

基于图信息的自监督多视角子空间聚类^①



吴峰, 刘改, 刘诗仪

(西安工程大学 计算机科学学院, 西安 710600)
通信作者: 吴峰, E-mail: 190721021@stu.xpu.edu.cn

摘要: 多视角子空间聚类方法通常用于处理高维度、复杂结构的数据。现有的大多数多视角子空间聚类方法通过挖掘潜在图信息进行数据分析与处理, 但缺乏对潜在子空间表示的监督过程。针对这一问题, 本文提出一种新的多视角子空间聚类方法, 即基于图信息的自监督多视角子空间聚类 (SMSC)。它将谱聚类与子空间表示相结合形成统一的深度学习框架。SMSC 首先通过挖掘多视角数据的一阶图和二阶图构成潜在图信息, 其次利用聚类结果监督多个视角的公共潜在子空间学习过程。通过在 4 个标准数据集上进行的广泛实验, 结果验证本文所提方法相较于传统的多视角子空间聚类方法更具有效性。

关键词: 多视角聚类; 子空间聚类; 自监督学习; 空间聚类算法; 机器学习算法

引用格式: 吴峰, 刘改, 刘诗仪. 基于图信息的自监督多视角子空间聚类. 计算机系统应用, 2022, 31(5): 377-381. <http://www.c-s-a.org.cn/1003-3254/8448.html>

Self-supervised Multi-view Subspace Clustering with Graph Information

WU Feng, LIU Gai, LIU Shi-Yi

(School of Computer Science, Xi'an Polytechnic University, Xi'an 710600, China)

Abstract: Multi-view subspace clustering methods are usually used to process high-dimensional and complex data. Most of the existing multi-view subspace clustering methods analyze and process data by mining potential graph information, with no supervision process for the representation of the potential subspace. To solve this problem, this study proposes a new multi-view subspace clustering method, namely self-supervised multi-view subspace clustering (SMSC) based on graph information. It combines spectral clustering with subspace representation to formulate a unified deep learning framework. SMSC constructs potential graph information by mining the first-order and second-order graphs of multi-view data and then uses clustering results to supervise the learning process of the common potential subspace of multi-view data. Extensive experiments on four standard datasets show that the proposed method is more effective than traditional multi-view subspace clustering methods.

Key words: multi-view clustering; subspace clustering; self-supervised learning; spatial clustering algorithm; machine learning algorithm

聚类分析是数据挖掘和人工智能领域中一种重要的数据处理技术, 其目的是根据未标记的数据点的内在相似性, 将其划分为相应的簇^[1]。聚类分析作为无监督学习中的一项基本任务, 在图像处理、数据挖掘等

众多领域具有广泛的应用前景。传统的聚类方法包含子空间聚类^[1]、谱聚类^[1]和 K-means 聚类^[2]等。近年来, 也有学者提出基于深度学习的聚类方法, 但目前这类方法通常用于处理传统的单视角数据。多视角数据

^① 基金项目: 西安市科技计划 (2020KJRC0027)

