Scalable Cross-view Sample Alignment for Multi-view Clustering with View Structure Similarity

Jun Wang[1], Zhenglai Li[2], Chang Tang[3], Suyuan Liu[1], Hao Yu[1]
Chuan Tang[1], Miaomiao Li[4*], Xinwang Liu[1*]
[1]National University of Defense Technology, [2]Shenzhen Institutes of Advanced Technology
[3]Huazhong University of Science and Technology, [4]Changsha College

Motivation

- Due to semantic discrepancies across views and the absence of supervision, establishing strict one-to-one correspondences is often difficult. In real-world scenarios, the sample relationships of different views are typically many-to-many, and enforcing strict matching may introduce noise and lead to sub-optimal alignment
- While some recent methods employ joint learning frameworks that integrate alignment with feature representation to enhance performance, they often fail to model explicit alignment relationships, thus limiting their scalability to other multi-view clustering methods that are not applicable in sample non-alignment Scenarios
- Clustering performance is heavily influenced by the choice of a benchmark view, yet selecting an appropriate one remains an open challenge in current approaches

Contribution

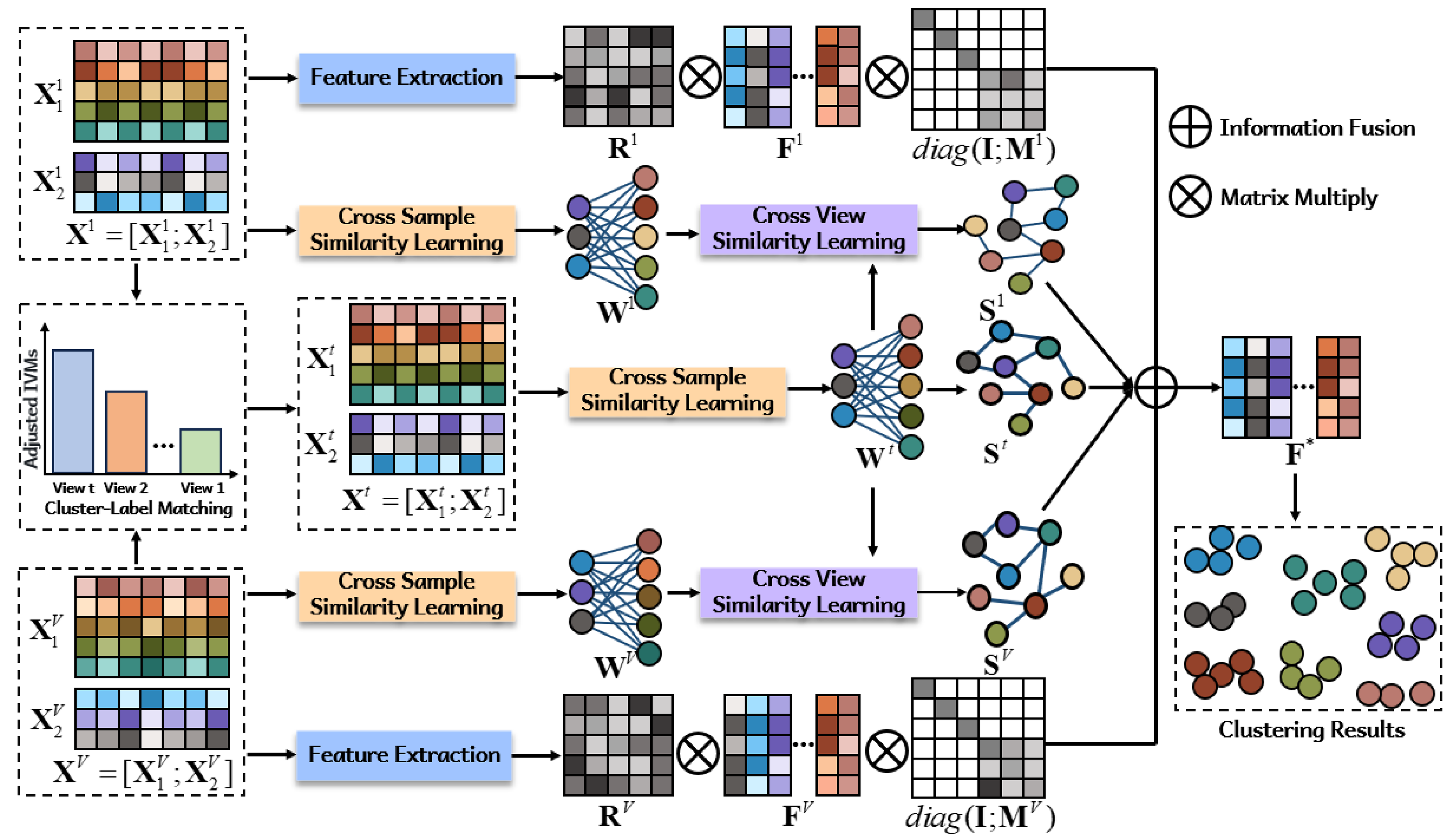
- We propose to select the baseline view by measuring the similarity between sample cluster distributions and their corresponding labels within each view, effectively minimizing the impact of irrelevant or noisy structural information on the alignment process.
- We propose a structural representation for each view based on the correlation between nonaligned and aligned samples. This representation guides cross-view alignment by integrating sample-level features with intrinsic structural information.
- An alternating optimization algorithm is proposed to efficiently solve the model. Its effectiveness is validated through extensive experiments on eight multi-view datasets.

