

# CMvSC: 知识迁移下的深度一致性多视图谱聚类网络\*



张熠玲, 杨燕, 周威, 欧阳小草, 胡节

(西南交通大学 计算机与人工智能学院, 四川 成都 611756)

通信作者: 杨燕, E-mail: yyang@swjtu.edu.cn

**摘要:** 谱聚类是聚类分析中极具代表性的方法之一, 由于其对数据结构没有太多假设要求, 受到了研究者的广泛关注. 但传统的谱聚类算法通常受到谱嵌入的可扩展性和泛化性的限制, 即: 无法应对大规模设置和复杂数据分布. 为克服以上缺陷, 旨在引入深度学习框架提升谱聚类的泛化能力与可扩展能力, 同时, 结合多视图学习挖掘数据样本的多样性特征, 从而提出一种知识迁移下的深度一致性多视图谱聚类网络(CMvSC). 首先, 考虑到单个视图的局部不变性, CMvSC 采用局部学习层独立学习每个视图的特有嵌入; 其次, 由于多视图具有全局一致性, CMvSC 引入全局学习层进行参数共享与特征迁移, 学习多视图间的共享嵌入; 同时, 考虑到邻接矩阵对谱聚类性能的重要影响, CMvSC 通过训练孪生网络和设计对比损失来学习成对数据间的近邻关系, 以替代传统谱聚类算法中的距离度量; 最后, 4 个数据集上的实验结果证明了 CMvSC 对多视图谱聚类任务的有效性.

**关键词:** 谱嵌入; 近邻学习; 知识迁移; 多视图聚类; 深度聚类

**中图法分类号:** TP181

中文引用格式: 张熠玲, 杨燕, 周威, 欧阳小草, 胡节. CMvSC: 知识迁移下的深度一致性多视图谱聚类网络. 软件学报, 2022, 33(4): 1373-1389. <http://www.jos.org.cn/1000-9825/6474.htm>

英文引用格式: Zhang YL, Yang Y, Zhou W, Ouyang XC, Hu J. CMvSC: Knowledge Transferring Based Deep Consensus Network for Multi-view Spectral Clustering. Ruan Jian Xue Bao/Journal of Software, 2022, 33(4): 1373-1389 (in Chinese). <http://www.jos.org.cn/1000-9825/6474.htm>

## CMvSC: Knowledge Transferring Based Deep Consensus Network for Multi-view Spectral Clustering

ZHANG Yi-Ling, YANG Yan, ZHOU Wei, OUYANG Xiao-Cao, HU Jie

(School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu 611756, China)

**Abstract:** Spectral clustering, which is one of the most representative methods in clustering analysis, receives much attention from scholars, because it does not constrain the data structure of the original samples. However, traditional spectral clustering algorithm usually contains two major limitations, i.e., it is unable to cope with the large-scale settings and complex data distribution. To overcome the above shortcomings, this study introduces a deep learning framework to improve the generalization and scalability of spectral clustering, and combines the multi-view learning to mine diverse features among data samples, finally proposes a knowledge transferring based deep consensus network for multi-view spectral clustering (CMvSC). First, considering the local invariance of single view, CMvSC adopts the local learning layer to learn the specific embedding of each view individually. Then, because of the global consistency among multiple views, CMvSC introduces the global learning layer to achieve parameter sharing and feature transferring, and learns the shared embedding in different views. Meanwhile, taking the effect of affinity matrix for spectral clustering into consideration, CMvSC learns the affinity correlation between the paired samples by training the Siamese network and designing the contrastive loss, which replaces the distance metric in traditional spectral clustering. Finally, the experimental results on four datasets demonstrate the effectiveness of the proposed

\* 基金项目: 国家自然科学基金(61976247)

本文由“面向开放场景的鲁棒机器学习”专刊特约编辑陈恩红教授、李宇峰副教授、邹权教授推荐.

收稿时间: 2021-05-29; 修改时间: 2021-07-16; 采用时间: 2021-08-27; jos 在线出版时间: 2021-10-26

CMvSC for multi-view clustering.

**Key words:** spectral embedding; affinity learning; knowledge transferring; multi-view clustering; deep clustering

聚类分析旨在从无标签数据中准确有效地发现样本类簇,其具有重要的科学意义与实际价值.大数据时代的到来和新媒体的发展,催生出了大规模的数据样本,但为其标记类别标签无疑需要耗费大量人力物力.聚类技术的发展,为这些数据的特征提取与潜在结构发现提供了强有力的帮助<sup>[1]</sup>.

目前,主流的机器学习研究框架大多关注样本数据的某一种形式,我们称其为单视图学习,其中,文本和图片等某种形式被称为视图<sup>[2]</sup>,不同的文本表示或不同的图片增广均可被视为不同视图.如:同一条新闻经由不同媒体报道可能有不同的文字描述,由此可生成不同的文本表示,这样的不同表示通常被视为多视图数据.对人类而言,我们的大脑往往可以同时处理多种视图的信息,提取不同视图间的互补信息和一致信息并融合,从而做出更准确、客观的决策.如:我们可以对不同媒体报道的同一新闻在大脑中进行整合和提取,从而推论出新闻事件的较全面的细节.换言之,多个视图的信息输入,使得我们人类可以更全面、客观、多样化地看待问题.真实世界中的多视图数据各种各样,不同视图可能有相同的表达形式,也可能不尽相同.例如,上述提到的同样新闻可以被不同的新闻媒体报道,因此产生了来自不同视图的文本特征;同样的情感极性既可以通过不同图片来表达,如微笑、大笑等不同表情图片均可表示开心的情绪,也可以直接用文字表达,由此产生了具有不同特征表示形式的视图数据.当视图之间具有不同表达形式时,通常被称为多模态数据.本篇工作主要关注具有相同表达形式的多视图数据.但不可否认的是,无论视图特征的形式如何变化,他们都只是同一样本对象的不同表示.因而,不同视图之间必然存在着某种程度的联系<sup>[3]</sup>.为了探索不同视图的相关性,同时挖掘多个视图的特征,发现其一致性与差异性知识,使得机器学习算法进一步接近人类智能,面向多视图学习的研究应运而生.

多视图聚类学习旨在从多个视图中探索并利用互补信息和一致信息,从而生成一个相比单视图聚类更准确、更鲁棒的数据划分结果<sup>[4]</sup>.谱聚类方法由于其对数据结构没有太多假设要求且学习性能优秀,一直是无监督数据分析中的极受欢迎的技术之一<sup>[5]</sup>.多视图谱聚类方法首先从所有视图的邻接矩阵中学习一个融合图,然后对融合图进行聚类,得到最后的划分结果.近年来,多视图谱聚类算法已经取得了较好的性能,但是仍存在一些不足:首先,其面向复杂数据分布时的时间效率以及性能均不理想;其次,大多数方法是基于传统机器学习或者只着眼于浅层模型,无法充分挖掘样本数据的深层特征,因此所构建的邻接图的知识表示能力有限<sup>[6]</sup>.为进一步提高多视图谱聚类算法解决真实场景问题的能力,同时考虑尽可能完善上述不足,众多学者考虑利用深度学习框架提升多视图谱聚类算法的深层特征学习能力.而多视图谱聚类算法与深度神经网络结合的关键性难点在于谱聚类损失函数与神经网络的联合优化问题.具体来说,谱聚类要求解决矩阵分解问题使得联合优化异常困难,因为矩阵分解的梯度无法被后向传播以达到优化神经网络的目的.谱网络(spectral clustering using deep neural networks, SpectralNet)<sup>[7]</sup>采用约束随机优化技术训练神经网络,同时优化谱聚类损失,有效解决了联合优化问题.但是 SpectralNet 仅考虑了单视图场景,忽略了多视图之间的互补信息和关联知识.多视图谱聚类网络(multi-view spectral clustering network, MvSCN)<sup>[6]</sup>将其扩展到多视图场景,挖掘视图内一致性与视图间差异性特征,大大提升了算法性能.与 MvSCN 类似,本篇工作同样关注深度多视图谱聚类网络,但与之不同的是:本文主要考虑多视图数据之间既存在单视图局部不变性,也隐含多视图全局一致性,因此分别针对单视图和多视图引入局部学习层与全局学习层以获得特有嵌入和共享嵌入,保证多视图特征的特有性和一致性,增强算法鲁棒性.

解决深度多视图谱聚类任务的关键挑战主要有以下两点.

1) 如何建立有效的多视图框架并获得包含足够丰富特征的一致性谱嵌入?与传统多视图谱聚类算法类似,深度多视图谱聚类利用神经网络层(如全连接层)学习原始数据样本的低秩表示,并对多视图谱嵌入进行融合;

2) 如何构建样本对间的近邻关系?邻接矩阵对谱聚类算法有重要影响,其不仅需要包含丰富的近邻关

系, 还需要体现样本对间的隐含关联.