收稿时间: 2021-07-07; 修改时间: 2021-08-11; 采用时间: 2021-08-17; csa 在线出版时间: 2022-04-11

对象由多个视角的数据实例组成,不同视角间存在一致性和互补性关系,仅直接采用传统的单视角聚类方法,无法有效挖掘多视角数据所包含的丰富信息.因而,近年来出现了一些多视角聚类方法.其中 Wang 等人^[3]考虑到不同视角的权重问题,提出一种基于图的多视角聚类算法(GMC),该方法首先将每个视角的数据矩阵转换为由相似度图矩阵构成的图矩阵,其次将它们进行融合生成统一的图矩阵.GMC自动为每个视角分配权重,共同学习每个视角的图矩阵和融合图,并在融合后直接生成最终的簇.Tang 等人^[4]提出基于联合潜在表示和相似性学习的多视角聚类方法(LALMVC),用于解决相似性矩阵无法有效表示数据内在几何结构与数据间邻域关系的问题.LALMVC首先学习各视角共享数据的潜在表示,然后在数据的潜在表示中利用流形学习自适应地获得相似性矩阵,从而较好描述数据几何关系.由于多视角的一致性与特定性通常无法共同用于子空间表示的学习过程,Luo 等人^[5]提出一致与特定多视角子空间聚类方法(CSMSC).该方法使用一个共享的一致性表示与一组特定性表示描述多视角自表示属性.其中一致性用于描述所有视角间的共同特性,而特定性用于捕获每个视角的内在差异.Zheng 等人^[6]通过挖掘局部和全局图信息并将其融合,用于约束子空间表示学习过程.由于该方法缺乏对自表达系数矩阵的监督过程,不能学到很好的特征表示.因此可以使用聚类标签监督自表达模块的训练,以获得更好的聚类效果.

基于上述问题,本文提出一种基于图信息的自监督子空间聚类方法,用来自监督式地挖掘多视角数据的潜在图信息.本文的主要贡献包含:(1)通过添加图正则化项获得更好的潜在空间表示;(2)通过在目标函数中添加自监督规范化项,实现聚类标签对自表达系数学习过程的监督.

1 相关工作

1.1 子空间聚类

子空间聚类旨在从多个子空间组成的数据中揭示固有的簇结构.现有的子空间聚类方法主要集中于构造好的自表达系数矩阵,然后学习良好的亲和矩阵并进行谱聚类,其中具有代表性的方法包括深度子空间聚类(DSCN)^[7]、稀疏子空间聚类(SSC)^[8]和超完全深度子空间聚类(ODSC)^[9].DSCN设计一种新的无监督

子空间聚类深度神经网络.该架构建立在深度自编码器上,非线性地将输入数据映射到潜在空间.在编码器和解码器之间引入一种新的自表达层模拟在传统子空间聚类中被证明有效的“自表达性”特性.SSC通过对自表达系数矩阵添加 L_1 范数,保证了自表达系数矩阵的稀疏性.该方法在处理噪声和缺失数据方面也起到了重要作用.ODSC首先融合不完全和过完全的自动编码器网络的特征,然后将它们通过自表达层提取一个更有意义和更鲁棒的输入数据表示进行聚类.

1.2 多视角子空间聚类

传统的单视角子空间聚类处理多视角数据只是进行简单的特征拼接,容易造成维度灾难等问题.多视角子空间聚类以单视角子空间聚类为基础,分别处理每个视角的特征.在互补性原则和一致性原则^[10]的辅助下,多视角子空间聚类建立不同视角间的联系并进行信息交互,克服了单视角子空间聚类划分多视角数据的问题.潜在多视角子空间聚类(LMSC)^[11]从多个视角探索潜在的互补信息,利用多个视角的互补性,潜在表示比每个视角更全面地描述数据,从而使子空间表示更准确和稳健.多视角深度子空间聚类(MDSC)^[12]由两个子网络组成,且分别完成所有视角的公共自表示矩阵和特定视角的自表示矩阵的学习.该方法通过引入希尔伯特·施密特独立标准作为多样性正则化项,将每个视角的自表示矩阵与公共的自表示矩阵对齐,从而取得了较好的聚类效果.

1.3 自监督学习

监督学习和无监督学习是机器学习两种基本的学习范式,而自监督学习是使用无标注数据自我监督的方式开展学习,其归类于无监督学习的范畴.自监督学习^[13]的思想较为简单,即输入的无监督数据通过数据本身的结构或者特性,构造伪标签(pseudo label)出来.有了标签以后就可以类似监督学习一样进行训练.近年来,有学者试图将自监督学习与多视角子空间聚类相结合,获得构造的伪标签后对整体网络进行训练.自监督卷积子空间聚类网络(S2ConvSCN)^[14]通过聚类标签监督每个特定视角的潜在表示学习过程和多个视角的公共潜在子空间表示学习过程,实现数据的自监督目的.多视角聚类中的自监督判别特征学习(SDMVC)^[15]应用深度自编码器独立学习每个视角的嵌入式特性,并利用多视角互补信息连接所有视角的嵌入特性,形成全局特性.以一种自监督的方式,获得伪

标签建立统一的目标分布并执行多视角判别特征学习. 这种统一的目标分布可实现多个视角一致性与多样性.

