



智能系统学报

CAAI TRANSACTIONS ON INTELLIGENT SYSTEMS

非对称卷积编码器的聚类算法

杨梦茵, 陈俊芬, 翟俊海

引用本文:

杨梦茵, 陈俊芬, 翟俊海. 非对称卷积编码器的聚类算法[J]. 智能系统学报, 2022, 17(5): 900–907.

YANG Mengyin, CHEN Junfen, ZHAI Junhai. A clustering method based on the asymmetric convolutional autoencoder[J]. *CAAI Transactions on Intelligent Systems*, 2022, 17(5): 900–907.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202107021>

您可能感兴趣的其他文章

结合地标点与自编码的快速多视图聚类网络

Fast multiview clustering network combining landmark points and autoencoder

智能系统学报. 2022, 17(2): 333–340 <https://dx.doi.org/10.11992/tis.202101011>

一种基于ELM-AE特征表示的谱聚类算法

A spectral clustering algorithm based on ELM-AE feature representation

智能系统学报. 2021, 16(3): 560–566 <https://dx.doi.org/10.11992/tis.202005021>

用户兴趣点耦合关系的兴趣点推荐方法

A POI recommendation approach based on user-POI coupling relationships

智能系统学报. 2021, 16(2): 228–236 <https://dx.doi.org/10.11992/tis.201907034>

一种深度自监督聚类集成算法

A deep self-supervised clustering ensemble algorithm

智能系统学报. 2020, 15(6): 1113–1120 <https://dx.doi.org/10.11992/tis.202006050>

加权PageRank改进地标表示的自编码谱聚类算法

An autoencoder spectral clustering algorithm for improving landmark representation by weighted PageRank

智能系统学报. 2020, 15(2): 302–309 <https://dx.doi.org/10.11992/tis.201904021>



微信公众平台



期刊网址

DOI: 10.11992/tis.202107021

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.tp.20220617.1648.006.html>

非对称卷积编码器的聚类算法

杨梦茵^{1,2}, 陈俊芬^{1,2}, 翟俊海^{1,2}

(1. 河北大学 数学与信息科学学院, 河北 保定 071002; 2. 河北省机器学习与计算智能重点实验室, 河北 保定 071002)

摘要: 基于深度神经网络的非监督学习方法通过联合优化特征表示和聚类指派, 大大提升了聚类任务的性能。但大量的参数降低了运行速度, 另外, 深度模型提取的特征的区分能力也影响聚类性能。为此, 提出一种新的聚类算法 (asymmetric fully-connected layers convolutional auto-encoder, AFCAE), 其中卷积编码器结合非对称全连接进行无监督的特征提取, 然后 K-means 算法对所得特征执行聚类。网络采用 3×3 和 2×2 的小卷积核, 大大减少了参数个数, 降低了算法复杂性。在 MNIST 上 AFCAE 获得 0.960 的聚类精度, 比联合训练的 DEC (deep embedding clustering) 方法 (0.840) 提高了 12 个百分点。在 6 个图像数据集上实验结果表明 AFCAE 网络有优异的特征表示能力, 能出色完成下游的聚类任务。

关键词: 无监督; 聚类; 深度神经网络; 卷积神经网络; 自编码器; 特征学习; 特征表示; 算法复杂性

中图分类号: TP181 **文献标志码:** A **文章编号:** 1673-4785(2022)05-0900-08

中文引用格式: 杨梦茵, 陈俊芬, 翟俊海. 非对称卷积编码器的聚类算法 [J]. 智能系统学报, 2022, 17(5): 900-907.

英文引用格式: YANG Mengyin, CHEN Junfen, ZHAI Junhai. A clustering method based on the asymmetric convolutional autoencoder [J]. CAAI transactions on intelligent systems, 2022, 17(5): 900-907.

A clustering method based on the asymmetric convolutional autoencoder

YANG Mengyin^{1,2}, CHEN Junfen^{1,2}, ZHAI Junhai^{1,2}

(1. College of Mathematics and Information Science, Hebei University, Baoding 071002, China; 2. Hebei Key Laboratory of Machine Learning and Computational Intelligence, Baoding 071002, China)

Abstract: Unsupervised learning methods based on deep neural networks have synergistically optimized the feature representation and clustering assignment, thus greatly improving the clustering performance. However, numerous parameters slow down the running speed, and the discriminative ability of the features extracted by deep models also influences their clustering performance. To address these two issues, a new clustering algorithm is proposed (asymmetric fully-connected layers convolutional autoencoder, AFCAE), where a convolutional autoencoder combined with several asymmetric fully-connected layers is used to extract the features, and the K-means algorithm is subsequently applied to perform clustering on the obtained features. AFCAE adopts 3×3 and 2×2 convolutional kernels, thereby considerably reducing the number of parameters and the computational complexity. The clustering accuracy of AFCAE on MNIST reaches 0.960, almost 12% higher than that of the jointly trained DEC method (0.840). Experimental results on six image data sets show that the AFCAE network has excellent feature representation ability and can finish the subsequent clustering tasks well.