为解决上述挑战, 本文提出了一种知识迁移下的深度一致性多视图谱聚类网络(knowledge transferring based deep consensus network for multi-view spectral clustering, CMvSC), 利用深度学习框架学习多视图数据的谱嵌入, 同时采用结合对比损失的孪生网络构建样本对间的近邻关系. 首先, 为挖掘每个视图的局部不变性, 局部学习层被提出用于获得单视图的特有嵌入, 挖掘单视图内的隐含关联; 同时, 为进一步学习多视图间的全局一致性, 全局学习层被设计以挖掘多视图的共享嵌入; 最后, 我们借鉴传统多视图谱聚类算法中的正交约束, 在 CMvSC 的顶层引入了正交约束层, 以避免“无效解”问题.

本文工作的主要贡献包括以下 5 个方面.

- 1) 设计了一种局部学习层, 用于挖掘单视图特有嵌入. 针对多视图数据中每个视图内具有局部不变性, 我们设计局部学习层深入挖掘单视图有效特征, 同时保证学习到的特有嵌入仍然保持原始相似结构;
- 2) 设计了一种全局学习层, 用于学习多视图共享嵌入. 考虑多视图数据是同一样本的不同特征表示, 不同视图间具有一定的全局一致性, 我们构建了一种全局学习层, 挖掘不同视图间的一致性特征;
- 3) 设计了一种基于对比学习的近邻学习框架. 为了同时考虑正负样本的相似性影响, 挖掘样本对间的隐含关联, 我们设计了一种孪生网络, 用于实现自适应的近邻学习, 结合基于距离的对比损失, 可获得具有更鲁棒的特征表征能力的邻接矩阵;
- 4) 引入了一种正交约束层, 对特征嵌入进行正交约束. 为避免特征嵌入出现“无效解”, 我们在一致性多视图谱聚类网络顶层引入了一种正交约束层, 保证谱嵌入的正交性;
- 5) 不同数据集上的实验结果表明, 我们提出的 CMvSC 可以有效解决多视图聚类问题.

## 1 相关工作

在本节中, 我们将讨论与所提出的 CMvSC 相关的工作, 主要包括多视图聚类和深度聚类.

### 1.1 多视图聚类

近年来, 多视图聚类方法被广泛研究. 多视图聚类旨在整合多个视图的特征, 学习一致性特征表示, 从而将样本划分为多个类簇. 传统的多视图聚类算法可以分为 4 类.

- 1) 基于协同训练的方法<sup>[4,8]</sup>: 该方法利用一个视图的先验知识或习得的知识来强化另一个视图的聚类性能, 迭代发现与所有视图都足够相近的一致性视图;
- 2) 多核学习<sup>[9,10]</sup>: 该方法根据不同视图定义不同的核函数, 然后线性或非线性地结合所有核以提升聚类性能;
- 3) 基于子空间的方法<sup>[11,12]</sup>: 该方法结合非负矩阵分解挖掘多视图样本之间的关系, 学习一个公共的低维特征表示作为一致性子空间用于聚类任务;
- 4) 基于图的方法<sup>[13,14]</sup>: 该方法旨在所有视图中寻找一个表示样本之间关联性的融合图, 然后利用谱聚类的方法对图进行划分, 从而获得最终聚类结果.

基于子空间的方法和基于图的方法由于其充分学习多视图数据间的补充信息和共享信息, 在数据划分任务上获得了较好的性能. Cao 等人提出了一种多样性诱导的多视图子空间聚类模型(diversity-induced multi-view subspace clustering, DiMSC), 其利用希尔伯特-施密特独立性指标(Hilbert-Schmidt independence criterion, HSIC)作为多样性项探索多视图数据间的潜在互补关联<sup>[15]</sup>. 进一步地, Zhang 等人提出了一种潜在多视图子空间聚类方法(latent multi-view subspace clustering, LMSC), 其学习多视图数据的潜在特征, 并且直接生成一个公共子空间表达<sup>[16]</sup>. 尽管基于子空间学习的方法在聚类等任务中已经获得了较好的性能, 但是其通常受原始数据的质量影响较大<sup>[16]</sup>. 基于图的方法由于对数据结构和数据类型没有太多限制, 受到了众多研究者的关注. 基于图的方法的主要代表是谱聚类算法, 其关键在于如何从数据中发现恰当的邻接矩阵. Kumar 等人利用协同训练技术迭代地优化邻接图<sup>[17,18]</sup>. Zhou 和 Burges 线性结合多个正则化的伴随矩阵<sup>[19]</sup>. Nie 等人将一个数据

样本属于另一个样本的邻居集的概率作为两个样本间的相似性度量,通过自适应邻居集构造邻接矩阵<sup>[20]</sup>.进一步地,假设所有邻接矩阵共享一个特征向量,Nie 等人提出了一种无参多视图聚类算法<sup>[21]</sup>.然而,该算法将邻接矩阵构建和谱聚类分为两个独立过程,因此无法联合优化.为解决上述问题,Nie 等人设计了一个同时实现邻接矩阵构建与多视图谱聚类的统一框架,基于自适应邻居集学习一个公共邻接矩阵<sup>[22]</sup>.为了提升邻接矩阵的特征表达能力和保留数据中隐含知识,Kang 等人提出了一种多图融合的多视图谱聚类算法,使得融合后的邻接矩阵尽可能接近每个视图的原始邻接矩阵,但增强了类簇结构<sup>[23]</sup>.Nie 等人提出了一种自加权的多视图聚类算法(self-weighted multiview clustering, SwMC),利用 Laplacian 秩约束项,近似地作为不同视图的图心.该算法的优势在于一旦获得目标邻接图,可以直接获得数据划分结果,而不需要任何传统算法(如  $K$ -means)进行后处理<sup>[24]</sup>.Zhang 等人关注二进制多视图数据(如图像数据),提出了一种二进制多视图聚类框架,其联合优化紧实的协同离散表示学习过程与二元聚类结构学习过程<sup>[25]</sup>.尽管传统的多视图谱聚类算法已经取得了较多成果和进展,但由于传统算法的限制性,大多现存的多视图谱聚类算法无法处理大规模复杂场景下的数据,且无法挖掘数据分布中的深层隐含特征.因此,本文考虑结合深度学习的非线性特征处理能力,通过构建神经网络实现多视图谱聚类,进一步提升多视图谱聚类算法的泛化能力与聚类性能.

## 1.2 深度聚类

现存的聚类算法虽然已经实现了较好的性能,但面向复杂数据分布场景时的数据划分能力还亟待提高.为克服现有工作的缺陷,结合深度学习的非线性特征处理能力来提升聚类算法的特征挖掘性能不失为一个好的解决方案<sup>[26,27]</sup>.Yang 等人设计了一种迭代学习数据表达和类簇划分的深度层次聚类模型(joint unsupervised learning of deep representations and image clusters, JULE),有效抓住了精细的图像类簇信息和可判别的特征表达<sup>[28]</sup>.类似地,Caron 等人提出了一种深度聚类框架(DeepCluster),首先采用  $K$ -means 聚集特征,然后根据类簇划分更新网络参数,迭代优化获得最佳聚类性能<sup>[29]</sup>.Shaham 等人引入深度学习方法来谱聚类算法中,提出了一种深度谱网络(spectral clustering using deep neural networks, SpectralNet)<sup>[7]</sup>,其旨在通过约束随机优化近似谱聚类算法,从而学习一个具有更多隐含知识的谱嵌入.

面向多视图学习场景的深度聚类算法也被研究者们争相探讨.在深度多视图聚类学习中,首先备受关注的是深度自动编码器,其旨在挖掘一个可重建出多视图输入的统一表达.详细来说,深度自动编码器采用共享编码器提取所有视图的统一特征表示,然后使用不同的解码器单独重建每一个视图.Andrew 等人提出了一种基于典型相关分析(canonical correlation analysis, CCA)的深度神经网络版本(deep canonical correlation analysis, DCCA)来学习两个视图的共享表达<sup>[30]</sup>.DCCA 利用两个子网络分别提取两个视图的非线性特征,同时最大化学到的特征的相关性以提高模型性能.进一步地,Wang 等人同时优化了所提取的多视图特征之间的线性相关分析以及自动编码器的重建误差,从而提出了一种深度线性相关自动编码器(deep canonically correlated auto encoders, DCCAE)<sup>[31]</sup>.Zhao 等人利用深度矩阵分解技术学习多视图样本的深层特征,提出了一种多视图深度矩阵分解方法(multi-view clustering via deep matrix factorization, MvDMF)<sup>[32]</sup>.Huang 等人拓展了谱聚类算法到深度学习版本,提出了一种多视图谱聚类网络(multi-view spectral clustering network, MvSCN)<sup>[6]</sup>.由上述可知:深度多视图聚类学习方面的研究尚为较少,还需深入探索深度神经网络与传统聚类算法之间的关联性,进一步提升多视图聚类算法在时间与性能上的效率.

## 2 多视图谱聚类算法

本节主要介绍传统的多视图谱聚类算法的基本原理.正如前面提到的,多视图谱聚类旨在融合多个邻接图以获得一致性指示矩阵,从而得到数据划分结果.假设  $X=\{X^1, X^2, \dots, X^V\}$  是一个包含  $V$  个视图的数据集,  $X^v \in \mathbb{R}^{N \times d_v}$  ( $1 \leq v \leq V$ ) 表示第  $v$  个视图的数据,其包含  $N$  个维度为  $d_v$  的样本.和谱聚类算法类似,多视图谱聚类算法首先针对每个视图构建邻接矩阵,使得视图中的每一个样本与它的  $k$  个最近邻居相连.通常来说,我们采用基于高斯核函数的欧氏距离来计算邻接矩阵,其数学表达式如下:

$$S_{ij}^v = \begin{cases} \exp\left(-\frac{\|x_i^v - x_j^v\|_2^2}{2\delta^2}\right), & x_i^v, x_j^v \text{互为邻居} \\ 0, & \text{其他} \end{cases} \quad (1)$$

其中,  $S_{ij}^v \in S^v$  表示第  $v$  个视图中第  $i$  和第  $j$  个样本的距离,  $\delta$  是控制高斯核函数作用范围的带宽.