2 基于图信息的自监督多视角子空间聚类

基于图信息的自监督多视角子空间聚类网络结构如图1所示. 该网络结构由多视角卷积自编码器、自表达模块、图正则化模块和谱聚类模块4部分组成, 其中多视角卷积自编码器又分为多视角卷积自编码器和多视角卷积自解码器. 假设 n 个样本共 k 个视角, 给定一个多视角数据集 $\mathbf{X} = \{\mathbf{X}^k\}_{k=1}^v$, $\mathbf{X}^k \in \mathbb{R}^{n \times d_k}$ 表示第 v 个视角的数据, d_k 表示第 k 个视角数据的维度.

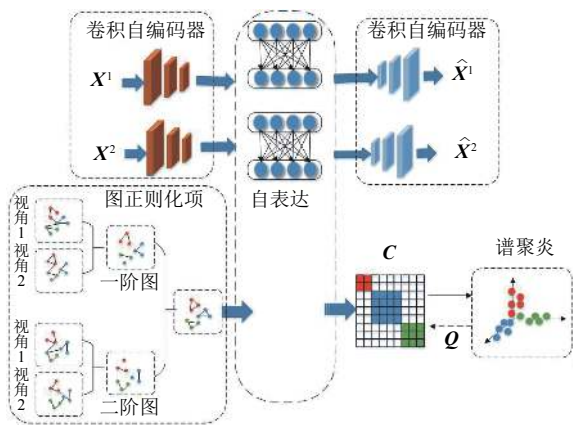


图1 基于图信息的自监督多视角子空间聚类网络结构

(1) 重构损失

数据在潜在空间中的表示应包含原始数据的主要信息, 将原始数据输入解码器网络以获得重构数据. 编码器—解码器网络的优化目标为最小化重构损失, 如下式:

$$\ell_1 = \frac{1}{2} \sum_{k=1}^v \|\mathbf{X}^k - \widehat{\mathbf{X}}^k\|_F^2 \quad (1)$$

其中, $\widehat{\mathbf{X}}^k$ 表示第 k 个视角的重构数据.

(2) 自表达损失

基于数据的自表达性质, 即同一子空间中的每个数据点都可以近似表示为其他数据点的线性组合, 为了寻找蕴含在数据中的子空间结构, 自表达模块的损失函数如下:

$$\ell_2 = \frac{1}{2} \sum_{k=1}^v \|\mathbf{z}^k - \mathbf{z}^k \mathbf{C}\|_F^2 \quad (2)$$

其中, \mathbf{z}^k 为第 k 个视角的潜在空间表示, $\mathbf{z}^k \mathbf{C}$ 表示初始数据经过自表达层的输出, 同时也是解码器网络的输入数据. \mathbf{C} 表示自表达系数矩阵.

(3) 图正则化损失

通过网络结构可以看到, 需要添加一个特定的约束来指导 \mathbf{C} 的聚类学习过程, Zheng 等人^[16] 通过添加一个图正则化项用来捕获多视角数据的局部信息和全局信息. 一阶相似度定义为两个数据点间的局部成对相似性, 而二阶相似度则定义为两个数据点的全局网络结构的相似性. 对于第 k 个视角来说, 第 i 个数据点和第 j 个数据点之间的相似度就是 $\exp\left(-\frac{\|\mathbf{X}_i^k - \mathbf{X}_j^k\|_2^2}{\sigma^2}\right)$, 其中, σ 代表中位欧几里得距离, 第 k 个视角的第 i 和第 j 数据点之间的一阶相似度如下:

$$\mathbf{A}^k = \begin{cases} \exp\left(-\frac{\|\mathbf{X}_i^k - \mathbf{X}_j^k\|_2^2}{\sigma^2}\right), & \text{if } \mathbf{X}_i^k \text{ and } \mathbf{X}_j^k \text{ are } mkNN \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