Keywords: unsupervised; clustering; deep neural network; convolutional neural network; autoencoder; feature learning; feature representation; algorithm complexity

收稿日期: 2021-07-09. 网络出版日期: 2022-06-20.

基金项目: 河北省引进留学人员资助项目 (C20200302); 河北省机器学习与计算智能重点实验室自主立项项目 (ZZ201909-202109-1); 河北省科技计划重点研发项目 (19210310D); 河北省自然科学基金项目 (F2021 201020); 河北省社会科学基金项目 (HB20TQ005).

通信作者: 陈俊芬. E-mail: chenjunfen2010@126.com.

无监督聚类是机器学习的重要分支, 根据数据自身的相似性揭示数据内部的隐藏结构。传统的聚类算法如 K 均值 (K-means)^[1-2]、高斯混合模型 GMM (Gaussian mixture model)^[3]、基于密度的聚类算法 DBSCAN (density-based spatial clustering of

applications with noise)^[4]等简单易实现,得到广泛应用。但在图像、文本、语音等高维且无结构的数据上容易发生维度灾难,且需要为每个数据集和任务人工设计特征,从而使聚类性能大打折扣。主成分分析法(principal component analysis, PCA)^[5]对高维数据进行降维并提取特征,其线性表征能力在应对某些特殊分布的数据时效果不佳。基于非监督学习的自编码器(auto-encoder, AE)^[6-7]通过多层复合映射能得到数据的非线性特征。自编码器将原始高维数据映射到低维特征空间,在该空间中更容易依据特征的分布形态执行聚类分析。

深度聚类是用深度神经网络进行表征学习和聚类指派的过程,通常卷积神经网络(convolutional neural networks, CNN)^[8]或堆叠自编码器自适应地学习特征表示,再使用传统聚类算法完成聚类指派。此类方法与非深度聚类算法相比,在基准测试图像数据集上都获得了较好的性能。而AE由此不断发展为稀疏自动编码器(sparse auto-encoder)^[9-10]、降噪自动编码器(denoising auto-encoder)^[11-12]以及卷积自动编码器(convolutional auto-encoder, CAE)^[13]。

现有的聚类方法大多侧重于建模实例之间的相似或相异关系,而忽略了提取更有效的表示,这在很大程度上影响了聚类性能。受此启发,本文提出了一种基于非对称全连接层的卷积自编码器的深度聚类算法进行图像聚类分析。非对称全连接层的卷积自编码器学习输入图片的特征表示,然后经典的K-means算法对特征进行聚类划分,对应成原图像的聚类结果。该方法的主要贡献包括:

- 1) 提出非对称全连接层的卷积网络;
- 2) 使用小卷积核,降低算法的复杂性,加快运行速度;
- 3) 在MNIST数据集上取得优于先进的深度聚类算法的聚类精度。

1 相关工作

自动编码器是无监督表示学习中重要算法之一,由于隐藏层维度通常比数据层小,它可以帮助提取更显著的特征。DEC(deep embedding clustering)^[14]先通过深度编码解码网络对数据进行降维,然后采用软分配确定样本点所属簇类,得到聚类结果。通过最小化软标签的分布和辅助目标分布之间的KL散度来迭代改善聚类。IDEC^[15]基于DEC的这种思路,在表征学习步骤时使用重构损失和聚类损失联合训练聚类网络。使用欠完

备自动编码器来学习嵌入特征。DCN^[16]结合了自动编码器和K-means算法。DCN预先训练自动编码器,而后优化重建损失和K-means损失。精心设计了网络结构,以避免琐碎和无意义的解决方案,并提出了一个有效的优化程序来处理挑战性问题。DEN^[17]利用自动编码器从原始数据中学习简化的表示。应用局部保留约束保留数据的局部结构属性,通过优化损失对网络进行微调实现聚类的精度的提高。

卷积自编码器中卷积核利用局部感受野抽取图像的局部特征和权值共享减少参数个数等优点,使得深度聚类方法DBC^[18]在图像数据集上获得很好的聚类结果。DBC尝试学习深度卷积自编码器以端到端的方式进行训练,设计卷积层(卷积层和反卷积层)和池化层(池化层和反池化层)组成的全连接卷积自编码器(FCAE)网络。使用t分布嵌入算法(t-SNE)^[19]分布去衡量特征点与特征聚类中心点的相似性。

CAE进行特征提取加快了网络训练的速度还提高了其下游应用任务的精度。基于深度特征表示的Softmax聚类算法(ASCAE-Softmax)^[20]算法设计非对称自编码器网络进行无监督聚类,其中随机初始化网络权值替代层层预训练和全连接层的重构误差作为目标函数的正则约束。该方法为无监督聚类分析和特征表达提供了新思路。