Optimization

Algorithm 1 The Algorithm of SSA-MVC.

Input: Unaligned multi-view data $\{\mathbf{X}^v\}_{v=1}^V$, the number of clusters k, the unified feature dimension d, and the hyper-parameter λ.
Construct the cross-view similarity graph $\{\mathbf{S}^v\}_{v=1}^V$ via Eqs. (3)-(6)).
Initialize $\{\mathbf{R}^v\}_{v=1}^V$, $\{\mathbf{M}^v\}_{v=1}^V$, $\{\alpha_v\}_{v=1}^V$.
while not converge **do**
  Update $\mathbf{F}^*$ via Eq. (8).
  Update $\{\mathbf{R}^v\}_{v=1}^V$ via Eq. (10).
  Update $\{\alpha_v\}_{v=1}^V$ via Eq. (11).
  Update $\{\mathbf{M}^v\}_{v=1}^V$ via Eq. (13).
end while
Conduct k-means clustering algorithm on the consensus partition $\mathbf{F}^*$.
Output: Clustering results $\mathbf{Y}$.

Methods



Baseline View Selection

$$H(Y,\mathbf{X},d^2)=\frac{\exp\left(\frac{1}{\sigma_{d^2}n}\sum_{\mathbf{x}\in\mathbf{X}}d^2(\mathbf{x},y)\right)}{\exp\left(\frac{1}{\sigma_{d^2}n}\sum_{i=1}^{k}\sum_{\mathbf{x}\in Y_i}d^2(\mathbf{x},y_i)\right)}\times\frac{\sum_{i=1}^{k}|Y_i|d^2(y_i,y)}{\sigma_{d^2}\,n\,(k-1)}$$

$$CLM(\mathbf{X})=\frac{1}{2\binom{k}{2}}\sum_{\substack{G\subseteq Y\\|G|=2}}\frac{1}{1+\exp\big(-\delta\cdot H(G,\mathbf{X},d^2)\big)}$$

Cross View Similarity Learning

$$\min_{\mathbf{S}^v}\sum_{\substack{v=1\\v\neq t}}^{V}\|\mathbf{w}_i^t-\mathbf{w}_j^v\|_2^2 s_{ij}^v+\beta(s_{ij}^v)^2\quad s.t.\ t=\arg\max_v CLM(\mathbf{X}^v),\mathbf{s}_i^\top\mathbf{1}=1,0\leq s_{ij}^v\leq 1,$$

Overall Objective Function

$$\max_{\mathbf{R}^v,\mathbf{F}^*,\mathbf{M}^v,\alpha_v}\mathrm{Tr}\left(\mathbf{F}^{*\top}\left(\alpha_t\mathbf{F}^t\mathbf{R}^t+\sum_{\substack{v=1\\v\neq t}}^{V}\alpha_v\begin{bmatrix}\mathbf{I}&0\\0&\mathbf{M}^v\end{bmatrix}\mathbf{F}^v\mathbf{R}^v\right)\right)+\lambda\sum_{v=1}^{V}\mathrm{Tr}(\mathbf{M}^{v\top}\mathbf{S}^v)$$

$$s.t.\ t=\arg\max_v CLM(\mathbf{X}^v),\mathbf{F}^{*\top}\mathbf{F}^*=\mathbf{I},\mathbf{R}^{v\top}\mathbf{R}^v=\mathbf{I},\sum_{v=1}^{V}\alpha_v^2=1,\mathbf{M}^{v\top}\mathbf{M}^v=\mathbf{I},$$