其中, \mathbf{A}^k 是第 k 个视角的一阶相似度矩阵, $mkNN$ 表示相互 k 最近邻策略. 第 k 个视角第 i 个数据点和第 j 个数据点之间的二阶相似度如下:

$$\widehat{\mathbf{A}}^k = \begin{cases} \exp\left(-\frac{\|\mathbf{A}_i^k - \mathbf{A}_j^k\|_2^2}{\sigma^2}\right), & \text{if } \mathbf{A}_i^k \text{ and } \mathbf{A}_j^k \text{ are } mkNN \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

其中, $\widehat{\mathbf{A}}^k$ 是第 k 个视角的二阶相似度矩阵. 通过将局部图信息和全局图信息进行结合, 得到原始数据的统一相似性矩阵:

$$\mathbf{A} = \bigodot_{k=1}^v \mathbf{A}^k + \bigodot_{k=1}^v \widehat{\mathbf{A}}^k \quad (5)$$

其中, \odot 表示哈达玛乘积, \mathbf{A} 包含多视角数据的局部和全局图信息. 通过添加一个图正则化项, 所需子空间表示的学习过程可以表示为式(6), 其中, \mathbf{L} 表示 \mathbf{A} 的拉普拉斯矩阵, Tr 表示矩阵的迹.

$$\ell_3 = \frac{1}{2} \sum_{i,j} \mathbf{A}_{ij} \|\mathbf{C}_i - \mathbf{C}_j\|_2^2 = Tr(\mathbf{C}^T \mathbf{L} \mathbf{C}) \quad (6)$$

(4) 自监督损失

获得自表达系数矩阵后, 可得到亲和矩阵 $\mathbf{W} = \frac{1}{2}(|\mathbf{C}| + |\mathbf{C}^T|)$. 在 \mathbf{W} 上应用谱聚类, 通过优化如下目标使相同类别的数据距离较近:

$$\min_{\mathbf{Q}} \sum_{i,j} w_{ij} \|\mathbf{q}_i - \mathbf{q}_j\|_2^2, \quad \text{s.t. } \mathbf{Q} \in \mathcal{Q} \quad (7)$$

其中, $\mathcal{Q} = \{\mathbf{Q} \in \{0, 1\}^{c \times n} : \mathbf{1}^T \mathbf{Q} = \mathbf{1}^T, \text{rank}(\mathbf{Q}) = c\}$ 是一组具有 c 组的有效分割矩阵, \mathbf{q}_i 和 \mathbf{q}_j 分别是第 i 和第 j 列,

表示每个指定簇的数据点的所属类别. 谱聚类产生了数据集的聚类标签, 尽管不一定是所有数据点的正确标签, 但它包含了关于数据集的有意义的信息, 这促使我们使用谱聚类的输出来监督自表达模块的训练. 自监督损失函数如下:

$$\ell_4 = \frac{1}{2} \sum_{i,j} w_{ij} \|q_i - q_j\|_2^2 \quad (8)$$

最后, 综合式 (1)、式 (2)、式 (6) 与式 (8), 得到 SMSC 网络的总体目标函数:

$$\ell = \ell_1 + \lambda_1 \ell_2 + \lambda_2 \ell_3 + \lambda_3 \ell_4 \quad (9)$$

其中, $\lambda_1, \lambda_2, \lambda_3$ 表示目标函数中的平衡参数.