计算得到每个视图的邻接矩阵后, 多视图谱聚类的目标函数可定义如下:

$$\left. \begin{aligned} \arg \min_{Y, a^v} \sum_{v=1}^V (a^v)^r \operatorname{tr}(Y^T L^v Y) \\ \text{s.t. } Y^T Y = I, \sum_{v=1}^V a^v = 1, a^v > 0 \end{aligned} \right\} \quad (2)$$

其中,  $Y \in \mathbb{R}^{N \times c}$  是聚类指示矩阵, 也可视为最终的谱嵌入,  $c$  是类簇数目;  $L^v$  是第  $v$  个视图的拉普拉斯矩阵, 定义为  $L^v = D^v - S^v$ ;  $D^v = \operatorname{diag}(S^v)$  是对角矩阵;  $a^v$  是反映每个视图贡献度的参数;  $r$  是控制每个视图不同权重分布的尺度. 详细细节见文献[16].

### 3 知识迁移下的深度一致性多视图谱聚类网络

本节将详细介绍提出的知识迁移下的深度一致性多视图谱聚类网络, 整体网络框架包含两部分(如图 1 所示): (1) 基于对比损失的近邻学习; (2) 基于知识迁移的谱嵌入学习.

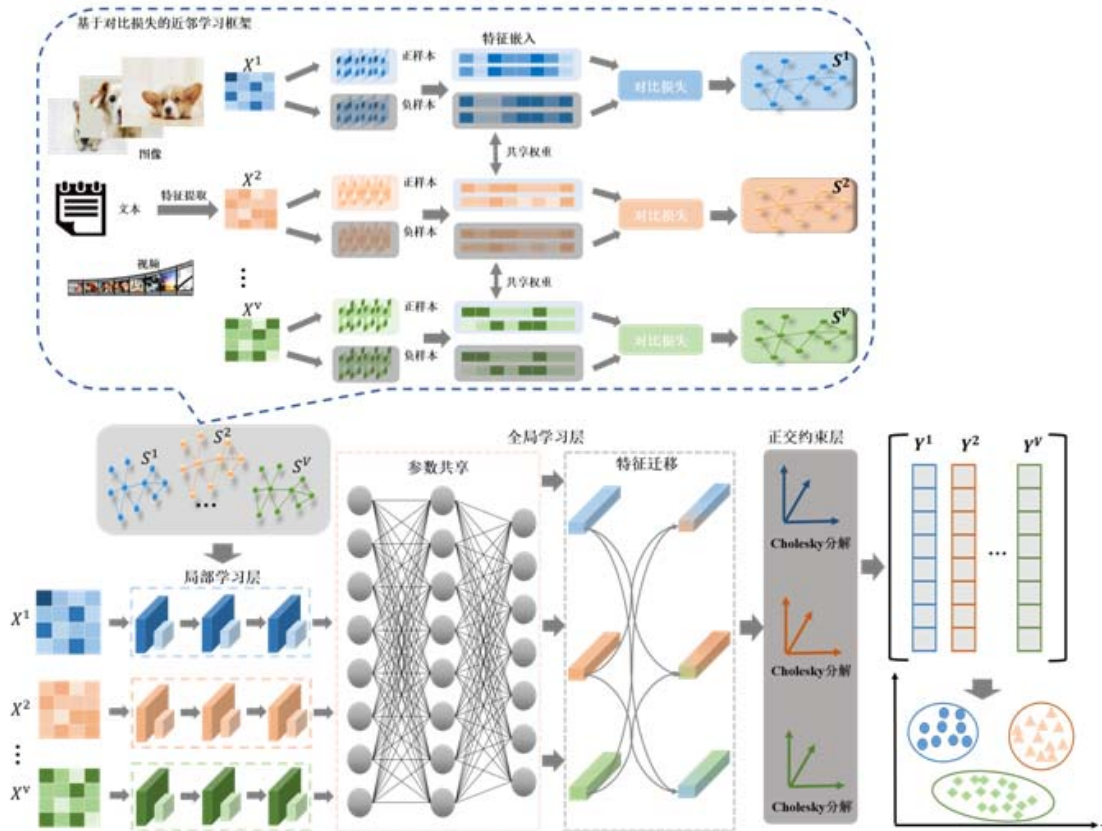


图 1 知识迁移下的深度一致性多视图谱聚类框架图

由于谱聚类算法需要引入样本对间的近邻关联, 在近邻学习部分, 不同于传统谱聚类中采用的距离度量,

CMvSC 设计了一个具有对比损失的孪生网络用于同时判别正负样本对, 自适应地学习包含样本隐含关联的近邻矩阵. 在谱嵌入学习部分, CMvSC 设计了一个深度学习框架用于多视图数据处理, 首先对每个视图设计网络分支, 将单视图数据输入到局部学习层以获得单视图特有嵌入; 其次考虑全局一致性, 结合知识迁移设计全局学习层, 其中分别包括参数共享与特征迁移, 通过动态融合获得一致性多视图谱嵌入; 最后, 采用 K-Means 算法得到最终的多视图数据划分结果.

### 3.1 基于对比损失的近邻学习

邻接矩阵对谱聚类算法的性能影响极大, 因此选择一个较好的近邻度量标准能够大大提升聚类效果. 尽管基于距离度量的近邻学习在一定程度上有效定义了样本关系, 但是其在应对数据分布较复杂的场景时效果不佳. 在本文中, CMvSC 引入了一种基于对比损失的近邻学习方式, 采用神经网络挖掘样本的非线性特征, 自适应地为每个视图学习邻接矩阵.

与直接采用样本邻居集或距离度量构造邻接矩阵的方法不同, 基于对比损失的近邻学习以孪生神经网络为框架, 引入对比学习原理, 同时考虑了正样本与负样本的影响, 使得最终输出的邻接矩阵既捕获了正样本之间的相似性, 也学习了负样本之间的差异性. 所提出的近邻学习的网络损失可表示为

$$L_a^v = P \|g_\theta(x_i^v) - g_\theta(x_j^v)\|_2^2 + (1-P) \max(\gamma - \|g_\theta(x_i^v) - g_\theta(x_j^v)\|_2, 0) \quad (3)$$

其中,

- $P \in \{0, 1\}$  表示判别标签, 即:  $P=1$  时表示  $(x_i^v, x_j^v)$  互为正样本; 反之,  $P=0$  时指  $(x_i^v, x_j^v)$  互为负样本;
- $g_\theta(\cdot)$  表示孪生网络中的映射函数, 用于嵌入样本数据到潜在空间中. 在我们的工作中, 其是指全连接层和 ReLu 函数组成的嵌入单元;
- $\gamma$  是指距离边界(通常定义为 1).

通过分析公式(3)中的两项, 不难看出, 其很好表达了成对样本的匹配程度. 当样本相似时(即  $P=1$ ), 损失只保留第 1 项, 即原本相似的样本经过映射后仍然相似; 当样本不相似时, 损失只剩第 2 项, 即样本不相似的样本经过映射后仍然不相似.

如何构建正负样本集合, 是基于对比损失的近邻学习的关键. 对于监督学习来说, 我们通常直接采用样本的类别标签来判断样本是否互为正/负样本: 若标签相同/不同, 则为正/负样本. 然而, 对于无监督学习来说, 我们无法得知样本标签, 这种情况下无法直接判断样本是否互为正/负样本. 为解决这个问题, 我们参考基于样本邻居集构造邻接矩阵的方法, 仍然采用最近  $k$  近邻法( $k$ -nearest neighbor,  $k$ NN)构建每个样本的邻居集, 然后根据邻居集划分正/负样本. 换言之, 当  $x_j^v$  属于  $x_i^v$  的邻居集时, 我们认为  $(x_i^v, x_j^v)$  互为正样本; 反之,  $(x_i^v, x_j^v)$  互为负样本. 为保证正负样本的比例均衡, 我们选择每个样本邻居集中的  $k$  个邻居作为正样本; 对于负样本, 我们则在样本邻居集外随机选择  $k$  个样本作为负样本.

