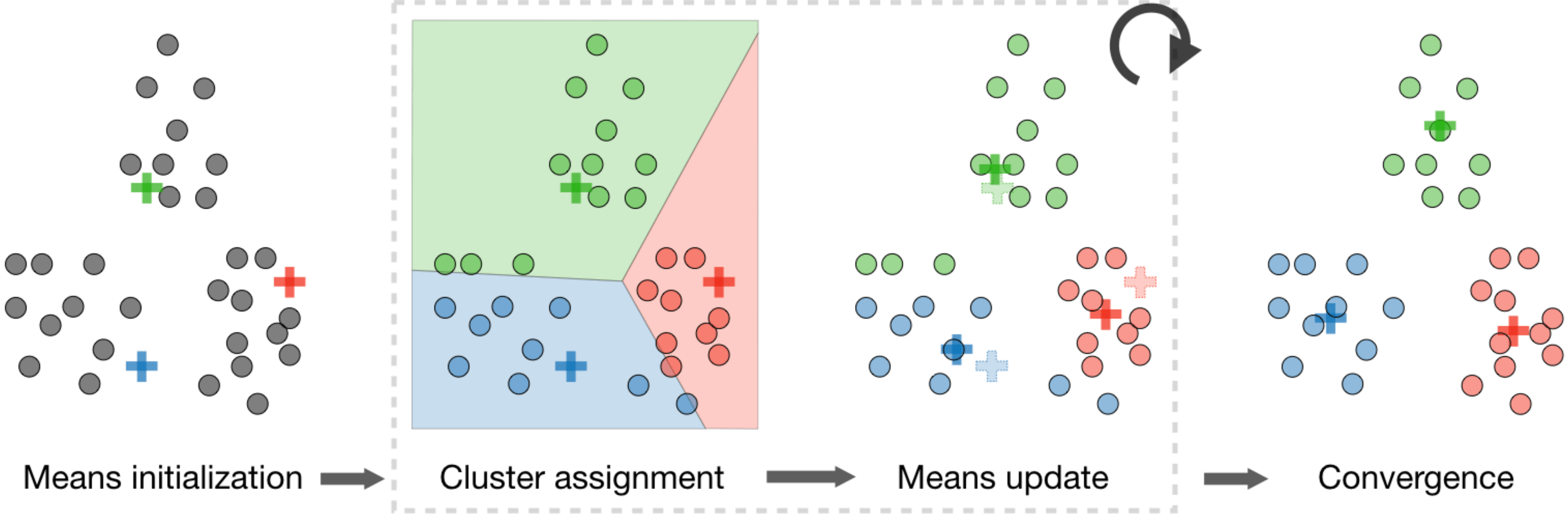
聚类：由简入繁

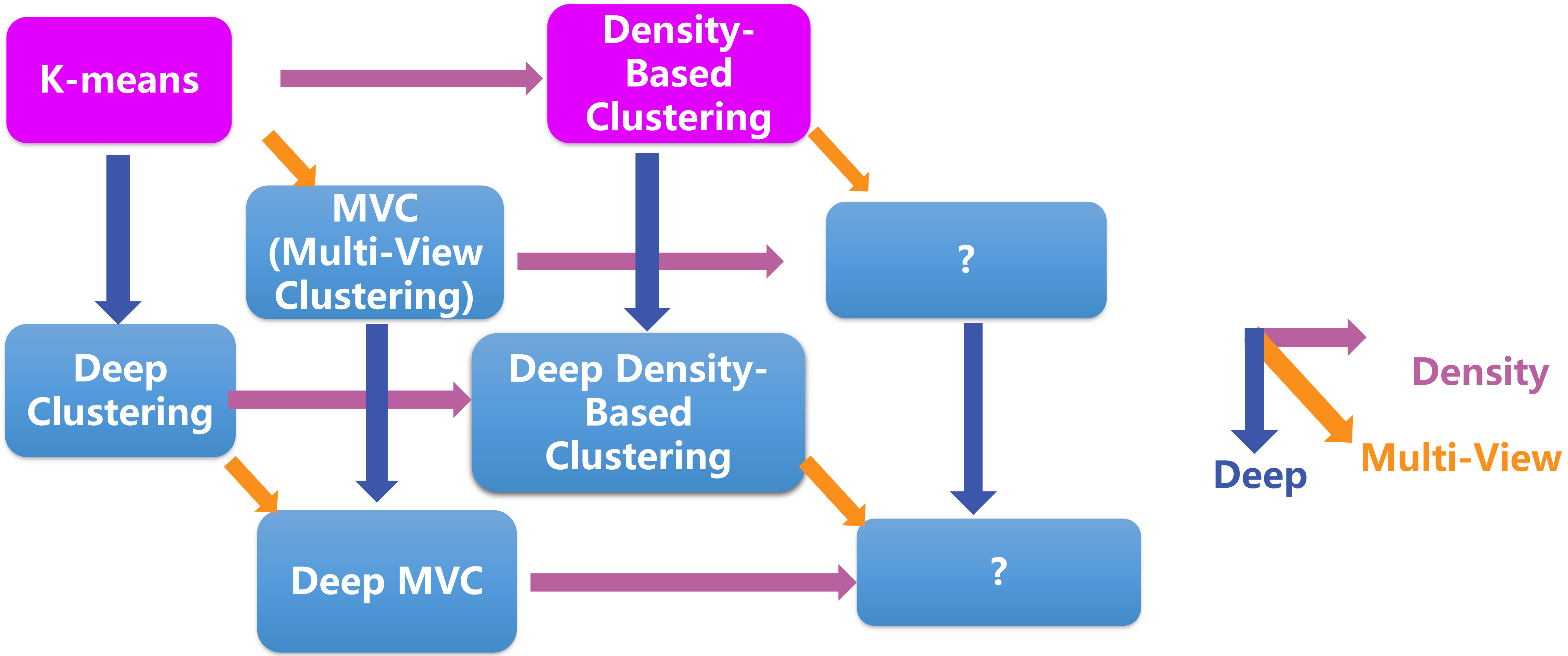
任亚洲
yazhou.ren@uestc.edu.cn
计算机科学与工程学院

无监督学习-k均值聚类

$$\underset{\mathbf{S}}{\arg\min} \sum_{i=1}^{k} \sum_{\mathbf{x} \in S_i} \left\| \mathbf{x} - \boldsymbol{\mu}_i \right\|^2$$

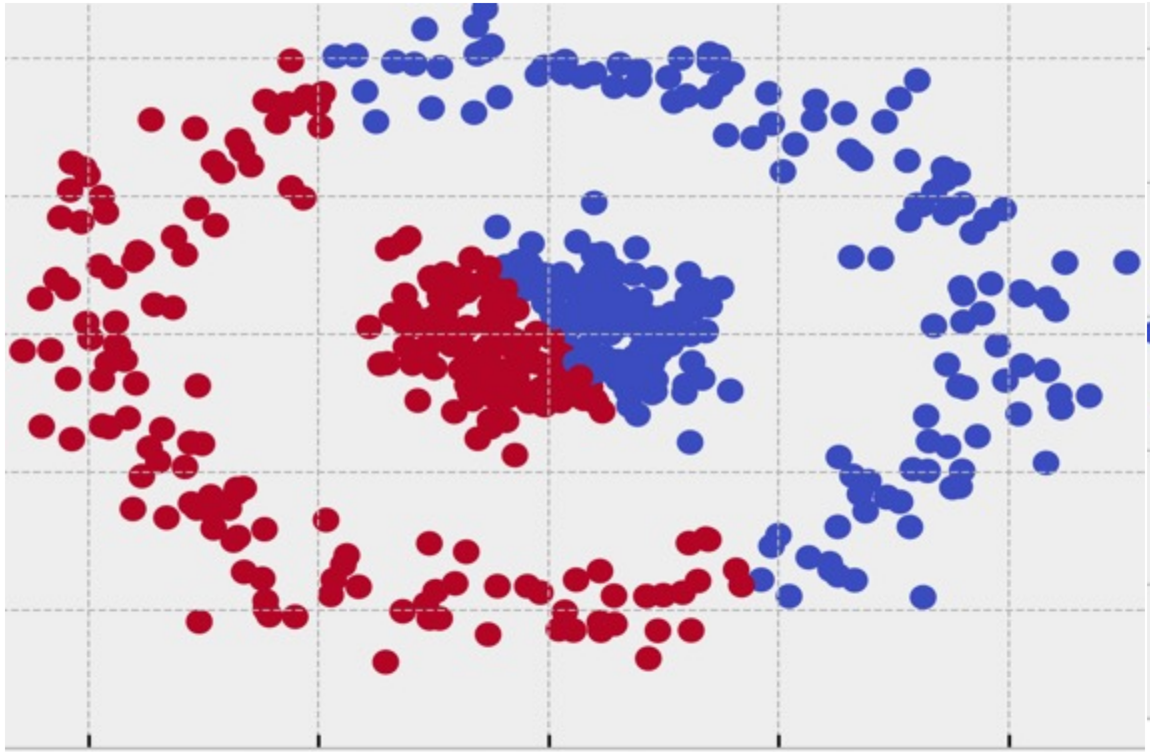


MacQueen, J. (1967, June). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability* (Vol. 1, No. 14, pp. 281-297).

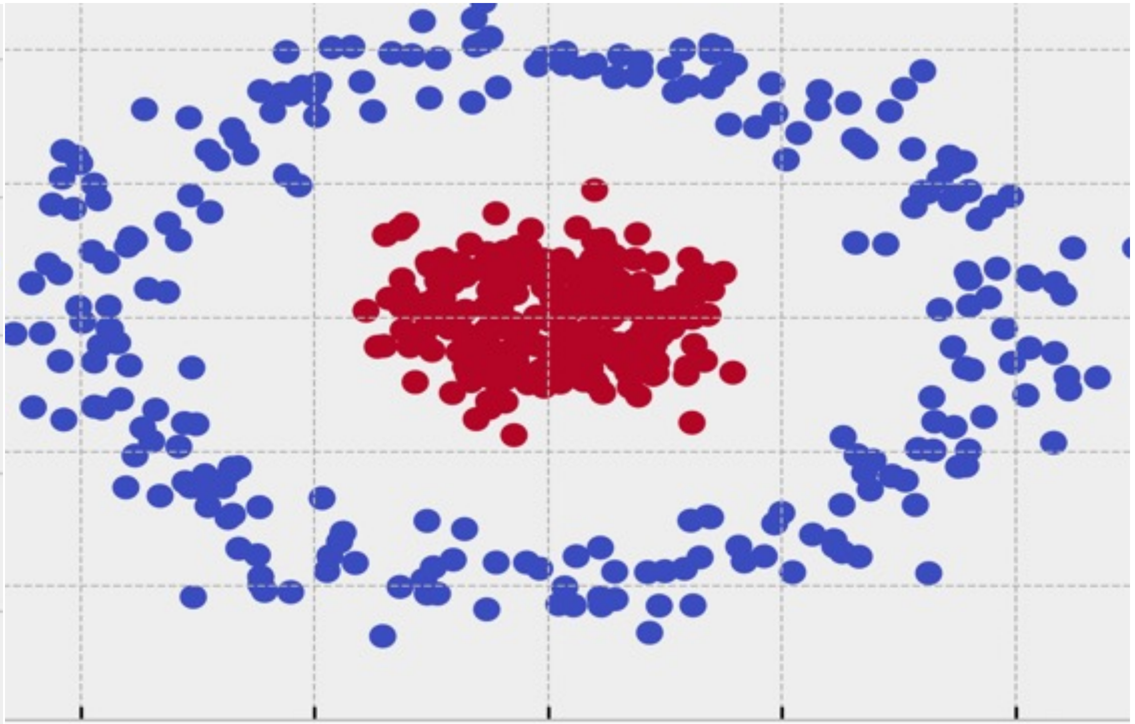

K-means
Density-Based Clustering
MVC (Multi-View Clustering)
?
Deep Clustering
Deep Density-Based Clustering
Deep MVC
?
Density
Multi-View
Deep

1996 : Density-based Clustering

K-Means



DBSCAN



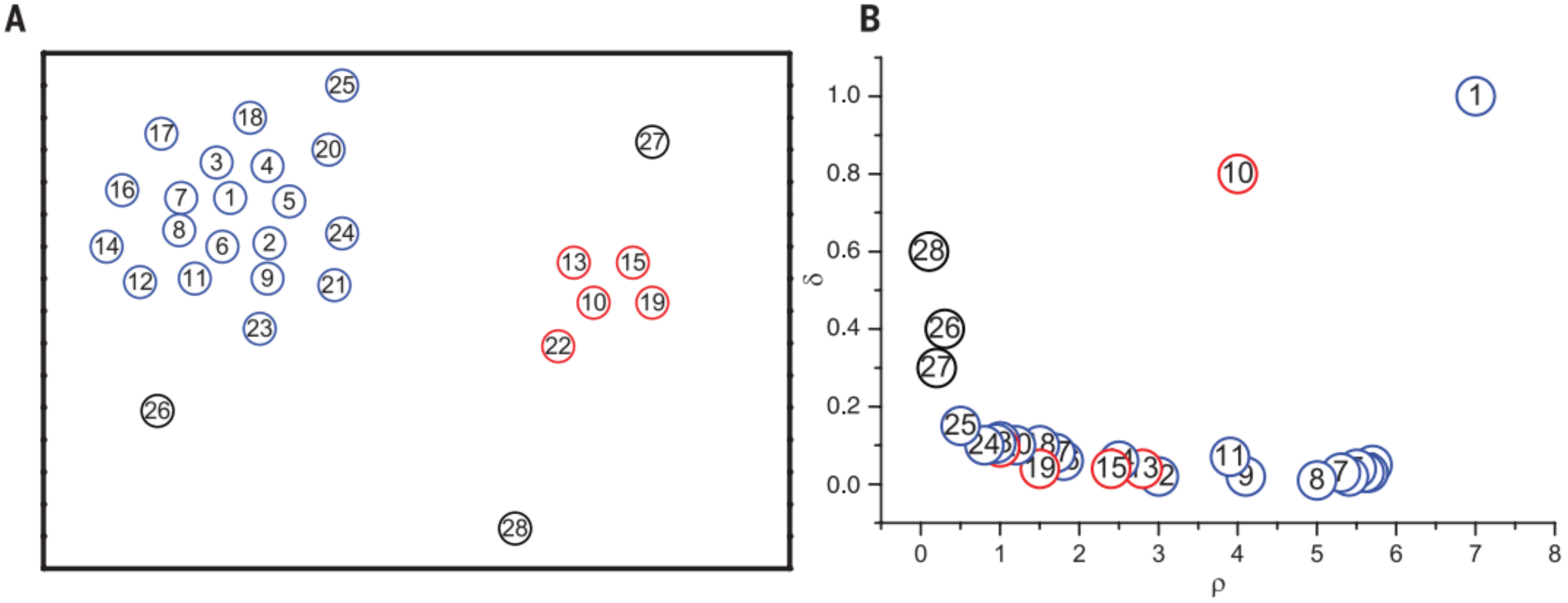
Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996, August). A density-based algorithm for discovering clusters in large spatial databases with noise. In Kdd (Vol. 96, No. 34, pp. 226-231).

Clustering by fast search and find of density peaks

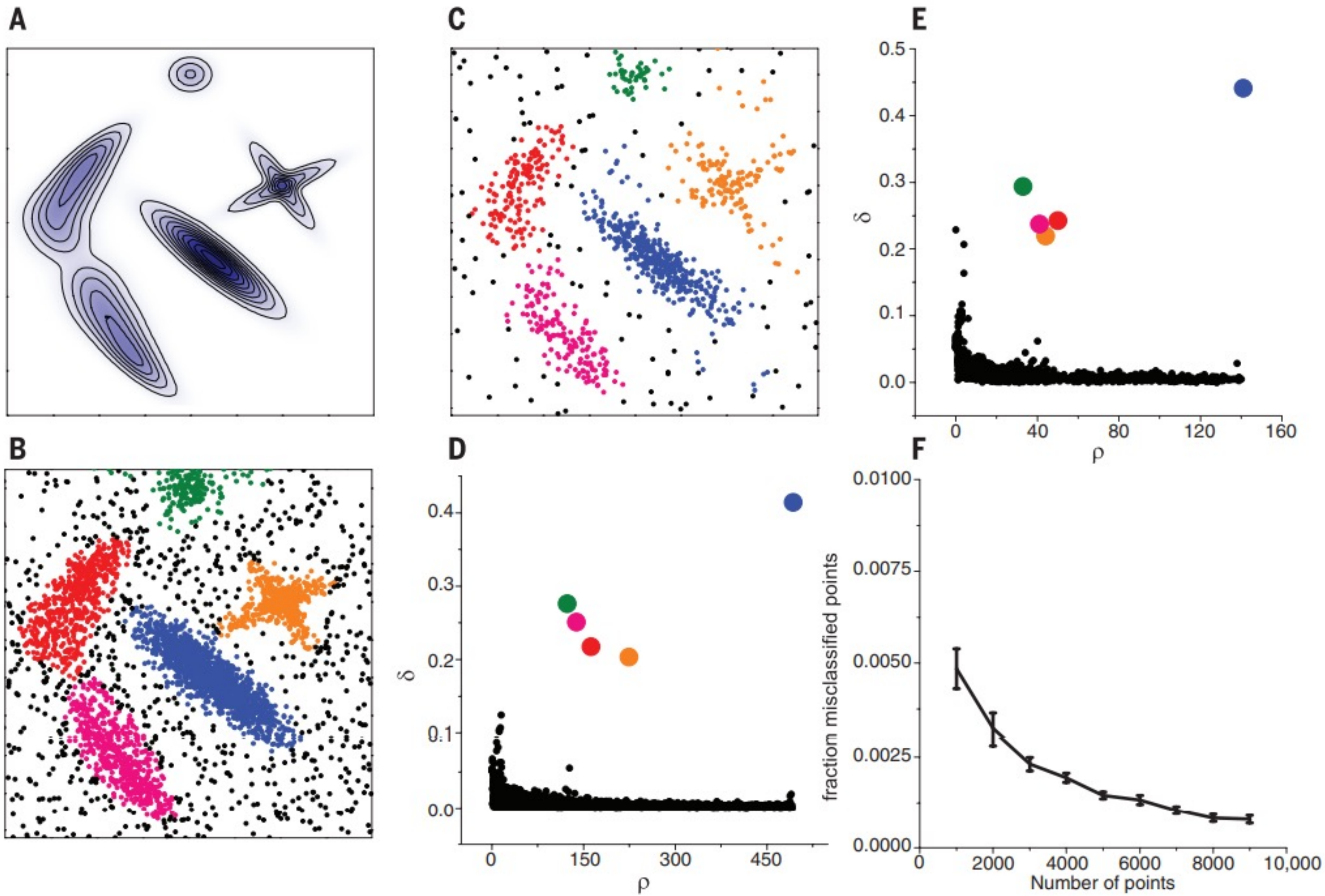
local density：$\rho_i = \sum_j \chi(d_{ij} - d_c)$

$$\chi(x) = 1 \text{ if } x < 0 \text{ and } \chi(x) = 0 \text{ otherwise},$$

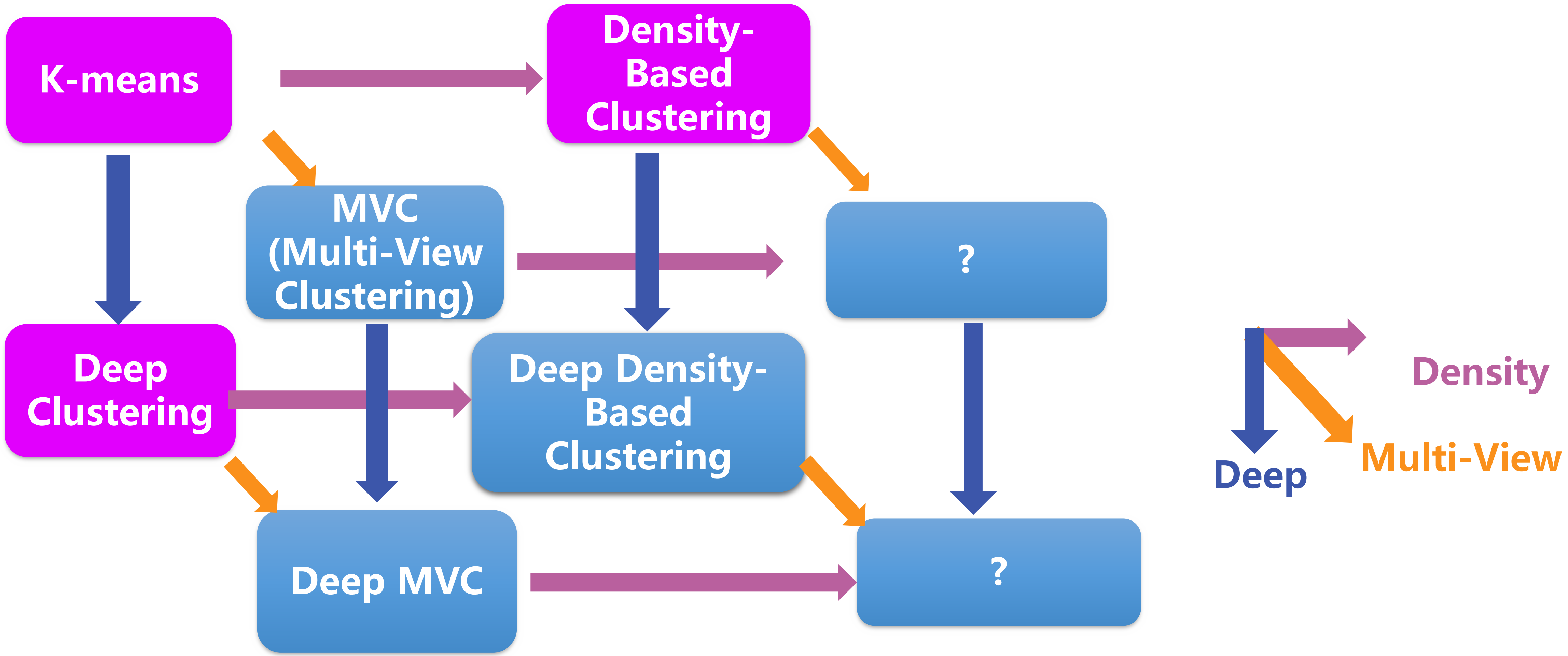
distance δ_i from points of higher density： $\delta_i = \min_{j:\rho_j > \rho_i} (d_{ij})$