3 实验

3.1 实验设置

本文所提出的方法在 4 个标准的真实数据集上进行了广泛的实验. 其中, Yale Face 包含 15 个人的面部图像, 共 165 张. ORL 包含 40 个人的面部图像, 共 400 张. Yale Face 和 ORL 数据集的每个图像都由 3 个特征来描述: intensity、LBP 和 Gabor. 而且每张图像的面部表情、光线强弱都是不同的. MSRCV1 数据集是由从 7 个簇中收集的 210 个图像样本组成, 每张图像由 6 个特征来描述: CENT、CMT、GIST、HOG、LBP 和 SIFT. BBC 数据集由英国广播公司的 685 份文本文件组成, 每份文件分为 4 个子部分.

本文采用了 4 种评价指标^[11]来评估算法的聚类效果, 分别是归一化互信息 (NMI)、准确率 (ACC)、F-度量值 (F-Measure) 和兰德指数 (RI).

3.2 对比实验

表 1-表 4 分别表示在 Yale Face、ORL、MSRCV1 和 BBC 数据集上不同算法的聚类实验结果, 其中加粗的数据表示在该组对比实验中得到的最优结果.

表 1 Yale Face 数据集对比实验结果

方法	NMI	ACC	F-Measure	RI
LMSC	0.705 6	0.668 9	0.507 4	0.935 1
DMSC	0.689 6	0.710 8	0.586 2	0.939 7
PMSC	0.640 2	0.625 0	0.459 8	0.928 9
MSCNLG	0.901 0	0.915 2	0.856 9	0.979 5
SMSC	0.911 2	0.927 3	0.876 8	0.981 1

结果表明, 本文所提出的 SMSC 相较于其他对比方法获得了更好的聚类结果. 与传统的多视角子空间聚类方法相比, SMSC 取得了显著的改进. 从表 1 中可以看出, 在 Yale Face 数据集上, SMSC 与 LMSC 相比,

NMI 和 ACC 指标分别提升了 20.56% 和 25.84%, 说明 SMSC 的效果是更好的一方. 在另外 3 个数据集上, 相比其他多视角算法, SMSC 也都有显著的提升. SMSC 首先挖掘数据的潜在信息, 其次通过自表达系数矩阵获得相似性矩阵并在其上应用谱聚类, 获得聚类标签信息后, 通过自监督规范化项对自表达系数矩阵进行监督, 获得更好的聚类效果.

表 2 ORL 数据集对比实验结果

方法	NMI	ACC	F-Measure	RI
LMSC	0.920 0	0.818 8	0.762 5	0.988 5
DMSC	0.953 4	0.865 8	0.888 8	0.992 4
PMSC	0.833 9	0.668 9	0.549 9	0.975 0
MSCNLG	0.955 9	0.893 9	0.900 1	0.993 8
SMSC	0.960 7	0.892 5	0.908 8	0.993 9

表 3 MSRCV1 数据集对比实验结果

方法	NMI	ACC	F-Measure	RI
LMSC	0.614 8	0.694 8	0.590 7	0.882 4
DMSC	0.699 8	0.800 2	0.695 2	0.903 3
PMSC	0.647 5	0.708 1	0.620 0	0.885 5
MSCNLG	0.872 1	0.938 0	0.886 7	0.966 7
SMSC	0.919 3	0.957 1	0.923 0	0.976 0

表 4 BBC 数据集对比实验结果

方法	NMI	ACC	F-Measure	RI
LMSC	0.559 9	0.739 0	0.628 9	0.833 4
DMSC	0.577 4	0.779 9	0.660 5	0.842 7
PMSC	0.008 4	0.321 0	0.381 1	0.243 8
MSCNLG	0.698 8	0.883 2	0.791 0	0.906 4
SMSC	0.736 7	0.914 2	0.866 4	0.931 2

3.3 收敛性实验

以 Yale Face 数据集上的实验为例. 图 2 显示了标准化损失和聚类结果指标 (NMI 和 ACC) 随着迭代次数增加的变化情况. 可以很明显的看出, SMSC 具有较快的收敛速度.