2 AFCAE 聚类算法

利用卷积自编码器进行无监督的特征提取,并应用在下游聚类任务中。与一些联合聚类算法不同,本文方法分阶段进行。

2.1 网络结构

本文提出了一种非对称全连接的卷积自编码器(asymmetric fully-connected layers convolutional auto-encoder, AFCAE),网络结构如图1上半部分所示。网络的输入是28×28的图片,然后通过卷积网络(C1-C4)进行特征提取。受经典联合聚类算法DBC^[18]的启发,卷积核多采用3×3。C1到C3的每个待提取特征图都使用3×3卷积核,而C4层上使用2×2的卷积核。在C1层使用步长为3的3×3卷积核,步长等于卷积核的尺寸相当于把图片分割成小片再做特征提取,便于在分辨率高的图像上捕捉丰富的局部特征信息。F1-F6是全连接层,其中F4层的神经元个数与数据集有关,其他层均设为50个。网络采用ReLU激活函数。AFCAE网络相对于对称式网络而言可以看成在F2-F6对称全连接部分前加入F1全连接层

进行数据的整合,从而形成非对称网络。改善网络的非线性特征表示的能力。

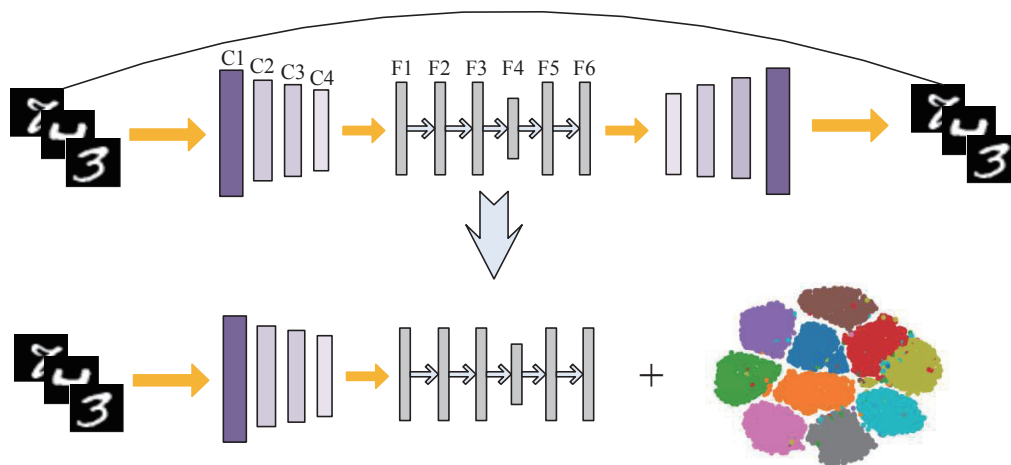


图 1 非对称全连接层卷积自编码器 (AFCAE) 网络框架

Fig. 1 Overall framework of the asymmetric fully-connected layers convolutional auto-encoder (AFCAE) network

这个 AFCAE 网络的端到端无监督预训练结束后,截取 C1 到 F6 层后接入 K-means 算法形成本文的深度聚类模型。为了减少名词的困扰,称其为 AFCAE 聚类算法,见图 1 的下半部分。后面

实验用到的 AFCAE 聚类算法详细化为图 2,对应的网络参数列于表 1,其中 $(k,n)/s$ 分别代表卷积核的大小、通道数和步长。

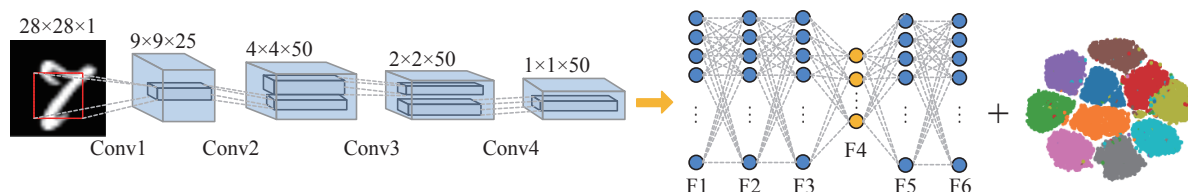


图 2 AFCAE 网络结构

Fig. 2 AFCAE network structure

表 1 AFCAE 网络参数详细表

Table 1 The detailed description of AFCAE network parameters

层	卷积核/步长
C1	(3,25)/3
C2	(3,50)/2
C3	(3,50)/1
C4	(2,50)/1
F1	50
F2	50
F3	50
F4	10
F5	50
F6	50

2.2 AFCAE 方法

AFCAE 网络损失函数为输入 x_i 与输出 \hat{x}_i 之间的误差平方和,为防止网络过拟合,加入 L_2 正则化

约束:

$$J(W; X; \hat{X}) = \frac{1}{N} \sum_{i=1}^N \|x_i - \hat{x}_i\|^2 + \alpha \sum \|W\|^2 \quad (1)$$

式中: $X = \{x_1, x_2, \dots, x_N\}$ 为图像数据集; N 为图片总个数。第 1 项损失为重构损失,输入 x_i 通过自编码器得到重构后的输出 \hat{x}_i 。第 2 项是 L_2 正则化约束, W 为网络参数。 α 为超参数,后面实验中设置 $\alpha = 0.01$ 。