构造好所有样本对应的正负样本集合后, 我们结合对比损失来构造具有相同网络分支的孪生网络进行近邻学习, 将样本数据进行映射, 最小化对比损失(如公式(3)所示), 并采用高斯核函数来计算最终的邻接矩阵, 如下:

$$S_{ij}^v = \begin{cases} \exp\left(-\frac{\|g_\theta(x_i^v) - g_\theta(x_j^v)\|_2^2}{2\delta^2}\right), & x_i^v, x_j^v \text{ 互为邻居} \\ 0, & \text{其他} \end{cases} \quad (4)$$

其中,  $\delta$  是高斯核函数中的宽度参数, 通常取值大于 0. 在本文工作中, 我们设置  $\delta$  为所有样本点与它们邻居集中最远邻居之间距离的中位数. 高斯核函数将数据映射到高维特征空间, 再计算样本间距离. 因上述孪生网络旨在最小化正样本间的距离而最大化负样本间的距离, 可能造成映射后的正样本对线性不可分, 高斯核函数的引入, 有利于避免线性不可分的问题. 最后, 将计算得到的邻接矩阵  $[S^1, S^2, \dots, S^v]$  用于基于知识迁移的谱嵌入学习(见第 3.2 节)中, 有效提升最终的特征表达效果, 从而提升聚类性能.

### 3.2 基于知识迁移的谱嵌入学习

尽管传统的多视图谱聚类算法已取得较好性能,但其仍存在一些限制:(1)谱嵌入的可扩展性有限,无法应对大规模数据的场景,这也是大多传统浅层模型的问题之一;(2)泛化能力不足,无法适应复杂数据分布,隐含特征学习能力有限.因此,本文设计了一个基于深度神经网络的参数模型,挖掘多视图数据的潜在分布结构,学习单视图的局部不变性和多视图的全局一致性,结合知识迁移将原始多视图数据映射为一致性谱嵌入,其数学表达式为

$$f_{\theta}(X^{N \times d_v}) \rightarrow Y^{N \times c} \quad (5)$$

其中,  $f_{\theta}(\cdot)$  表示 CMvSC 的网络参数模型;  $X^{N \times d_v}$  表示包含  $N$  个样本的原始第  $v$  个视图数据,且数据维度为  $d_v$ ;  $Y^{N \times c} = [Y^1, Y^2, \dots, Y^V]$  是指经模型映射后的一致性谱嵌入,每个视图的特征维度均为  $c$  (通常地,我们设定  $c$  等于类簇数目).

如何挖掘多视图数据的特有特征和共享特征以获得蕴含丰富信息的一致性谱嵌入,是 CMvSC 的一个关键性任务.考虑到每个视图的数据特征结构不尽相同,但每个独立视图存在局部不变性,即同一视图中的不同数据样本满足相同数据分布,因此,我们首先设计局部学习层分别学习每个视图的特有嵌入;其次,由于不同视图都是相同数据对象的不同特征表示,因此多视图在潜在分布上存在全局一致性,同时,不同视图之间存在重叠、互补的特征,因而我们提出一种全局学习层,用于同时挖掘多个视图中的隐含共享知识,该层先采用参数共享学习多视图间的潜在分布一致性,再利用特征迁移动态融合不同视图间的互补特征,以输出一致性多视图谱嵌入.

此外,传统谱聚类算法中引入了正交约束  $Y^T Y = I$ ,因此 CMvSC 设计一种正交约束层来对网络输出进行限制,保证多视图谱嵌入矩阵的正交性.最后得到经网络映射后的多视图一致性特征表示  $Y^{N \times c}$ ,再采用传统的聚类算法(如  $K$ -means)获得最终的多视图聚类结果.以下我们将详细介绍 CMvSC 中的局部学习层和全局学习层,以及网络中的其他关键部分.

#### 3.2.1 局部学习层

不同视图之间往往存在尺度差异性和特征差异性,但每个视图内部的所有数据样本却具有分布一致性,我们称其具有局部不变性.因此,我们首先考虑独立学习每个视图的特有嵌入,以挖掘单视图特有特征.本节提出了一种局部学习层,采用全连接层(dense)以及 ReLu 函数学习每个视图的特有嵌入,以获得新的单视图特征表示.

新的单视图隐含特征表示需要保证与原始数据结构具有一定一致性,即:原始空间中距离较近的样本经过特征映射后,在新的特征表示空间中距离仍然较近.因此,所设计的局部学习层需要尽可能地挖掘原始数据样本的有效特征,去除冗余特征,同时不丢失原始数据的隐含结构.与传统的特征流型学习不同,全连接层利用每一个网络节点对原始单视图数据进行特征嵌入,能够更充分地抓住原始数据结构与潜在分布特征,因此全连接层可以被视为分类器用于融合不同维度的特征;同时,当数据样本为图像等形式时,可以采用卷积层等预处理方式将原始数据(维度为  $w \times h$ )转换为  $d$  维数据,再使用全连接层;此外,当前一层网络层为全连接层时,此时使用全连接层等价于卷积核为  $1 \times 1$  的一维卷积层.最后,我们结合 ReLu 函数引入非线性因素,将每个网络节点的映射进行非线性输出,以获得具有更佳特征表示能力的表达.

经过局部学习层后,我们获得新的单视图特有嵌入.综上可知,局部学习层的目标是保证每个视图中相似度较高的数据样本在特征嵌入后仍然相似,因此局部学习层的损失可表示为

$$L_1^v = \frac{1}{N^2} \sum_{i,j} S_{ij}^v \|h_{i,i}^v - h_{i,j}^v\|_2^2 \quad (6)$$

其中,  $h_{i,i}^v \in H_i^v$  表示局部学习层输出的第  $v$  个视图中第  $i$  个样本的隐含特征表示;  $S_{ij}^v$  是第  $v$  个视图的邻接图矩阵,在本篇工作中,采用第 3.1 节中基于对比损失的近邻学习框架获得每个视图的邻接图矩阵;  $N$  是样本数目.该损失函数的设计意义在于:如果邻接图矩阵中第  $i$  个样本节点和第  $j$  个样本节点相连且具有较高的相似度,那么第  $i$  个样本和第  $j$  个样本的隐含特征表示应该仍然相似,即具有较小的距离.

### 3.2.2 全局学习层

多视图数据是对相同数据样本的不同特征表示, 不同视图具有相同的簇标签, 因此可以认为多视图间往往存在一定的分布一致性, 即不同视图的数据划分应该是相同的, 我们称其为全局一致性. 同时, 不同视图的特征空间通常具有重叠性和互补性, 因此, 有效融合多视图特征能够在一定程度上强化特征的表达能力, 从而提升最终的聚类性能. 本节提出了一种全局学习层挖掘多视图一致性特征, 采用参数共享与特征迁移分别实现共享知识挖掘与特征动态融合, 从而获得融合多个视图信息的一致性谱嵌入.

全局学习层与局部学习层类似, 同样采用全连接层(dense)对数据表达进行非线性变换, 但为了得到具有分布一致性与特征互补性的一致性多视图谱嵌入, 其借鉴了迁移学习技术, 首先采用参数共享的网络层挖掘所有视图的一致性特征, 然后设计视图特有的网络层从一致性特征中进行知识迁移, 有效融合其他视图信息与当前视图特有信息. 两种学习方式如下.

- 参数共享

神经网络中的参数空间往往是决定最终特征表达的关键因素, 换言之, 每个视图的参数空间隐含了数据样本的分布特征与潜在结构. 考虑到多视图间存在分布一致性, 本节采用参数共享的全连接层对多个视图进行特征嵌入, 即级联所有视图的局部学习层输出, 再采用全连接层进行学习, 从而挖掘出多个视图共享的分布特征, 其数学表达式如下:

$$H_p = \Theta(\text{concat}[H_1^1; \dots; H_1^v; \dots; H_1^V]) \quad (7)$$

其中,  $\Theta$ 表示一致性网络层的参数空间,  $\text{concat}$ 表示级联操作,  $H_1^v$ 表示第  $v$  个视图的局部学习层输出.

- 特征迁移

众所周知, 多视图数据是同一样本集的多源/多通道特征表示. 因此, 不同视图之间的特征可能存在重叠、互补等特性. 本节考虑多视图特征表示之间的互补原则, 融合其他视图信息能够有效增强当前视图的特征表达能力. 由此, 我们采用视图特有的权重对包含不同视图特征的一致性特征表达进行映射, 通过非线性映射对不同视图信息进行迁移, 并与当前视图的局部信息进行融合, 以实现不同视图间的特征迁移, 形成新的视图表示, 其数学表达式如下:

$$Y^v = W_p^v \circ H_p + H_1^v \quad (8)$$

其中,  $\circ$ 表示 Hadamard 积, 即逐元素乘积;  $W_p^v$ 是可学习参数, 用于自适应地调整每个视图的贡献. 针对每个视图, 我们都学习了一个最终的指示矩阵, 其主要原因在于: 多视图不仅包含共享知识, 每个视图也具有特有信息, CMvSC 旨在使得最终的谱嵌入既保留共享特征, 又尽可能保持特有性质. 因此, 通过全局学习层的特征变换后, 我们为每一个视图独立学习一个指示矩阵, 在一定程度上能够提升聚类的准确性.

全局学习层旨在最小化不同视图之间的差异性, 以获得多视图共享的特征嵌入. 因此, 多视图全局学习层的目标函数可表示为

$$L_2 = \frac{1}{NV^2} \sum_{v,v'} \sum_i^N \|y_i^v - y_i^{v'}\|_2^2 \quad (9)$$

其中,  $v$  和  $v'$  分别指代不同视图.