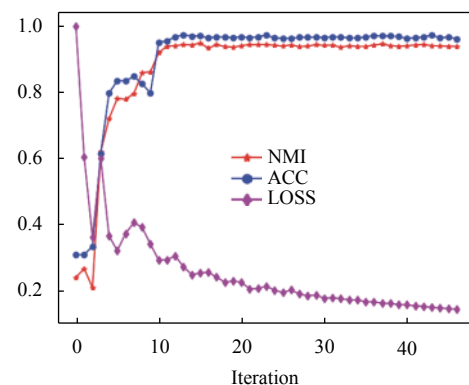


图 2 收敛性结果示意图

4 结论

本文提出一种基于图信息的自监督多视角子空间聚类方法,该方法自监督式地挖掘多视角数据的潜在信息,提升聚类效果.通过在4个标准数据集上的大量实验验证,结果表明所提方法优于几种常用的多视角聚类方法.

参考文献

- 1 章永来,周耀鉴. 聚类算法综述. 计算机应用, 2019, 39(7): 1869–1882. [doi: 10.11772/j.issn.1001-9081.2019010174]
- 2 杨俊闯,赵超. K-Means 聚类算法研究综述. 计算机工程与应用, 2019, 55(23): 7–14, 63. [doi: 10.3778/j.issn.1002-8331.1908-0347]
- 3 Wang H, Yang Y, Liu B. GMC: Graph-based multi-view clustering. IEEE Transactions on Knowledge and Data Engineering, 2020, 32(6): 1116–1129. [doi: 10.1109/TKDE.2019.2903810]
- 4 Tang C, Zhu XZ, Liu XW, *et al.* Learning a joint affinity graph for multiview subspace clustering. IEEE Transactions on Multimedia, 2019, 21(7): 1724–1736. [doi: 10.1109/TMM.2018.2889560]
- 5 Luo SR, Zhang CQ, Zhang W, *et al.* Consistent and specific multi-view subspace clustering. Proceedings of the 32nd AAAI Conference on Artificial Intelligence. New Orleans: AAAI, 2018. 3730–3737.
- 6 Zheng QH, Zhu JH, Ma YY, *et al.* Multi-view subspace clustering networks with local and global graph information. Neurocomputing, 2021, 449: 15–23. [doi: 10.1016/j.neucom.2021.03.115]
- 7 Ji P, Zhang T, Li HD, *et al.* Deep subspace clustering networks. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: ACM, 2017. 23–32.
- 8 Elhamifar E, Vidal R. Sparse subspace clustering: Algorithm, theory, and applications. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(11): 2765–2781. [doi: 10.1109/TPAMI.2013.57]
- 9 Valanarasu JMJ, Patel VM. Overcomplete deep subspace clustering networks. 2021 IEEE Winter Conference on Applications of Computer Vision. Waikoloa: IEEE, 2021. 746–755.
- 10 何士豪,张玉龙,唐启凡. 互补与一致的多视角子空间聚类网络. 西安交通大学学报, 2021, 55(7): 166–178.
- 11 Zhang CQ, Hu QH, Fu HZ, *et al.* Latent multi-view subspace clustering. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017. 4333–4341.
- 12 Zhu PF, Hui BY, Zhang CQ, *et al.* Multi-view deep subspace clustering networks. arXiv: 1908.01978, 2019.
- 13 Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv: 1511.06434, 2015.
- 14 Zhang JJ, Li CG, You C, *et al.* Self-supervised convolutional subspace clustering network. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019. 5468–5477.
- 15 Xu J, Ren YZ, Tang HY, *et al.* Self-supervised discriminative feature learning for Multi-view clustering. arXiv: 2103.15069, 2021.
- 16 Zheng QH, Zhu JH, Li ZY, *et al.* Consistent and complementary graph regularized multi-view subspace clustering. arXiv: 2004.03106, 2020.