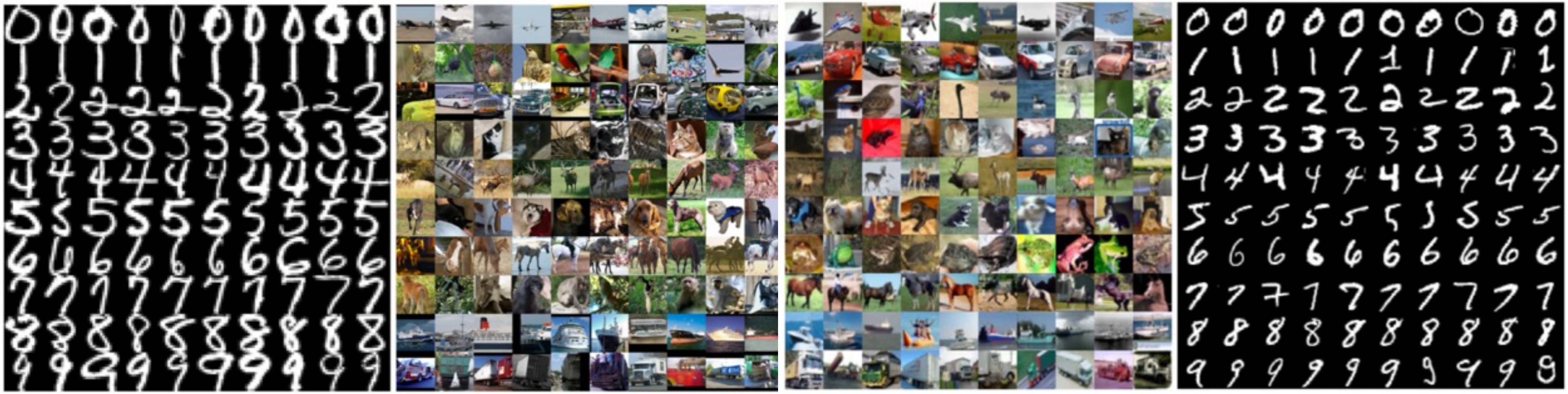
Rodriguez, A., & Laio, A. (2014). Clustering by fast search and find of density peaks. Science, 344(6191), 1492-1496.



Rodriguez, A., & Laio, A. (2014). Clustering by fast search and find of density peaks. science, 344(6191), 1492-1496.

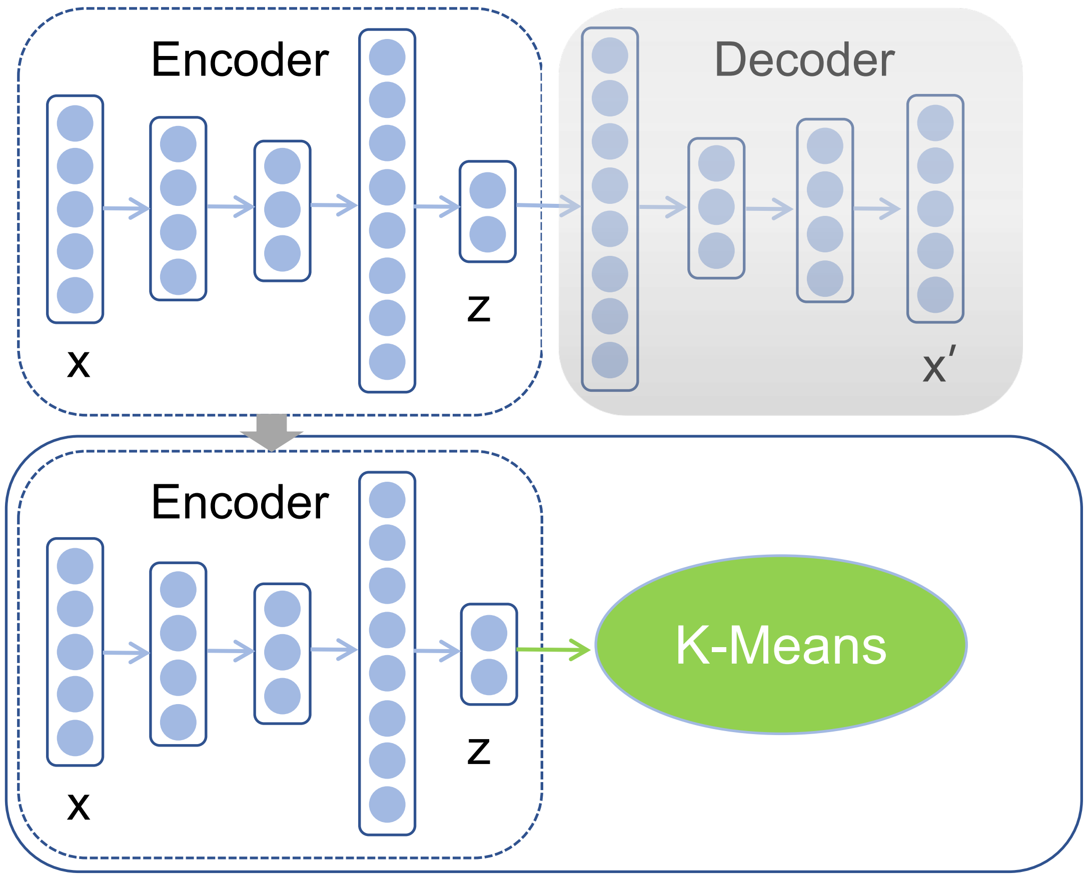

K-means
Density-Based Clustering
MVC (Multi-View Clustering)
?
Deep Clustering
Deep Density-Based Clustering
Deep MVC
?
Density
Multi-View
Deep

Limitations of shallow models



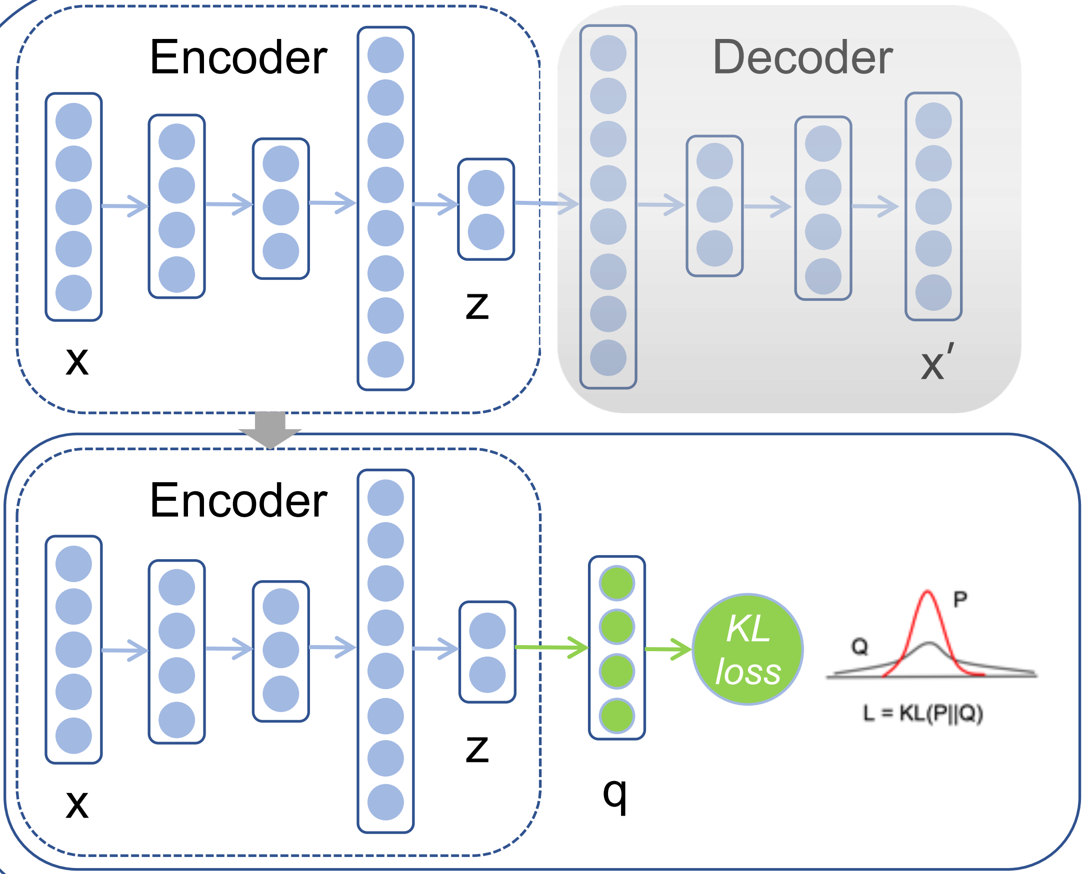
(a) USPS (b) STL-10 (c) CIFAR-10 (d) MNIST

Deep clustering via AE



Perform clustering on learned features

Deep clustering via AE

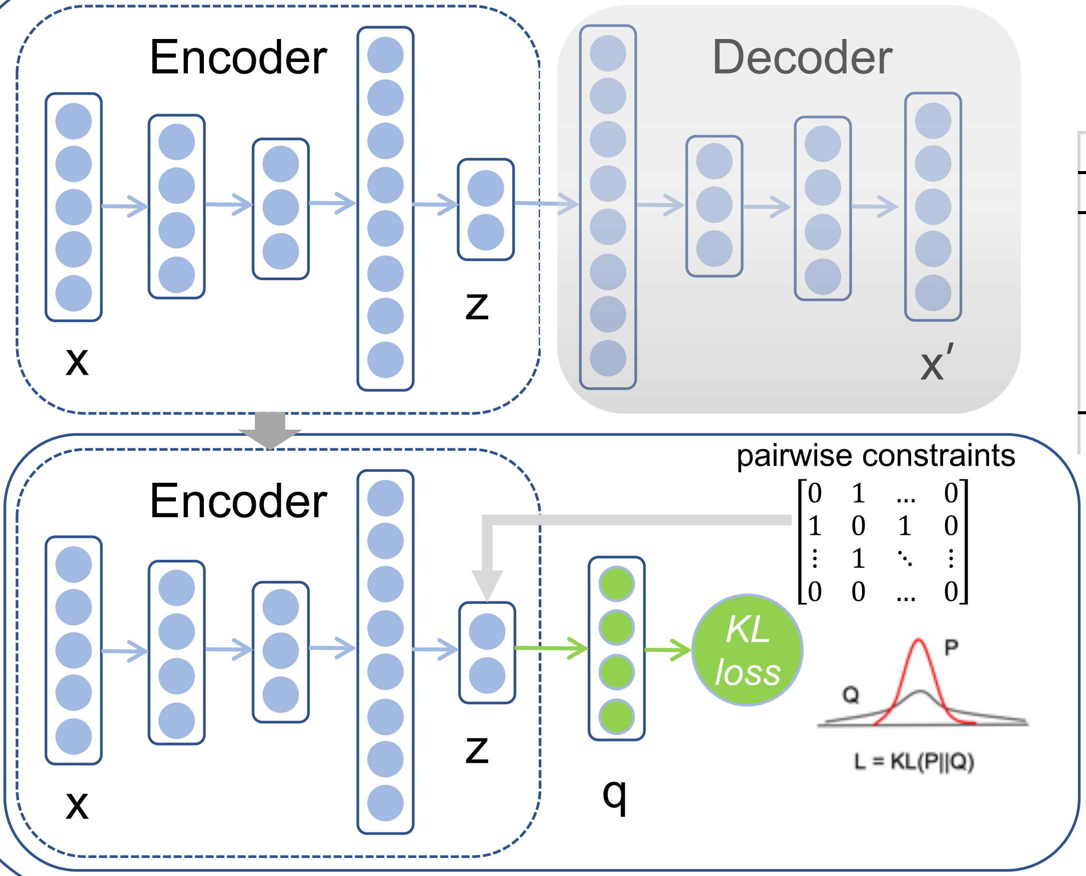


- DEC：

Xie, J., Girshick, R., & Farhadi, A. Unsupervised Deep Embedding for Clustering Analysis. ICML, 2016 (pp. 478-487).

[0.6, 0.4] → [0.8, 0.2]
Q P

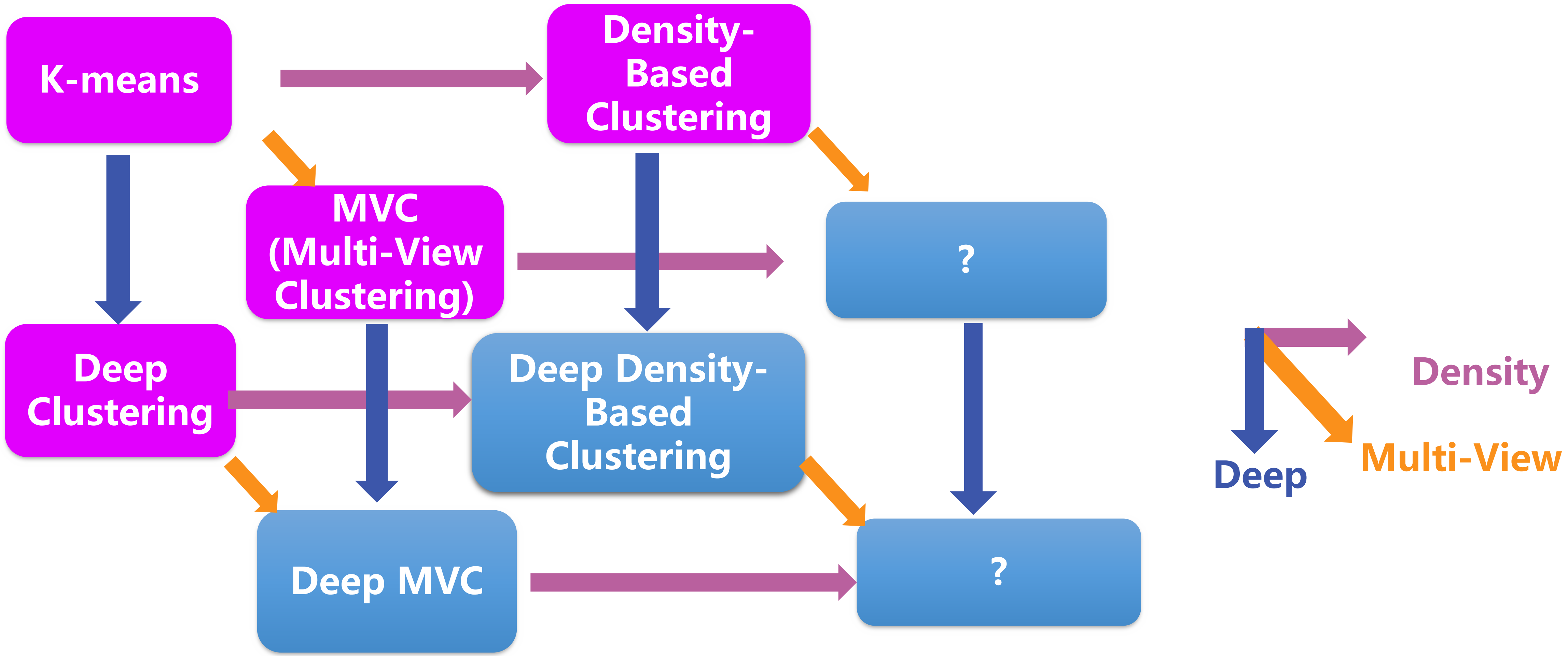
Deep clustering via AE



Clustering results measured by ACC(%).

