AE^2-Nets: Autoencoder in Autoencoder Networks for Learning Multi-view Representation

张长青

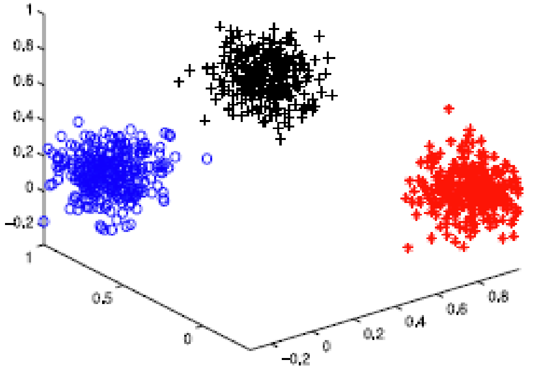
天津大学 智能与计算学部

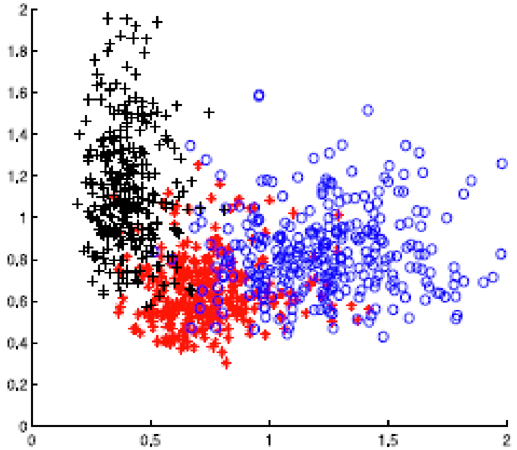
2019-04-02

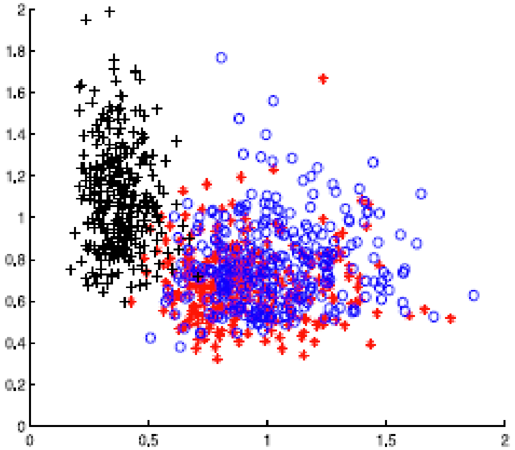
[CVPR'19] Changqing Zhang (张长青), Yeqing Liu, Huazhu Fu,
AE^2-Nets: Autoencoder in Autoencoder Networks,
IEEE Conference on Computer Vision and Pattern Recognition (CVPR, Oral Paper) 2019.

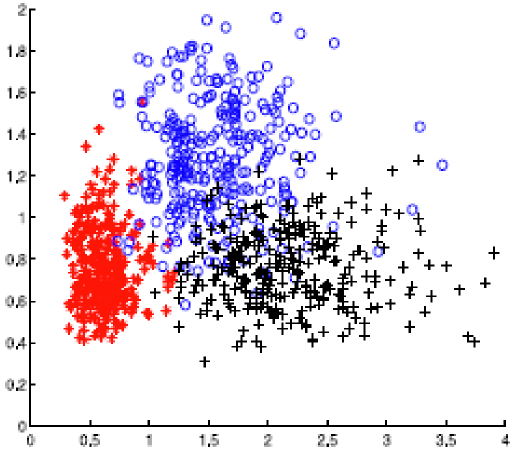
多视图学习

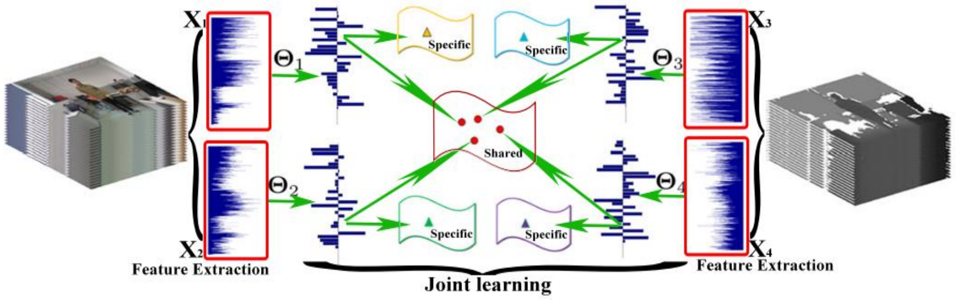
为什么多视图学习重要?

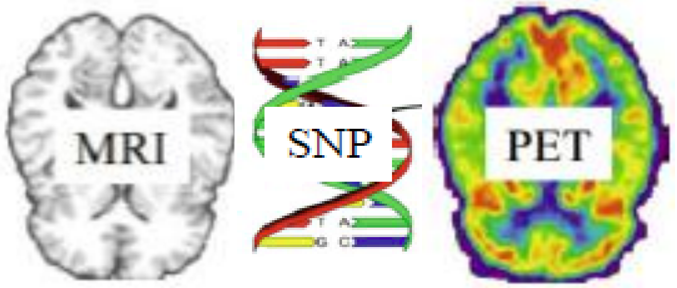


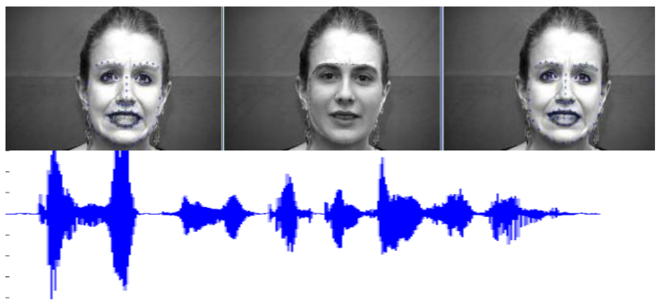






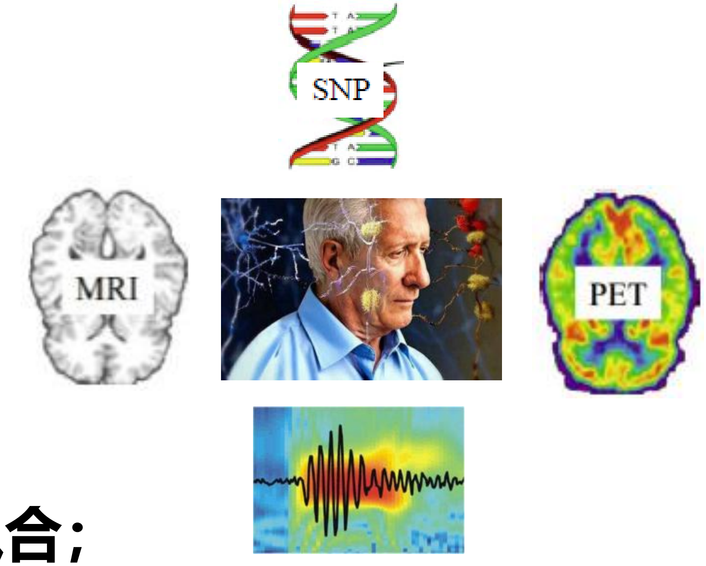






多视图学习本质问题思考

- **特性：视图之间关系**
  - **基本特性1：一致性（相关性），相同对象；**
  - **基本特性2：互补性（独立性），不同角度；**
  - **衍生特性：完备性-多视图的信息是相对完备的。**
- **挑战：**
  - **高维：维度灾难带来类可分性差、难以揭示聚类结构/过拟合；**
  - **度量：预先定义的距离/相似度计算；**
  - **平衡：一致性与互补性如何平衡；**
  - **缺失：视图残缺；**
  - **高度异构：如非数值属性混杂；**
  - **…**