Experiments

Clustering Results

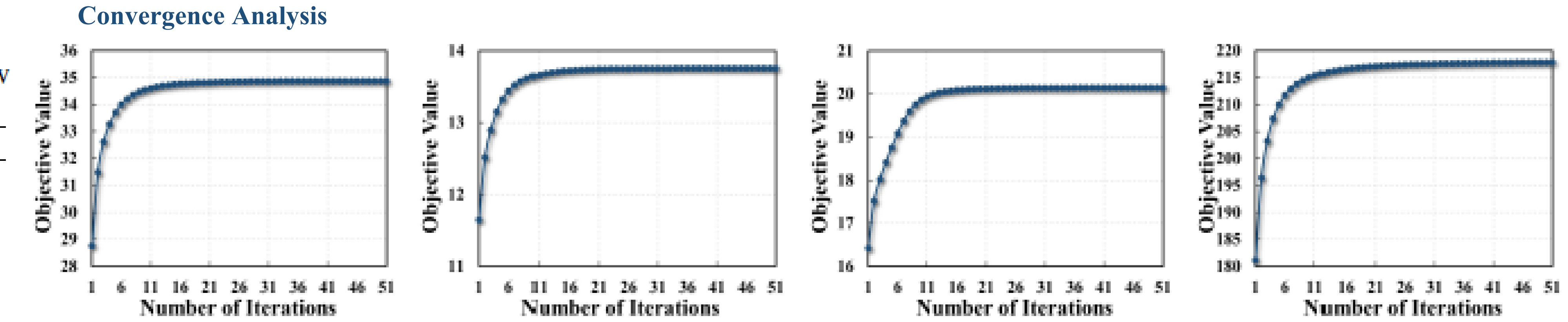
Table 1: ACC comparison of all methods with and without Hungarian alignment on eight multi-view datasets under a sample alignment ratio $\rho = 50\%$.

Method	Yale	3sources	MSRCV	100leaves	HW	Scene	EMNIST	Hdigit
EEOMVC	58.18±0.00	59.76±0.00	72.86±0.00	65.44±0.00	93.85±0.00	26.91±0.00	46.11±0.00	65.93±0.00
EEOMVC + Hungarian	52.73±0.00	48.52±0.00	71.90±0.00	67.19±0.00	67.30±0.00	26.00±0.00	44.24±0.00	62.28±0.00
DealMVC	33.94±0.00	31.01±0.00	28.10±0.00	7.69±0.00	47.94±0.00	22.54±0.00	45.62±0.00	65.66±0.00
DealMVC + Hungarian	24.85±0.00	29.11±0.00	28.29±0.00	9.34±0.00	39.97±0.00	21.86±0.00	37.83±0.00	82.72±0.98
MVCAN	32.48±2.50	31.12±2.31	58.14±1.76	49.51±1.24	50.62±0.42	33.30±0.41	45.06±1.02	65.50±2.98
MVCAN + Hungarian	33.09±1.56	49.37±4.60	51.20±2.80	40.71±1.53	47.08±3.56	29.75±0.54	49.56±9.49	57.37±3.76
EBMGC	39.39±0.00	38.46±0.00	42.86±0.00	33.94±0.00	56.75±0.00	21.58±0.00	33.50±0.00	50.91±0.00
EBMGC + Hungarian	32.73±0.00	40.24±0.00	47.14±0.00	33.94±0.00	51.70±0.00	26.33±0.00	41.00±0.00	59.46±0.00
Vsc_mH	53.94±0.00	62.13±0.00	64.29±0.00	38.56±0.00	42.50±0.00	28.03±0.00	46.47±0.00	65.19±0.00
OpVuC	53.94±0.00	57.40±0.00	30.00±0.00	53.13±0.00	30.10±0.00	31.82±0.00	51.09±0.00	62.90±0.00
DCMVC	27.15±1.01	46.51±4.52	45.71±2.29	48.83±0.83	69.34±1.01	26.02±0.60	59.55±3.60	65.74±2.31
DCMVC + Hungarian	23.88±1.19	35.15±1.10	44.57±2.46	39.34±0.95	50.47±0.99	24.07±0.45	40.11±1.28	35.80±0.99
LMTC	52.58±3.76	48.28±3.49	53.29±3.15	35.58±0.94	64.99±1.16	28.96±0.92	41.91±0.80	59.25±0.30
LMTC + Hungarian	54.61±4.74	48.05±1.18	55.43±3.29	35.30±1.45	54.24±2.72	28.53±0.86	41.69±0.70	55.46±2.13
TMSL	24.82±1.70	42.25±2.66	43.98±0.97	47.47±1.40	62.61±0.61	29.13±0.49	OOM	OOM
TMSL + Hungarian	68.79±2.78	56.21±1.05	44.45±1.45	47.12±1.11	53.12±0.05	27.02±0.24	OOM	OOM
DSTL	35.91±1.93	61.54±0.00	39.48±3.81	36.87±1.42	47.61±1.57	20.45±0.59	28.87±0.40	40.90±0.68
DSTL + Hungarian	37.73±2.76	59.76±0.19	43.71±1.43	30.57±1.03	43.33±0.51	19.31±0.79	30.36±0.33	50.09±0.00
Ours	64.24±3.62	64.44±1.29	83.52±0.39	70.63±1.29	96.55±0.00	35.91±0.27	77.38±3.14	71.78±1.26

