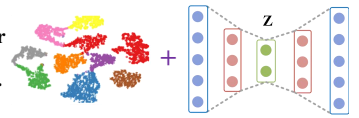
Deep Fusion Clustering Network

Wenxuan Tu[1], Sihang Zhou[1], Xinwang Liu[1], Xifeng Guo[1], ZhipingCai[1], En Zhu[1], Jieren Cheng[2]

[1] National University of Defense Technology, Changsha, China; [2] Hainan University, Haikou, China

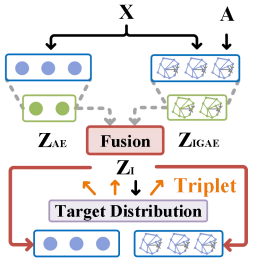
Background

The goal of deep (graph) clustering is to group similar data points together without supervision or prior knowledge of nature of the clusters using deep learning techniques.



Challenges & Motivation

- ✓ Information from two sources is simply aligned or concatenated, leading to insufficient information interaction and merging.
- ✓ Existing methods seldom use information from both sources for target distribution generation, making the guidance of network training less comprehensive and accurate.



1. Topology provides diverse structure information, while node contains abundant attributes. The negotiation between two kinds of information sources is essential for intrinsic data structure revealing.

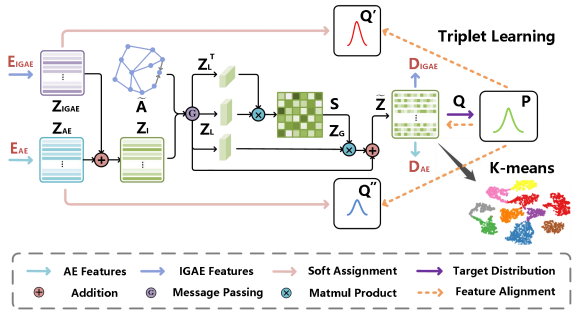
2. A strong tendency that clustering-oriented representation learning benefits from a reliable and robust self-supervised-based guidance.

Contributions

- ✓ We propose a deep fusion clustering network (DFCN). In this network, a structure and attribute information fusion (SAIF) module is designed for better both-source information interaction.
- ✓ We develop an improved graph autoencoder (IGAE) to further improve the generalization capability of the DFCN.
- ✓ Extensive experiments on six datasets verify the effectiveness of our proposed method.

Method

Structure and Attribute Information Fusion Module



Cross-modality Dynamic Fusion Mechanism

① $\mathbf{Z}_I = \alpha \mathbf{Z}_{AE} + (1-\alpha)\mathbf{Z}_{IGAE}$

Initial fused embedding

② $\mathbf{Z}_L = \widetilde{\mathbf{A}}\mathbf{Z}_I$

Local structure enhanced Z_I

③ $\mathbf{S}_{ij} = \frac{e^{(\mathbf{Z}_L \mathbf{Z}_L^{\mathbf{T}})_{ij}}}{\sum_{k=1}^{N} e^{(\mathbf{Z}_L \mathbf{Z}_L^{\mathbf{T}})_{ik}}}$

Normalized self-correlation matrix

④ $\mathbf{Z}_G = \mathbf{S}\mathbf{Z}_L$

Global structure enhanced Z_L

⑤ $\widetilde{\mathbf{Z}} = \beta \mathbf{Z}_G + \mathbf{Z}_L$

Clustering embedding

✓ Fuse and refine the information from both AE and IGAE for learning consensus latent representations.

Triplet Self-supervised Strategy

$$q_{ij} = \frac{(1+\|\tilde{z}_i - u_j\|^2/v)^{-\frac{v+1}{2}}}{\sum_{j'}(1+\|\tilde{z}_i - u_{j'}\|^2/v)^{-\frac{v+1}{2}}}$$

Soft assignment distribution

$$p_{ij} = \frac{q_{ij}^2 / \sum_i q_{ij}}{\sum_{j'}(q_{ij'}^2 / \sum_i q_{ij'})}$$

Target distribution

$$L_{KL} = \sum_i \sum_j p_{ij} log \frac{p_{ij}}{(q_{ij} + q_{ij}^{'} + q_{ij}^{''})/3}$$

Triplet clustering loss

✓ Generate more reliable guidance for the training of AE, IGAE, and the fusion part, making them benefit from each other.

Optimization Target

$$L = \underbrace{L_{AE} + L_{IGAE}}_{Reconstruction} + \underbrace{\lambda L_{KL}}_{Clustering}$$

✓ λ is a pre-defined coefficient set to 10 which balances the importance of two procedures.