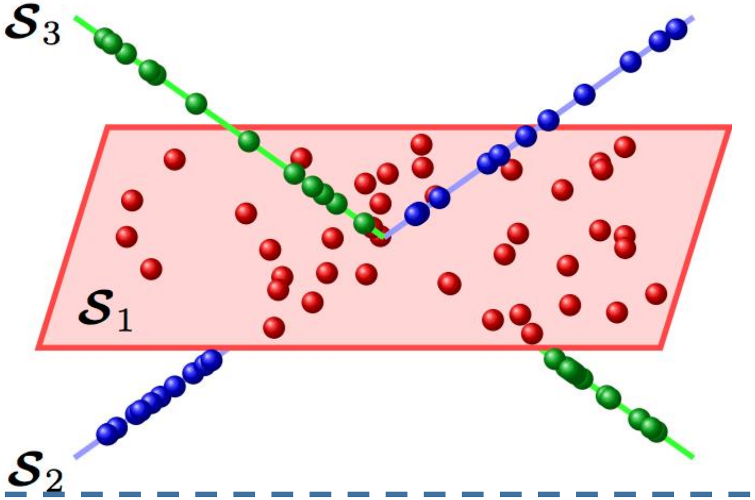


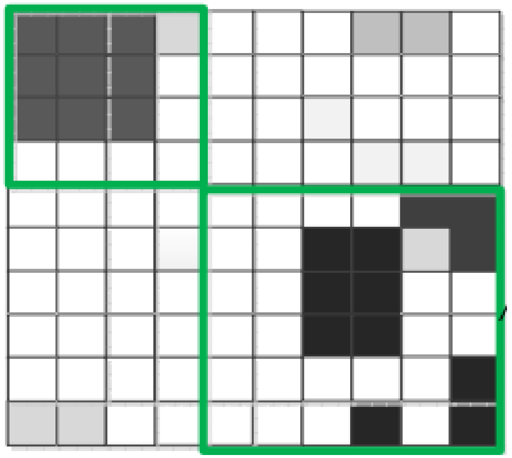
数据自表示的子空间聚类

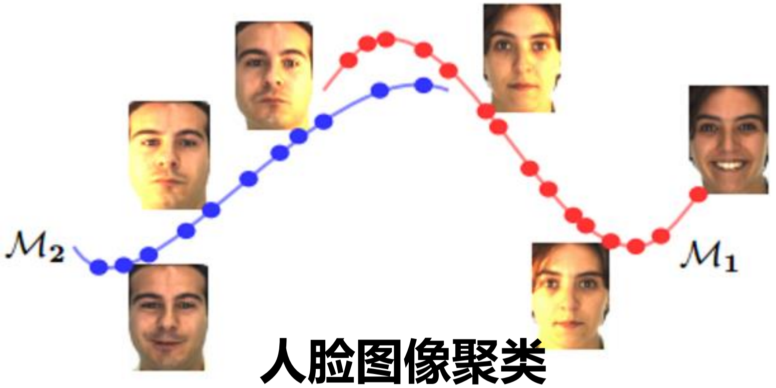
1. 维度灾难->维度福音 [高维]
2. 预先定义相似度函数->自适应学习相似度 [度量]