训练好非对称全连接层卷积自编码器后,保留 C1-F6 层网络结构和参数,使用 F6 层输出作为特征信息进行聚类。

2.3 复杂性分析

卷积层的理论时间复杂性表示^[21]为

$$O\left(\sum_{l=1}^d n_{l-1} k_l^2 n_l m_l^2\right) \quad (2)$$

式中: d 为卷积层的数量; l 是卷积层的索引; n_l 是第 l 层中卷积核的数量,也称为输出通道数; n_{l-1} 称

为第 l 层的输入通道数; k_l 是卷积核的尺寸; m_l 为卷积核输出的特征图大小。

全连接层的时间复杂性为输入通道和输出通道的乘积, 表示为

$$O\left(\sum_{i=1}^d n_{l-1} n_l\right) \quad (3)$$

3 实验结果与分析

本节主要通过 K-means 对特征表示进行聚类分析来验证 AFCAE 网络的特征表示能力。所有实验均在 6 个图像数据集上进行, 3.1 节详细介绍不同数据集。为了减少随机初始化对 K-means 算法性能的影响, 每组实验都重复 50 次, 选取最好的聚类精度。

实验环境: IntelCorei5-6300HQ 处理器, NVIDIA 2.0GB 显存, 8.0GB RAM 显卡; 基于开源的 Keras 库搭建 AFCAE 网络。

3.1 实验数据集

1) MNIST: 由 70 000 个手写数字组成的灰度图像数据集。图像尺寸为 28×28 , 属于 10 个不同的类 (<http://yann.lecun.com/exdb/mnist/>)。

2) CAS-PEAL-R1: 属于 40 个不同类的 200 幅灰度图像数据集。每张图片尺寸为 480×360 。是纯色背景下人脸有表情变化的正面视图。且每人有 5 幅图片 (<http://www.jdl.ac.cn/peal/JDL-PEAL-Release.htm>)。

3) COIL-20: 由日常生活物品组成的 1 440 张图片尺寸为 128×128 的数据集。数据集类别为 20 类, 是有角度、无形变的灰度图片 (<https://www.dazhu.anlan.com/2019/10/06/5d999ded06295/>)。

4) BioID-Face: 数据集为 23 位人物组成的 1 521 幅灰度图像, 图片尺寸为 384×286 。每张图片为正面视图, 具有较大的姿态变化和表情变化 (<https://www.bioid.com/facedb/>)。

5) IMM-Face: 属于 40 个类簇的 240 幅图片, 图片尺寸为 640×48 , 视图为纯色背景下有侧面和正面、有表情变化的图片, 每人 6 幅彩色/灰度图片 (<http://www.imm.dtu.dk/~aam/aamexplorer/>)。

6) UMISTS: 包括 20 个人共 564 幅图像, 图片尺寸为 220×220 , 纯色背景下每个人具有不同角度、不同姿态的灰度图像 (https://see.xidian.edu.cn/vips/database_Face.html)。

3.2 评价指标

本文使用评价聚类性能的数值指标是聚类精度 (accuracy, ACC) 和标准互信息 (normalized mutual information, NMI)。两个指标值越近 1, 说明聚类准确度越高。

聚类精度 (ACC): 聚类精度定义为聚类指派对的数据个数与全部数据个数之比:

$$ACC = \frac{\sum_{i=1}^N \delta(s_i, r_i)}{N} \quad (4)$$

$$\delta(s_i, r_i) = \begin{cases} 1, & s_i = r_i \\ 0, & \text{其他} \end{cases}$$

式中: s_i 代表真实标签; r_i 是聚类指派的标签; N 为总的的数据个数。经典匈牙利算法对聚类类标和真实类标进行匹配, 通过最佳类别指派得到最优类别结果。ACC 值越接近 1, 说明聚类准确度越高。

标准互信息 (NMI): 将互信息归一化到 $[0, 1]$, 若互信息为 0 表示两者毫无关联; 若为 1 表示完全相关。NMI 可定义为

$$NMI = \frac{2I(s_i, r_i)}{H(s_i) + H(r_i)} \quad (5)$$

式中: H 为信息熵; I 是真实标签和聚类标签之间的互信息。NMI 衡量了样本标签的预测分布与真实分布的相关程度。

3.3 消融实验

为证实 AFCAE 网络具有优异的特征提取能力, 在 MNIST 和 COIL-20 数据集上从全连接层是否对称, 瓶颈层参数选择以及聚类层数的选择这 3 方面进行详尽的实验分析。

3.3.1 全连接层的选择

AFCAE 网络可以看成在 F2-F6 对称全连接部分前加入全连接层 F1。卷积自动编码器中添加的全连接层是将提取的特征进行整合。为了证实合理增加全连接层数有助于提高网络的聚类精度。保持编码器和解码器网络结构不变, 对全连接层部分尝试了不同深度以及对称和非对称式的设计。为了对比的公平性, 选取全连接层 F6 层的特征输出进行聚类分析并比较聚类精度, 实验结果列于表 2, 其中“ $d-50-c-50$ ”中的数字代表全连接层的神经元个数, d 是输入全连接层数据的维度, c 是瓶颈层神经元的个数, 括号中数值为进行 20 次实验取得的方差。黑体为本文选取网络全连接层结构以及对应的聚类精度。