Experimental Results

Overall Results

Data	Metric	K-means	AE	DEC	IDEC	GAE	VGAE	ARGA	DAEGC	SDCN$_Q$	SDCN	DFCN
USPS	ACC	66.8±0.0	71.0±0.0	73.3±0.2	76.2±0.1	63.1±0.3	56.2±0.7	66.8±0.7	73.6±0.4	77.1±0.2	78.1±0.2	79.5±0.2
	NMI	62.6±0.0	67.5±0.0	70.6±0.3	75.6±0.1	60.7±0.6	51.1±0.4	61.6±0.3	71.1±0.2	77.7±0.2	79.5±0.3	82.8±0.3
	ARI	54.6±0.0	58.8±0.1	63.7±0.3	67.9±0.1	50.3±0.6	41.0±0.6	51.1±0.6	63.3±0.3	70.2±0.2	71.8±0.2	75.3±0.2
	F1	64.8±0.0	69.7±0.0	71.8±0.2	74.6±0.1	61.8±0.4	53.6±1.1	66.1±1.2	72.5±0.5	75.9±0.2	77.0±0.2	78.3±0.2
HHAR	ACC	60.0±0.0	68.7±0.3	69.4±0.3	71.1±0.4	62.3±1.0	71.3±0.4	63.3±0.8	76.5±2.2	83.5±0.2	84.3±0.2	87.1±0.1
	NMI	58.9±0.0	71.4±1.0	72.9±0.4	74.2±0.4	55.1±1.4	63.0±0.4	57.1±1.4	69.1±2.3	78.8±0.3	79.9±0.1	82.2±0.1
	ARI	46.1±0.0	60.4±0.9	61.3±0.5	62.8±0.5	42.6±1.6	51.5±0.7	44.7±1.0	60.4±2.2	71.8±0.2	72.8±0.1	76.4±0.1
	F1	58.3±0.0	66.4±0.3	67.3±0.3	68.6±0.3	62.6±1.0	71.6±0.3	61.1±0.9	76.9±2.2	81.5±0.1	82.6±0.1	87.3±0.1
REUT	ACC	54.0±0.0	74.9±0.2	73.6±0.1	75.4±0.1	54.4±0.3	60.9±0.2	56.2±0.2	65.6±0.1	79.3±0.1	77.2±0.2	77.7±0.2
	NMI	41.5±0.5	49.7±0.3	47.5±0.3	50.3±0.2	25.9±0.4	25.5±0.2	28.7±0.3	30.6±0.3	56.9±0.3	50.8±0.2	59.9±0.4
	ARI	28.0±0.4	49.6±0.4	48.4±0.1	51.3±0.2	19.6±0.2	26.2±0.4	24.5±0.4	31.1±0.2	59.6±0.3	55.4±0.4	59.8±0.4
	F1	41.3±2.4	61.0±0.2	64.3±0.2	63.2±0.1	43.5±0.4	57.1±0.2	51.1±0.2	61.8±0.1	66.2±0.2	65.5±0.1	69.6±0.1
ACM	ACC	67.3±0.7	81.8±0.1	84.3±0.8	85.1±0.5	84.5±1.4	84.1±0.2	86.1±1.2	86.9±2.8	87.0±0.1	90.5±0.2	90.9±0.2
	NMI	32.4±0.5	49.3±0.2	54.5±1.5	56.6±1.2	55.4±1.9	53.2±0.5	55.7±1.4	56.2±4.2	58.9±0.2	68.3±0.3	69.4±0.4
	ARI	30.6±0.7	54.6±0.2	60.6±1.9	62.2±1.5	59.5±3.1	57.7±0.7	62.9±2.1	59.4±3.9	65.3±0.2	73.9±0.4	74.9±0.4
	F1	67.6±0.7	82.0±0.1	84.5±0.7	85.1±0.5	84.7±1.3	84.2±0.2	86.1±1.2	87.1±2.8	86.8±0.1	90.4±0.2	90.8±0.2
DBLP	ACC	38.7±0.7	51.4±0.4	58.2±0.6	60.3±0.6	61.2±1.2	58.6±0.1	61.6±1.0	62.1±0.5	65.7±1.3	68.1±1.8	76.0±0.8
	NMI	11.5±0.4	25.4±0.2	29.5±0.3	31.2±0.5	30.8±0.9	26.9±0.1	26.8±1.0	32.5±0.5	35.1±1.1	39.5±1.3	43.7±1.0
	ARI	7.0±0.4	12.2±0.4	23.9±0.4	25.4±0.6	22.0±1.4	17.9±0.1	22.7±0.3	21.0±0.5	34.0±1.8	39.2±2.0	47.0±1.5
	F1	31.9±0.3	52.5±0.4	59.4±0.5	61.3±0.6	61.4±2.2	58.7±0.1	61.8±0.9	61.8±0.7	65.8±1.2	67.7±1.5	75.7±0.8
CITE	ACC	39.3±3.2	57.1±0.1	55.9±0.2	60.5±1.4	61.4±0.8	61.0±0.4	56.9±0.7	64.5±1.4	61.7±1.1	66.0±0.3	69.5±0.2
	NMI	16.9±3.2	27.6±0.1	28.3±0.3	27.2±2.4	34.6±0.7	32.7±0.3	34.5±0.8	36.4±0.9	34.4±1.2	38.7±0.3	43.9±0.2
	ARI	13.4±3.0	29.3±0.1	28.1±0.4	25.7±2.7	33.6±1.2	33.1±0.5	33.4±1.5	37.8±1.2	35.5±1.5	40.2±0.4	45.5±0.3
	F1	36.1±3.5	53.8±0.1	52.6±0.2	61.6±1.4	57.4±0.8	57.7±0.5	54.8±0.8	62.2±1.3	57.8±1.0	63.6±0.2	64.3±0.2