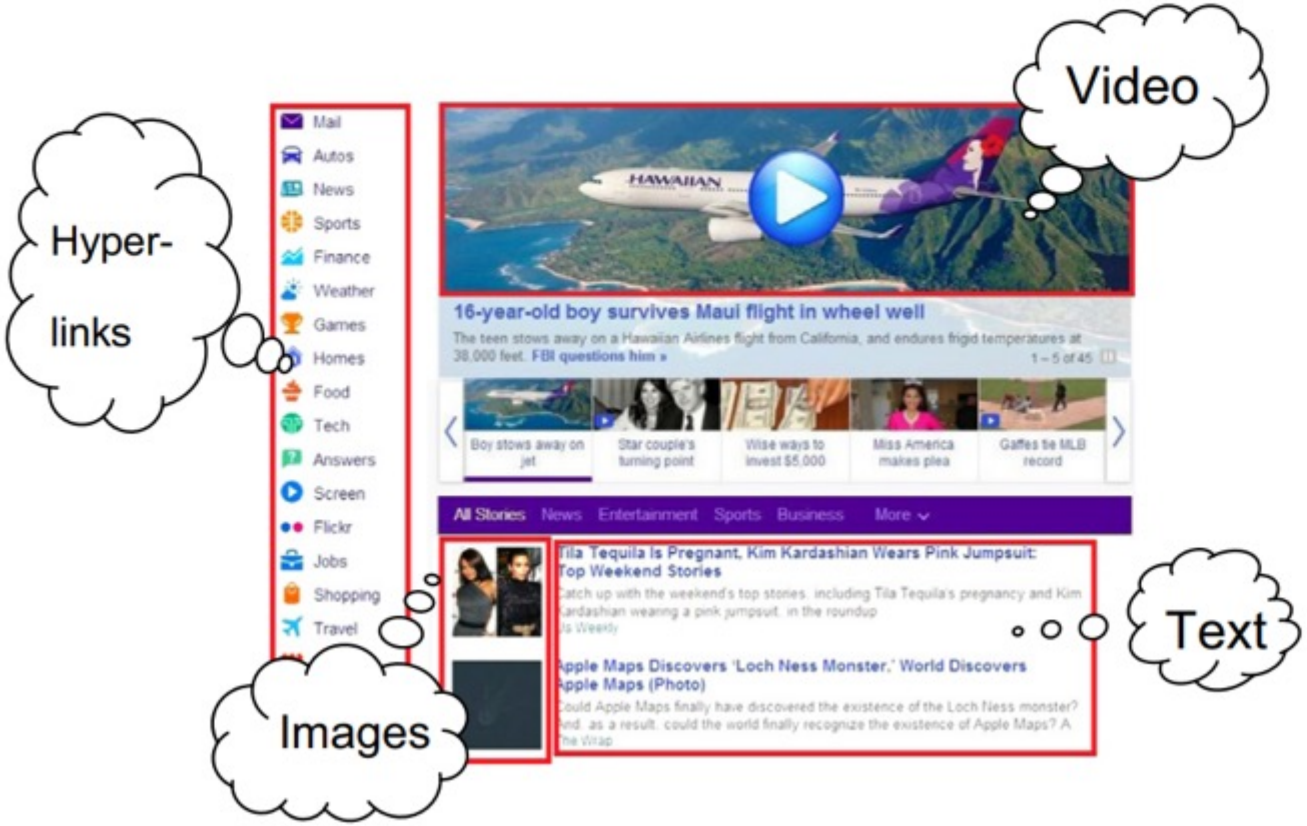
Data	k-means	KM-cst	AE+KM	AE+KM-cst	DEC	IDEC	SDEC
USPS	65.67	68.18	70.28	71.87	75.81	75.86	**76.39**
STL-10	28.31	29.09	34.00	35.15	37.40	36.99	**38.86**
CIFAR-10	23.75	23.91	23.89	24.36	26.26	25.02	**27.26**
MNIST	52.98	54.27	74.09	75.98	84.94	83.85	**86.11**
20NG	33.77	33.89	40.81	47.71	50.11	53.63	**78.12**

Yazhou Ren et al., Semi-supervised Deep Embedded Clustering. Neurocomputing , 2019


K-means
Density-Based Clustering
MVC (Multi-View Clustering)
?
Deep Clustering
Deep Density-Based Clustering
Deep MVC
?
Density
Deep
Multi-View

Multi-View Clustering

While in real-world, an object can be always describe by multiple views.
Conventional clustering methods only work on single-view data.



$$\min_{C^v,B} \sum_{v=1}^{m} ||X^v - C^v B||_F^2$$

2011
Co-regularized multi-view spectral clustering
Kumar et al. in NIPS

2013
Multi-view k-means clustering on big data
Cai et al. in IJCAI

2015
Multi-view subspace clustering
Gao et al. in ICCV

2017
Self-weighted multi-view clustering with multiple graphs
Nie et al. in IJCAI

2021
Dual self-paced multi-view clustering，**Ren et al.** in Neural Networks

Self-Paced Learning

The difficulty levels of learning different instances vary extensively

Simple samples:

Complex samples:

Self-Paced Learning

The difficulty levels of learning different instances vary extensively

Simple samples:

Result in the non-convex issue!

Complex samples:

Self-Paced Learning

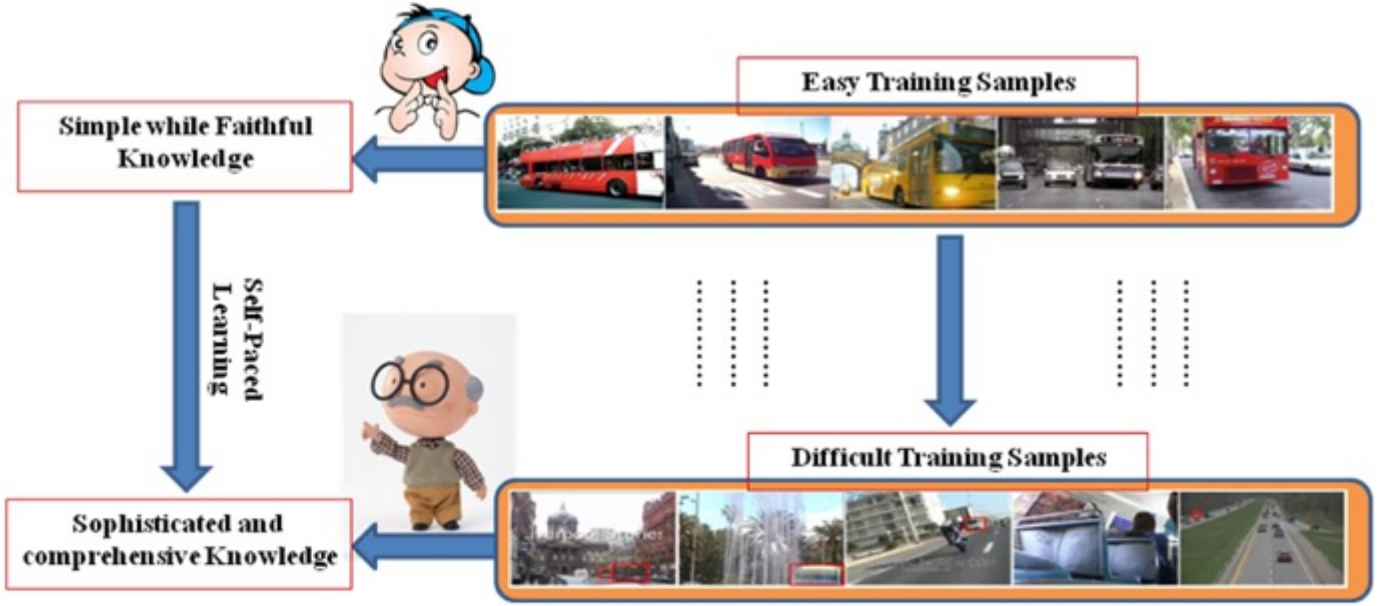
Self-Paced Learning: **Train the model from simplicity to complexity**

Increasing λ progressively

Model：

$$\min_{\theta,\ v_i} \sum_{i=1}^{n} v_i f(x_i, y_i, \theta) - \lambda \sum_{i=1}^{n} v_i$$

$$s.t.\ \ v_i \in \{0,1\}$$



Kumar et al., Self-paced learning for latent variable models, NIPS, 2010

L. Jiang, D. Meng et al., Easy samples first: Self-paced reranking for zero-example multimedia search. ACM MM, 2014

Y. Ren et al., Robust Softmax Regression for Multi-class Classification with Self-Paced Learning, IJCAI, 2017

L. Pan, S. Ai**, Y. Ren**, Z. Xu, Self-Paced Deep Regression Forests with Consideration on Underrepresented Samples, ECCV, 2020

Self-Paced Learning in Multi-View Clustering

2015
Multi-View Self-Paced Learning for Clustering
Xu et. al. in IJCAI

2017
Self-Paced Learning Based Multi-view Spectral Clustering
Yu et.al. in ICTAI

2018
Self-Paced Multi-Task Multi-View Capped-norm Clustering
Ren et. al. in ICONIP

2019
Deep Self-Paced Learning for Semi-Supervised Person Re-Identification Using Multi-View Self-Paced Clustering
Xin et.al. in ICIP

2020
Self-paced and auto-weighted multi-view clustering
Ren et.al. in Neurocomputing

Dual self-paced multi-view clustering, Z. Huang, **Y. Ren**, et al. Neural Networks, 2021

- Introduce the self-paced idea to both instance selection and feature selection

The fundamental form of multi-view clustering:

$$\min_{C^v,B} \sum_{v=1}^{m} ||X^v - C^v B||_F^2$$

Applying self-paced learning $\longrightarrow$

Conventional self-paced multi-view clustering:

$$\min_{C^v,B,W^v} \sum_{v=1}^{m} ||(X^v - C^v B)W^v||_F^2 + f(W)$$

- Self-paced learning: gradually increase the number of samples to train the model.

Similarity $\longleftrightarrow$

- Feature selection: gradually decrease the number of features to train the model.

The proposed model:

$$\min_{C^v,B,w^v_{fea},w^v_{sam}} \sum_{v=1}^{m} \eta(v) \| diag(\sqrt{w^v_{fea}})(X^v - C^v B) diag(\sqrt{w^i_{sam}}) \|_F^2$$

$$+ \sum_{v=1}^{m} \sum_{i=1}^{d^v} \eta(v) f(w^v_{i\_fea}, \lambda^v_{fea})$$

$$+ \sum_{v=1}^{m} \sum_{i=1}^{n} \eta(v) f(w^v_{i\_sam}, \lambda^v_{sam})$$

$$s.t. \quad w^v_{i\_fea}, w^v_{i\_sam} \in [0,1],$$

$$b_{ij} \in \{0,1\}, \sum_{i=1}^{k} b_{ij} = 1, \forall j = 1,2,\ldots,n$$

For simplicity, in following parts, we denote:

$diag(\sqrt{w^v_{sam}}) \longrightarrow W^v_{sam}$ $\qquad$ $diag(\sqrt{w^v_{fea}}) \longrightarrow W^v_{fea}$

Algorithm 1 The DSMVC algorithm.

Input: Data set X^v, $v = 1,2,\ldots,m$; Cluster number k.

Output: The final cluster center matrix C^v, assignment matrix B, $v = 1,2,\ldots,m$.

Initialize C^v and B by optimizing the fundamental MVC model.
Obtain $\eta(v)$.
repeat
  for each view **do**
    Update λ^v_{sam} to let more samples join the training.
    Fix C^v, W^v_{fea} and B, update W^v_{sam}.
  end for
  repeat
    for each view **do**
      Fix W^v_{sam}, W^v_{fea} and B, update C^v.
    end for
    update B.
  until convergence or exceed the maximal number of iterations.
  for each view **do**
    Update λ^v_{fea} to let less features join the training.
    Fix C^v, W^v_{sam} and B, update W^v_{fea}.
  end for
  repeat
    for each view **do**
      Fix W^v_{sam}, W^v_{fea} and B, update C^v.
    end for
    update B.
  until convergence or exceed the maximal number of iterations.
until all data points are selected
return C^v and B, $v = 1,2,\ldots,m$.

Experimental results：

Table 2: Results on Handwritten numerals.

Methods	ACC(%)	purity(%)	NMI(%)
KM(1)	57.71±4.98	63.71±3.69	60.27±2.41
KM(2)	62.99±6.72	65.38±5.03	64.38±2.76
KM(3)	70.53±7.27	73.38±5.84	70.93±3.59
KM(4)	38.09±1.55	43.80±0.97	47.76±0.24
KM(5)	70.05±6.89	72.56±6.36	70.38±3.99
KM(6)	52.10±2.93	55.95±2.33	50.01±1.82
KM(Allfea)	50.72±4.17	56.01±2.44	57.37±1.64
Co-train	73.28±5.87	74.92±3.89	71.04±2.15
Co-reg	78.09±6.89	80.63±5.36	75.50±2.91
MVKKM	62.18±3.34	65.56±2.40	65.80±1.19
RMVK	60.41±5.50	63.04±4.55	64.82±2.10
AMGL	81.22±6.53	84.24±4.99	**86.89±2.66**
CAMVC	74.98±7.96	78.84±6.90	78.07±4.25
MSPL	80.26±3.93	83.60±3.22	82.80±2.25
SPMVC	**80.15±8.90**	**84.15±6.57**	**87.17±3.90**
DSMVC-N	**82.07±9.35**	**84.98±6.67**	84.64±3.89
DSMVC	**83.27±5.69**	**86.42±2.53**	**87.53±4.69**

Table 3: Results on Cornell.

Methods	ACC(%)	purity(%)	NMI(%)
KM(1)	42.70±2.14	44.96±1.00	8.60±3.36
KM(2)	45.56±5.87	48.56±3.58	12.34±5.42
KM(Allfea)	47.47±6.42	49.69±5.01	13.54±6.96
Co-train	40.62±1.27	46.41±0.88	14.48±1.40
Co-reg	42.39±1.09	44.10±0.36	5.65±2.45
MVKKM	41.64±3.72	44.72±1.03	7.27±1.83
RMVK	43.09±1.72	45.13±0.87	10.28±3.61
AMGL	42.68±0.40	43.81±0.26	3.74±0.37
CAMVC	44.10±2.74	49.16±2.12	9.81±4.95
MSPL	44.09±3.18	46.19±2.65	8.78±4.62
SPMVC	42.36±2.91	48.53±2.52	15.23±2.42
DSMVC-N	**48.48±6.03**	**53.62±6.39**	**20.42±7.34**
DSMVC	**50.51±7.43**	**56.92±6.37**	**23.20±6.28**

Table 4: Results on Texas.

Methods	ACC(%)	purity(%)	NMI(%)
KM(1)	55.51±1.63	57.18±1.25	7.90±4.50
KM(2)	55.56±6.07	60.04±4.61	16.18±10.90
KM(AllFea)	**56.63±5.84**	60.41±5.04	14.63±10.38
Co-train	48.13±2.75	58.00±0.89	14.22±1.78
Co-reg	53.37±2.67	56.04±0.22	4.57±1.89
MVKKM	52.84±5.63	56.84±0.89	7.70±3.71
RMVK	**57.47±2.27**	59.16±2.77	16.05±5.95
AMGL	56.13±0.43	56.84±0.25	5.43±0.51
CAMVC	**59.23±4.49**	60.71±4.42	14.92±9.54
MSPL	56.68±3.85	58.97±2.94	11.30±6.30
SPMVC	**56.90±4.31**	62.34±2.96	19.91±3.64
DSMVC-N	**57.77±7.41**	**65.99±5.64**	**25.72±6.63**
DSMVC	**58.13±5.72**	**68.59±3.65**	**28.64±5.55**

Table 5: Results on Washington.

Methods	ACC(%)	purity(%)	NMI(%)
KM(1)	49.80±6.84	51.46±6.86	8.44±7.84
KM(2)	**57.54±9.77**	61.39±9.64	25.36±14.50
KM(Allfea)	**58.75±9.40**	66.29±8.68	26.23±12.89
Co-train	53.99±2.25	62.93±1.22	19.30±1.78
Co-reg	55.97±2.95	58.64±4.19	16.68±3.63
MVKKM	48.39±1.85	49.49±1.83	8.65±4.53
RMVK	**58.41±8.11**	60.10±8.19	18.92±9.40
AMGL	47.26±0.20	48.26±0.00	3.58±0.32
CAMVC	**58.97±10.57**	60.81±10.62	22.53±14.99
MSPL	52.67±7.99	54.16±7.95	13.75±11.15
SPMVC	**59.10±5.41**	66.14±1.30	25.08±3.55
DSMVC-N	56.78±4.61	**69.20±1.21**	**32.98±2.47**
DSMVC	**60.09±5.66**	**69.96±1.99**	**35.01±2.93**

Table 6: Results on Wisconsin.

Methods	ACC(%)	purity(%)	NMI(%)
KM(1)	46.44±2.18	48.67±1.82	5.69±2.30
KM(2)	**59.81±7.92**	62.60±8.70	28.97±12.41
KM(Allfea)	58.57±6.93	60.33±7.68	25.95±11.35
Co-train	42.58±1.89	52.57±1.10	8.28±0.83
Co-reg	47.35±0.24	47.76±0.21	4.06±0.37
MVKKM	45.62±2.88	48.03±1.34	6.29±2.22
RMVK	47.95±2.98	50.44±2.27	7.62±2.65
AMGL	47.09±0.16	47.55±0.00	4.03±0.31
CAMVC	56.49±7.30	59.58±7.88	21.57±9.36
MSPL	55.50±6.85	57.66±6.79	21.87±9.27
SPMVC	47.90±4.31	59.04±2.53	20.37±3.36
DSMVC-N	**63.64±4.99**	**72.69±1.75**	**40.44±2.77**
DSMVC	**64.45±4.61**	**73.82±2.18**	**41.03±2.79**

Our recent work

Non-Linear Fusion for Self-Paced Multi-View Clustering (NSMVC, ACM MM, 2021)

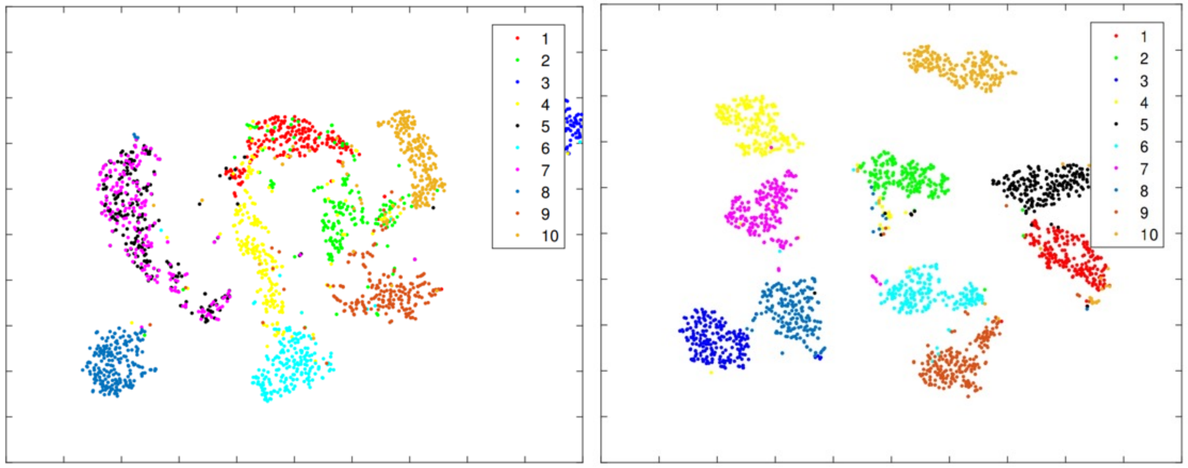
Motivation

1. The view quality issue in Multi-View Clustering (MVC) task.

2. The conventional linear-weighting approach.

3. The effectiveness of non-linear fusion in instance-level.

$$\|A\|_{2,1} = \sum_{i=1}^{n} \sqrt{\sum_{j=1}^{p} A_{ji}^2} = \sum_{i=1}^{n} \|a_i\|$$



Corrupt view

Informative view

Proposed Method

- In Non-Linear Fusion for Self-Paced Multi-view Clustering (NSMVC). We directly assigning different exponents to each view. By this way, our method alleviate the negative impact from the corrupt views while maintain the availability of more reliable views.

Model:

$$\sum_{v=1}^{m} \phi(v)^{\eta(v)}$$

$$s.t. \quad \phi(v) \geq 0, \quad \eta(v) \in (0, 1]$$

In NSMVC, a more reliable view will be assigned with a larger η(v), through this approach, the model is more sensitive to the variation to its corresponding Ø(v).

On the contrast, the corrupt views will be granted with much smaller η(v) and even close to 0, which means its contribution to the loss function will close to the constant 1, thus it has little influence on the clustering result.

