Learn from View Correlation: An Anchor Enhancement Strategy for Multi-view Clustering

Supplementary Material

7. Comparison Results with Other Methods on F-score

Due to space limitations, we only present comparison results on ACC, NMI and Purity in main body. In Table 6, we further provide F-score comparison on all datasets. From the results, in terms of f-score index, the proposed method is still better than the comparative method on most datasets. Among them, AEVC outperformed the suboptimal method by 5.9% on Dermatology dataset. This fully verifies the superiority of the proposed algorithm in clustering performance.

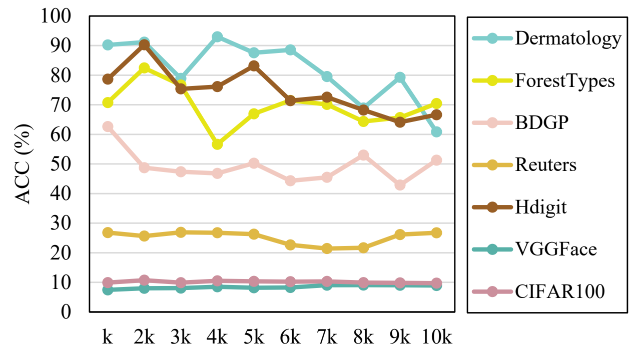


Figure 5. Sensitivity analysis of m and γ on all datasets.

8. Detailed Experimental Results of Sec. 4.3

Due to space limitations, we only present partial results of the ablation study on the anchor enhancement module. In Table 7, we further provide experimental results on all datasets and include the method of enhancement without alignment for comparison. From the experimental results, it can be observed that the clustering performance with both alignment and anchor enhancement consistently outperforms other comparative methods, validating the effectiveness of the proposed enhancement module. It is worth noting that directly enhancing without pre-alignment actually leads to negative effects, resulting in a decrease in clustering performance. This is because the initially generated anchors are not aligned, and the new anchors generated using the unaligned anchor graphs from the neighboring views are unreliable, which also demonstrates the necessity of alignment before enhancement.

9. Detailed Experimental Results of Sec. 4.4

To further validate the effectiveness of each component in the revised graph construction module, we compare the proposed method with the approaches that remove view weights and regularized term in Table 8. From the experimental results, it can be observed that AEVC outperforms the comparison groups without view weights on all datasets, indicating that the strategy of allocating fusion weights based on the importance of views is effective. Additionally, incorporating the regularization term leads to improved clustering performance of the proposed method on all datasets, validating the effectiveness of learning consistent anchor graphs guided by $\hat{\mathbf{Z}}$.

10. Detailed Experimental Results of Sec. 4.6

Sensitivity analysis of λ and γ. We further investigate the impact of the balancing hyperparameter λ and the enhancement rate γ on the final clustering performance on other datasets, and the results are presented in Fig. 6. It is evident that larger values of γ yield better results across all datasets. This observation indicates that higher proportions of newly added anchors during anchor enhancement lead to improved performance, thereby reflecting the effectiveness of anchor augmentation. Moreover, variations in λ also influence the clustering performance. This can be attributed to the guidance of $\hat{\mathbf{Z}}$, which helps adjust the sparsity of $\mathbf{Z}$ and enhances the representation capability of the learned anchor graph.

Sensitivity analysis of m. Fixing λ and γ, we analyze the influence of different numbers of anchors on the clustering performance. We traverse from k to $10k$ for m, where k is the number of clusters, and the corresponding results are shown in Figure 5. From the results, it can be observed that the clustering performance is significantly affected by the variation in the number of anchors in small datasets. Specifically, in the ForestTypes and Hdigit datasets, the clustering performance is optimal when m is set to $2k$. However, in large-scale datasets, the clustering performance of AEVC is less affected by variations in m.

11. Time Comparison

We compare the computational efficiency of different anchor-based MVC algorithms on each benchmark dataset and report their execution time in Table 9. Compared to LMVSC, AEVC has a shorter runtime primarily due to directly learning a consistent anchor graph without additional post-fusion processes. Compared to iteration-optimization based methods like UDBGL and FDAGF, AEVC is a single-step update method that has higher efficiency. In summary, AEVC maintains high efficiency while enhancing the quality of anchors, making it capable of handling the challenges posed by large-scale data.

Table 6. F-score comparison of eight anchor-based MVC methods on seven datasets. The best is marked in bold and underlined, the second best is marked in bold.