Table 1 Clustering performance on six datasets (mean±std). The red and blue values indicate the best and the runner-up results, respectively.

Ablation Study

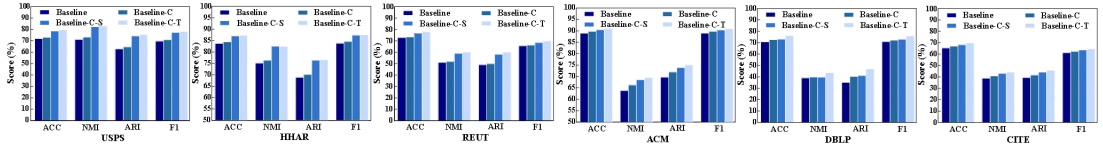


Figure 1 Ablation comparisons of cross-modality dynamic fusion mechanism and triplet self-supervised strategy in SAIF.

Dataset	Model	ACC	NMI	ARI	F1
USPS	+AE	78.3±0.3	81.3±0.1	73.6±0.3	76.8±0.3
	+IGAE	76.9±0.4	77.1±0.4	68.8±0.6	74.8±0.5
	DFCN	**79.5±0.2**	**82.8±0.3**	**75.3±0.2**	**78.3±0.2**
HHAR	+AE	75.2±1.4	**82.8±1.0**	71.7±1.2	72.6±0.9
	+IGAE	82.8±0.1	79.6±0.1	72.3±0.1	83.4±0.1
	DFCN	**87.1±0.1**	82.2±0.1	**76.4±0.1**	**87.3±0.1**
REUT	+AE	69.3±0.8	48.5±1.6	44.6±1.1	58.3±0.6
	+IGAE	71.4±1.7	52.5±1.0	49.1±2.2	61.5±2.9
	DFCN	**77.7±0.2**	**59.9±0.4**	**59.8±0.4**	**69.6±0.1**
ACM	+AE	90.2±0.3	67.5±0.8	73.2±0.8	90.2±0.3
	+IGAE	89.6±0.2	65.6±0.4	71.8±0.4	89.6±0.2
	DFCN	**90.9±0.2**	**69.4±0.4**	**74.9±0.4**	**90.8±0.2**
DBLP	+AE	64.2±2.9	30.2±3.2	29.4±3.4	64.6±2.8
	+IGAE	67.5±1.0	34.2±1.1	31.5±1.1	67.6±1.0
	DFCN	**76.0±0.8**	**43.7±1.0**	**47.0±1.5**	**75.7±0.8**
CITE	+AE	69.3±0.3	42.9±0.4	44.7±0.4	**64.4±0.3**
	+IGAE	67.9±0.9	41.8±1.0	43.0±1.4	63.7±0.7
	DFCN	**69.5±0.2**	**43.9±0.2**	**45.5±0.3**	64.3±0.2

Table 2 Ablation comparisons of the target distribution generation with signle- or both-source information.

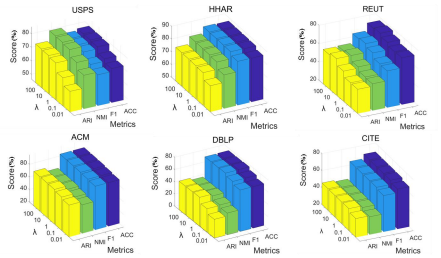


Figure 2 The sensitivity of DFCN with the variation of λ.

- ✓ Extensive experiments on six benchmark datasets have demonstrated that the proposed DFCN consistently outperforms the state-of-the-art deep clustering methods.
- ✓ The results of ablation studies have demonstrated the superiority of each component of our method.