Proposed Method

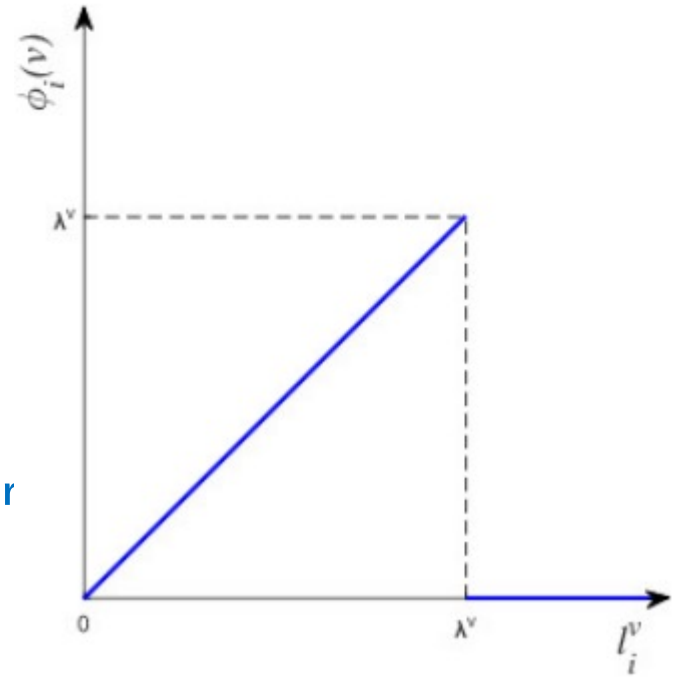
In NSMVC, we also apply SPL to further enhance the clustering performance. However, as the values of η(v) in different views are not consistent, we cannot solve the proposed non-linear model by conventional SPL scenario that requires a regularizer.

To this end, we design a novel regularizer-free modality of SPL to define Ø(v):

$$\phi(v) = \sum_{i=1}^{n} \phi_i(v) = \sum_{i=1}^{n} \lceil max(1 - l_i^v/\lambda^v, 0) \rceil \times l_i^v.$$

l_i^v is computed by the following formula and its relationship with $\emptyset_i(v)$ can be described by the figure on the right side:

$$l_i^v = ||x_i^v - C^v b_i||_2^2.$$



Therefore, Ø(v) actually has the same selection result as the following formula in the conventional SPL manner:

$$\min_{w^v} \sum_{i=1}^{n} w_i^v l_i^v - w_i^v \lambda^v$$
$$s.t. \quad w_i^v \in \{0, 1\}.$$

$w_i^v = 1$ only when $l_i^v \leq \lambda^v$.

Compared with the conventional SPL scenario, the novel regularizer-free SPL modality in our method is totally constituted with the loss from the samples and thus confirms the non-negativity of Ø(v), which is essential to the non-linear model.

Proposed Method

∅(v) can be written in the form of the squared F-norm:

$$\phi(v) = ||(X^v - C^v B) diag(w^v)||_F^2,$$

As the value of λ^v controls the participation of the samples in the v^{th} view, it actually reflects the quality of the view. Generally, the better views have smaller values of λ^v and vice versa. With the constraint that $\eta(v) \in (0, 1]$, we simply set $\eta(v)$ as:

$$\eta(v) = \frac{\min_u \lambda^u}{\lambda^v}.$$

Therefore, the model of NSMVC and its corresponding solving algorithm are:

$$\min_{C^v, B, w^v} \sum_{v=1}^{m} ||(X^v - C^v B) diag(w^v)||_F^{2\min_u \lambda^u / \lambda^v}$$

Algorithm 1 The NSMVC Algorithm.

Input: Data set X^v, $v = 1, 2, \ldots, m$; Cluster number k.

Output: The final cluster center matrix C^v, assignment matrix B, $v = 1, 2, \ldots, m$.

Initialize C^v and B randomly.
repeat
  for each view v **do**
    Update λ^v to let more samples join the training.
    for each sample i **do**
      Update $w_i^v = 1$ if $l_i^v <= \lambda^v$, otherwise $w_i^v = 0$.
    end for
  end for
  Update $\eta(v)$ for each view.
  repeat
    for each view v **do**
      Fix $\eta(v)$, w^v and B, update C^v.
    end for
    Fix $\eta(v)$, C^v and w^v, update B.
  until convergence or exceed the maximal number of iterations
until all data points are selected
return C^v and B, $v = 1, 2, \ldots, m$.

Results

Table 1: Summary of the data sets used in the experiments.

View	Handwritten Numerals	MSRCv1	Cornell	Texas	Washington	Wisconsin
1	Profile correlations (216)	Color Moments (24)	Citation (195)	Citation (187)	Citation (230)	Citation (265)
2	Fourier coefficients (76)	HOG (576)	Content (1703)	Content (1398)	Content (2000)	Content (1703)
3	Karhunen coefficients (64)	GIST (512)	-	-	-	-
4	Morphological (6)	LBP (256)	-	-	-	-
5	Pixel averages (240)	Centrist (254)	-	-	-	-
6	Zernike moments (47)	-	-	-	-	-
# Samples	2000	210	195	187	230	265
# Classes	10	7	5	5	5	5

* Numbers in parentheses are the number of features in each view.

Table 2: Results on Handwritten Numerals.

Methods	ACC(%)	Purity(%)	NMI(%)
KM(1)	57.71±4.98	63.71±3.69	60.27±2.41
KM(2)	62.99±6.72	65.38±5.03	64.38±2.76
KM(3)	70.53±7.27	73.38±5.84	70.93±3.59
KM(4)	38.09±1.55	43.80±0.97	47.76±0.24
KM(5)	70.05±6.89	72.56±6.36	70.38±3.99
KM(6)	52.10±2.93	55.95±2.33	50.01±1.82
KM(All)	50.72±4.17	56.01±2.44	57.37±1.64
Co-train	73.28±5.87	74.92±3.89	71.04±2.15
Co-reg	78.09±6.89	80.63±5.36	75.50±2.91
MVKKM	62.18±3.34	65.56±2.40	65.80±1.19
AMGL	81.22±6.53	84.24±4.99	86.89±2.66
CAMVC	74.98±7.96	78.84±6.90	78.07±4.25
MSPL	80.26±3.93	83.60±3.22	82.80±2.25
SAMVC	75.37±12.71	79.74±11.91	82.62±12.49
DMVC	79.91±8.56	83.77±6.72	85.18±4.01
NSMVC	**88.52±6.40**	**90.53±4.63**	**89.10±2.25**

Table 3: Results on MSRCv1.

Methods	ACC(%)	Purity(%)	NMI(%)
KM(1)	35.76±2.38	37.88±2.45	24.25±2.50
KM(2)	62.69±6.60	64.60±5.59	54.16±4.45
KM(3)	62.00±5.52	64.90±4.17	57.03±3.74
KM(4)	47.29±1.55	49.55±0.97	41.38±0.24
KM(5)	54.64±6.79	55.55±5.25	45.18±3.19
KM(All)	46.29±3.10	46.29±3.03	42.07±2.18
Co-train	66.55±5.77	69.33±4.46	58.18±3.46
Co-reg	41.52±4.31	44.21±3.89	35.36±3.61
MVKKM	70.19±3.73	70.95±3.31	61.61±3.07
AMGL	69.74±7.04	71.81±5.01	**68.35±3.24**
CAMVC	67.88±5.18	71.14±3.49	62.86±2.73
MSPL	50.19±6.01	52.29±4.94	43.30±2.66
SAMVC	65.31±8.82	68.19±8.01	61.57±6.28
DMVC	62.57±10.49	63.67±10.48	58.75±8.67
NSMVC	**74.65±5.62**	**77.14±4.44**	**66.65±5.00**

Results

Table 4: Results on Cornell.

Methods	ACC(%)	Purity(%)	NMI(%)
KM(1)	42.70±2.14	44.96±1.00	8.60±3.36
KM(2)	45.56±5.87	48.56±3.58	12.34±5.42
KM(All)	47.47±6.42	49.69±5.01	13.54±6.96
Co-train	40.62±1.27	46.41±0.88	14.48±1.40
Co-reg	42.39±1.09	44.10±0.36	5.65±2.45
MVKKM	41.64±3.72	44.72±1.03	7.27±1.83
AMGL	42.68±0.40	43.81±0.26	3.74±0.37
CAMVC	44.10±2.74	49.16±2.12	9.81±4.95
MSPL	44.09±3.18	46.19±2.65	8.78±4.62
SAMVC	43.43±0.82	44.69±0.50	6.82±3.11
DMVC	44.30±2.74	46.72±2.82	12.42±3.99
NSMVC	**50.62±6.45**	**60.21±2.76**	**26.56±3.94**

Table 5: Results on Texas.

Methods	ACC(%)	Purity(%)	NMI(%)
KM(1)	55.51±1.63	57.18±1.25	7.90±4.50
KM(2)	55.56±6.07	60.04±4.61	16.18±10.90
KM(All)	**56.63±5.84**	60.41±5.04	14.63±10.38
Co-train	48.13±2.75	58.00±0.89	14.22±1.78
Co-reg	53.37±2.67	56.04±0.22	4.57±1.89
MVKKM	52.84±5.63	56.84±0.89	7.70±3.71
AMGL	56.13±0.43	56.84±0.25	5.43±0.51
CAMVC	**59.23±4.49**	60.71±4.42	14.92±9.54
MSPL	**56.68±3.85**	58.97±2.94	11.30±6.30
SAMVC	56.81±1.38	57.68±1.35	8.54±4.54
DMVC	**56.90±4.29**	59.84±3.24	16.32±7.15
NSMVC	**57.81±4.93**	**67.17±1.45**	**25.23±2.60**

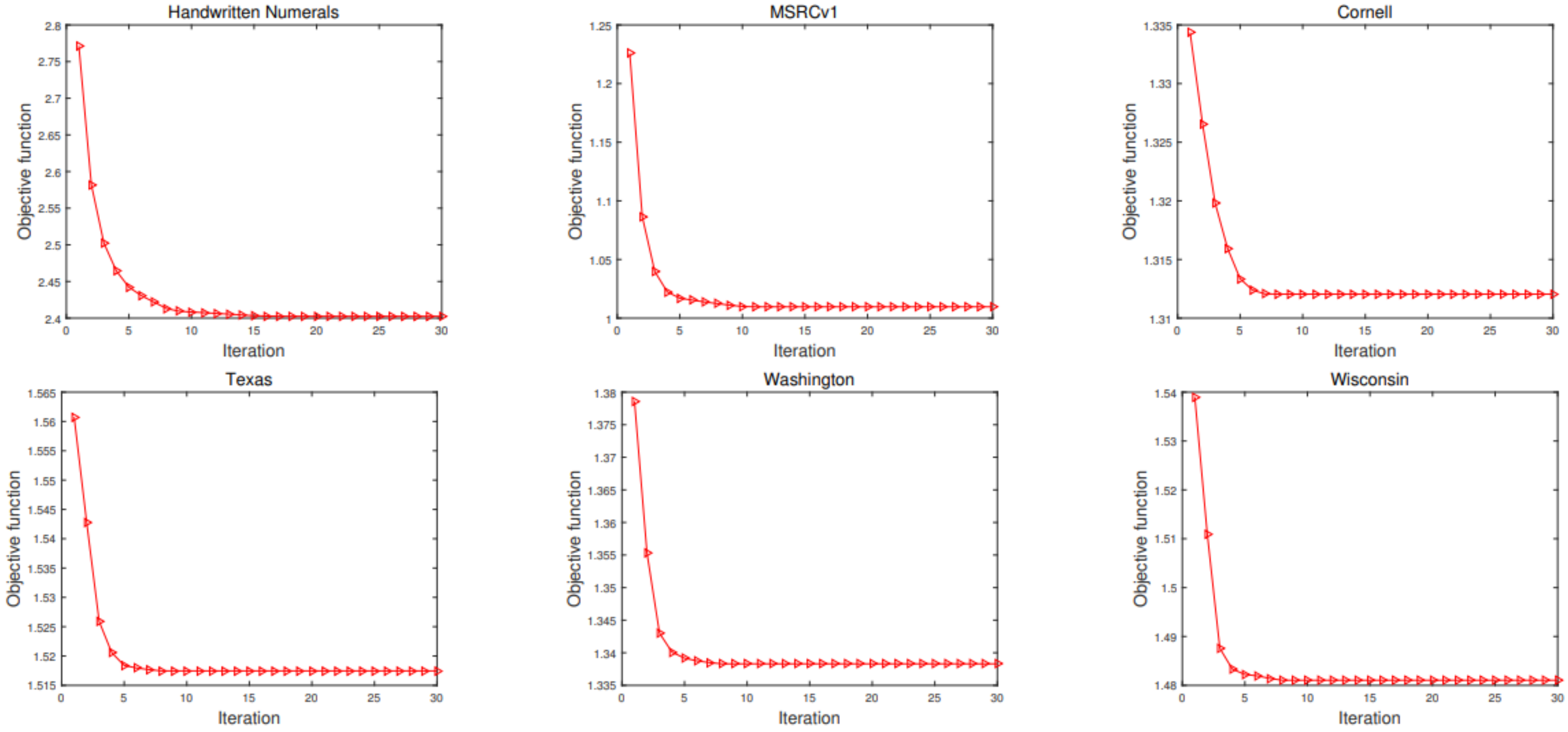
Table 6: Results on Washington.

Methods	ACC(%)	Purity(%)	NMI(%)
KM(1)	49.80±6.84	51.46±6.86	8.44±7.84
KM(2)	**57.54±9.77**	61.39±9.64	25.36±14.50
KM(All)	**58.75±9.40**	66.29±8.68	26.23±12.89
Co-train	53.99±2.25	62.93±1.22	19.30±1.78
Co-reg	55.97±2.95	58.64±4.19	16.68±3.63
MVKKM	48.39±1.85	49.49±1.83	8.65±4.53
AMGL	47.26±0.20	48.26±0.00	3.58±0.32
CAMVC	**58.97±10.57**	60.81±10.62	22.53±14.99
MSPL	52.67±7.99	54.16±7.95	13.75±11.15
SAMVC	52.93±7.95	53.70±8.12	11.39±9.43
DMVC	**58.45±6.96**	62.44±7.54	22.09±9.38
NSMVC	**57.96±5.82**	**71.13±3.20**	**36.20±3.79**

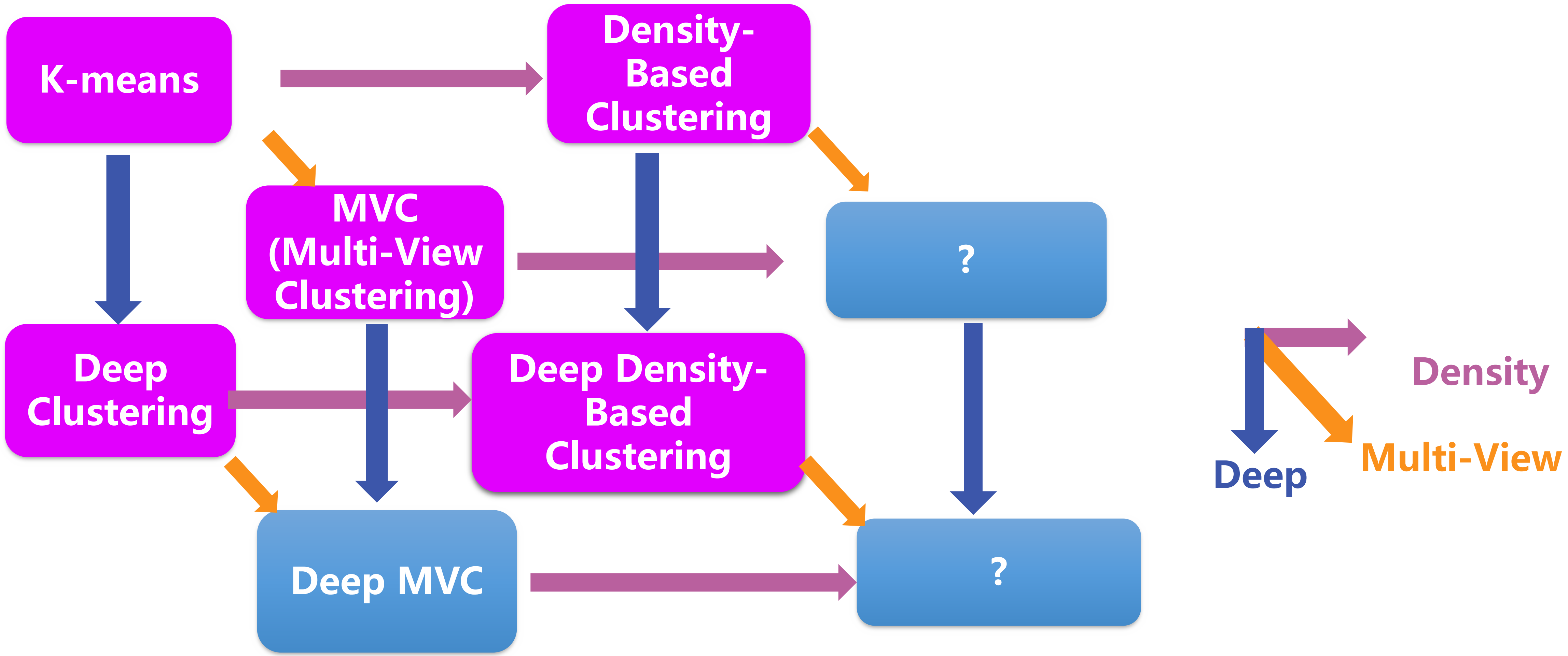
Table 7: Results on Wisconsin.