### 3.2.3 正交约束层

传统谱聚类算法中为避免出现“无效解”(trivial solution), 通常会引入正交约束. “无效解”是指为了最小化谱聚类的目标损失函数, 模型将所有数据输入在潜在空间中映射为同一表示, 以使得目标损失函数值最小. 但这样的结果对于聚类任务来说是无效的, 我们希望得到包含足够丰富信息的特征嵌入, 才能获得较好的数据划分结果. 在 CMvSC 中, 同样为了避免出现“无效解”, 我们设计了正交约束层用于引入正交约束, 使得:

$$(Y^v)^T Y^v = I_{N \times N} \quad (10)$$

其中,  $Y^v$ 表示模型输出的第  $v$  个视图的谱嵌入,  $I$ 是单位矩阵.

正交约束层的目的与传统谱聚类算法中的正交约束作用一致, 正交化每一列特征嵌入, 以保证最终输出



结果的有效性. 在本文中, 我们采用 Cholesky 分解实现正交约束, 并将正交约束层搭建在网络最顶层, 严格保证最终输出的有效性.

### 3.2.4 目标函数

综上所述, CMvSC 主要考虑同时挖掘单视图的局部特有信息与多视图间的全局共享知识, 因此, 网络的目标函数可分为局部学习损失与全局学习损失两部分, 如下所示:

$$L_c = (1 - \lambda) \sum_{v=1}^V L_1^v + \lambda L_2 \quad (11)$$

其中,  $L_1$  表示局部学习层的损失,  $L_2$  表示全局学习层的损失,  $\lambda$  是调整两种损失的平衡参数.  $L_1$  主要用于控制单个视图内的损失, 最小化类内差异, 同时最大化类间差异, 使得原始空间中较相似的样本的隐含表示仍然相似, 保证局部不变性.  $L_2$  的目的是控制不同视图间的学习损失, 使得不同视图间的对应样本的隐含表示尽可能一致, 保证多视图全局一致性.  $\lambda$  用于平衡局部学习层与全局学习层的损失在目标函数中的重要性, 当  $\lambda$  为 0 时, 表示只考虑单视图内的损失;  $\lambda$  取 1 时, 网络只考虑多视图间的损失.

## 4 网络训练与优化

### 4.1 网络训练

基于知识迁移的谱嵌入学习和基于对比损失的近邻学习是两个独立过程, 因此其训练过程也是相互独立的. 另外, 由于所提出的 CMvSC 适用于深度聚类任务, 谱嵌入学习与近邻学习的网络训练过程均为无监督, 不需要构建训练集与测试集, 可直接输入样本数据获得最终数据划分结果. 值得注意的是: 谱嵌入学习网络的输出维度为  $c$  (类簇数目), 以获得聚类指示矩阵.

### 4.2 基于代码空间的训练

就多视图数据而言, 通常每个视图具有不同的特征分布. 此时, 如果直接进行多视图间的特征提取和知识迁移, 可能会因为度量差异和特征结构差异的影响得到冗余信息, 无法获得有效的一致性谱嵌入. 因此, CMvSC 采用基于代码空间 (code space) 的训练方法. 基于代码空间的表达是指原始数据经过预训练好的模型转换后的低维数据表达, 其通常包含更少的冗余信息和更丰富的有效特征. 部分学者直接采用自动编码器用于网络初始化<sup>[33-35]</sup>, 但是这样的网络初始化往往会在训练过程中改变代码空间的低维数据表达. 因此, 在本篇工作中, 我们借鉴 SpectralNet<sup>[7]</sup>和 MvSCN<sup>[6]</sup>中直接采用基于代码空间的表达作为输入数据的训练方式, 然后应用 CMvSC 进行处理. 在实验部分, 我们采用文献[34]中的深度嵌入聚类模型 (deep embedded clustering, DEC) 对原始数据进行处理, 获得的低维表达作为基于代码空间的表达用于 CMvSC 的输入.

## 5 实验

本节将评估 CMvSC 与 11 个对比算法在 4 个数据集上性能. 此外, 我们将给出网络详细结构设计与实验过程中的详细参数设置.

### 5.1 数据集

实验部分, 我们选择了 4 个多视图数据集用于性能评估, 包括:

- Noisy MNIST: 我们采用原始数据集作为视图 1, 然后随机选择类内样本添加噪声来构造视图 2. 由此, 我们得到一个包含 70 000 个样本的双视图数据集 (下载链接: <http://ttic.uchicago.edu/~wwang5/dccae.html>, createMNIST.m);
- Caltech101-20: 我们提取 6 种手工特征作为 6 个视图, 包括 Gabor 特征、Wavelet Moments 特征、CENTRIST 特征、HOG 特征、GIST 特征和 LBP 特征, 其包含了 20 个类别, 共 2 386 张图片 (下载链接: [http://www.vision.caltech.edu/Image\\_Datasets/Caltech101/](http://www.vision.caltech.edu/Image_Datasets/Caltech101/));

- Reuters: 我们使用 Reuters 数据库中的英语版本和 4 种翻译版本(包括法语、德语、西班牙语和意大利语)来构造多视图数据集, 其包含了 6 个类别, 共 18 758 个样本(下载链接: <https://archive.ics.uci.edu/ml/datasets.html>);
- NUSWIDEOBJ: 我们选择 NUS 提供的 5 种特征(包括 Color Histogram, Color Moments, Color Correlation, Edge Distribution 和 Wavelet texture)构建多视图数据集, 其包含了 31 个类别, 共 30 000 张图片(下载链接: <http://lms.comp.nus.edu.sg/research/NUS-WIDE.htm>).

大多数对比算法是基于传统机器学习方法, 面对大规模数据场景往往会有内存溢出、矩阵分解失败等问题, 因此, 对于 Noisy MNIST, Reuters 和 NUSWIDEOBJ 数据集, 我们随机选择 10k 数据进行实验. 所有实验数据集的详细信息见表 1.

表 1 数据集详细信息

数据集	视图数	样本数	特征数	类簇数
Noisy MNIST	2	10 000	784/784	10
Caltech101-20	6	2 386	48/40/254/1984/512/928/2386	20
Reuters	5	10 000	21531/24892/34251/15506/11547	6
NUSWIDEOBJ	5	10 000	65/226/145/74/129	31

## 5.2 对比算法

本文工作选择了 11 个算法与我们的 CMvSC 进行比较, 其中包含 3 个单视图聚类算法和 8 个多视图聚类算法.

- 单视图聚类算法: (1) Spectral clustering(SC)<sup>[36]</sup>; (2) Low rank representation(LRR)<sup>[12]</sup>; (3) Spectral clustering using deep neural networks (SpectralNet)<sup>[7]</sup>;
- 多视图聚类算法: (1) Deep canonical correlation analysis (DCCA)<sup>[30]</sup>; (2) Deep canonically correlated autoencoders (DCCA)<sup>[31]</sup>; (3) Diversity-induced multi-view subspace clustering (DiMSC)<sup>[15]</sup>; (4) Latent multi-view subspace clustering (LMSC)<sup>[16]</sup>; (5) Multi-view clustering via deep matrix factorization (MvDMF)<sup>[32]</sup>; (6) Self-weighted multiview clustering with multiple graphs (SwMC)<sup>[24]</sup>; (7) Binary multi-view clustering (BMVC)<sup>[25]</sup>; (8) Multi-view spectral clustering network (MvSCN)<sup>[6]</sup>.

为采用多视图数据集评估单视图聚类算法, 我们均分别在每个视图上进行实验, 然后选择性能最佳的视图结果作为最终评估结果. 此外, DCCA 和 DCCA 只适用于两个视图的数据集, 因此我们选择性能最佳的两个视图作为最终的评估结果.

对于所有的对比算法和 CMvSC, 我们均使用基于代码空间的低维表达作为数据输入. 所有实验在标准 Ubuntu-16.04 操作系统中实现, 利用 Keras 和 TensorFlow 进行模型搭建, 采用 1 块 NVIDIA 1080Ti GPU 加速. 值得一提的是: CMvSC 采用 K-means 算法对网络输出的特征表达进行数据划分, 为尽可能降低 K-means 算法因随机初始化对聚类结果产生较大影响, 在实验过程中, 我们对网络输出进行 10 次 K-means 聚类, 且每次选择不同的初始化种子(seed), 最终选择最好的结果作为数据划分结果.

## 5.3 性能评价指标

本篇工作采用准确率(ACC)、F-measure(*F-meas*)和标准化互信息值(NMI)这 3 个评价指标评估所提出的 CMvSC 与所有对比算法的性能, 以上 3 种指标均为值越大, 表明聚类效果越好.

ACC 表示聚类结果的准确率, 正确的结果越多, ACC 值越大, 其计算方法如下:

$$ACC = \frac{\sum_{i=1}^N \delta(\text{map}(\tilde{y}_i), y_i)}{N} \quad (12)$$

其中,  $N$  是样本总数,  $\tilde{y}_i$  和  $y_i$  分别表示第  $i$  个样本的网络输出的类簇标签与真实标签.  $\text{map}(\tilde{y}_i)$  用于将网络分配类簇标签映射到等价的真实标签.  $\delta(\cdot, \cdot)$  表示映射后的类簇标签是否与真实标签相同: 若相同, 则  $\delta(\cdot, \cdot)=1$ ; 反之为 0.