[自表示子空间假设]高维空间中的数据来自**多个低维子空间，**同一个子空间的数据可**由其他数据线性表示。**

线性子空间　　　　**子空间表示**





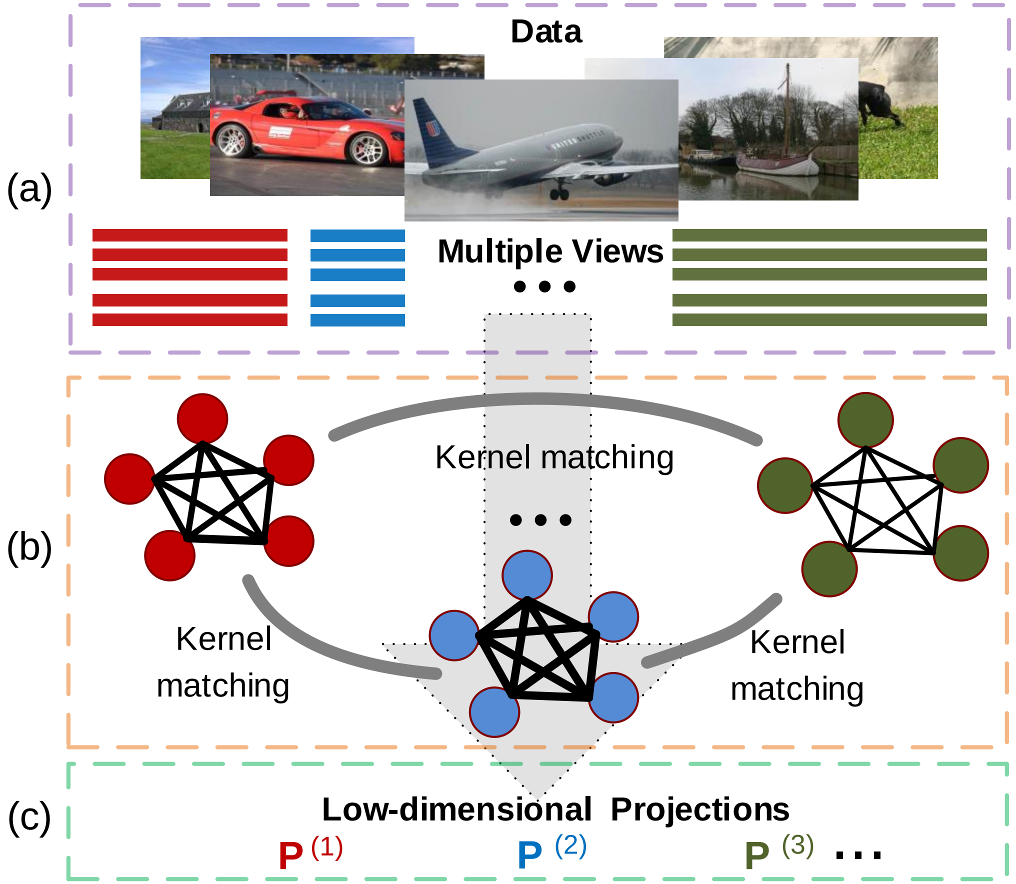


人脸图像聚类

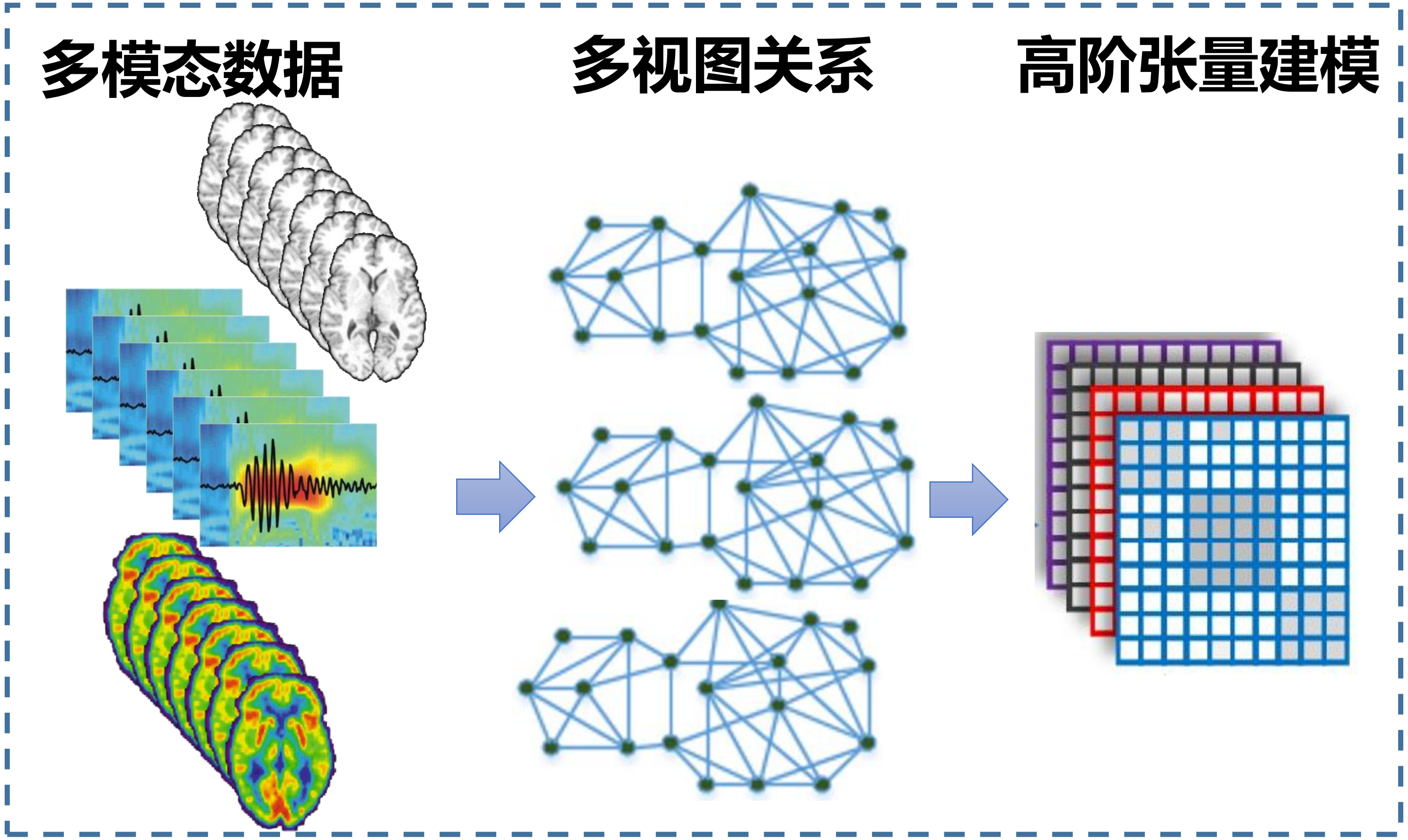
运动图像分割

增强一致性：全局视角-高阶关联

成对视图关系约束



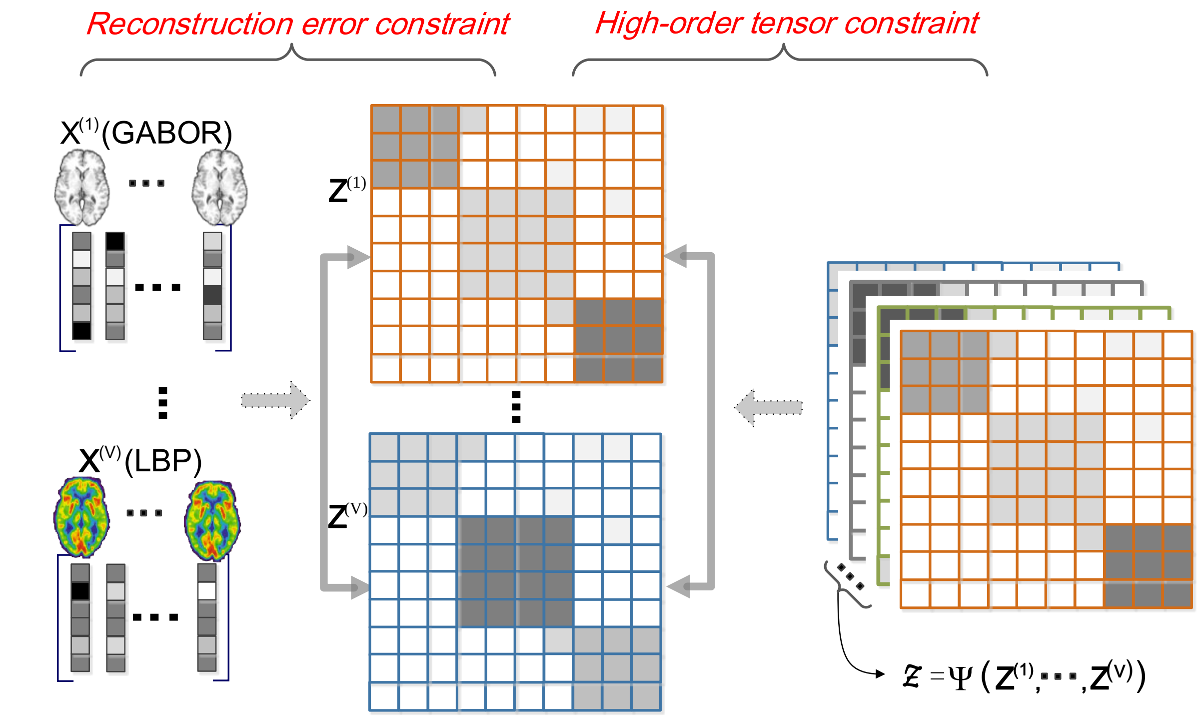
高阶关联



[ICCV'15] Changqing Zhang（张长青）, Huazhu Fu, Si Liu, Guangcan Liu, Xiaochun Cao, *Low-Rank Tensor Constrained Multiview Subspace Clustering*, **ICCV 2015**.