Table 2: NMI comparison of all methods with and without Hungarian alignment on eight multi-view datasets under a sample alignment ratio $\rho = 50\%$.

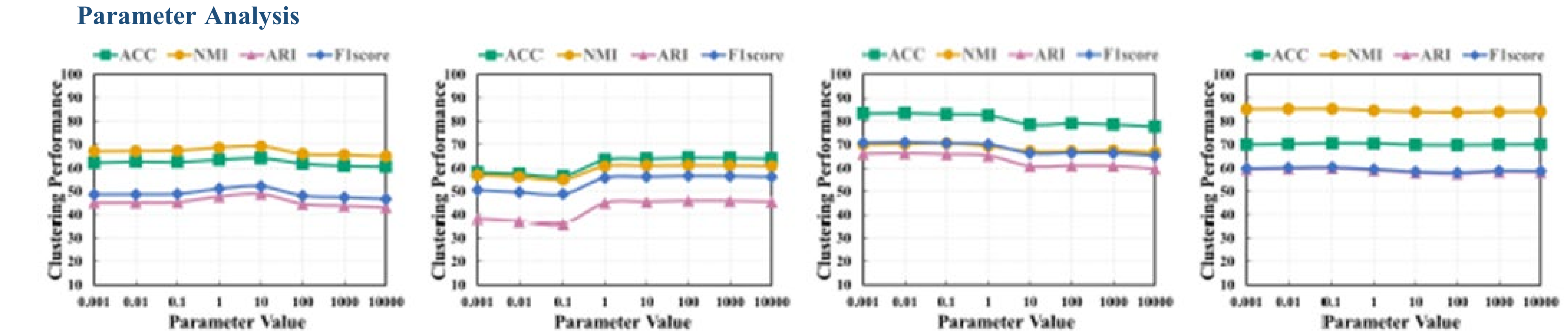
Method	Yale	3sources	MSRCV	100leaves	HW	Scene	EMNIST	Hdigit
EEOMVC	62.23±0.00	39.32±0.00	56.08±0.00	75.62±0.00	88.20±0.00	16.59±0.00	32.54±0.00	70.96±0.00
EEOMVC + Hungarian	57.37±0.00	31.87±0.00	56.93±0.00	77.03±0.00	62.65±0.00	18.26±0.00	29.18±0.00	53.11±0.00
DealMVC	38.08±0.00	6.69±0.74	14.00±3.92	25.34±0.44	27.20±0.89	11.89±1.09	31.39±0.59	39.90±1.52
DealMVC + Hungarian	23.65±0.00	7.31±0.48	13.08±0.17	27.34±3.63	26.08±2.49	16.31±3.23	22.80±0.48	65.46±1.81
MVCAN	38.43±1.87	12.87±1.67	46.19±1.94	69.50±0.95	32.31±0.48	30.96±0.89	20.14±0.13	60.89±1.68
MVCAN + Hungarian	38.53±1.22	47.72±3.00	35.16±4.14	62.08±1.45	45.05±5.03	25.76±0.33	37.94±15.14	53.82±3.54
EBMGC	43.31±0.00	23.68±0.00	22.79±0.00	58.22±0.00	35.99±0.00	11.14±0.00	17.99±0.00	25.07±0.00
EBMGC + Hungarian	38.18±0.00	23.98±0.00	24.79±0.00	58.22±0.00	29.52±0.00	15.08±0.00	19.44±0.00	39.70±0.00
Vsc_mH	62.00±0.00	48.81±0.00	56.01±0.00	68.53±0.00	30.67±0.00	25.41±0.00	36.92±0.00	55.77±0.00
OpVuC	55.77±0.00	36.86±0.00	13.63±0.00	78.00±0.00	17.74±0.00	29.69±0.00	45.94±0.00	47.51±0.00
DCMVC	31.40±1.27	26.22±2.46	33.61±2.59	67.15±0.42	60.71±3.05	15.64±0.32	61.11±2.62	59.83±1.83
DCMVC + Hungarian	27.63±1.07	16.49±1.72	22.27±1.69	60.41±0.47	29.39±0.68	12.99±0.31	20.68±0.19	16.65±0.20
LMTC	57.39±2.89	40.01±4.69	33.29±3.81	58.75±0.55	45.75±0.55	23.24±0.45	24.01±0.25	47.31±0.25
LMTC + Hungarian	57.79±3.67	43.85±2.17	37.60±2.84	58.98±0.97	37.95±0.58	23.16±0.58	24.59±0.87	47.02±2.49
TMSL	28.68±1.25	12.02±1.10	23.71±1.08	69.33±0.61	48.62±0.54	21.34±0.36	OOM	OOM
TMSL + Hungarian	68.40±1.91	31.68±0.92	25.28±0.81	69.46±0.58	28.16±0.07	19.77±0.27	OOM	OOM
DSTL	39.59±1.18	37.00±0.00	23.41±1.99	60.53±0.55	29.05±0.50	15.54±0.33	11.64±0.22	20.40±0.15
DSTL + Hungarian	41.17±1.62	40.46±0.38	24.58±1.13	54.13±0.51	26.08±0.23	14.05±0.50	15.32±0.29	42.34±0.00
Ours	69.31±1.32	61.20±0.83	70.28±0.55	85.34±0.42	92.09±0.00	30.27±0.20	74.84±0.94	75.50±0.14