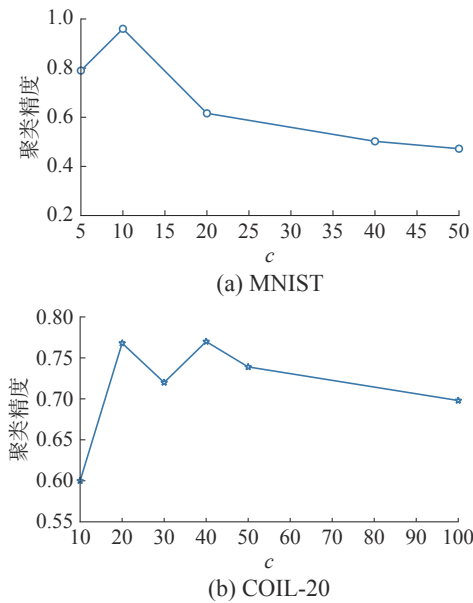
根据表 2 可知, 随着全连接层数的增加, 聚类精度先增长后有所下降, 说明全连接层的增加有助于网络提取有效的特征, 使得 K-means 算法在 MNIST 上有 0.960 的聚类精度。实验发现不断地增加全连接层的深度, 聚类精度呈下降趋势, 说明不能盲目增加网络深度。故本文全连接部分选取 $d-50-50-50-c-50-50$ 结构。

接着在 MNIST 和 COIL-20 数据集上验证瓶颈层神经元个数 c 的选择, 结果见图 3。

表 2 全连接层部分的设计和对应的聚类精度

Table 2 Design of fully connected layers and corresponding clustering accuracies

全连接层的设计结构	MNIST(ACC)	COIL-20(ACC)
$d-50-c-50$	0.923(7.5×10^{-4})	0.688(6.4×10^{-4})
$d-50-50-c-50$	0.935(6.9×10^{-4})	0.714(4.3×10^{-3})
$d-50-50-c-50-50$	0.934(5.8×10^{-4})	0.737(6.5×10^{-3})
$d-50-50-50-c-50-50$	0.960 (4.6×10^{-4})	0.768 (3.6×10^{-4})
$d-50-50-50-c-50-50-50$	0.910(6.8×10^{-4})	0.751(5.7×10^{-4})
$d-50-50-50-50-c-50-50-50$	0.896(7.9×10^{-4})	0.693(6.5×10^{-4})

图 3 瓶颈层神经元个数 c 影响聚类精度Fig. 3 Clustering accuracy affected by neuron numbers c in bottleneck layer

观察图 3(a) 和 (b) 不难发现, 在简单的 MNIST 数据集上, 随着神经元个数的增加, 聚类精度先上升后一直呈下降趋势, $c=10$ (基准类别数) 时出现最高精度; 而在相对复杂的 COIL-20 数据集上, 随着神经元个数的增加, 聚类精度并不稳定, c 为 20(基准类别数) 和 40 时均出现最高聚类精度。可见瓶颈层神经元个数影响网络的抽象表达能力, 也最终影响聚类性能。综合考虑, 后续实验中设定 AFCAE 网络的瓶颈层神经元个数 c 为聚类簇数。

3.3.2 聚类输入层的选择

本组实验在 MNIST 数据集上测试 AFCAE 网络中不同的全连接层的输出特征在聚类性能上的差异。在某一全连接层后接 K-means 算法, 进行深度聚类分析, 所得聚类性能如图 4 所示。显然, F6 层的特征使得聚类性能 ACC(0.960) 和 NMI(0.916) 均达到最优。在卷积层 C4 上获取的局部

特征, 全连接层通过权值矩阵将局部特征进行组合。随着全连接层数的增加, 特征表示不断抽象组合, 更趋向于全局特征的表示, 因此 F6 层所提取的特征可看作聚类的输入, 实验也证明了 F6 层上的聚类精度最高。

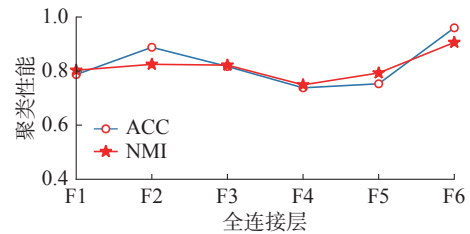


图 4 MNIST 数据集上每个全连接层的聚类精度

Fig. 4 Clustering accuracy of each fully connected layer on MNIST dataset

3.3.3 卷积部分的选择

本组实验在 MNIST 上验证卷积核与卷积层数对聚类性能的影响。全连接部分与前面的设置相同, 只改变卷积核大小、数量和卷积层数, 从而建立 A、B、C、D 和 E 共 5 个卷积部分, 参数的详细信息见表 3, 其中 $(k,n)/s$ 分别为卷积核的大小、数目和步长。