增强一致性：低秩张量多视角子空间聚类

关键创新：构建相似度矩阵->学出低秩高阶张量



(a) Multiview Data (b) Multiview Subspace Representations (c) Low-rank Tensor

- 相似度矩阵自动学出，而非预先定义
- 挖掘多视图高阶关，联而非传统成对约束
- 结构化误差，噪声在多视图间的一致性

$$\min_{\mathbf{Z}^{(v)},\mathbf{E}^{(v)}} ||\mathcal{Z}||_* + \lambda||\mathbf{E}||_{2,1}$$

$$\text{s.t. } \mathbf{X}^{(v)} = \mathbf{X}^{(v)}\mathbf{Z}^{(v)} + \mathbf{E}^{(v)},\ v = 1, 2, ..., V,$$

$$\mathcal{Z} = \Psi(\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, ..., \mathbf{Z}^{(V)}),$$

$$\mathbf{E} = [\mathbf{E}^{(1)}; \mathbf{E}^{(2)}; ...; \mathbf{E}^{(V)}],$$

[ICCV'15] Changqing Zhang (张长青), Huazhu Fu, Si Liu, Guangcan Liu, Xiaochun Cao, *Low-Rank Tensor Constrained Multiview Subspace Clustering*, **ICCV 2015**.

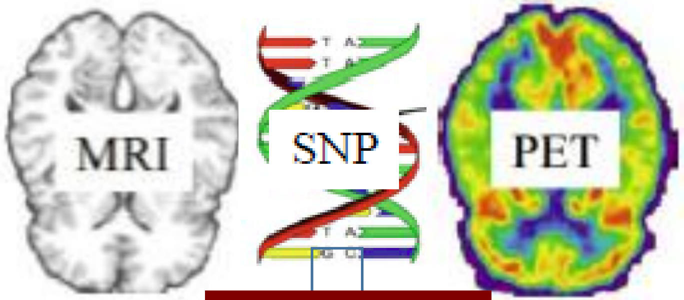
增强互补性：多样性诱导的多视图聚类

映射之后，互补性是否依然保持?

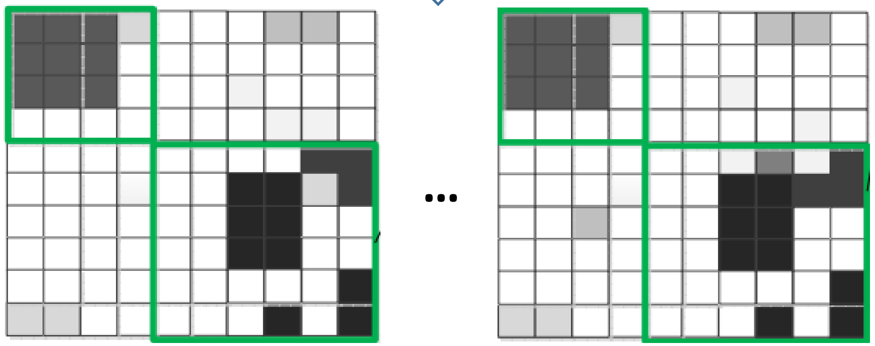
输入多视图
具有一定互补性

↓ **映射/变换**

输出[相似度矩阵/分类标记/回归结果]
是否仍然具有互补性



↓ **映射/变换**



[CVPR’15] Xiaochun Cao, **Changqing Zhang (张长青) ***, Huazhu Fu, Si Liu, Hua Zhang, *Diversity-induced Multiview Subspace Clustering*, **CVPR 2015**.

增强互补性：多样性诱导的多视图聚类

关键创新：增强多视图间的多样性（第一个多视图子空间聚类方法）

- 自适应学习相似度矩阵
- 显式增强样本自表示的多样性
- 自动加权不同视图

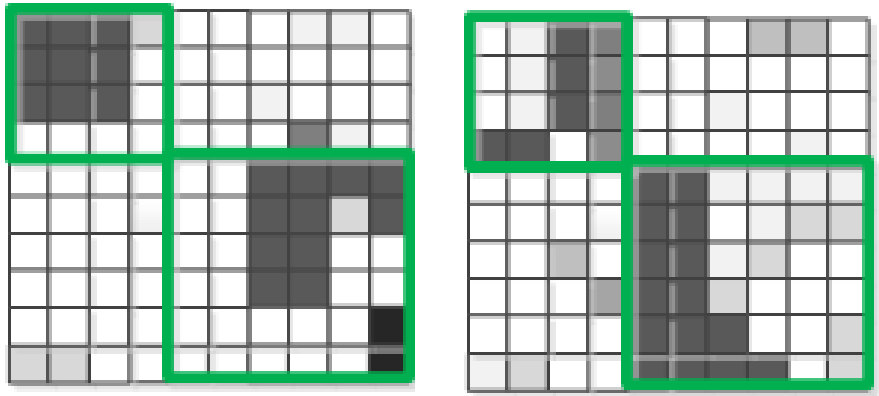
$$\mathcal{O}(\mathbf{Z}^{(1)},...,\mathbf{Z}^{(V)};\mathbf{P}^{(1)},...,\mathbf{P}^{(V)};\alpha)$$

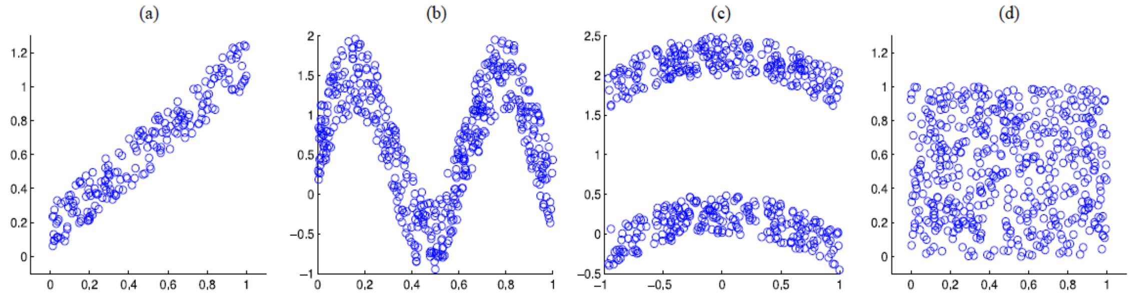
$$= \underbrace{\sum_{v=1}^{V}\Big\{\alpha_v^r\|\mathbf{P}^{(v)}\mathbf{X}^{(v)}-\mathbf{P}^{(v)}\mathbf{X}^{(v)}\mathbf{Z}^{(v)}\|_F^2}_{\text{reconstruction error}}$$

$$+\underbrace{\lambda_S tr(\mathbf{Z}^{(v)}\mathbf{L}^{(v)}\mathbf{Z}^{(v)^T})}_{\text{smoothness}}+\underbrace{\lambda_P\|\mathbf{X}^{(v)}-(\mathbf{P}^{(v)})^T\mathbf{P}^{(v)}\mathbf{X}^{(v)}\|_F^2}_{\text{information preservation}}\Big\}$$

$$+\underbrace{\lambda_V\sum_{v\neq w}\mathrm{HSIC}(\mathbf{Z}^{(v)},\mathbf{Z}^{(w)})}_{\text{diversity}}$$