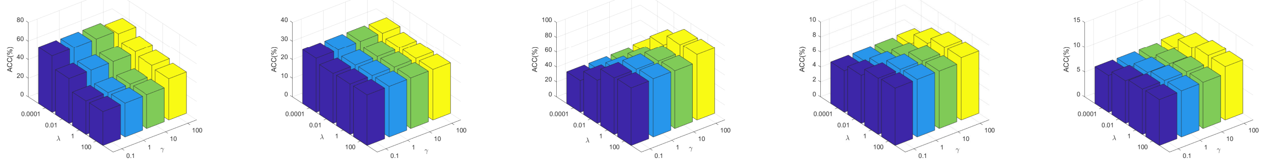
Datasets	LMVSC	SMVSC	FMCNOF	FPMVS-CAG	UDBGL	FastMICE	FDAGF	FMVACC	Proposed
					F-score				
Dermatology	70.50±0.04	70.06±0.04	57.89±0.00	77.37±0.07	84.26±0.00	85.41±0.00	80.70±6.40	78.01±3.85	**90.50±4.18**
ForestTypes	64.24±0.02	57.27±0.01	46.38±0.00	58.28±0.04	62.03±0.00	66.54±0.00	64.40±4.23	64.61±0.25	**69.72±0.13**
BDGP	38.51±0.01	28.81±0.00	28.89±0.00	28.79±0.01	29.43±0.00	39.01±0.00	34.60±1.46	44.58±3.27	**45.42±0.56**
Reuters	21.78±0.00	22.41±0.01	**23.41±0.00**	17.39±0.00	17.98±0.00	23.31±0.00	22.30±0.31	21.71±0.07	22.20±0.01
Hdigit	**85.62±0.07**	50.93±0.02	24.40±0.00	49.71±0.05	20.03±0.00	83.85±0.00	68.50±5.47	77.55±4.42	79.06±2.63
VGGFace	2.47±0.00	3.19±0.00	2.39±0.00	3.14±0.00	2.27±0.00	2.07±0.00	2.76±0.15	2.66±0.06	**3.21±0.03**
CIFAR100	3.52±0.00	4.46±0.00	2.74±0.00	3.74±0.00	2.69±0.00	4.00±0.00	3.44±0.13	3.00±0.09	**4.48±0.05**

Table 7. Ablation study of anchor enhancement module on all datasets.

Methods	Dermatology	ForestTypes	BDGP	Reuters	Hdigit	VGGFace	CIFAR100
Baseline	90.87±3.60	70.76±0.05	54.38±0.10	26.29±0.10	57.31±0.04	5.79±0.07	8.08±0.08
+Align	91.82±3.78	75.51±0.17	57.97±0.09	27.78±0.04	59.61±1.83	5.70±0.07	7.49±0.10
+Ehn.	75.80±2.49	62.92±0.88	45.23±0.27	25.82±0.15	54.49±3.83	5.66±0.12	7.44±0.09
+Align+Ehn.	**93.20±4.50**	**82.53±0.09**	**60.65±0.62**	**29.91±0.02**	**88.01±2.97**	**8.35±0.07**	**10.74±0.08**

Table 8. Ablation study of revised anchor graph construction module on all datasets.

Methods	Dermatology	ForestTypes	BDGP	Reuters	Hdigit	VGGFace	CIFAR100
w/o. view weight	91.60±6.26	81.45±0.00	60.50±0.46	26.84±0.04	86.87±2.05	8.05±0.13	9.20±0.10
w/o. regularized term	89.98±3.40	79.76±0.10	60.65±0.62	29.04±0.00	77.17±2.08	7.14±0.16	8.88±0.09
Proposed	**93.20±4.50**	**82.53±0.09**	**60.65±0.62**	**29.91±0.02**	**88.01±2.97**	**8.35±0.07**	**10.74±0.08**



(a) BDGP (b) Reuters (c) Hdigit (d) VGGFace (e) CIFAR100

Figure 6. Sensitivity analysis of λ and γ on other five datasets.

Table 9. Time comparison of compared algorithms on benchmark datasets (s).

Datasets	LMVSC	SMVSC	FMCNOF	FPMVS-CAG	UDBGL	FastMICE	FDAGF	FMVACC	Proposed
Dermatology	0.60	0.12	0.16	0.13	1.97	0.71	1.34	0.46	0.42
ForestTypes	1.08	0.15	0.08	0.15	2.92	0.90	1.81	1.02	0.54
BDGP	6.25	1.69	0.40	1.25	69.42	8.38	8.63	5.26	4.05
Reuters	40.51	16.21	7.59	11.62	93.42	49.32	67.19	96.45	22.91
Hdigit	39.00	28.36	1.46	21.78	112.94	15.08	36.46	42.53	31.88
VGGFace	849.54	365.28	33.93	341.81	2416.00	54.17	712.75	920.23	497.03
CIFAR100	1111.40	728.38	60.39	527.31	3526.10	71.66	1093.50	1402.40	930.20