$F$ - $mea$  是通过计算查全率和查准率来评估聚类性能的指标, 查全率( $R$ )与查准率( $P$ )的定义如下:

$$R = \frac{TP}{TP + FN} \tag{13}$$

$$P = \frac{TP}{TP + FP} \tag{14}$$

其中,  $TP$  表示真实标签为同类的样本被聚为同一类簇,  $FN$  是真实标签为同类的样本被聚为不同类簇,  $FP$  表示真实标签为不同类的样本被聚为同一类簇.

$F$ - $mea$  综合考虑查全率与查准率, 其定义如下:

$$F\text{-}mea = \frac{2 \times P \times R}{P + R} \tag{15}$$

NMI 是一种利用信息论中的信息熵来计算聚类结果与真实标签间的相似程度的指标, 其计算公式如下:

$$NMI = 2 \times \frac{I(\tilde{Y}, Y)}{H(\tilde{Y}) \times H(Y)} \tag{16}$$

其中,  $\tilde{Y}$  和  $Y$  分别表示网络输出的聚类结果和真实标签,  $I(\tilde{Y}, Y)$  表示聚类结果与真实标签间的互信息,  $H(\cdot)$  表示熵. 互信息与熵的计算公式如下:

$$I(\tilde{Y}, Y) = \sum_{i=1}^c \sum_{j=1}^c \frac{m_{ij}}{N} \log \left( \frac{m_{ij}}{|\tilde{Y}_i| \times |Y_j|} \right) \tag{17}$$

$$H(Y) = - \sum_{i=1}^c \frac{|Y_i|}{N} \log \left( \frac{|Y_i|}{N} \right) \tag{18}$$

其中,  $m_{ij}$  表示网络输出为第  $i$  类而真实标签为第  $j$  类的样本数,  $N$  是样本总数,  $c$  是类簇数.

### 5.4 参数设置及网络结构设计

CMvSC 中的关键参数主要包括近邻学习中的邻居数  $k$  和谱嵌入学习中控制损失函数的平衡参数  $\lambda$ . 因此, 本节主要讨论以上两个参数的影响. 如图 2 所示, 我们在 4 个数据集上探讨了邻居数  $k$  的影响.

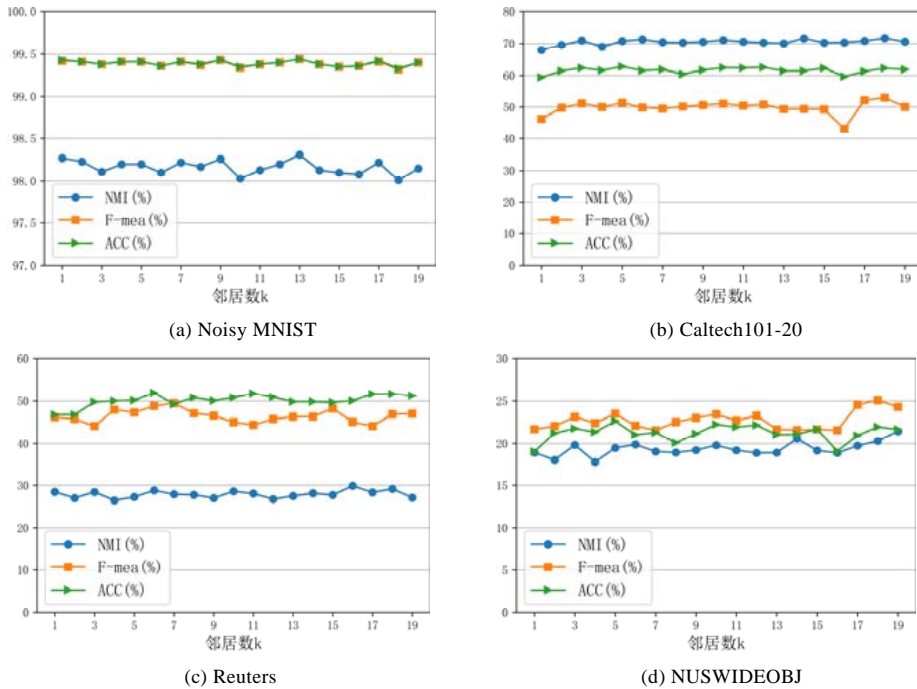


图 2 参数  $k$  的影响比较

不难看出, CMvSC 对于邻居数  $k$  的设置并不算敏感, 整体性能相对稳定. 在实际实验中, 相较于 Noisy MNIST, 其他 3 个数据集的邻居数  $k$  取值范围更大. 原因在于: 邻居数越少, 邻接矩阵会更稀疏. 而其他 3 个数据集的视图数较多, 当数据经过全局学习层处理后, 所获得的特征嵌入可能包含较多极小值, 如果此时邻接矩阵同样太稀疏, 后续的正交约束层将无法完成正交分解.

如图 3 所示, 我们在 4 个数据集上讨论了不同平衡参数  $\lambda$  下的性能变化, 可以得出, 不同数据集在参数  $\lambda$  大于 0.0005 时性能均趋于稳定.

此外, 我们给出了 4 个数据集的谱嵌入学习的训练损失变化曲线图, 如图 4 所示. 可以看出, 所有数据集均在前 20 次 Epochs 中迅速收敛.

最后, 为保证实验结果的可复现性, 我们给出了所有必要参数的设置以及网络结构的设置, 分别见表 2 和表 3.

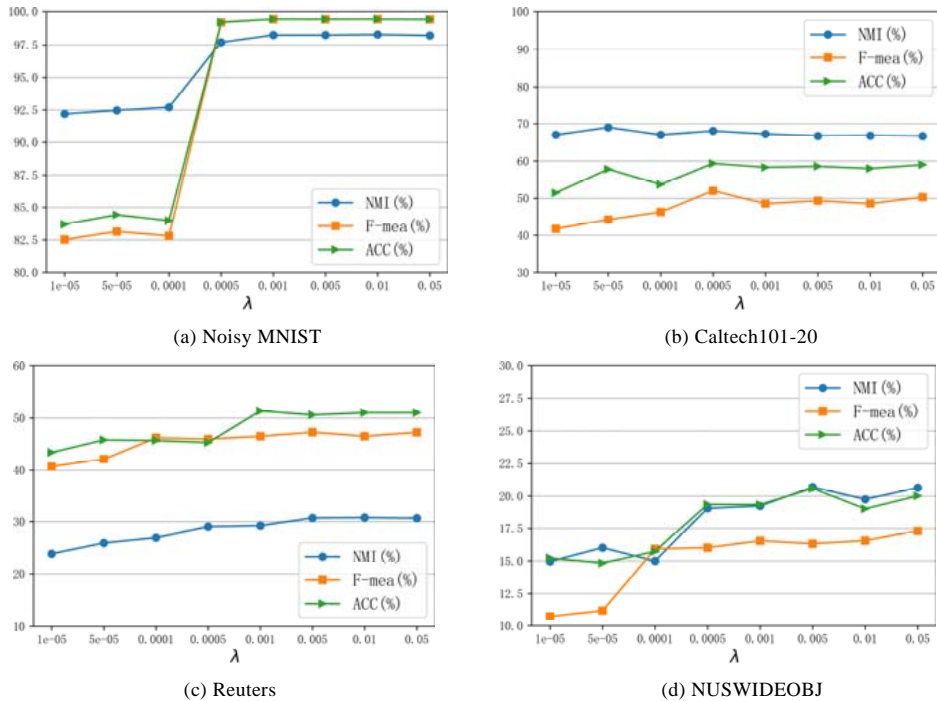


图 3 参数  $\lambda$  的影响比较

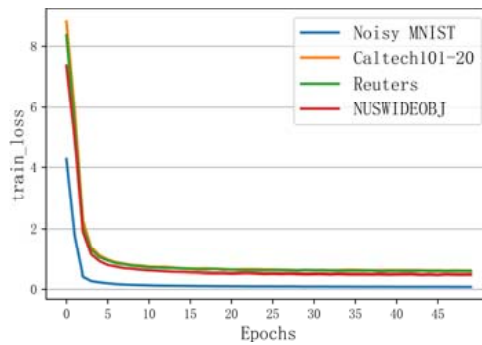


图 4 4 个数据集上的训练损失变化曲线

表 2 参数设置详细信息

数据集	子网络	邻居数 $k$	平衡参数 $\lambda$	批次(batch)	正交批次(orthogonal batch)	学习率
Noisy MNIST	近邻学习 谱嵌入学习	13	0.000 5	128	-	0.001
				512	512	0.001
Caltech101-20	近邻学习 谱嵌入学习	15	0.000 5	128	-	0.001
				512	512	0.001
Reuters	近邻学习 谱嵌入学习	20	0.000 5	128	-	0.001
				512	512	0.001
NUSWIDEOBJ	近邻学习 谱嵌入学习	20	0.000 5	128	-	0.001
				512	512	0.001