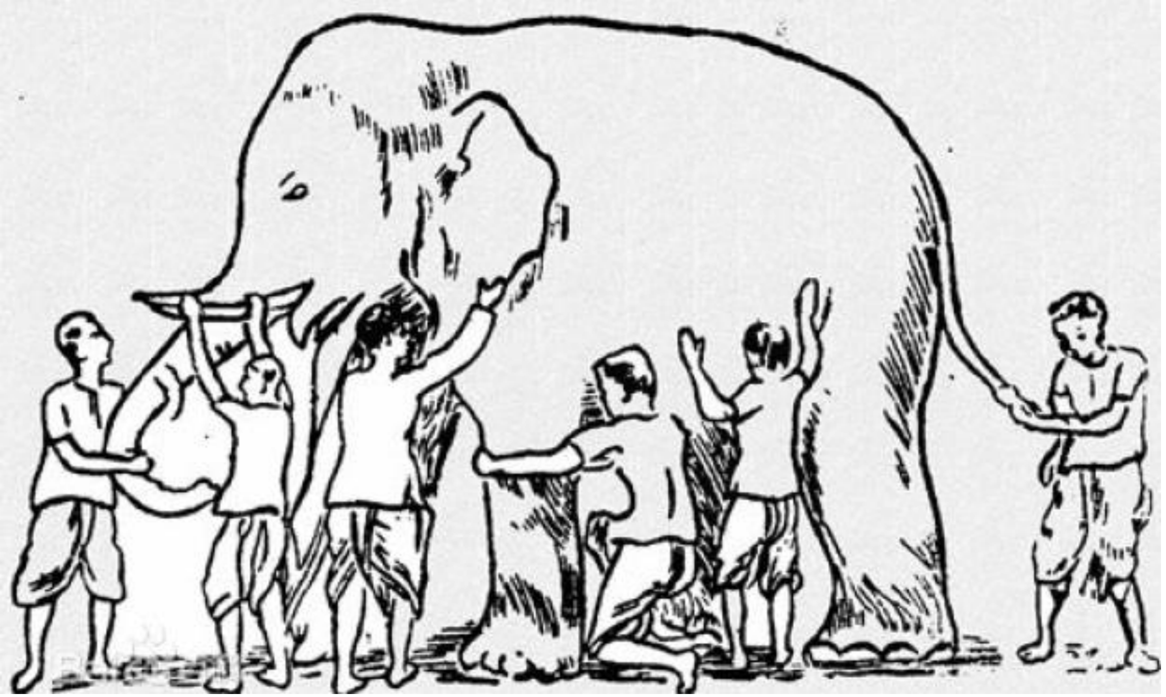
weighted errors of different views





[CVPR'15] Xiaochun Cao, **Changqing Zhang (张长青) ***, Huazhu Fu, Si Liu, Hua Zhang, *Diversity-induced Multiview Subspace Clustering*, **CVPR 2015**.

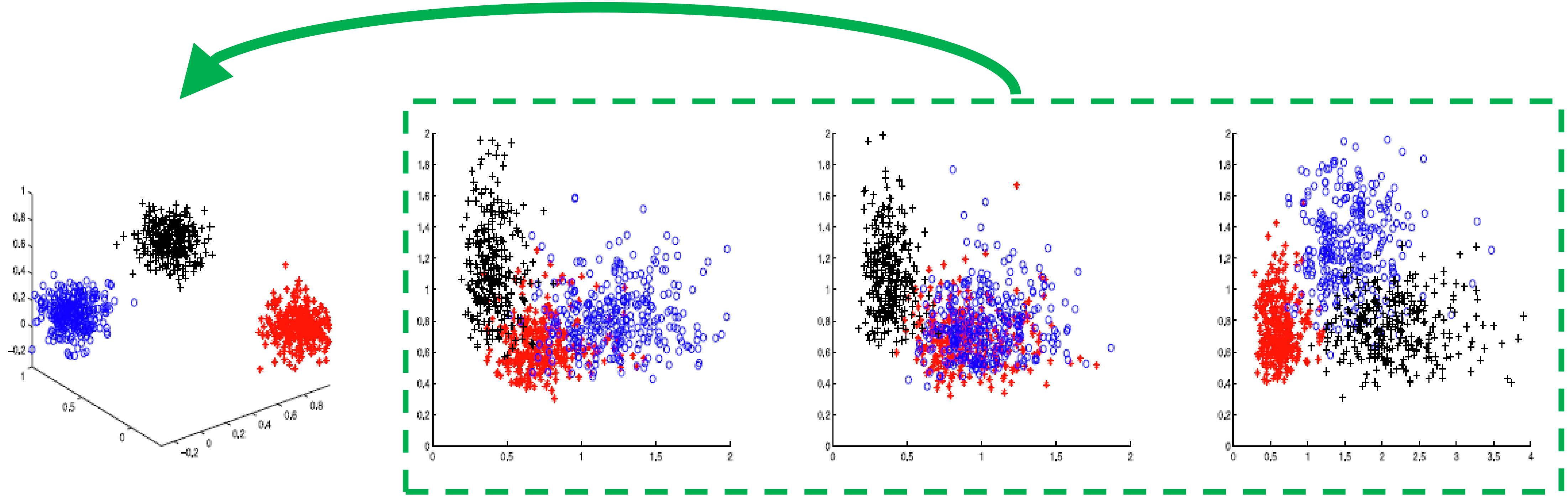
重新思考：多视图融合



- 传统集成方式：每个视图**根据各自信息决策**，然后投票决定最后结果；【***Results voting***】
- 新思路：面向任务、将各视图信息汇总、提炼。【***Task-oriented representation learning***】
- 挑战：视图之间的关联复杂，一致性（**共性**）与互补性（**个性**）如何共存？

增强完备性：隐表达多视角子空间聚类

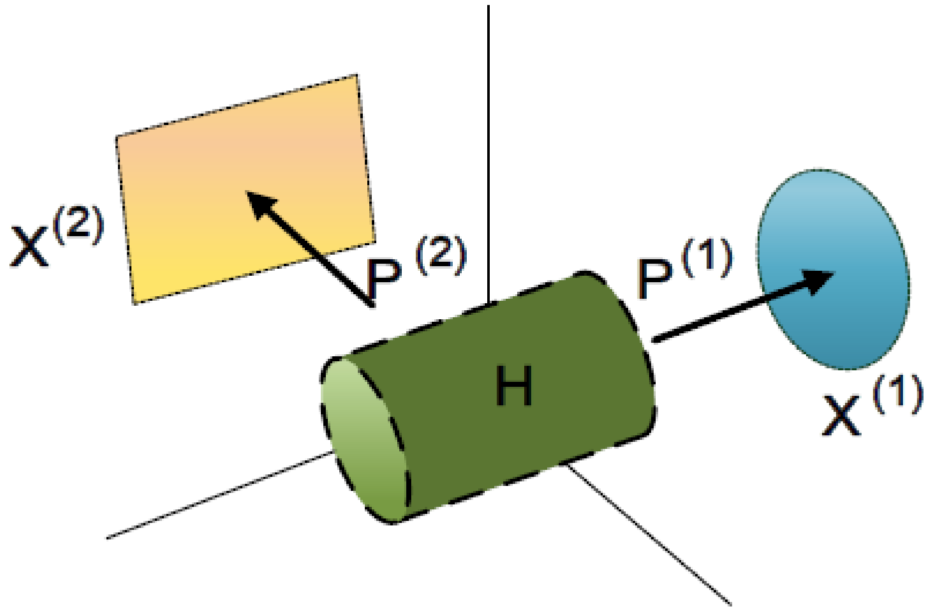
能否自适应恢复出完备的多视图表示?