表 3 5 个不同卷积网络结构对比

Table 3 Comparison 5 different convolutional structures

卷积层网络	C1	C2	C3	C4	C5	C6
A	(25,3)/3	(50,3)/2	(50,3)/1	(50,3)/2		
B	(25,3)/3	(50,3)/2	(50,3)/1	(50,2)/2		
C	(25,3)/3	(25,2)/2	(25,3)/1	(50,2)/1		
D	(25,3)/3	(50,2)/2	(50,2)/1	(50,2)/1	(50,2)/1	
E	(25,3)/2	(25,3)/1	(25,3)/1	(25,3)/2	(50,3)/1	(50,2)/1

由于不同数据集的图片尺寸不统一, 不考虑输入图像尺寸, 仅考虑输入通道、输出通道以及卷积核大小时, 理论时间复杂性用式子 $n_{i-1}k_i^2n_i$ 来计算。比如网络 B 上的理论运行时间为: $1 \times 3^2 \times 25 +$

$25 \times 3^2 \times 50 + 50 \times 3^2 \times 50 + 50 \times 2^2 \times 50 = 43\,950$ 。

以网络 B 的理论时间为基准, 定义网络的复杂性, 计算公式为

$$\text{Comp}(\text{net}) = \frac{\text{net 的理论运行时间}}{\text{B 的理论运行时间}} \quad (6)$$

这 5 个网络的运行时间、网络复杂性以及聚类精度列于表 4。

表 4 卷积网络的运行时间, 复杂性和聚类精度

Table 4 Running time, complexity and clustering accuracy of the convolutional structures

网络	运行时间/s	复杂性	精度
A	56 475	1.284	0.900
B	43 950	1	0.960
C	13 350	0.303	0.930
D	35 225	0.801	0.942
E	38 350	0.872	0.935

由表 3 和表 4 可知, B 网络 C4 的 2×2 卷积核被 A 网络的 3×3 替代, 复杂性和运行时间增加了, 而聚类精度却大大降低; 而 C 网络只将 B 网络的 C2 和 C3 的卷积核数量减少为 25 个, 复杂性和理论运行时间大大减少, 同时聚类精度也降低; 对比 B 网络, D 网络多了一个 2×2 的卷积层, 但是 C2、C3 的卷积核也替换成 2×2 , 使得运行时间减少, 复杂性降低, 同时聚类精度也降低了; E 网络变化比较大, 将 C2、C3 和 C4 的卷积核个数减少为 25 个, 又增加了 C5 和 C6 卷积层, 运行时间和复杂性没有提高反而继续下降, 聚类精度也降低了。

通过对比卷积部分的复杂性和聚类精度, 发现小卷积核有利于网络提取适合聚类的特征, 犹如网络 A 到 B 的转换, 精度也随之增长。卷积核的数量对网络的特征提取也起到积极作用。但是一味地增加网络深度, 会导致网络过拟合, 从而精度下降。综上分析, 本文选取卷积网络 B 作为 AFCAE 的卷积部分进行后续实验。

3.4 对比实验与分析

AFCAE 与其他聚类方法在 MNIST 和 COIL-20 上的对比实验分析, 包括经典将 K-means 算法应用于原始图像聚类方法 KMS; 使用深度自编码器进行特征提取后, 使用 K-means 进行后续聚类的 DAE-KMS 算法, 在此基础上同时优化了数据重建误差和表示紧凑性的 AEC 算法; 以及深度表示和图像聚类的联合无监督学习 DEC、IDEC、DBC、JULE 和 ASCAE-softmax 算法。

AFCAE 算法的最优聚类性能见表 5 和表 6。同时, 选取文献 [20] 中的 DBC 在 MNIST 和 COIL-

20 上的聚类精度分别列于表 5 和表 6。其中黑体字表示最优的性能。

表 5 MNIST 数据集上各类聚类方法的对比

Table 5 Comparison clustering performances of several clustering methods on MNIST dataset

方法	ACC	NMI
KMS ^[18]	0.535	0.531
AEC ^[22]	0.760	0.669
IEC ^[23]	0.609	0.542
DEC ^[14]	0.840	—
DBC ^[20]	0.766	0.759
DAE-KMS ^[14]	0.818	—
IDEC ^[15]	0.880	0.867
JULE ^[24]	—	0.913
ASCAE ^[20]	0.925	0.854
ASCAE-softmax ^[20]	0.960	0.910
AFCAE	0.960	0.916

表 6 COIL-20 数据集上 6 个聚类算法的对比

Table 6 Comparison clustering performances with six methods on COIL-20 dataset

方法	ACC	NMI
KMS ^[18]	0.592	0.767
DEC ^[20]	0.731	0.813
DBC ^[20]	0.724	0.822
ASCAE ^[20]	0.740	0.823
ASCAE-Softmax ^[20]	0.755	0.833
DEN ^[17]	0.725	0.870
AFCAE	0.768	0.868