表 3 网络结构详细信息

数据集	近邻学习	谱嵌入学习	
Noisy MNIST	Dense (ReLU, $size=1024$ )	局部学习层	Dense (ReLU, $size=1024$ )
	Dense (ReLU, $size=1024$ )		Dense (ReLU, $size=1024$ )
	Dense (ReLU, $size=1024$ )		Dense (ReLU, $size=1024$ )
	Dense (tanh, $size=10$ )	全局学习层	Dense (ReLU, $size=1024$ )
		正交约束层	Dense (ReLU, $size=10$ )
Caltech101-20	Dense (ReLU, $size=2048$ )	局部学习层	Dense (ReLU, $size=2048$ )
	Dense (ReLU, $size=1024$ )		Dense (ReLU, $size=1024$ )
	Dense (ReLU, $size=512$ )		Dense (ReLU, $size=512$ )
	Dense (tanh, $size=10$ )	全局学习层	Dense (ReLU, $size=512$ )
		正交约束层	Dense (ReLU, $size=20$ )
Reuters	Dense (ReLU, $size=512$ )	局部学习层	Dense (ReLU, $size=512$ )
	Dense (ReLU, $size=256$ )		Dense (ReLU, $size=256$ )
	Dense (tanh, $size=128$ )		全局学习层
			正交约束层
NUSWIDEOBJ	Dense (ReLU, $size=2048$ )	局部学习层	Dense (ReLU, $size=2048$ )
	Dense (ReLU, $size=1024$ )		Dense (ReLU, $size=1024$ )
	Dense (ReLU, $size=512$ )		Dense (ReLU, $size=512$ )
	Dense (tanh, $size=100$ )	全局学习层	Dense (ReLU, $size=1024$ )
		正交约束层	Dense (ReLU, $size=31$ )

## 5.5 性能评估

我们评估了 CMvSC 和 11 个对比算法的性能,性能比较结果见表 4, 最优结果已加粗表示。

表 4 对比算法与 CMvSC 的性能对比(ACC,  $F$ -mea, NMI)

数据集 对比算法\评估指标	Noisy MNIST			Caltech101-20			Reuters			NUSWIDEOBJ		
	ACC	$F$ -mea	NMI	ACC	$F$ -mea	NMI	ACC	$F$ -mea	NMI	ACC	$F$ -mea	NMI
SC	66.26	66.42	61.36	42.50	34.15	62.41	45.94	38.17	22.26	15.32	10.33	15.58
LRR	56.96	55.06	65.84	39.15	29.83	59.53	41.52	27.26	26.37	13.94	10.73	14.16
SpectralNet	84.68	82.21	90.14	51.05	36.91	64.55	46.64	29.45	24.66	15.38	11.52	15.19
DCCA	95.50	95.46	89.47	42.83	37.60	62.03	29.40	25.54	6.73	16.00	8.83	11.34
DCCAE	94.92	94.87	88.45	44.76	38.87	61.19	30.28	25.21	8.87	14.76	8.55	11.65
DiMSC	47.24	50.25	34.84	21.46	16.59	24.70	40.50	37.38	13.51	9.28	7.49	7.53
LMSC	66.88	66.79	61.94	38.14	30.06	57.02	40.06	33.20	28.89	15.40	12.14	16.30
MvDMF	75.26	75.00	67.12	35.96	26.23	47.25	45.78	24.93	24.69	12.04	7.49	7.53
SwMC	98.98	98.96	97.14	49.87	35.53	62.32	32.84	20.59	23.00	13.84	4.53	9.58
BMVC	90.40	85.98	93.47	36.55	25.70	56.19	46.96	34.82	22.10	14.12	9.95	12.57
MvSCN	99.16	98.96	97.66	57.94	43.15	67.92	48.86	43.35	25.85	16.86	11.88	15.97
CMvSC	<b>99.42</b>	<b>99.44</b>	<b>98.15</b>	<b>61.26</b>	<b>52.65</b>	<b>70.15</b>	<b>49.98</b>	<b>45.66</b>	<b>28.92</b>	<b>18.37</b>	<b>14.65</b>	<b>18.88</b>

从表 4 可以看出:

- 在所有数据集上,多视图聚类算法整体优于单视图聚类算法,说明了相较于单视图,多视图具有更强的特征表示能力,蕴含更丰富的有效信息;
- 极个别多视图聚类算法在某些数据集上性能弱于单视图聚类算法,如: DCCA 和 DCCA<sub>E</sub> 在多个数据集上的 NMI 值低于 SpectralNet,由此说明不同视图的特征融合可能会产生负面影响,需要仔细斟酌融合方式与融合时机;
- 与多视图聚类算法相比,可以看出 CMvSC 取得了较好的结果,由此说明了 CMvSC 通过同时挖掘单视图特有特征和多视图共享知识,能够学习到更准确有效的特征表达;
- 值得一提的是: SpectralNet 和 MvSCN 与本文提出的 CMvSC 均采用自适应的邻接图矩阵,与其他采用传统的邻接图矩阵构造方法的模型(如 SC 和 SwMC)相比, SpectralNet 在单视图聚类算法中具有明显的优越性, MvSCN 与 CMvSC 在多视图聚类算法中同样表现优秀,由此说明了自适应邻接图更能抓住数据的潜在关联;
- CMvSC 相较于 MvSCN 的主要不同在于引入了全局学习层,从而进行参数共享和特征迁移. 从对比结果中可知, CMvSC 具有较优的性能. 证明了多视图间的知识迁移,能进一步提升特征嵌入的表达能力.

## 5.6 消融实验

为了证明 CMvSC 网络中各个部分的有效性,我们设计了不同的变体方法进行消融实验,评估结果如表 5 所示. 变体设计如下.

- CMvSC-nc: 不使用基于代码空间的低维输入,直接采用原始数据作为输入;
- CMvSC-na: 不使用基于对比损失的近邻学习,直接采用公式(4)的方法构造邻接矩阵;
- CMvSC-nl: 不使用局部学习层,直接将基于代码空间的多视图低维输入级联,再作为全局学习层的输入;
- CMvSC-np: 不使用参数共享,直接将局部学习层的输出作为特征迁移的输入;
- CMvSC-nf: 不使用特征迁移,直接将参数共享的输出作为正交约束层的输入.

表 5 变体算法与 CMvSC 的性能对比(ACC,  $F$ -mea, NMI)

数据集 算法评估指标	Noisy MNIST			Caltech101-20			Reuters			NUSWIDEOBJ		
	ACC	$F$ -mea	NMI	ACC	$F$ -mea	NMI	ACC	$F$ -mea	NMI	ACC	$F$ -mea	NMI
CMvSC-nc	85.03	82.92	90.55	55.24	44.69	66.10	43.78	40.05	22.16	13.21	10.01	13.95
CMvSC-na	88.28	84.95	90.73	56.48	46.83	66.82	45.37	41.22	25.55	15.79	11.86	14.66
CMvSC-nl	89.13	90.25	90.88	56.35	47.34	68.03	45.88	42.31	25.64	15.67	12.21	14.89
CMvSC-np	93.66	92.89	92.57	58.64	49.67	69.11	47.97	43.86	27.02	17.11	13.89	16.58
CMvSC-nf	92.57	93.06	92.19	57.68	49.86	69.21	48.57	44.21	27.13	17.64	13.58	15.99
<b>CMvSC</b>	<b>99.42</b>	<b>99.44</b>	<b>98.15</b>	<b>61.26</b>	<b>52.65</b>	<b>70.15</b>	<b>49.98</b>	<b>45.66</b>	<b>28.92</b>	<b>18.37</b>	<b>14.65</b>	<b>18.88</b>

如表 5 所示,可以得出以下 5 点结论.

- CMvSC 在不同数据集上的性能远远优于 CMvSC-nc,由此说明了基于代码空间的低秩表达相比原始数据样本包含更少的冗余信息、更多的有效知识. 因此,采用基于代码空间的训练方法,模型能够快速挖掘数据中的隐含关联,从而提高数据划分性能;
- 与 CMvSC-na 相比, CMvSC 有更好的聚类性能,证明了基于对比损失的近邻学习比传统的邻接矩阵构造方法更能抓住数据间的深层关联. 此外,由于基于对比损失的近邻学习同时考虑了正样本与负样本的影响,这也为性能提升提供了帮助;
- 当不使用局部学习层时, CMvSC-nl 与 CMvSC 相比的聚类效果有所降低,说明局部学习层抓住了每个视图的特有性质,对于学习单视图局部不变性具有一定优势,在一定程度上提升了聚类效果;
- CMvSC-np 在各个数据集上的性能都稍逊于 CMvSC,由此证明了参数共享能够有效发现多视图间的

共享特征, 学习多视图分布一致性, 增强谱嵌入的表达能力, 从而提高聚类性能;

- CMvSC 与 CMvSC-nf 相比具有更好的聚类性能, 表明了特征迁移能够发现多视图特征表示之间的互补性质, 进一步强化谱嵌入表征能力, 在一定程度上提升聚类性能。

## 6 结论与展望

本文提出了一种知识迁移下的深度一致性多视图谱聚类网络(CMvSC), 利用多视图学习挖掘数据样本的多样性特征, 同时结合深度学习框架提升算法的特征学习能力与可扩展能力。针对每个视图的局部不变形, CMvSC 引入局部学习层挖掘单视图特有嵌入; 考虑到多视图在数据划分上的全局一致性, CMvSC 构建全局学习层进行知识迁移, 从而学习多视图共享嵌入; 此外, 为深入发现数据样本中的隐含关联, CMvSC 引入孪生网络和对比损失用于近邻学习, 得到同时考虑正负样本影响的邻接矩阵; 最后, 在 4 个数据集上的实验结果有效证明了 CMvSC 的优越性。