[CVPR'17] **Changqing Zhang (张长青)**, Qinghua Hu, Huazhu Fu, Pengfei Zhu, Xiaochun Cao,
Latent Multi-View Subspace Clustering, **CVPR 2017** **(Spotlight Paper)**.

增强完备性：隐表达多视图聚类

信息融合的完备性：信息反向传输+迭代修正



- 传统的正向传输，难以保证完备性；
- 完备信息可以退化为各视角信息，模拟信息传输过程，灵活融合多源数据。

正向传输 $\sum_{v=1}^{V}||f(\mathbf{x}_n^{(v)};\boldsymbol{\theta}^{(v)})-\mathbf{h}_n||^2$

反向传输 $\sum_{v=1}^{V}||f(\mathbf{h}_n;\boldsymbol{\theta}^{(v)})-\mathbf{x}_n^{(v)}||^2$

[CVPR'17] **Changqing Zhang (张长青)**, Qinghua Hu, Huazhu Fu, Pengfei Zhu, Xiaochun Cao,
Latent Multi-View Subspace Clustering, **CVPR 2017 (Spotlight Paper)**.

增强完备性：嵌套式自编码机

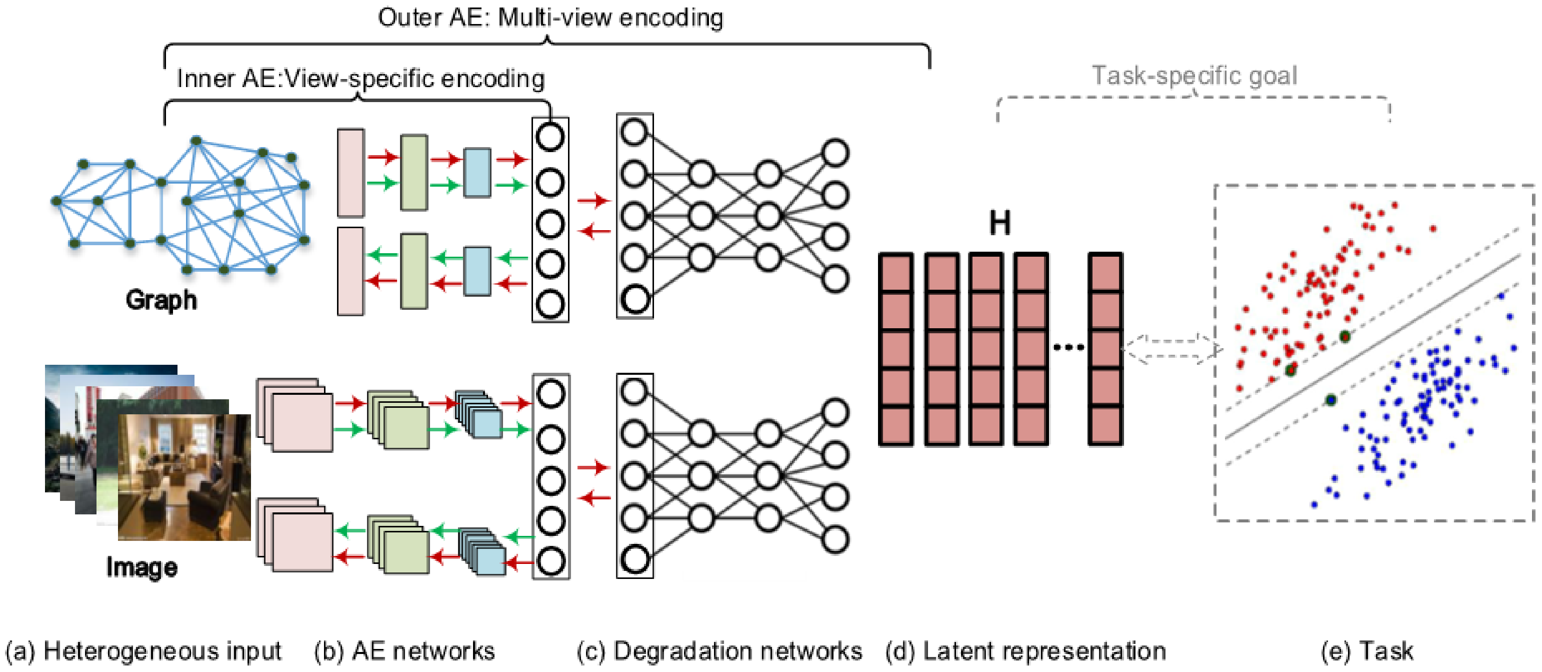
Motivation

- **Completeness (完备性):** Model multi-view integration with *information degradation mimicking data transmission*, which brings great flexibility for balancing the consistence and complementarity across different views.
- **Robustness (鲁棒性)**: *Nested encoding* to integrate intrinsic information of each view from view-specific *autoencoder*.
- **End-to-End (端到端):** Flexible for heteronomous data by using *modal-specific convolutional neural networks*.

[CVPR'19] Changqing Zhang (张长青), Yeqing Liu, Huazhu Fu,
AE^2-Nets: Autoencoder in Autoencoder Networks,
IEEE Conference on Computer Vision and Pattern Recognition (CVPR, Oral Paper) 2019.

增强完备性：嵌套式自编码机

关键创新：内层提取本质信息+外层融合多源信息



- 无监督多视图完备、紧致统一表示
- 应对多视图复杂性关联
- 避免直接对高维数据进行重建，融合各视图本质信息；
- 端到端异构数据编码。

$$\min_{\{\Theta_{ae}^{(v)},\Theta_{dg}^{(v)},\mathbf{H}\}} \frac{1}{2}\sum_{v=1}^{V}\left(\left\|\mathbf{X}^{(v)}-\mathbf{Z}^{(M,v)}\right\|_F^2 + \lambda\left\|\mathbf{Z}^{(\frac{M}{2},v)}-\mathbf{G}^{(v)}\right\|_F^2\right)$$

[CVPR'19] Changqing Zhang (张长青), Yeqing Liu, Huazhu Fu,
AE^2-Nets: Autoencoder in Autoencoder Networks,
IEEE Conference on Computer Vision and Pattern Recognition (CVPR, Oral Paper) 2019.