Methods	ACC(%)	Purity(%)	NMI(%)
KM(1)	46.44±2.18	48.67±1.82	5.69±2.30
KM(2)	**59.81±7.92**	62.60±8.70	28.97±12.41
KM(All)	**58.57±6.93**	60.33±7.68	25.95±11.35
Co-train	42.58±1.89	52.57±1.10	8.28±0.83
Co-reg	47.35±0.24	47.76±0.21	4.06±0.37
MVKKM	45.62±2.88	48.03±1.34	6.29±2.22
AMGL	47.09±0.16	47.55±0.00	4.03±0.31
CAMVC	56.49±7.30	59.58±7.88	21.57±9.36
MSPL	55.50±6.85	57.66±6.79	21.87±9.27
SAMVC	49.93±3.13	48.93±3.15	7.04±4.69
DMVC	53.01±6.93	58.73±7.24	19.68±10.17
NSMVC	**60.30±4.89**	**73.70±1.73**	**40.48±1.85**

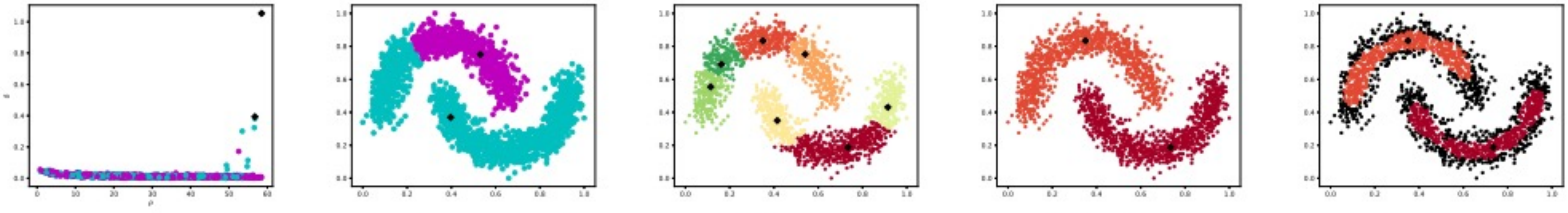
Results



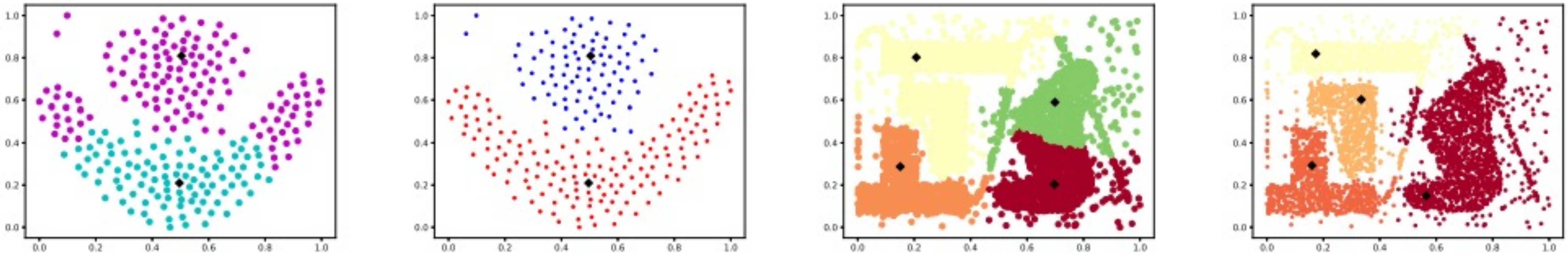
Handwritten Numerals
MSRCv1
Cornell
Texas
Washington
Wisconsin
Objective function
Iteration


K-means
Density-Based Clustering
MVC (Multi-View Clustering)
?
Deep Clustering
Deep Density-Based Clustering
Deep MVC
?
Density
Multi-View
Deep

Y. Ren et al, Deep Density-based Image Clustering, KBS, 2020



(a) Decision graph (b) Clustering result (c) Initial result (d) Final result (e) Border points



(a) Result of DenPeak (b) Result of DDC (c) Result of DenPeak (d) Result of DDC

Deep Density-based Image Clustering

Deep feature learning

$$\arg\min_{\Theta,\Omega} \frac{1}{n}\sum_{i=1}^{n} \|x_i - g_\Omega(f_\Theta(\tilde{x}_i))\|_2^2$$

$$\arg\min_{\Theta,\Omega} \frac{1}{n}\sum_{i=1}^{n} \|\bar{x}_i - g_\Omega(f_\Theta(\bar{x}_i))\|_2^2$$

Local clusters generation

Density：

$$\rho_i = \sum_{z_j \in \mathcal{Z}\setminus\{z_i\}} \exp\left(-\left(\frac{d_{ij}}{d_c}\right)^2\right)$$

$$\delta_i = \min_{j:\rho_j>\rho_i}(d_{ij})$$

Local cluster centers：$\delta_i > d_c \text{ and } \rho_j > \bar{\rho}$

Merging local clusters

Definition 2. ***(Core and border points of a cluster)*** *Suppose a point z_i is from local cluster $\mathcal{C}^{(k)}$, it is defined as a core point if the following condition holds:*

$$\rho_j > \bar{\rho}^{(k)} \tag{6}$$

where $\bar{\rho}^{(k)} = \frac{1}{n_k}\sum_{z_j\in\mathcal{C}^{(k)}} \rho_j$ is the average density of all the points in $\mathcal{C}^{(k)}$ and n_k is the number of points in $\mathcal{C}^{(k)}$. Otherwise, z_i is considered as a border point.

Definition 3. ***(Density directly-connectable of clusters)*** *A local cluster $\mathcal{C}^{(k)}$ is density directly-connectable from a local cluster $\mathcal{C}^{(l)}$ if:*

$$\exists \text{ core points } z_i \in \mathcal{C}^{(k)} \text{ and } z_j \in \mathcal{C}^{(l)}, \text{ such that } d_{ij} < d_c.$$

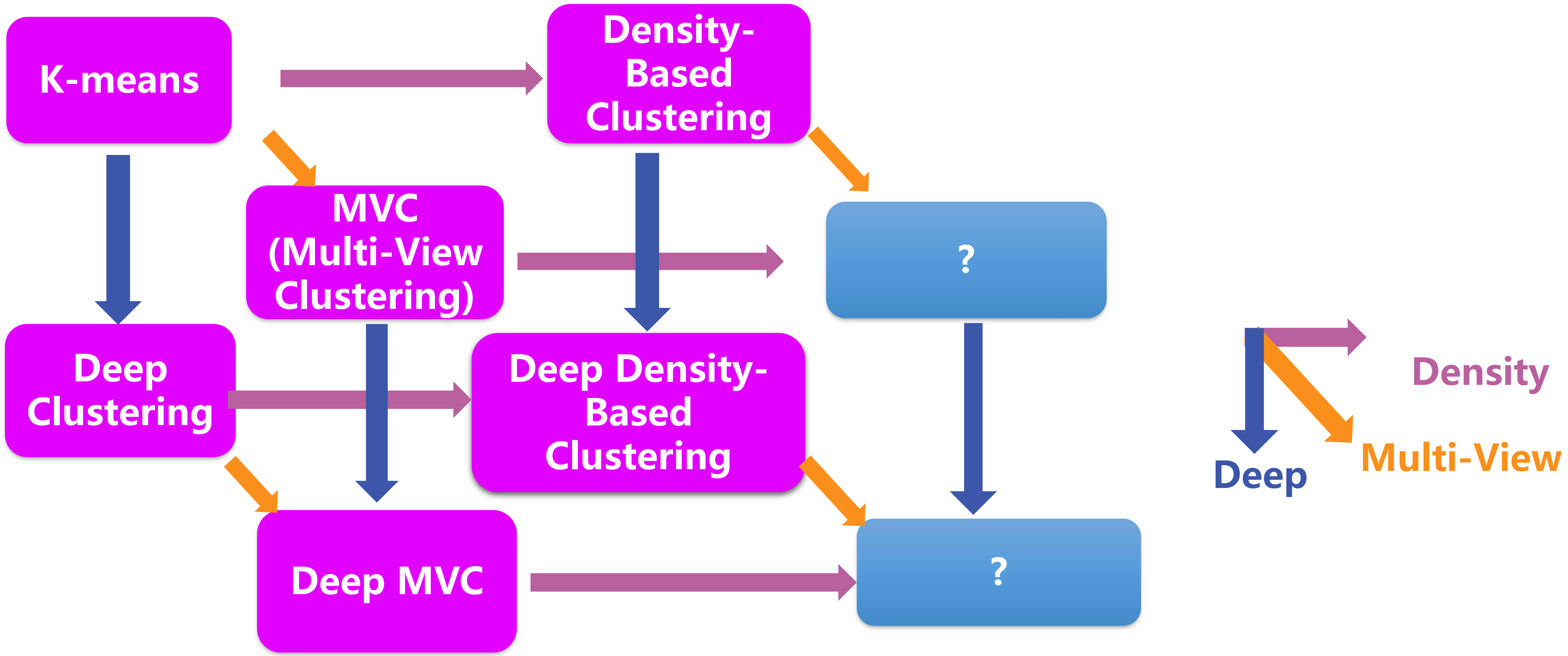
Definition 4. ***(Density connectable of clusters)*** *A local cluster $\mathcal{C}^{(k)}$ is density-connectable to a local cluster $\mathcal{C}^{(l)}$ if:*

$$\exists \text{ a path } \mathcal{C}^{(k)} = \mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_m = \mathcal{C}^{(l)} \tag{8}$$

where cluster $\mathcal{C}_j$ is density directly-connectable from cluster $\mathcal{C}_{j-1}$ $(j = 2, \ldots, m)$ and m is the path length.

Deep Density-based Image Clustering

	MNIST		MNIST-test		USPS		Fashion		LetterA-J	
	ACC	NMI	ACC	NMI	ACC	NMI	ACC	NMI	ACC	NMI
k-means	0.485	0.470	0.563	0.510	0.611	0.607	0.554	0.512	0.354	0.309
DBSCAN	- -	- -	0.114	0	0.167	0	0.1	0	0.1	0
DenPeak	- -	- -	0.357	0.399	0.390	0.433	0.344	0.398	0.300	0.211
DEC	0.849	0.816	0.856	0.830	0.758	0.769	0.591	0.618	0.407	0.374
IDEC	0.881*	0.867*	0.846	0.802	0.759	0.777	0.523	0.600	0.381	0.318
DCN	0.830*	0.810*	0.802*	0.786*	0.688*	0.683*	-	-	-	-
JULE	0.964*	0.913*	0.961*	0.915*	0.950*	0.913*	-	-	-	-
DCC	0.963*	-	-	-	-	-	-	-	-	-
DED	- -	- -	0.690	0.818	0.781	0.855	0.473	0.617	0.371	0.440
ConvDEC	0.940	0.916	0.861	0.847	0.784	0.820	0.514	0.588	0.517	0.536
ConvDEC-DA	**0.985**	**0.961**	0.955	**0.949**	**0.970**	**0.953**	0.570	0.632	0.571	**0.608**
DDC	0.965	0.932	**0.965**	0.916	0.967	0.918	**0.619**	**0.682**	**0.573**	0.546
DDC-DA	**0.969**	**0.941**	**0.970**	**0.927**	**0.977**	**0.939**	**0.609**	**0.661**	**0.691**	**0.629**


K-means
Density-Based Clustering
MVC (Multi-View Clustering)
?
Deep Clustering
Deep Density-Based Clustering
Deep MVC
?
Density
Deep
Multi-View

Deep Multi-view Clustering

Multi-view clustering + Deep learning techniques

- Representation learning
- Self-supervised learning
- Adversarial learning
- Contrastive learning
- ……

Representation learning

Learn the low-dimensional nonlinear representations using DAE, VAE, etc.

Deep Multimodal Clustering with Cross Reconstruction

Xianchao Zhang[1,2], Xiaorui Tang[1,2], Linlin Zong[1,2(✉)], Xinyue Liu[1,2], and Jie Mu[1,2]

Self-supervised learning

Learn the pseudo labels from some views to supervise other views' learning.

Self-supervised Discriminative Feature Learning for Multi-view Clustering

Jie Xu, Yazhou Ren*, Huayi Tang, Zhimeng Yang, Lili Pan, Yang Yang, *Senior Member, IEEE*, Xiaorong Pu

Adversarial learning

Apply generative adversarial networks (GAN) to learn the latent features of multiple views.

Deep Adversarial Multi-view Clustering Network

Zhaoyang Li[1], **Qianqian Wang**[1*], **Zhiqiang Tao**[2], **Quanxue Gao**[1†] and **Zhaohua Yang**[3]
[1]State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China.
[2]Department of Electrical and Computer Engineering, Northeastern University, USA.
[3]School of Instrumentation Science and Opto-electronics Engineering, Beihang University, China.

Contrastive learning

Contrastive learning can be used to learn high-level and consistent multi-view features.

COMPLETER: Incomplete Multi-view Clustering via Contrastive Prediction

Yijie Lin[1], Yuanbiao Gou[1], Zitao Liu[2], Boyun Li[1], Jiancheng Lv[1], Xi Peng[1*]
[1] College of Computer Science, Sichuan University, China.
[2] TAL Education Group, China.
{linyijie.gm,gouyuanbiao,zitao.jerry.liu,liboyun.gm,pengx.gm}@gmail.com; lvjiancheng@scu.edu.cn

Our recent work

Deep Embedded Multi-view Clustering with Collaborative Training (DEMVC, Information Sciences, 2021)

Deep Embedded Multi-view Clustering with Collaborative Training

Jie Xu[a], Yazhou Ren[a,*], Guofeng Li[a], Lili Pan[b,c], Ce Zhu[d] and Zenglin Xu[e,f]

[a] School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
[b] School of Information and Communication, University of Electeonic Science and Technology of China, Chengdu, Sichuan, 611731, China
[c] Yangtze Delta Region Institute (Quzhou), University of Electronic Science and Technology of China, Quzhou, Zhejiang 324000, China
[d] School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
[e] Harbin Institute of Technology, Shenzhen, China
[f] Peng Cheng Lab, Shenzhen, China

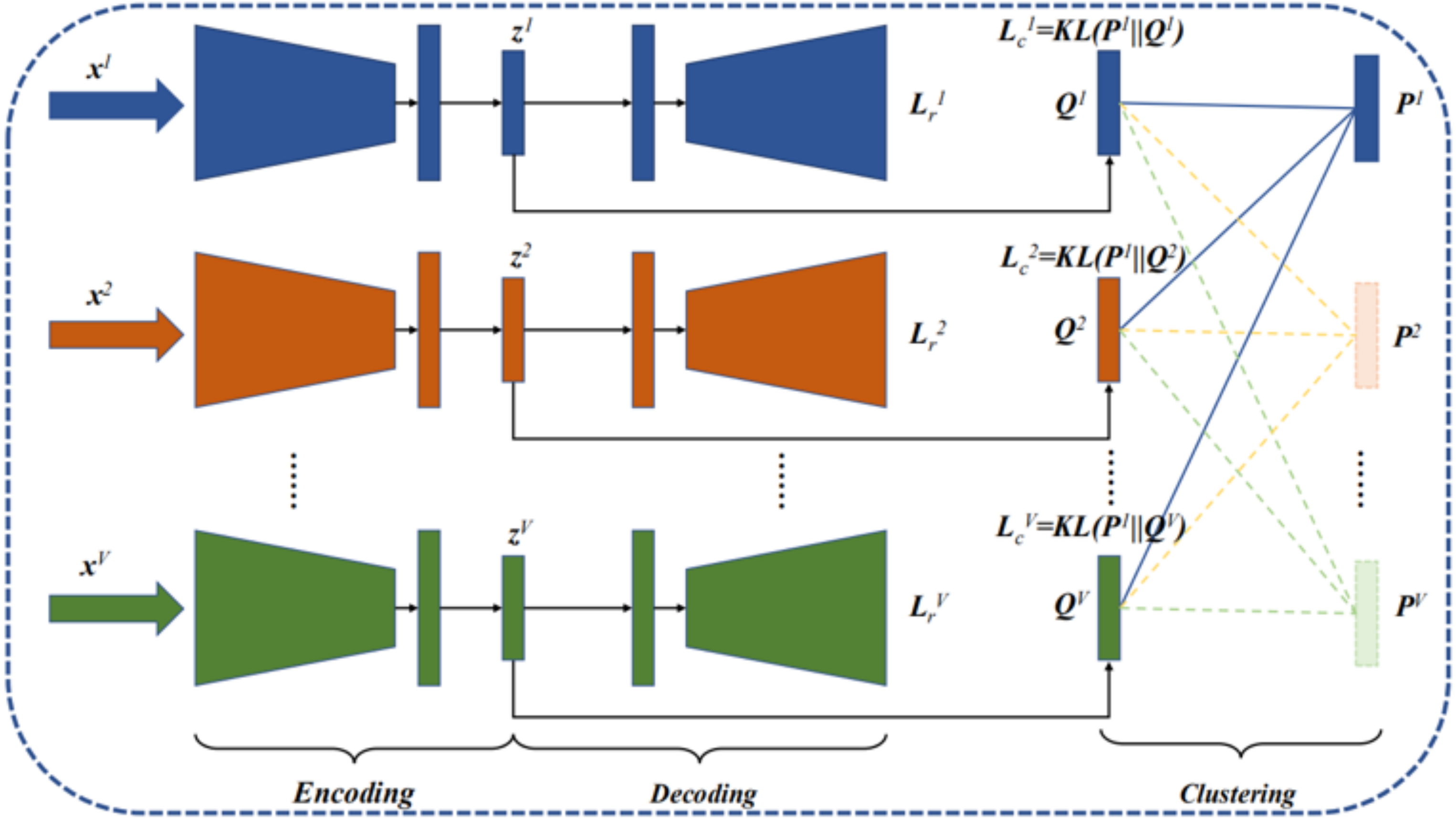
ARTICLE INFO

Keywords:
Deep embedded clustering
Multi-view clustering
Unsupervised learning
Collaborative training

ABSTRACT

Multi-view clustering has attracted increasing attentions recently by utilizing information from multiple views. However, existing multi-view clustering methods are either with high computation and space complexities, or lack of representation capability. To address these issues, we propose deep embedded multi-view clustering with collaborative training (DEMVC) in this paper. Firstly, the embedded representations of multiple views are learned individually by deep autoencoders. Then, both consensus and complementary principles of multiple views are taken into account and a novel collaborative training scheme is proposed. Concretely, the feature representations and cluster assignments of all views are learned collaboratively. A new consistency strategy for cluster centers initialization is further developed to improve the multi-view clustering performance with collaborative training. Experimental results on several popular multi-view datasets show that DEMVC achieves significant improvements over state-of-the-art methods.

The framework of DEMVC:

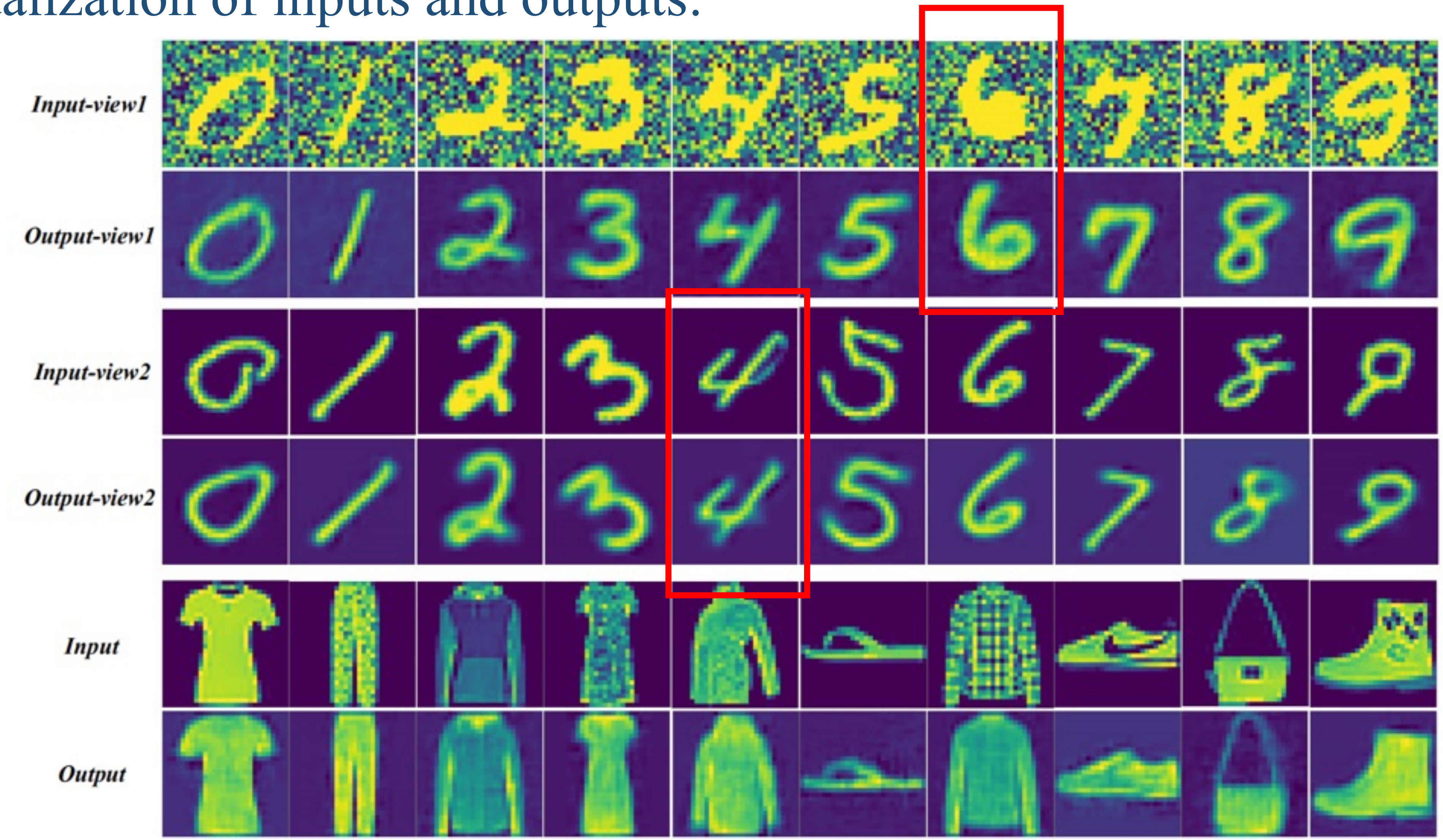


DEMVC applies k-means on one view (the referred view) to obtain an auxiliary target distribution.