然而, CMvSC 的不足在于其近邻学习与谱嵌入学习是两个独立的过程。因此在未来工作中, 我们考虑设计一个端到端的模型, 同时进行邻接图构建与多视图聚类。

### References:

- [1] Kang Z, Zhao XJ, Peng C, *et al.* Partition level multiview subspace clustering. *Neural Networks*, 2020: 279–288.
- [2] Yang Y, Wang H. Multi-view clustering: A survey. *Bia Data Mining and Analytics*, 2018, 1(2): 83–107
- [3] Zhang YP, Zhou J, Deng ZH, *et al.* Multi-view fuzzy clustering approach based on medoid invariant constraint. *Ruan Jian Xue Bao/Journal of Software*, 2019, 30(2): 282–301 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5625.htm> [doi: 10.13328/j.cnki.jos.005625]
- [4] Bickel S, Scheffer T. Multi-view clustering. In: *Proc. of the 4th IEEE Int'l Conf. on Data Mining*. 2004. 19–26.
- [5] Xie JY, Ding LJ, Wang MZ. Spectral clustering based unsupervised feature selection algorithms. *Ruan Jian Xue Bao/Journal of Software*, 2020, 57(9): 1928–1938 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5927.htm> [doi: 10.13328/j.cnki.jos.005927]
- [6] Huang ZY, Zhou TJ, Peng X, *et al.* Multi-view spectral clustering network. In: *Proc. of of the 28th Int'l Joint Conf. on Artificial Intelligence*. 2019. 2563–2569.
- [7] Shaham U, Stanton K, Li H, *et al.* SpectralNet: Spectral clustering using deep neural networks. In: *Proc. of the 6th Int'l Conf. on Learning Representations*. arXiv: 1801.01587, 2018.
- [8] Appice A, Malerba D. A co-training strategy for multiple view clustering in process mining. *IEEE Trans. on Services Computing*, 2016, 9(6): 832–845.
- [9] Du L, Zhou P, Shi L, *et al.* Robust multiple kernel  $k$ -means using  $l_{2,1}$ -norm. In: *Proc. of the 24th Int'l Joint Conf. on Artificial Intelligence*. 2015. 3476–3482.
- [10] Lu YT, Wang LT, Lu JF, *et al.* Multiple kernel clustering based on centered kernel alignment. *Pattern Recognition*, 2014, 47(11): 3656–3664.
- [11] Luo SR, Zhang CQ, Zhang W, *et al.* Consistent and specific multi-view subspace clustering. In: *Proc. of the AAAI Conf. on Artificial Intelligence*. 2018. 3730–3737.
- [12] Liu GC, Lin ZC, Yan SC, *et al.* Robust recovery of subspace structures by low-rank representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2013, 35(1): 171–184.
- [13] Hou CP, Nie FP, Tao H, *et al.* Multi-view unsupervised feature selection with adaptive similarity and view weight. *IEEE Trans. on Knowledge and Data Engineering*, 2017, 29(9): 1998–2011.
- [14] Wang H, Yang Y, Liu B. GMC: Graph-based multi-view clustering. *IEEE Trans. on Knowledge and Data Engineering*, 2019, 32(6): 1116–1129.
- [15] Cao XC, Zhang CQ, Fu HZ, *et al.* Diversity-induced multi-view subspace clustering. In: *Proc. of the 2015 IEEE Conf. on Computer Vision and Pattern Recognition*. 2015. 586–594.

- [16] Zhang CQ, Hu QH, Fu HZ, *et al.* Latent multi-view subspace clustering. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition. 2017. 4279–4287.
- [17] Kumar A, Daumé H. A co-training approach for multi-view spectral clustering. In: Proc. of the 28th Int'l Conf. on Machine Learning. 2011. 393–400.
- [18] Kumar A, Rai P, Daumé H. Co-regularized multi-view spectral clustering. In: Proc. of the 24th Int'l Conf. on Neural Information Processing Systems. 2011. 1413–1421.
- [19] Zhou DY, Burges CJ. Spectral clustering and transductive learning with multiple views. In: Proc. of the 24th Int'l Conf. on Machine Learning. 2007. 1159–1166.
- [20] Nie FP, Wang XQ, Huang H. Clustering and projected clustering with adaptive neighbors. In: Proc. of the 20th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. 2014. 977–986.
- [21] Nie FP, Li J, Li XL. Parameter-free auto-weighted multiple graph learning: A framework for multiview clustering and semi-supervised classification. In: Proc. the 25th Int'l Joint Conf. on Artificial Intelligence. 2016. 1881–1887.
- [22] Nie FP, Cai GH, Li XL. Multi-view clustering and semi-supervised classification with adaptive neighbours. In: Proc. of the 31st AAAI Conf. on Artificial Intelligence. 2017. 2408–2414.
- [23] Kang Z, Shi GX, Huang SD, *et al.* Multi-graph fusion for multi-view spectral clustering. Knowledge-Based Systems, 2020, 189: 105102.
- [24] Nie FP, Li J, Li XL. Self-weighted multiview clustering with multiple graphs. In: Proc. of the 26th Int'l Joint Conf. on Artificial Intelligence. 2017. 2564–2570.
- [25] Zhang Z, Liu L, Shen FM, *et al.* Binary multi-view clustering. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2019, 41(7): 1774–1782.
- [26] Li XL, Zhang R, Wang Q, *et al.* Autoencoder constrained clustering with adaptive neighbors. IEEE Trans. on Neural Networks and Learning Systems, 2021, 32(1): 443–449.
- [27] Xie JY, Girshick R, Farhadi A. Unsupervised deep embedding for clustering analysis. In: Proc. of the 33rd Int'l Conf. on Machine Learning, Vol.48. 2016. 478–487.
- [28] Yang JW, Parikh D, Batra D. Joint unsupervised learning of deep representations and image clusters. In: Proc. of the 2016 IEEE Conf. on Computer Vision and Pattern Recognition. 2016. 1063–6919.
- [29] Caron M, Bojanowski P, Joulin A, *et al.* Deep clustering for unsupervised learning of visual features. In: Proc. of the European Conf. on Computer Vision. 2018. 132–149.
- [30] Andrew G, Arora R, Bilmes J, *et al.* Deep canonical correlation analysis. In: Proc. of the 30th Int'l Conf. on Machine Learning. 2013. 1247–1255.
- [31] Wang WR, Arora R, Livescu K, *et al.* On deep multi-view representation learning: Objectives and optimization. In: Proc. of the 32nd Int'l Conf. on Machine Learning. 2015. 1083–1092.
- [32] Zhao HD, Ding ZM, Fu Y. Multi-view clustering via deep matrix factorization. In: Proc. of the 31st AAAI Conf. on Artificial Intelligence. 2017. 2921–2927.
- [33] Yang B, Fu X, Sidiropoulos ND, *et al.* Towards  $k$ -means-friendly spaces: Simultaneous deep learning and clustering. In: Proc. of the 34th Int'l Conf. on Machine Learning. 2017. 3861–3870.
- [34] Xie JY, Girshick R, Farhadi A. Unsupervised deep embedding for clustering analysis. In: Proc. of the 33rd Int'l Conf. on Machine Learning. 2016. 478–487.
- [35] Jiang ZX, Zheng Y, Tan HC, *et al.* Variational deep embedding: An unsupervised and generative approach to clustering. In: Proc. of the 26th Int'l Joint Conf. on Artificial Intelligence. 2017. 1965–1972.
- [36] Ng A, Jordan M, Weiss Y. On spectral clustering: Analysis and an algorithm. In: Proc. of the 15th Int'l Conf. on Neural Information Processing Systems. 2002. 849–856.

#### 附中文参考文献:

- [3] 张远鹏, 周洁, 邓赵红, 等. 代表点一致性约束的多视角模糊聚类算法. 软件学报, 2019, 30(2): 282–301. <http://www.jos.org.cn/1000-9825/5625.htm> [doi: 10.13328/j.cnki.jos.005625]



- [5] 谢娟英, 丁丽娟, 王明钊. 基于谱聚类的无监督特征选择算法. 软件学报, 2020, 31(4): 1009–1024. <http://www.jos.org.cn/1000-9825/5927.htm> [doi: 10.13328/j.cnki.jos.005927]



张熠玲(1994—), 女, 博士生, 主要研究领域为多视图学习, 多任务学习, 时空数据挖掘.



欧阳小草(1994—), 女, 博士生, CCF 学生会员, 主要研究领域为深度学习, 时空数据挖掘.



杨燕(1964—), 女, 博士, 教授, 博士生导师, CCF 杰出会员, 主要研究领域为人工智能, 大数据分析, 集成学习, 多视图学习, 聚类分析, 时空挖掘.



胡节(1978—), 女, 博士, 副教授, CCF 专业会员, 主要研究领域为人工智能, 机器学习, 数据挖掘, 聚类分析, 聚类集成.



周威(1994—), 男, 博士生, 主要研究领域为深度学习, 时空数据挖掘, 多视图学习.