Convergence Analysis



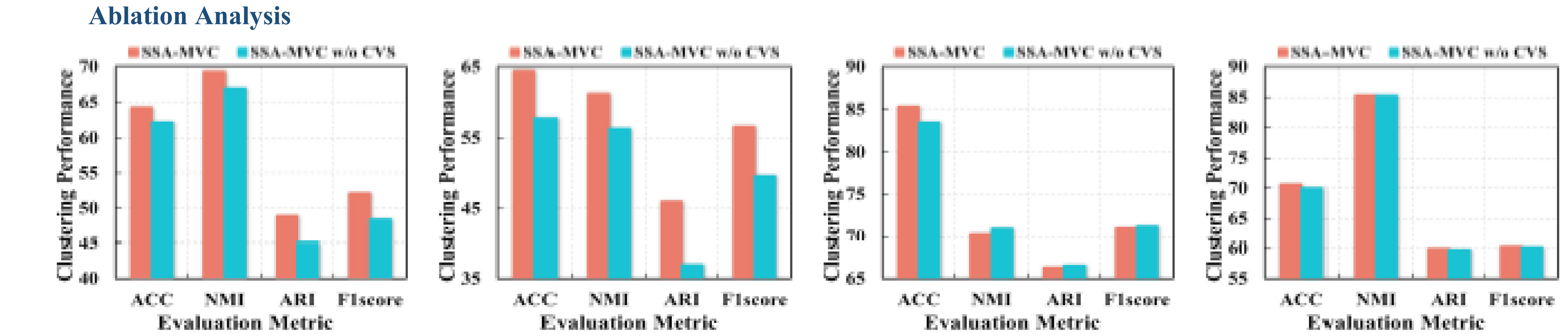
(a) Yale (b) 3sources (c) MSRCV (d) 100leaves

Parameter Analysis



(a) Yale (b) 3sources (c) MSRCV (d) 100leaves

Ablation Analysis



(a) Yale (b) 3sources (c) MSRCV (d) 100leaves