This auxiliary distribution is used to refine the deep autoencoders and clustering soft assignments for all views.

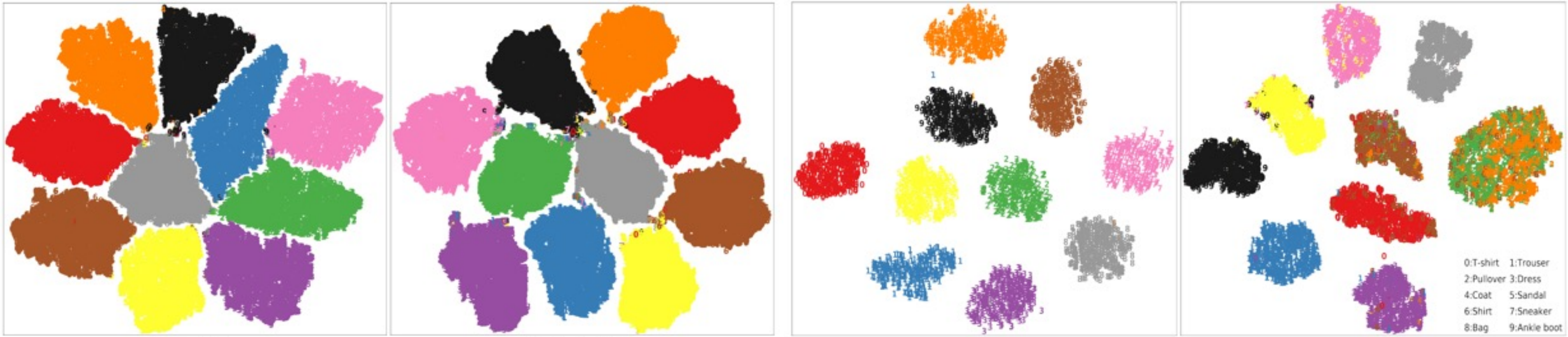
Each view will become the referred view in sequence to ensure that the multi-view clustering takes full advantage of all views.

Visualization of inputs and outputs:



This indicates DEMVC's good representation capability of sample features and reconstruction capability, which is the premise to improve clustering performance.

Visualization by *t*-sne:



(a) NoisyMnist (b) RotatingMnist (c) USPS (d) Fashion-10K

Linear complexity make it easy
to handle large-scale dataset
(e.g. 70,000 examples)

Quantitative comparison:

Multi-view methods

Methods	Noisy-Rotated		MNIST-USPS		Caltech101-20	
	ACC	NMI	ACC	NMI	ACC	NMI
DCCA (ICML 2013)	97.00†	92.00†	97.42*	93.60*	-	46.48*
DCCAE (ICML 2015)	**97.50**†	**93.40**†	**98.00***	**94.70***	-	45.56*
DiMSC (CVPR 2015)	/	/	48.34*	36.02*	-	29.05*
LMSC (CVPR 2017)	/	/	78.60*	78.49*	-	**63.55***
BMVC (TPAMI 2018)	85.61	81.48	88.68	89.93	47.44	60.28
COMIC (ICML 2019)	/	/	47.76	64.16	**62.32**	60.56
DEMVC (ours)	**99.87**	**99.53**	**99.83**	**99.49**	**56.05**	**68.87**

Single-view methods

Methods	MNIST-10K		Fashion-10K	
	ACC	NMI	ACC	NMI
DEC (ICML 2016)	83.41	79.22	56.70	61.29
IDEC (IJCAI 2017)	84.25	82.77	57.43	61.55
DCN (ICML 2017)	83.31‡	80.86‡	58.67‡	59.40‡
DEC-DA (ACML 2018)	97.93	95.81	53.55	59.91
k-SCN (ACCV 2018)	87.14‡	78.15‡	63.78‡	62.04‡
NCSC (ICML 2019)	94.09‡	86.12‡	72.14‡	68.60‡
DEMVC-2 views (ours)	**99.87**	**99.60**	**84.75**	**87.14**
DEMVC-3 views (ours)	**99.99**	**99.96**	**78.99**	**90.88**

Our recent work

Self-supervised Discriminative Feature Learning for Multi-view Clustering (SDMVC, arXiv, 2021)

Self-supervised Discriminative Feature Learning for Multi-view Clustering

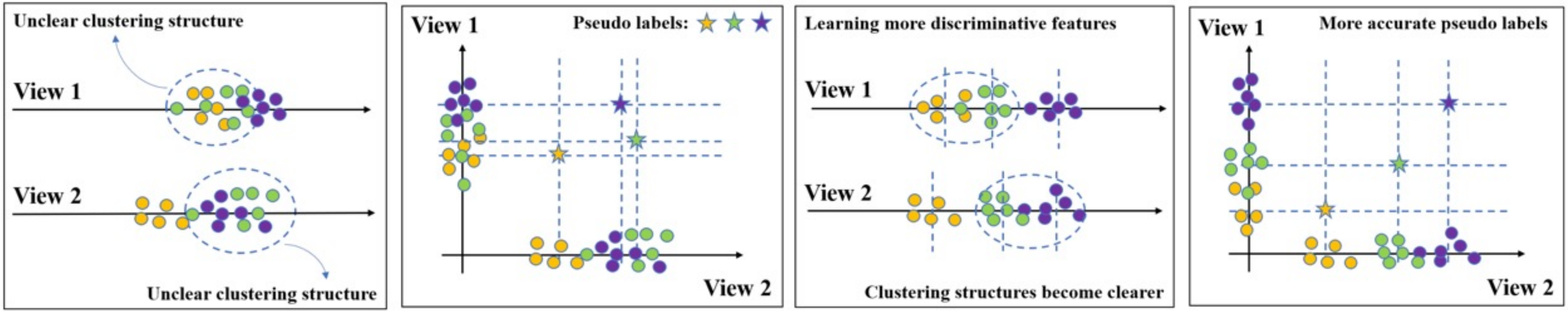
Jie Xu, Yazhou Ren*, Huayi Tang, Zhimeng Yang, Lili Pan, Yang Yang, *Senior Member, IEEE*, Xiaorong Pu

Abstract—Multi-view clustering is an important research topic due to its capability to utilize complementary information from multiple views. However, there are few methods to consider the negative impact caused by certain views with unclear clustering structures, resulting in poor multi-view clustering performance. To address this drawback, we propose self-supervised discriminative feature learning for multi-view clustering (SDMVC). Concretely, deep autoencoders are applied to learn embedded features for each view independently. To leverage the multi-view complementary information, we concatenate all views' embedded features to form the global features, which can overcome the negative impact of some views' unclear clustering structures. In a self-supervised manner, pseudo-labels are obtained to build a unified target distribution to perform multi-view discriminative feature learning. During this process, global discriminative information can be mined to supervise all views to learn more discriminative features, which in turn are used to update the target distribution. Besides, this unified target distribution can make SDMVC learn consistent cluster assignments, which accomplishes the clustering consistency of multiple views while preserving their features' diversity. Experiments on various types of multi-view datasets show that SDMVC achieves state-of-the-art performance. The code is available at https://github.com/SubmissionsIn/SDMVC.

Index Terms—Deep clustering; Multi-view clustering; Unsupervised learning; Self-supervised learning

How to improve multi-view clustering ?

How to extract complementary information from multiple views?

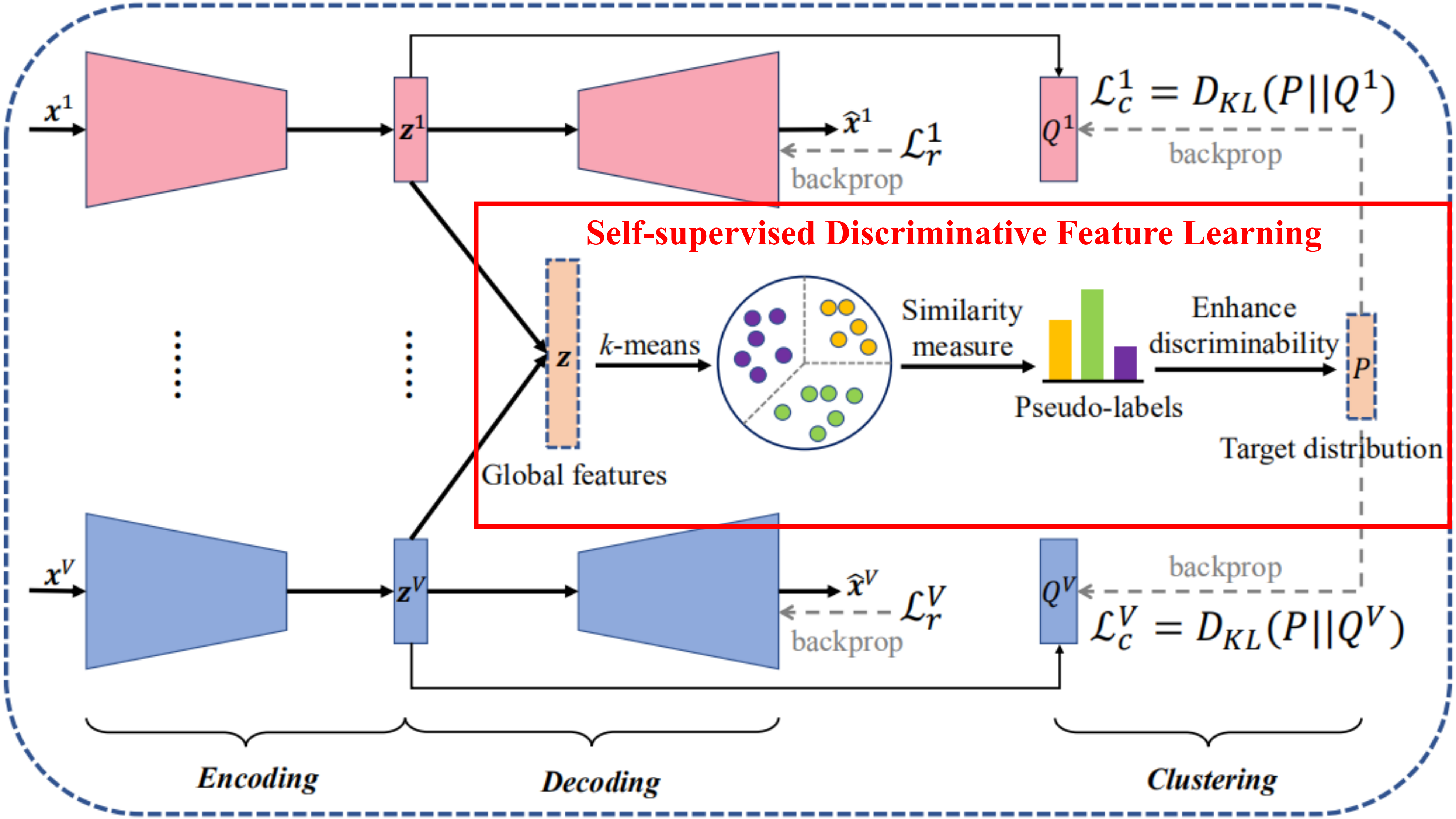


(a) (b) (c) (d)

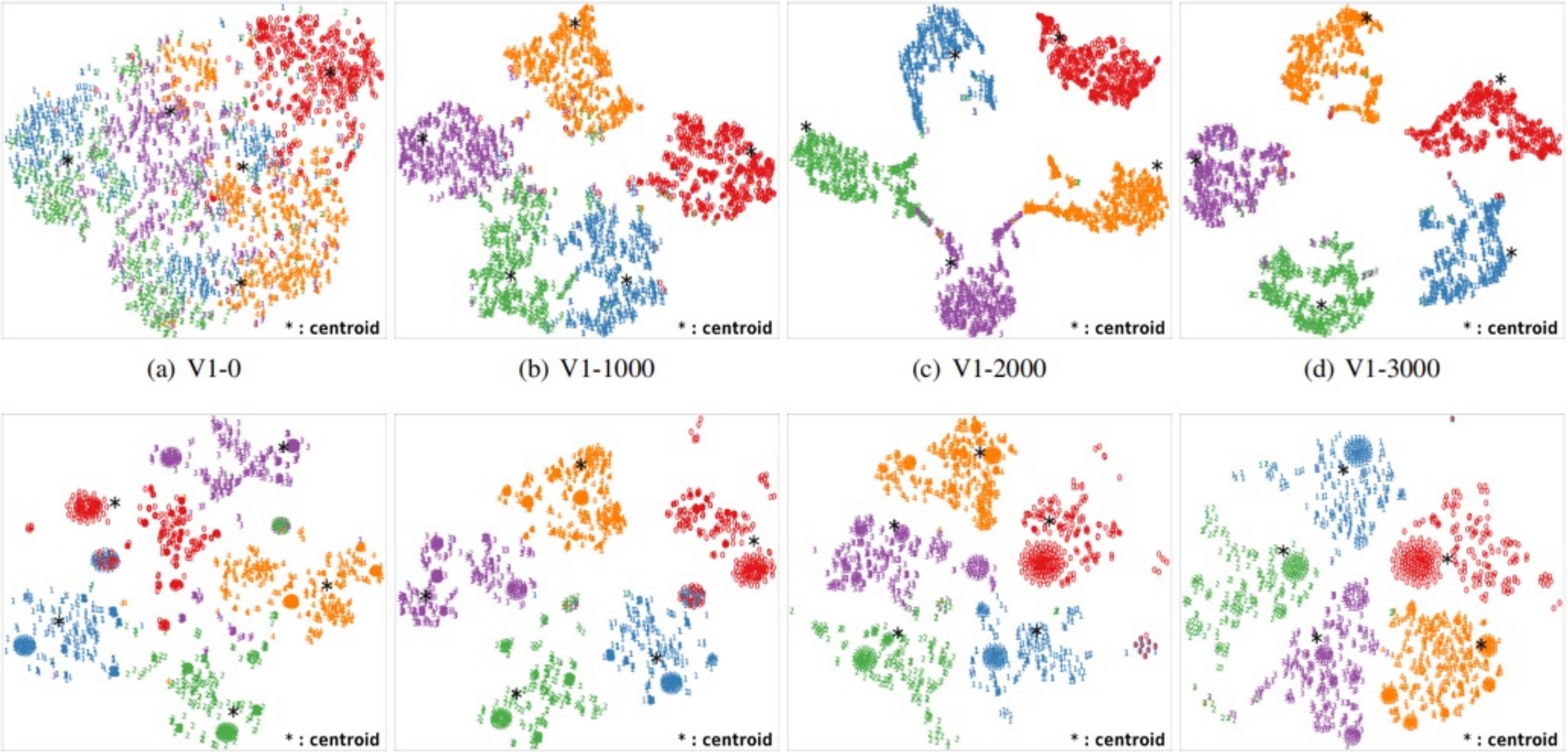
The discriminability of different views' clustering structures is different.

The clustering structures of different views can correct each other.

The framework of SDMVC:



Visualization the features in learning process (BDGP data set) :



(a) V1-0 (b) V1-1000 (c) V1-2000 (d) V1-3000

(e) V2-0 (f) V2-1000 (g) V2-2000 (h) V2-3000

The clustering structures of embedded features become clearer and clearer while their centroids are gradually separated.

Quantitative comparison:

	MNIST-USPS 2 views, $K = 10$ 5,000 examples			Fashion-MV 3 views, $K = 10$ 10,000 examples			BDGP 2 views, $K = 5$ 2,500 examples			Caltech101-20 6 views, $K = 20$ 2,386 examples		
Methods	ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI
k-means (1967)	0.7678	0.7233	0.6353	0.7093	0.6561	0.5689	0.4324	0.5694	0.2604	0.4179	0.3351	0.2605
SC (2002)	0.6596	0.5811	0.4864	0.5354	0.5772	0.4261	0.5172	0.5891	0.3156	0.4620	0.4589	0.3933
DEC (2016)	0.7310	0.7146	0.6323	0.6707	0.7234	0.6291	0.9478	0.8662	0.8702	0.4268	0.6251	0.3767
IDEC (2017)	0.7658	0.7689	0.6801	0.6919	0.7501	0.6522	0.9596	0.8940	0.9025	0.4318	0.6253	0.3773
BMVC (2018)	0.8802	0.8945	0.8448	0.7858	0.7488	0.6835	0.3492	0.1202	0.0833	0.5553	0.6203	0.5038
MVC-LFA (2019)	0.7678	0.6749	0.6092	0.7910	0.7586	0.6887	0.5468	0.3345	0.2881	0.4221	0.5846	0.2994
COMIC (2019)	0.4818	0.7085	0.4303	0.5776	0.6423	0.4361	0.5776	0.6423	0.4361	0.6232	0.6838	0.6931
SAMVC (2020)	0.6965	0.7458	0.6090	0.6286	0.6878	0.5665	0.5386	0.4625	0.2099	0.5218	0.5961	0.4653
PVC (2020)	0.6500	0.6118	0.4964	–	–	–	0.4724	0.2972	0.2520	–	–	–
DEMVC (2021)	0.9976	0.9939	0.9948	0.7864	0.9061	0.7793	0.9548	0.8720	0.8901	0.5748	0.6781	0.5068
SDMVC (ours)	**0.9982**	**0.9947**	**0.9960**	**0.8626**	**0.9215**	**0.8405**	**0.9816**	**0.9447**	**0.9548**	**0.7158**	**0.7176**	**0.7265**

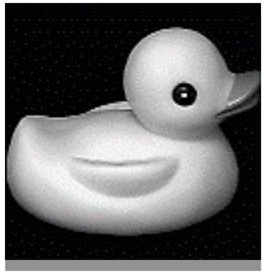
The clustering performance of SDMVC is better than other methods.

