 2013, pp. 1181–1186.

[24] X. Liu, M. Li, L. Wang, Y. Dou, J. Yin, and E. Zhu, "Multiple kernel k-means with incomplete kernels," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*. AAAI Press, 2017, pp. 2259–2265.

[25] X. Zhu, X. Liu, M. Li, E. Zhu, L. Liu, Z. Cai, J. Yin, and W. Gao, "Localized incomplete multiple kernel k-means." in *International Joint Conferences on Artificial Intelligence*, 2018, pp. 3271–3277.

[26] X. Liu, X. Zhu, M. Li, C. Tang, E. Zhu, J. Yin, and W. Gao, "Efficient and effective incomplete multi-view clustering," in *AAAI Conference on Artificial Intelligence*, 2019, pp. 1–8.

[27] H. Gao, Y. Peng, and S. Jian, "Incomplete multi-view clustering," in *International Conference on Intelligent Information Processing*. Springer, 2016, pp. 245–255.

[28] W. Zhou, H. Wang, and Y. Yang, "Consensus graph learning for incomplete multi-view clustering," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 2019, pp. 529–540.

[29] L. Zhao, Z. Chen, Y. Yang, Z. J. Wang, and V. C. Leung, "Incomplete multi-view clustering via deep semantic mapping," *Neurocomputing*, vol. 275, pp. 1053–1062, 2018.

[30] Q. Wang, Z. Ding, Z. Tao, Q. Gao, and Y. Fu, "Partial multi-view clustering via consistent gan," in *IEEE International Conference on Data Mining*. IEEE, 2018, pp. 1290–1295.

[31] C. Xu, Z. Guan, W. Zhao, H. Wu, Y. Niu, and B. Ling, "Adversarial incomplete multi-view clustering," in *International Joint Conference on Artificial Intelligence*. AAAI Press, 2019, pp. 3933–3939.