由表 5 可知, AFCAE 网络的聚类精度达到 0.96, 超过联合训练的其他深度聚类算法, 进一步证实了 F6 层的聚类效果。图 5 是 AFCAE 网络的 F1 层和 F6 层聚类簇的二维可视化图, F1 全连接层的聚类簇大致可以区分开, 但各簇类间距较小且分布杂乱。F6 中仅有少量特征散乱分布, 各簇间有明显的分界线。可视化图直观地证实了全连接层帮助卷积自编码器整合所提取的特征。

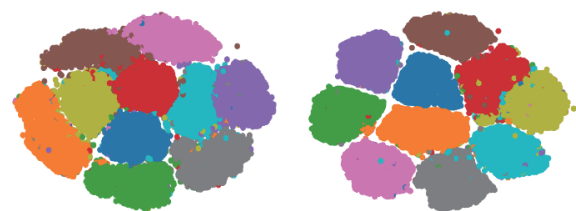


图 5 MNIST 数据集的聚类簇可视化图

Fig. 5 Visualization of clustering results on MNIST dataset

在 COIL-20 上, 对 AFCAE 的 F6 层特征进行 K-means 聚类, 仅得到 0.624 的聚类精度。在 C1-C4 层加入 BN 层改善网络梯度的变化范围, 进而改善网络抽取局部特征的能力, 微调结构后聚类精度达到了 0.768。聚类可视化如图 6 所示。



图 6 带 BN 层的 AFCAE 在 COIL-20 上的聚类可视化
Fig. 6 Visualization of clustering results of AFCAE with BN layer on COIL-20 dataset

根据表 6 不难发现, AFCAE 的聚类精度 0.768 高于联合训练的 DEC 方法的 0.731, 也略高于 ASCAE-softmax 方法的 0.755。但是 NMI 不及 DEN 方法的。由于 COIL-20 数据集由外形简单但不同角度的物品图片组成, 实验过程中发现一些物品轮廓相似, 导致的特征之间区分度较小, 无法清晰地分成不同的类簇。因此对于轮廓相似的物品图像仍需进一步研究可辨识的特征。

AFCAE 网络在 4 个人脸数据集上也进行了 2 组实验。第 1 组验证了每个全连接层的聚类精度, 见图 7; 第 2 组把 F6 层的特征输出后进行 K-means 聚类分析, 聚类的 ACC 和 NMI 列于表 7。

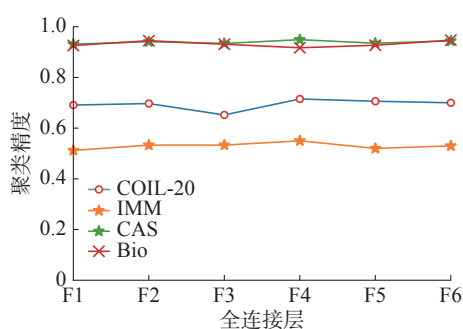


图 7 全连接层的特征输出的聚类精度
Fig. 7 Clustering accuracy of feature output of fully connected layer

根据图 7 可知 AFCAE 网络具有较好的整体稳定性。表 7 显示在 CAS-PEAL-R1 和 BioID-Face 上 AFCAE 的聚类性能不错, 而在 IMM 和 UMISTS 上不太令人满意, 可能是这两个数据集的图片均有不同程度的表情和姿态, AFCAE 网络没能捕获到合适的特征。这将是未来工作之一。

表 7 AFCAE 算法在 4 个人脸数据集上的聚类性能
Table 7 Clustering performances of AFCAE algorithm on four face datasets

数据集	ACC	NMI
CAS-PEAL-R1	0.930	0.962
BioID-Face	0.882	0.960
IMM-Face	0.580	0.795
UMISTS	0.488	0.686

4 结束语

本文提出一个非对称全连接层聚类网络 AFCAE, 结合 K-means 并由此提出一个 AFCAE 深度聚类方法。该方法通过改善网络结构, 提取更有辨识力的聚类特征来提高聚类性能。在 MNIST 和 COIL-20 上通过详细对比和分析全连接部分, 卷积部分包括卷积核大小和数目、卷积层数, 特征输出层 F6 的选择。验证了本文的 AFCAE 方法降低运行时间的基础上提高聚类性能, 而且还优于深度聚类算法 DEC 和 ASCAE 的聚类性能。但实验中也发现小卷积核网络对形变较大或类别不确定性较大的数据集的聚类效果不令人满意, 这将是未来工作之一。

参考文献:

- [1] LLOYD S. Least squares quantization in PCM[J]. *IEEE transactions on information theory*, 1982, 28(2): 129–137.
- [2] HAEUSSER P, PLAPP J, GOLKOV V, et al. Associative deep clustering: training a classification network with no labels[C]//German Conference on Pattern Recognition. Cham: Springer, 2019: 18–32.
- [3] REYNOLDS D. Gaussian mixture models[M]//Encyclopedia of Biometrics. Boston, MA: Springer US, 2009: 659–663.
- [4] LIU Peng, ZHOU Dong, WU Naijun. VDBSCAN: varied density based spatial clustering of applications with noise[C]//2007 International Conference on Service Systems and Service Management. Chengdu, China. IEEE, 2007: 1–4.
- [5] ABDI H, WILLIAMS L J. Principal component analysis[J]. *Wiley interdisciplinary reviews:computational statistics*, 2010, 2(4): 433–459.
- [6] ALQAHTANI A, XIE X, DENG J, et al. A deep convolutional auto-encoder with embedded clustering[C]//2018 25th IEEE International Conference on Image Processing. Athens, Greece. IEEE, 2018: 4058–4062.
- [7] YU Tianqi, WANG Xianbin, SHAMI A. UAV-enabled spatial data sampling in large-scale IoT systems using denoising autoencoder neural network[J]. *IEEE Internet of*