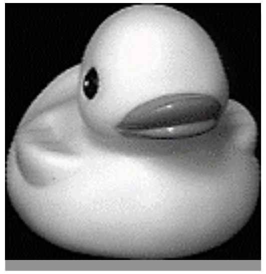
Our recent work

Multi-VAE: Learning Disentangled View-common and View-peculiar Visual Representations for Multi-view Clustering (ICCV, 2021)

Motivation

View-1





View-2

Peculiar information: **size, angle, background, etc.**

$+$

Common information: **category, semantic, etc.**

$+$

Peculiar information: **size, angle, background, etc.**

Proposed Method

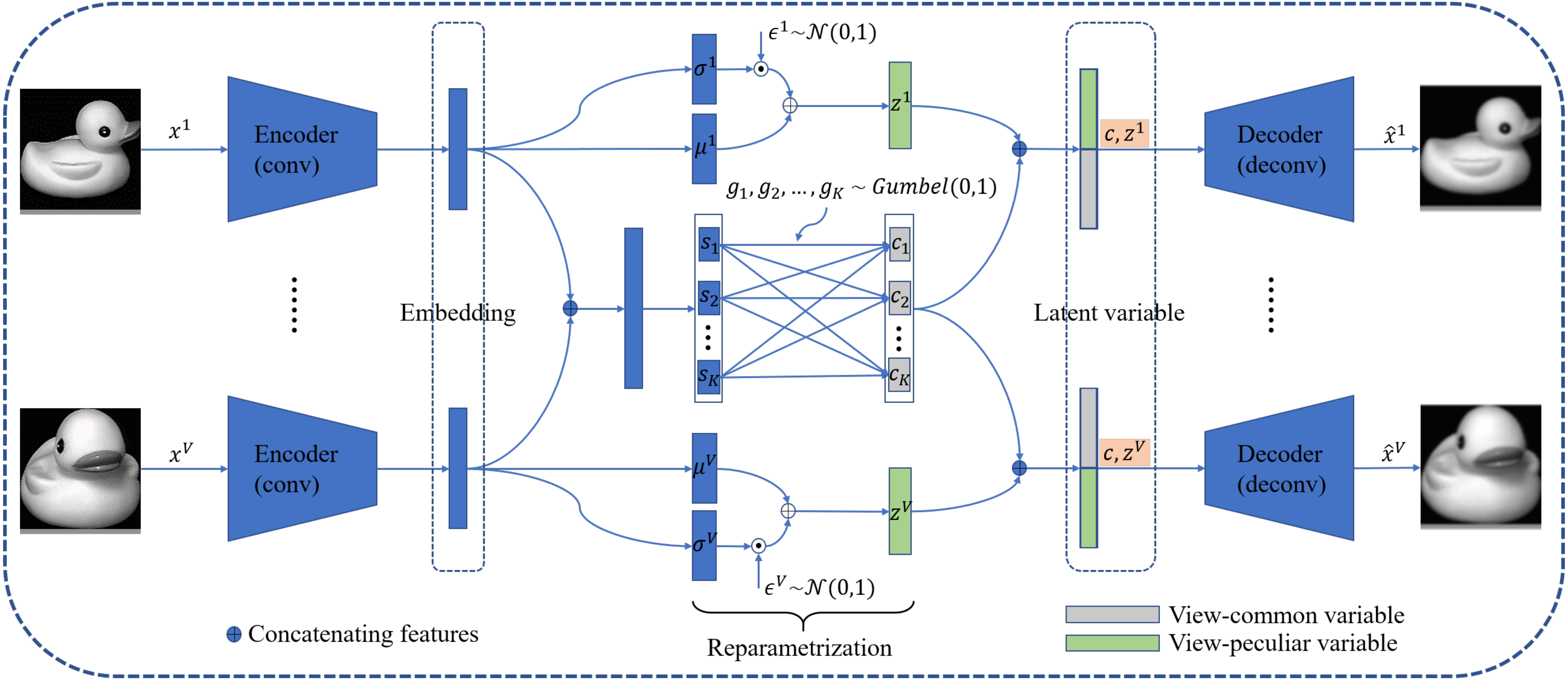
- Fusing the representations of multiple views may lead to their negative interference.
- Learning a common feature space may result in the decays of diversity among multiple views.
- So, our motivation is to **disentangle** the multi-view features into different feature space and learn explainable multi-view representations.

Disentangled Representation Learning
解离/解耦/解纠缠表示学习

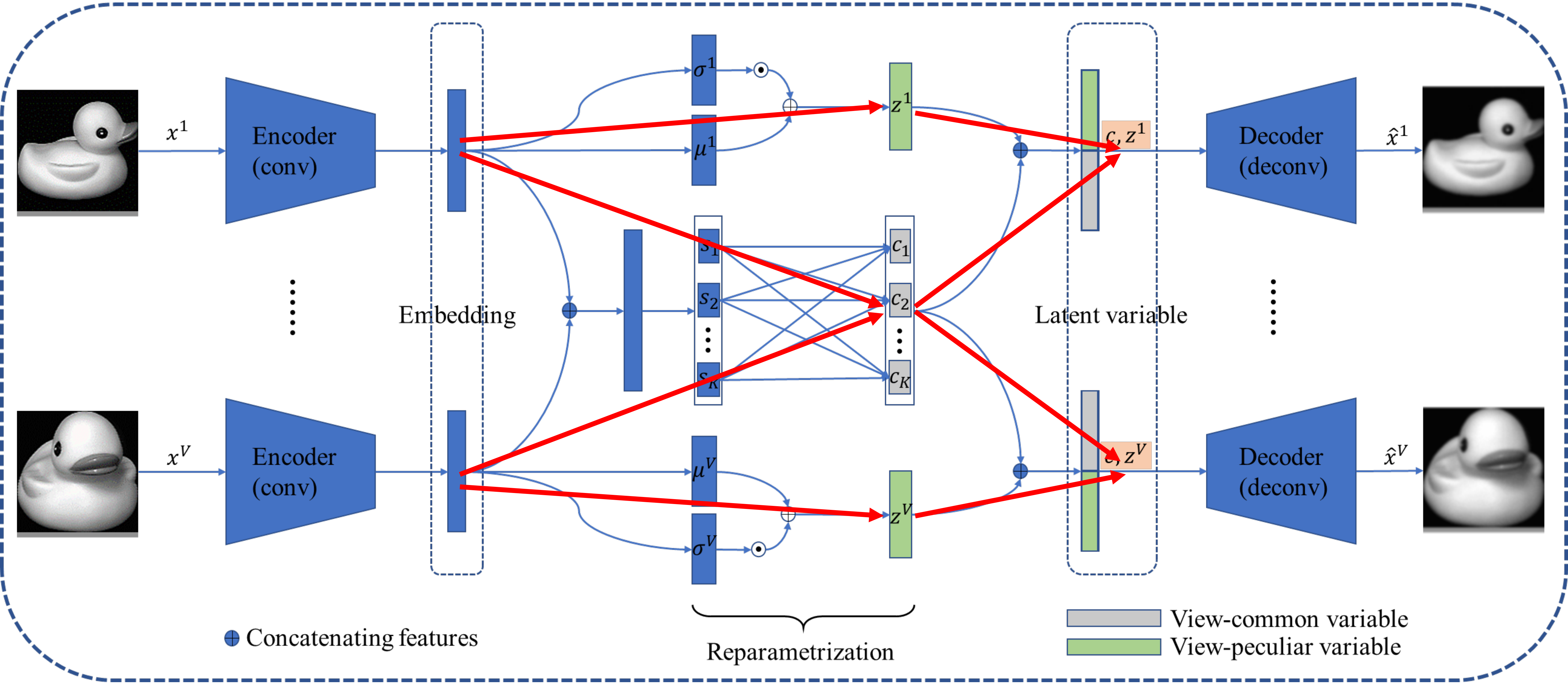
view-common variable c view-peculiar variables z^v

$$p(\boldsymbol{x}^v, \boldsymbol{z}^v, \boldsymbol{c}) = p(\boldsymbol{x}^v | \boldsymbol{z}^v, \boldsymbol{c}) p(\boldsymbol{z}^v, \boldsymbol{c})$$
$$= p(\boldsymbol{x}^v | \boldsymbol{z}^v, \boldsymbol{c}) p(\boldsymbol{z}^v) p(\boldsymbol{c})$$

Proposed Method



Proposed Method



Results

	Datasets	Multi-COIL-10				Multi-COIL-20				Object-Digit-Product			
	Size	720 samples, 3 views				1,440 samples, 3 views				720 samples, 3 views			
	Metrics	ACC	NMI	ARI	Purity	ACC	NMI	ARI	Purity	ACC	NMI	ARI	Purity
Single-view	K-means (1967)	0.733	0.769	0.648	0.757	0.415	0.645	0.384	0.415	0.326	0.297	0.143	0.337
	DEC (2016)	0.740	0.774	0.656	0.765	0.651	0.784	0.587	0.677	0.317	0.344	0.168	0.334
	IDEC (2017)	0.736	0.772	0.651	0.763	0.657	0.784	0.591	0.679	0.327	0.343	0.167	0.337
	β-VAE (2018)	0.598	0.685	0.514	0.632	0.531	0.667	0.450	0.573	0.297	0.278	0.111	0.321
	JointVAE (2018)	0.649	0.724	0.553	0.681	0.537	0.678	0.456	0.548	0.320	0.254	0.126	0.331
Multi-view	BMVC (2018)	0.678	0.681	0.530	0.678	0.834	0.900	0.813	0.881	0.810	0.661	0.634	0.810
	RMSL (2019)	0.964	0.925	0.921	0.964	0.665	0.763	0.587	0.691	0.950	0.917	0.906	0.953
	MVC-LFA (2019)	0.860	0.868	0.799	0.871	0.801	0.852	0.738	0.802	0.926	0.880	0.849	0.926
	COMIC (2019)	0.796	0.916	0.729	0.799	0.496	0.770	0.309	0.500	0.201	0.419	0.146	0.203
	SAMVC (2020)	0.667	0.826	0.621	0.729	0.570	0.791	0.554	0.610	0.770	0.826	0.702	0.801
	DEMVC (2021)	0.891	0.948	0.897	0.900	0.850	0.965	0.860	0.850	0.801	0.901	0.784	0.801
	Multi-VAE-C (ours)	0.900	0.967	0.897	0.900	0.845	0.943	0.842	0.876	0.897	0.942	0.873	0.897
	Multi-VAE-CZ (ours)	0.993	0.989	0.985	0.993	0.980	0.976	0.961	0.980	0.977	0.971	0.954	0.977

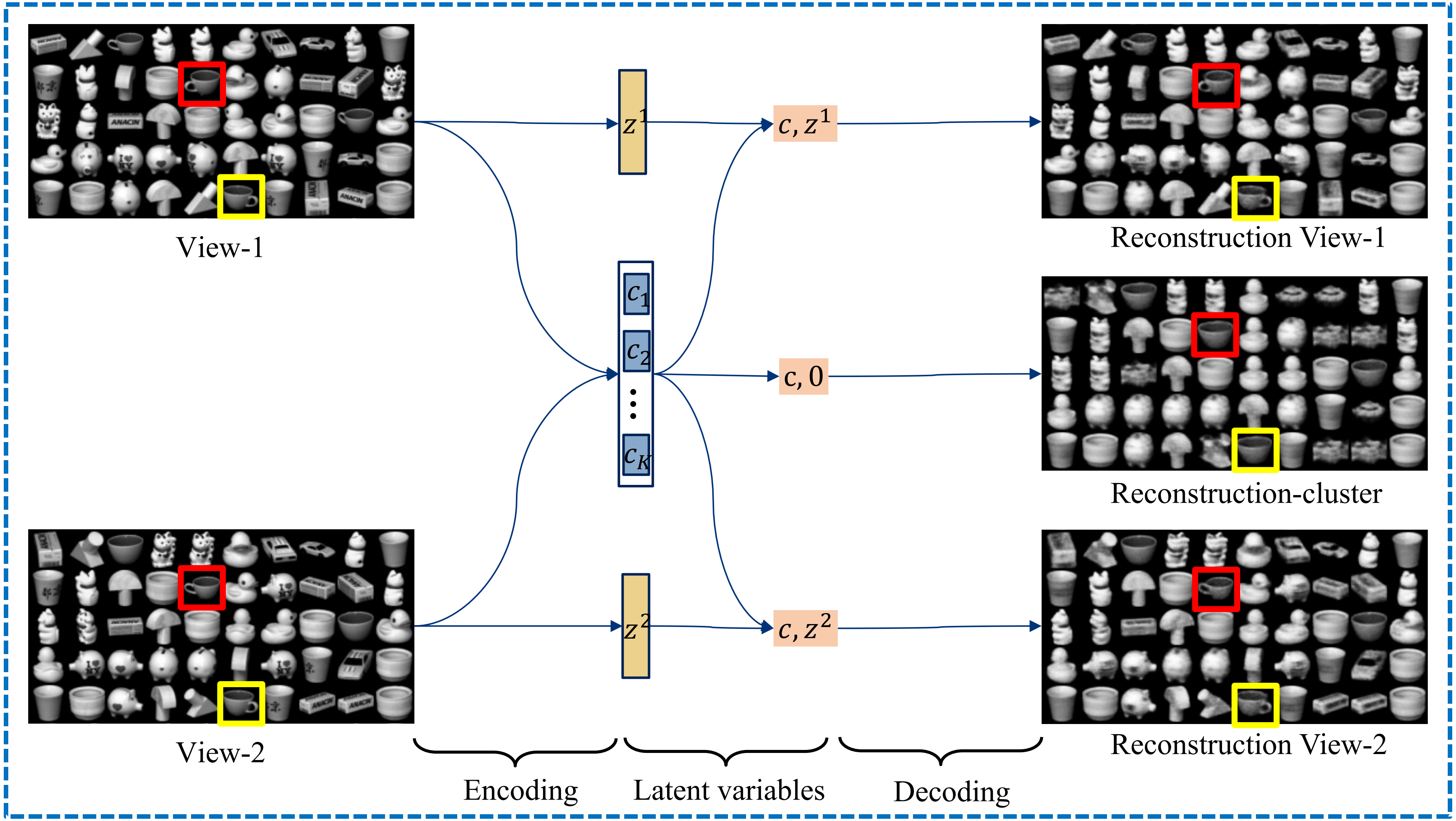
Table 1. Comparison results on small-scale datasets. The best and the second best values are highlighted in red and blue, respectively.

Results

	Datasets	Multi-MNIST				Mult-Fashion				Digit-Product			
	Size	70,000 samples, 2 views				10,000 samples, 3 views				30,000 samples, 2 views			
	Metrics	ACC	NMI	ARI	Purity	ACC	NMI	ARI	Purity	ACC	NMI	ARI	Purity
Single-view	K-means (1967)	0.539	0.482	0.360	0.577	0.476	0.513	0.348	0.551	0.349	0.346	0.187	0.390
	DEC (2016)	0.875	0.849	0.803	0.875	0.563	0.617	0.451	0.609	0.396	0.408	0.226	0.422
	IDEC (2017)	0.884	0.868	0.826	0.884	0.569	0.625	0.461	0.615	0.402	0.442	0.233	0.433
	β-VAE (2018)	0.493	0.436	0.291	0.519	0.513	0.510	0.337	0.513	0.343	0.317	0.174	0.385
	JointVAE (2018)	0.641	0.614	0.490	0.651	0.393	0.368	0.246	0.415	0.471	0.435	0.289	0.479
Multi-view	BMVC (2018)	0.893	0.902	0.856	0.897	0.779	0.756	0.682	0.782	0.548	0.442	0.379	0.570
	RMSL (2019)	–	–	–	–	0.376	0.342	0.204	0.391	–	–	–	–
	MVC-LFA (2019)	–	–	–	–	0.782	0.748	0.685	0.784	–	–	–	–
	COMIC (2019)	–	–	–	–	0.578	0.642	0.436	0.608	–	–	–	–
	SAMVC (2020)	–	–	–	–	0.622	0.688	0.557	0.661	0.649	0.619	0.499	0.674
	DEMVC (2021)	0.982	0.989	0.986	0.982	0.786	0.903	0.772	0.791	0.798	0.896	0.833	0.798
	Multi-VAE-C (ours)	0.989	0.996	0.989	0.989	0.816	0.856	0.762	0.818	0.853	0.832	0.810	0.853
	Multi-VAE-CZ (ours)	0.999	0.998	0.999	0.999	0.907	0.883	0.839	0.907	0.925	0.934	0.907	0.923

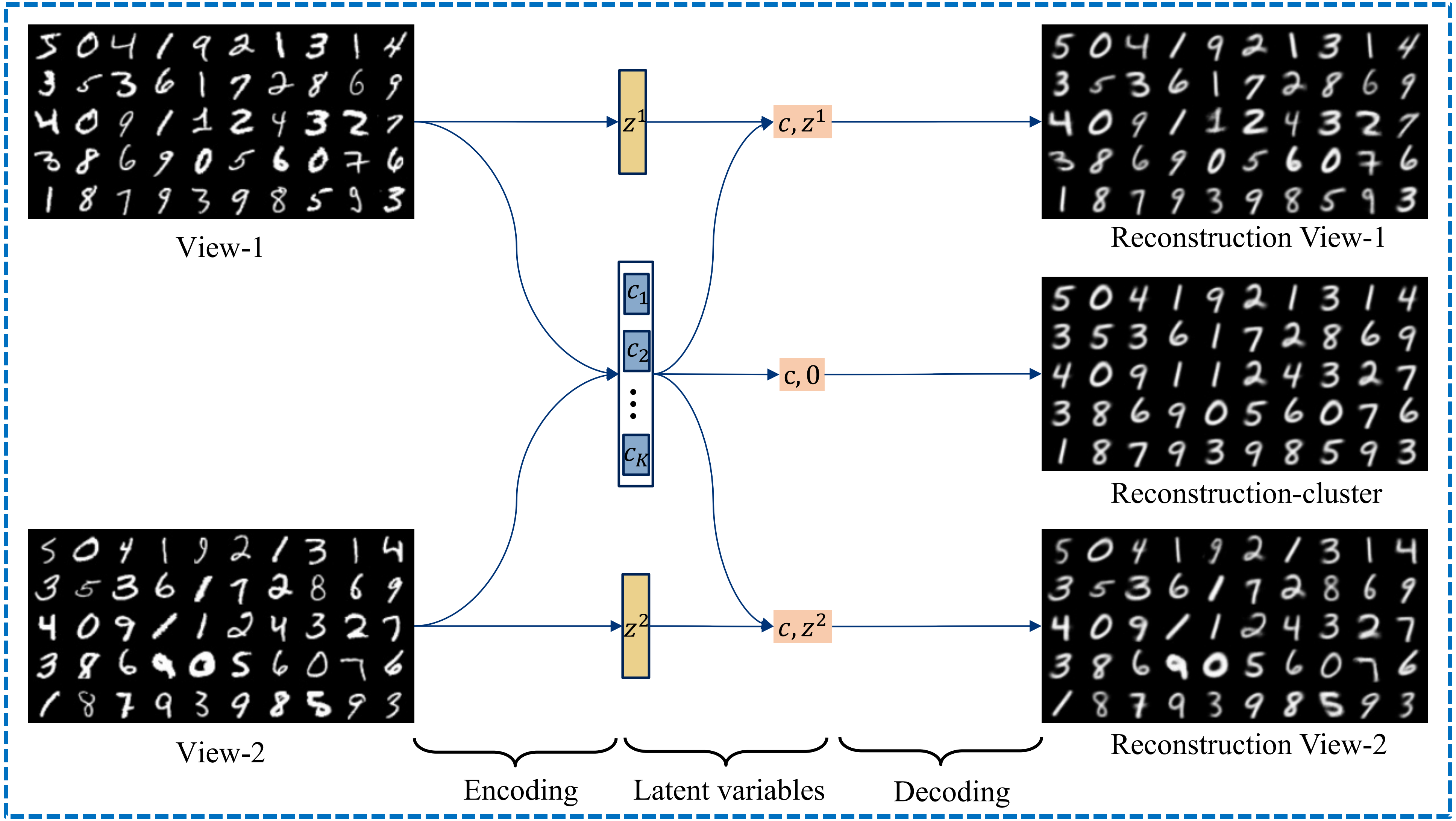
Table 2. Comparison results on large-scale datasets. "–" denotes the unknown result due to high complexity of the corresponding method.