增强完备性：嵌套式自编码机

Framework

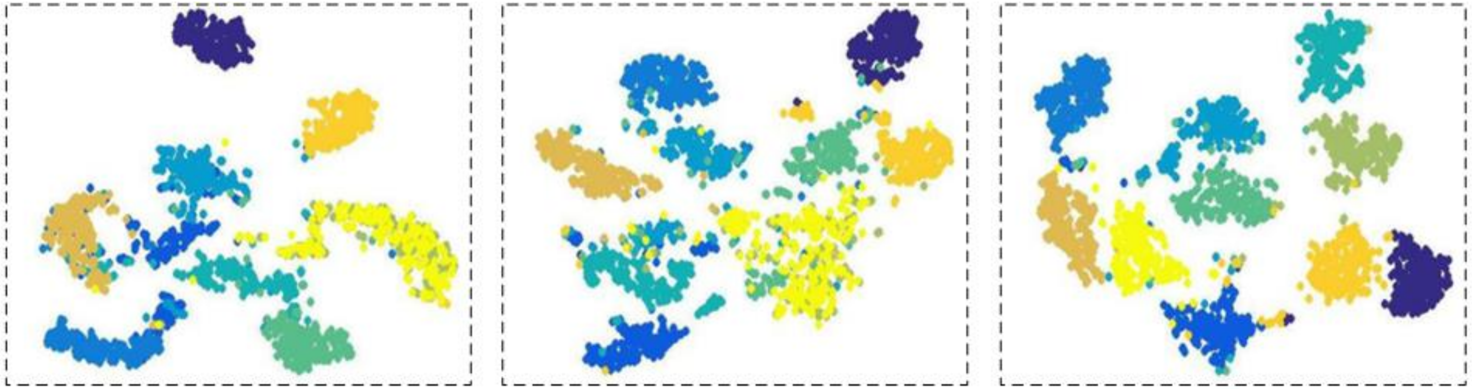
- **Inner AE (内层编码器)**: *<u>View-specific autoencoder networks</u>* under the constraints of reconstructions based on the learned latent representation for better view-specific representation.

- **Outer AE (外层编码器)**: *<u>Degradation networks</u>* for better latent representation to reconstruct learned view-specific representation.

- **AE in AE (嵌套编码器)**: Integrate intrinsic information from all views into an *<u>intact latent representation</u>*.

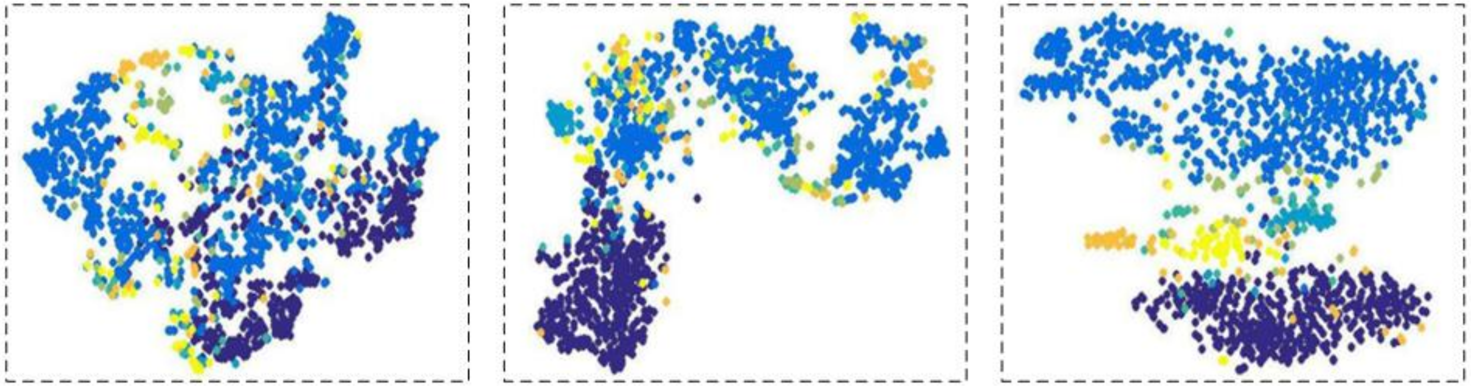
[CVPR'19] Changqing Zhang (张长青), Yeqing Liu, Huazhu Fu,
AE^2-Nets: Autoencoder in Autoencoder Networks,
<u>IEEE Conference on Computer Vision and Pattern Recognition (CVPR, Oral Paper) 2019.</u>

增强完备性：嵌套式自编码机



(1) View 1 (2) View 2 (3) Ours

多视图融合有效性实验验证



(1) View 1 (2) View 2 (3) Ours

[CVPR’19] Changqing Zhang (张长青), Yeqing Liu, Huazhu Fu,
AE^2-Nets: Autoencoder in Autoencoder Networks,
IEEE Conference on Computer Vision and Pattern Recognition (CVPR, Oral Paper) 2019.

增强完备性：嵌套式自编码机

统一表示在分类及聚类任务上验证

Table 1: Performance comparison on clustering task.

Datasets	Methods	ACC	NMI	F_score	RI
handwritten	FeatConcate	76.04 ± 2.28	75.70 ± 1.44	70.96 ± 2.05	93.93 ± 0.42
	CCA [14]	66.43 ± 7.62	69.62 ± 6.06	62.05 ± 7.70	91.83 ± 1.79
	DCCA [2]	66.26 ± 0.16	66.01 ± 0.45	59.05 ± 0.39	91.39 ± 0.06
	DCCAE [27]	69.17 ± 1.02	66.96 ± 0.91	60.50 ± 1.10	91.77 ± 0.21
	MDcR [33]	76.72 ± 2.77	**76.68** ± 0.93	**71.93** ± **2.22**	**94.11** ± **0.48**
	DMF-MVC [36]	71.86 ± 4.25	73.09 ± 3.23	66.66 ± 4.69	92.85 ± 1.13
	Ours	**81.52** ± **1.62**	71.39 ± 1.50	68.57 ± 1.86	93.68 ± 0.38
Caltech101	FeatConcate	47.23 ± 0.22	57.19 ± 0.61	52.15 ± 0.28	73.45 ± 0.16
	CCA [14]	45.37 ± 0.09	50.53 ± 0.03	52.15 ± 0.19	73.27 ± 0.09
	DCCA [2]	56.71±10.50	57.61 ± 6.78	62.32±12.75	76.34 ± 6.86
	DCCAE [27]	62.11 ± 2.78	**64.38** ± **4.11**	65.43 ± 4.24	79.31 ± 2.06
	MDcR [33]	46.51 ± 0.67	56.43 ± 0.56	51.55 ± 0.56	73.27 ± 0.30
	DMF-MVC [36]	55.75 ± 5.67	45.52 ± 2.28	55.67 ± 5.50	73.43 ± 2.33
	Ours	**66.46** ± **4.55**	60.60 ± 1.93	**73.42** ± **4.91**	**83.14** ± **2.33**
ORL	FeatConcate	61.10 ± 1.51	79.28 ± 0.70	47.03 ± 2.21	97.10 ± 0.25
	CCA [14]	56.98 ± 2.06	76.03 ± 0.79	45.13 ± 1.83	97.32 ± 0.09
	DCCA [2]	59.68 ± 2.04	77.84 ± 0.83	47.72 ± 2.05	97.42 ± 0.13
	DCCAE [27]	59.40 ± 2.20	77.52 ± 0.86	46.71 ± 2.22	97.39 ± 0.14
	MDcR [33]	61.70 ± 2.19	79.45 ± 1.20	48.48 ± 2.59	97.28 ± 0.22
	DMF-MVC [36]	65.38 ± 2.86	82.87 ± 1.26	52.01 ± 3.43	97.29 ± 0.30
	Ours	**68.85** ± **2.11**	**85.73** ± **0.78**	**59.93** ± **1.31**	**97.94** ± **0.11**

Table 2: Performance comparison on classification task.