- things journal, 2019, 6(2): 1856–1865.
- [8] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278–2324.
- [9] MASCI J, MEIER U, CIREŞAN D, et al. Stacked convolutional auto-encoders for hierarchical feature extraction[M]//Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011: 52–59.
- [10] LEE Honglak, EKANADHAM C, NG A Y. Sparse deep belief net model for visual area V2[C]//Proc of Conf on Advances in Neural Information Processing Systems. Washington D. C., USA: MIT Press, 2007: 873–880.
- [11] VINCENT P, LAROCHELLE H, BENGIO Y, et al. Extracting and composing robust features with denoising autoencoders[C]//Proceedings of the 25th international conference on Machine learning-ICML '08. Helsinki, Finland. New York: ACM Press, 2008: 1096–1103.
- [12] BENGIO Y, LAMBLIN P, POPOVICI D, et al. Greedy layer-wise training of deep networks[C]//Proc of Advances in Neural Information Processing Systems. Washington, USA: MIT Press, 2006: 153–160.
- [13] MA Xiaolei, DAI Zhuang, HE Zhengbing, et al. Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction[J]. *Sensors*, 2017, 17(4): 818.
- [14] XIE Junyuan, ROSS G, ALI F. Unsupervised deep embedding for clustering analysis[C]//Proc of ICML'16 Proc of the 33rd Int Conf on Int Conf on Machine Learning. New York City, NY: Semantic Scholar, 2016: 478–487.
- [15] GUO Xifeng, GAO Long, LIU Xinwang, et al. Improved deep embedded clustering with local structure preservation[C]//IJCAI'17: Proceedings of the 26th International Joint Conference on Artificial Intelligence. New York: ACM, 2017: 1753–1759.
- [16] YANG Bo, FU Xiao, NICHOLAS D S, et al. Towards K-means-friendly spaces: simultaneous deep learning and clustering[C]//Proc of ICML'17 Proc of the 34th Int Conf on Machine Learning. Sydney, Australia: TonyJebara, 2016: 3861–3870.
- [17] HUANG Peihao, HUANG Yan, WANG Wei, et al. Deep embedding network for clustering[C]//2014 22nd International Conference on Pattern Recognition. Stockholm, Sweden. IEEE, 2014: 1532–1537.
- [18] LI Fengfu, QIAO Hong, ZHANG Bo. Discriminatively boosted image clustering with fully convolutional auto-encoders[J]. *Pattern recognition*, 2018, 83: 161–173.
- [19] VAN L, MAATEN D, GEOFFREY H. Visualizing data using t-SNE[J]. *Journal of machine learning research*, 2008, 9(2605): 2579–2605.
- [20] 陈俊芬, 赵佳成, 韩洁, 等. 基于深度特征表示的 Softmax 聚类算法 [J]. *南京大学学报 (自然科学版)*, 2020, 56(4): 533–540.
- CHEN Junfen, ZHAO Jiacheng, HAN Jie, et al. Softmax clustering algorithm based on deep features representation[J]. *Journal of Nanjing university (natural science edition)*, 2020, 56(4): 533–540.
- [21] HE Kaiming, SUN Jian. Convolutional neural networks at constrained time cost[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA. IEEE, 2015: 5353–5360.
- [22] SONG Chunfeng, LIU Feng, HUANG Yongzhen, et al. Auto-encoder based Data clustering[C]//Iberoamerican Congress on Pattern Recognition. Berlin, Heidelberg: Springer, 2013: 117–124.
- [23] LIU Hongfu, SHAO Ming, LI Sheng, et al. Infinite ensemble for image clustering[C]//KDD'16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2016: 1745–1754.
- [24] YANG Jianwei, PARIKH D, BATRA D. Joint unsupervised learning of deep representations and image clusters[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. IEEE, 2016: 5147–5156.

作者简介:



杨梦茵, 硕士研究生, 主要研究方向为图像聚类和机器学习。



陈俊芬, 副教授, 博士, CCF 会员, 主要研究方向为数据挖掘、机器学习和图像处理。主持河北省留学回国基金 1 项。发表学术论文 10 余篇。



翟俊海, 教授, 博士生导师, 博士, 河北大学学术委员会委员, 中国人工智能学会知识工程与分布智能专业委员会委员、粒计算与知识发现专业委员会委员, 主要研究方向为大数据处理、机器学习、深度学习。主持省重点自然科学基金项目 1 项和省自然科学基金项目 2 项, 近 3 年发表学术论文 10 余篇。