Results



View-1
View-2
z^1
c_1
c_2
c_K
z^2
c, z^1
c, 0
c, z^2
Reconstruction View-1
Reconstruction-cluster
Reconstruction View-2
Encoding
Latent variables
Decoding

Results


View-1
View-2
z^1
z^2
c_1
c_2
c_K
c, z^1
c, 0
c, z^2
Reconstruction View-1
Reconstruction-cluster
Reconstruction View-2
Encoding
Latent variables
Decoding

Our recent work

Deep Incomplete Multi-view Clustering via Mining Cluster Complementarity (AAAI, 2022)

Introduction

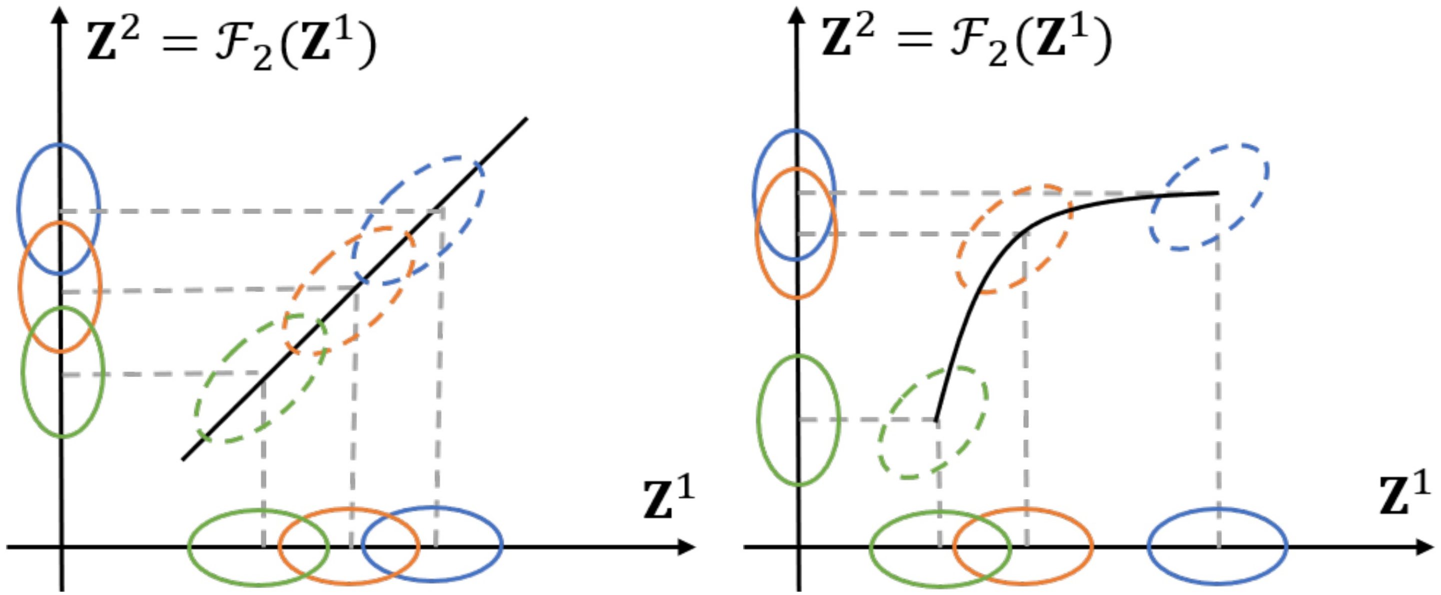
Incomplete multi-view clustering groups the multi-view data containing missing data in some views. Previous methods suffer from the two issues:

(1) the inaccurate imputation or padding for missing data negatively affects the performance.

(2) the quality of features after fusion might be interfered by the low-quality views, especially for the inaccurate imputed views.

Motivation

The complementary information across multiple views can be described by nonlinear mappings. To avoid the two issues, we establish a novel framework without imputation and fusion to handle incomplete multi-view data. That is, the missing data does not need to be imputed or padded and the cluster assignments do not depend on the fusion process of multiple views.



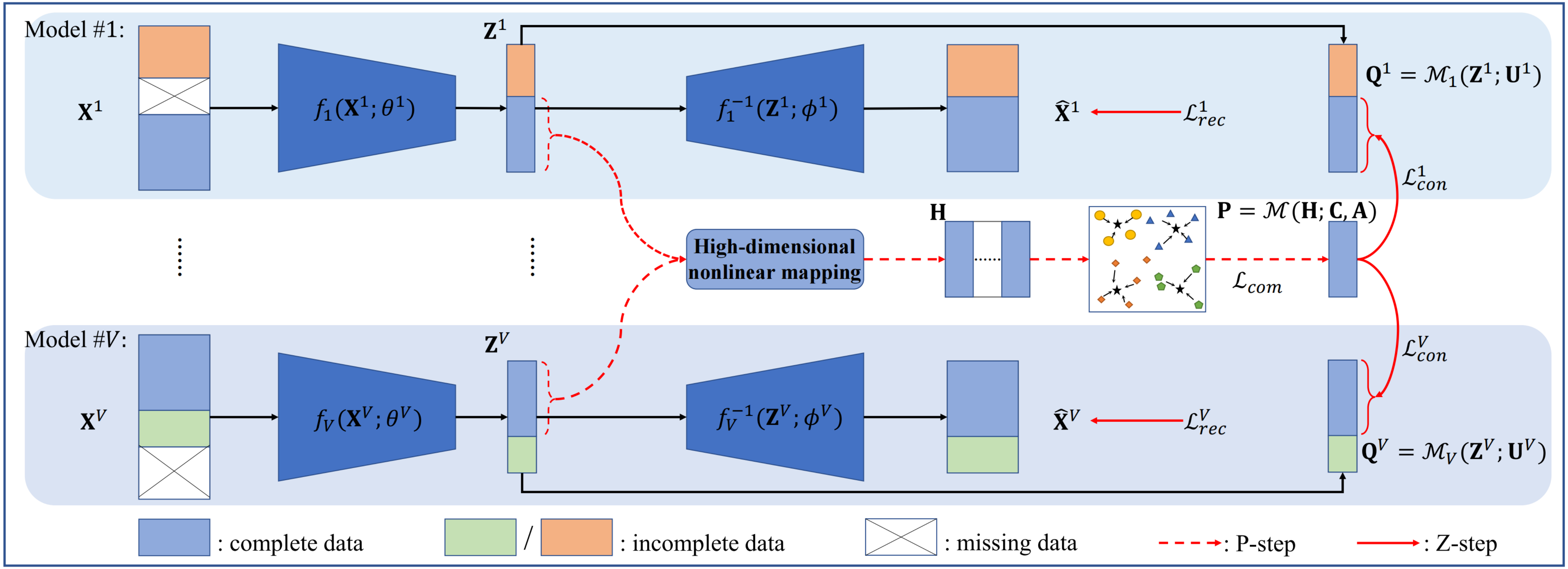
(a) $\mathcal{F}_2$ is linear mapping (b) $\mathcal{F}_2$ is nonlinear mapping

Loss function

$\mathcal{L}_{rec}$ is the reconstruction objective
$\mathcal{L}_{com}$ is the complementarity objective
$\mathcal{L}_{con}$ is the consistency objective

$$
\begin{aligned}
\mathcal{L} &= \mathcal{L}_{rec} + \mathcal{L}_{com} + \mathcal{L}_{con} \\
&= \min_{\mathbf{C},\mathbf{A},\{\mathbf{Z}^v,\mathbf{U}^v\}_{v=1}^V} \sum_{v=1}^{V} \left\| \mathbf{X}^v - f_v^{-1}(\mathbf{Z}^v) \right\|_F^2 \\
&+ \sum_{i\in\mathcal{X}} \sum_{j=1}^{K} \|\mathbf{h}_i - \mathbf{c}_j\|_2^2 + \sum_{v=1}^{V} H(\mathbf{P},\mathbf{Q}^v), \\
s.t.\ & \mathbf{P} = \mathcal{M}(\mathbf{H};\mathbf{C},\mathbf{A}), \mathbf{A}\mathbf{A}^T = \mathbf{I}_K, \mathbf{Q}^v = \mathcal{M}_v(\mathbf{Z}^v;\mathbf{U}^v)
\end{aligned}
$$

Framework



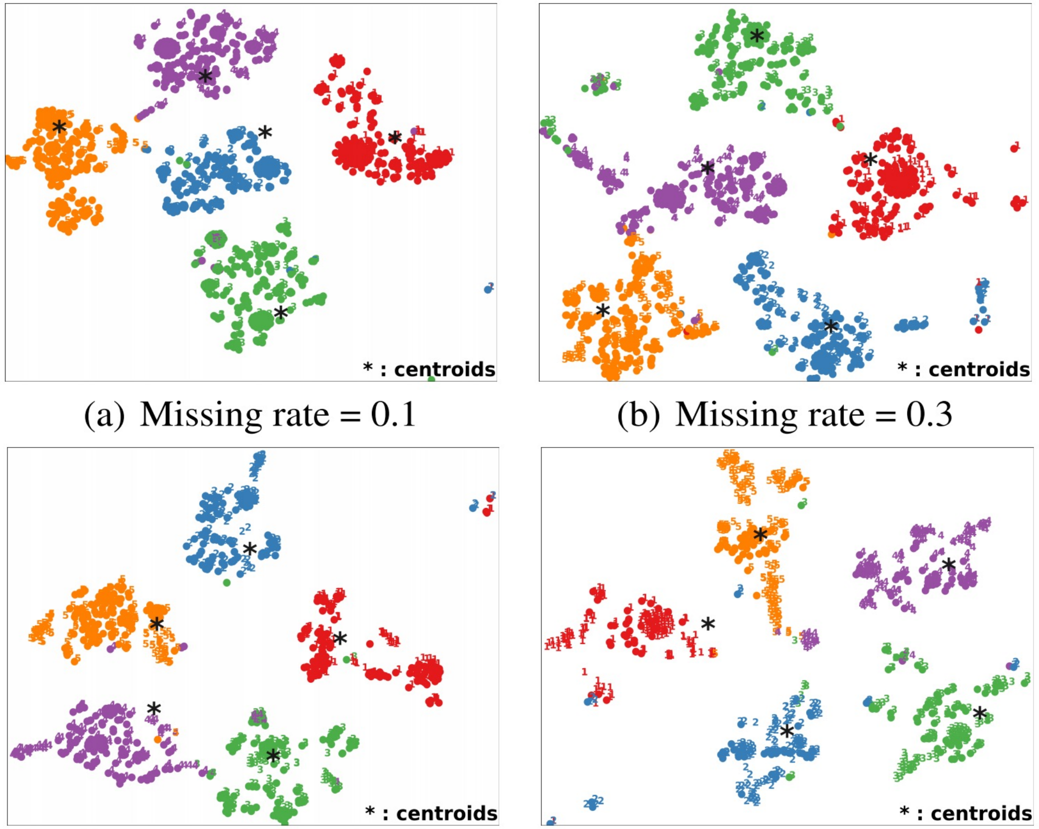
Model #1:
$\mathbf{X}^1$
$f_1(\mathbf{X}^1;\theta^1)$
$\mathbf{Z}^1$
$f_1^{-1}(\mathbf{Z}^1;\phi^1)$
$\hat{\mathbf{X}}^1$
$\mathcal{L}_{rec}^1$
$\mathbf{Q}^1=\mathcal{M}_1(\mathbf{Z}^1;\mathbf{U}^1)$
$\mathcal{L}_{con}^1$
High-dimensional nonlinear mapping
$\mathbf{H}$
$\mathbf{P}=\mathcal{M}(\mathbf{H};\mathbf{C},\mathbf{A})$
$\mathcal{L}_{com}$
Model #V:
$\mathbf{X}^V$
$f_V(\mathbf{X}^V;\theta^V)$
$\mathbf{Z}^V$
$f_V^{-1}(\mathbf{Z}^V;\phi^V)$
$\hat{\mathbf{X}}^V$
$\mathcal{L}_{rec}^V$
$\mathcal{L}_{con}^V$
$\mathbf{Q}^V=\mathcal{M}_V(\mathbf{Z}^V;\mathbf{U}^V)$
: complete data
/ : incomplete data
: missing data
: P-step
: Z-step

Results

The results show that our method achieves superior performance.

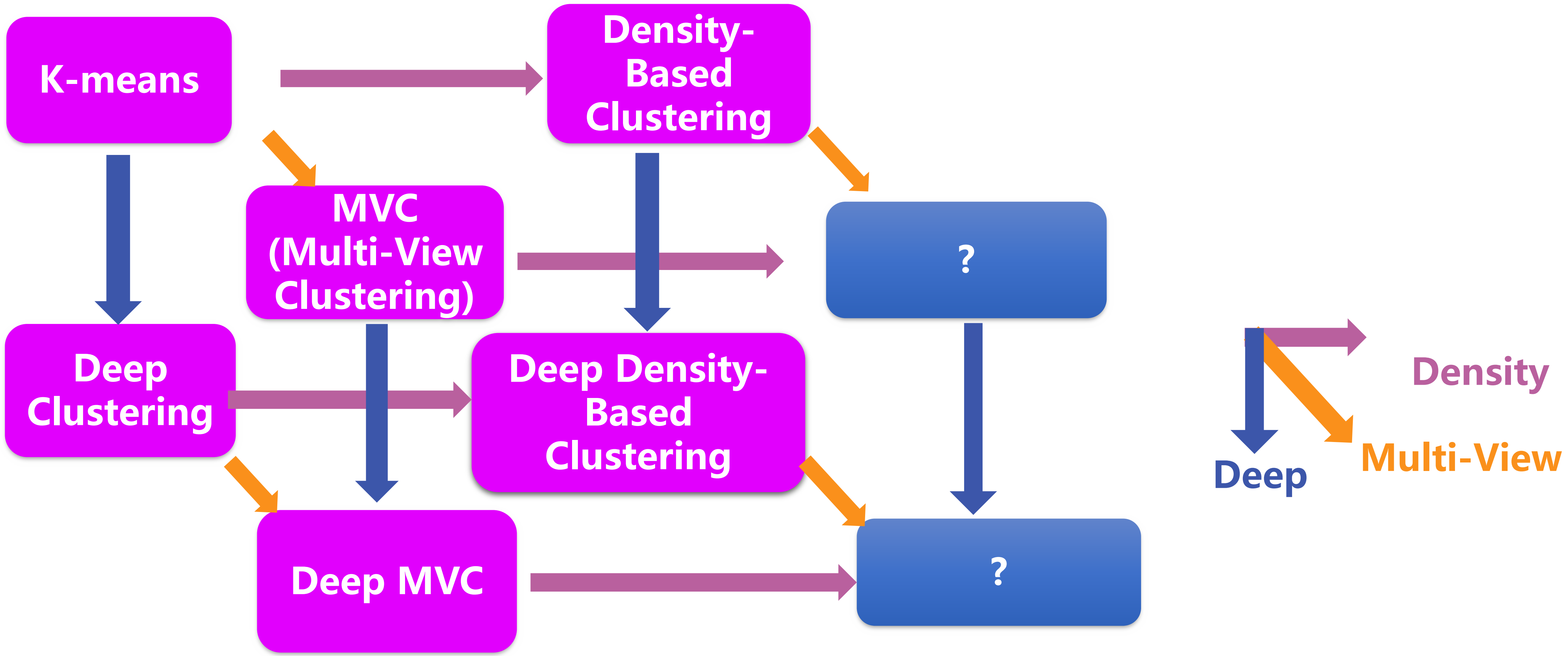
	Missing rates	0.3		
	Evaluation metrics	ACC	NMI	ARI
BDGP	SRLC (Zhuge et al. 2019)	0.697	0.458	0.430
	APMC (Guo and Ye 2019)	0.814	0.589	0.594
	TMBSD (Li et al. 2021)	0.714	0.597	0.546
	IMVTSC-MVI (Wen et al. 2021)	0.934	0.816	0.844
	DiMVMC (Wei et al. 2020)	0.730	0.677	0.565
	CDIMC-net (Wen et al. 2020)	0.757	0.692	0.467
	COMPLETER (Lin et al. 2021)	0.552	0.511	0.255
	DIMVC (**ours**)	**0.954**	**0.866**	**0.889**
Caltech	SRLC (Zhuge et al. 2019)	0.453	0.566	0.311
	APMC (Guo and Ye 2019)	0.437	0.600	0.302
	TMBSD (Li et al. 2021)	0.415	0.615	0.281
	IMVTSC-MVI (Wen et al. 2021)	0.590	0.668	0.445
	DiMVMC (Wei et al. 2020)	0.358	0.521	0.301
	CDIMC-net (Wen et al. 2020)	0.443	0.439	0.265
	COMPLETER (Lin et al. 2021)	0.741	0.690	0.835
	DIMVC (**ours**)	**0.761**	**0.697**	**0.842**

The learned features of all available data have the similar cluster structures.



(a) Missing rate = 0.1 (b) Missing rate = 0.3

(c) Missing rate = 0.5 (d) Missing rate = 0.7


K-means
Density-Based Clustering
MVC (Multi-View Clustering)
?
Deep Clustering
Deep Density-Based Clustering
Deep MVC
?
Density
Deep
Multi-View

相关工作

1. Jie Xu, Huayi Tang, **Yazhou Ren***, Liang Peng, Xiaofeng Zhu, Lifang He. Multi-level Feature Learning for Contrastive Multi-view Clustering. CVPR (CCF A 类会议), 2022.
2. Jie Xu, Chao Li, **Yazhou Ren***, Liang Peng, Yujie Mo, Xiaoshuang Shi, Xiaofeng Zhu. Deep Incomplete Multi-view Clustering via Mining Cluster Complementarity. AAAI (CCF A 类会议), 2022.
3. Jie Xu, **Yazhou Ren***, Huayi Tang, Xiaorong Pu, Xiaofeng Zhu, Ming Zeng, Lifang He. Multi-VAE: Learning Disentangled View-common and View-peculiar Visual Representations for Multi-view Clustering. ICCV (CCF A 类会议), 2021, pages 9234-9243.
4. Zongmo Huang, **Yazhou Ren*,** Xiaorong Pu, Lifang He. Non-Linear Fusion for Self-Paced Multi-View Clustering. ACM MM (CCF A 类会议), 2021, pages 3211-3219.
5. Zongmo Huang, **Yazhou Ren***, Xiaorong Pu, Lili Pan, Dezhong Yao, Guoxian Yu. Dual self-paced multi-view clustering. Neural Networks (SCI二区), 2021, 140: 184-192.
6. Jie Xu, **Yazhou Ren***, Guofeng Li, Lili Pan, Ce Zhu, Zenglin Xu. Deep Embedded Multi-view Clustering with Collaborative Training. Information Sciences (SCI升级版一区), 2021, 573: 279-290.
7. Jie Xu, **Yazhou Ren***, Huayi Tang, Zhimeng Yang, Lili Pan, Yang Yang, Xiaorong Pu. Self-supervised Discriminative Feature Learning for Multi-view Clustering (SDMVC, arXiv, 2021)
8. **Yazhou Ren***, Ni Wang, Mingxia Li, Zenglin Xu. Deep density-based image clustering, Knowledge-Based Systems (SCI一区), 2020, 197: 105841.
9. **Yazhou Ren***, Kangrong Hu, Xinyi Dai, Lili Pan, Steven CH Hoi, Zenglin Xu. Semi-supervised deep embedded clustering , Neurocomputing (SCI 二区), 2019, 325: 121-130.
10. **Y. Ren***, P. Zhao et al., Robust Softmax Regression for Multi-class Classification with Self-Paced Learning, IJCAI, 2017
11. L. Pan, S. Ai, **Y. Ren**, Z. Xu, Self-Paced Deep Regression Forests with Consideration on Underrepresented Samples, ECCV, 2020

Tutors

yazhou.ren@uestc.edu.cn

Yazhou Ren
任亚洲 电子科技大学

Xiaorong Pu
蒲晓蓉 电子科技大学

Students

Jie Xu
徐杰

Zongmo Huang
黄宗漠

Zhimeng Yang
杨之蒙

Zirui Wu
吴之锐

Huayi Tang
唐华镱

谢　　谢！