Datasets	Methods	$G_{80\%}/P_{20\%}$	$G_{70\%}/P_{30\%}$	$G_{50\%}/P_{50\%}$	$G_{20\%}/P_{80\%}$
handwritten	FeatConcate	89.60 ± 1.40	88.97 ± 0.73	88.87 ± 0.44	85.68 ± 0.53
	CCA [14]	93.78 ± 0.82	93.47 ± 0.93	93.28 ± 0.66	91.12 ± 0.74
	DCCA [2]	95.18 ± 0.55	94.62 ± 0.64	94.35 ± 0.46	92.79 ± 0.51
	DCCAE [27]	95.78 ± 0.46	95.10 ± 0.64	94.79 ± 0.58	92.63 ± 0.54
	MDcR [33]	92.33 ± 0.73	91.55 ± 0.39	91.41 ± 0.68	88.11 ± 0.61
	DMF-MVC [36]	94.68 ± 0.71	93.72 ± 0.60	93.33 ± 0.46	88.23 ± 0.57
	Ours	**96.93** ± **0.71**	**96.55** ± **0.66**	**95.88** ± **0.71**	**93.38** ± **0.49**
Caltech101	FeatConcate	87.88 ± 0.67	87.47 ± 0.56	87.17 ± 0.49	87.10 ± 0.45
	CCA [14]	91.10 ± 0.96	90.07 ± 1.03	89.82 ± 0.49	89.08 ± 0.71
	DCCA [2]	92.12 ± 0.58	91.46 ± 0.70	91.30 ± 0.48	90.73 ± 0.38
	DCCAE [27]	91.58 ± 1.02	90.91 ± 0.75	90.54 ± 0.44	89.44 ± 0.43
	MDcR [33]	90.14 ± 0.74	89.45 ± 0.76	88.95 ± 0.41	88.46 ± 0.35
	DMF-MVC [36]	85.51 ± 1.05	84.67 ± 0.82	81.88 ± 0.73	74.19 ± 0.99
	Ours	**93.77** ± **1.35**	**92.98** ± **1.37**	**92.49** ± **0.72**	**91.36** ± **0.69**
ORL	FeatConcate	79.13 ± 2.36	74.58 ± 1.32	68.00 ± 2.23	48.28 ± 2.27
	CCA [14]	77.13 ± 3.96	73.83 ± 4.89	67.95 ± 2.77	49.00 ± 1.84
	DCCA [2]	83.25 ± 2.71	78.92 ± 1.93	71.15 ± 1.86	51.69 ± 1.75
	DCCAE [27]	81.62 ± 2.95	80.00 ± 1.47	72.80 ± 2.04	51.25 ± 1.90
	MDcR [33]	92.00 ± 1.58	90.83 ± 2.08	83.35 ± 1.08	57.38 ± 2.08
	DMF-MVC [36]	93.13 ± 1.21	91.75 ± 1.64	85.45 ± 1.85	56.44 ± 2.50
	Ours	**97.88** ± **1.19**	**96.00** ± **2.18**	**92.20** ± **1.18**	**70.16** ± **2.54**

[CVPR'19] Changqing Zhang (张长青), Yeqing Liu, Huazhu Fu,
AE^2-Nets: Autoencoder in Autoencoder Networks,
IEEE Conference on Computer Vision and Pattern Recognition (CVPR, Oral Paper) 2019.

增强完备性：嵌套式自编码机

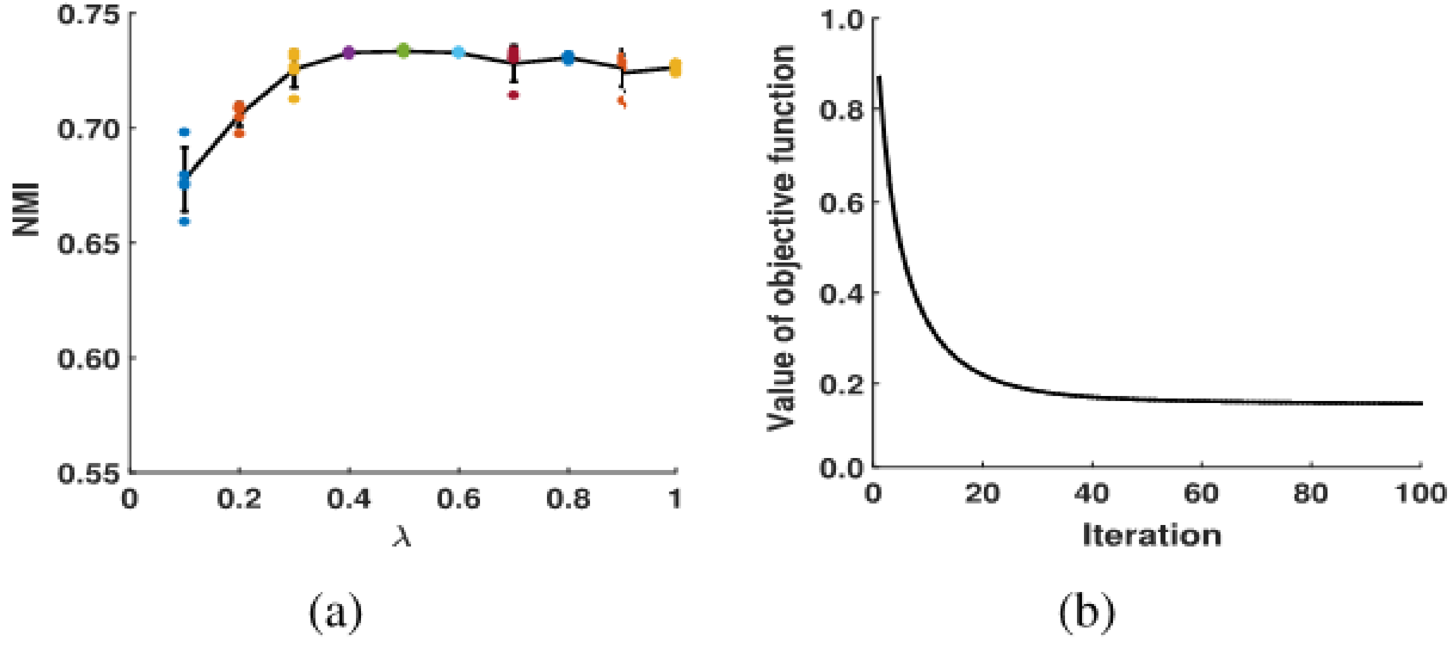


Figure 3: Parameter tuning (a) and convergence curve (b).

- 选择相对较大的lambda值，增强将各视图信息编码到统一表示；
- 迭代下降，具有较好收敛特性。注：一次iteration包含更新ae，dg和h，其中，每个更新又分别包含多次batch更新。

[CVPR'19] Changqing Zhang (张长青), Yeqing Liu, Huazhu Fu,
AE^2-Nets: Autoencoder in Autoencoder Networks,
IEEE Conference on Computer Vision and Pattern Recognition (CVPR, Oral Paper) 2019.

谢谢大家！

张长青

天津大学 智能与计算学部