

DOI:10.13232/j.cnki.jnju.2020.04.011

基于深度特征表示的 Softmax 聚类算法

陈俊芬*, 赵佳成, 韩 洁, 翟俊海

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

摘 要: 图像和语音已成为日常生活和科研的常见数据类型, 图像的聚类分析是数据挖掘和图像处理领域的重要任务之一. 基于自编码器的深度聚类方法具有表征能力有限的缺点, 并且特征的生成与聚类指派是分步进行的. 为此, 提出一种基于新颖卷积自编码器的深度 Softmax 聚类算法 (Asymmetric Convolutional Auto-encoder Based Softmax Clustering, ASCAE-Softmax). 首先设计一种非对称的卷积自编码器网络结构 (ASCAE), 通过优化卷积和添加全连接层, 使整个网络呈非对称; 接着使用 Softmax 聚类器把特征映射成聚类概率分布, 构造辅助目标概率分布, 将特征学习与聚类判别联合在一起. 通过迭代最小化 KL (Kullback-Leibler) 散度损失达到清晰的聚类划分. 实验结果表明, 该方法能够学习出使同类更加紧凑、异类更加稀疏的特征表示, 且聚类结果优于经典的深度聚类算法.

关键词: 无监督学习, 特征表示, 卷积自编码器, 图像聚类, Softmax 分类器

中图分类号: TP391

文献标识码: A

Softmax clustering algorithm based on deep features representation

Chen Junfen*, Zhao Jiacheng, Han Jie, Zhai Junhai

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

Abstract: Image and speech have been common data in daily life and academic research. Therefore, image clustering analysis becomes one of the vital tasks in data mining and image processing fields. The deep clustering methods based on auto-encoders have the limited representation ability. Moreover, feature extraction and clustering assignment are carried out separately. A new deep Softmax clustering algorithm (ASCAE-Softmax) based on a novel convolutional auto-encoder is proposed. Firstly, an asymmetric convolutional auto-encoder network structure (ASCAE) is designed. The whole network is asymmetric with an optimizing convolution operation and adding fully connected layers. Secondly, Softmax clustering is proposed that is composed of mapping features into clustering probability distribution, making auxiliary target probability distribution, and combining features learning with clustering assignment. Then, clustering divisions become clearer by iteratively minimizing KL (Kullback-Leibler) divergence. Experimental results showed that the proposed deep clustering algorithm can achieve the optimal features representation which makes the intra-clusters more compact and the inter-clusters more dispersive, and the clustering result is better than the state-of-the-art deep clustering algorithms.

Key words: unsupervised learning, features representation, convolutional auto-encoder, image clustering, Softmax classifier

由于实际问题中存在大量无标签数据, 迫使众多研究的对象从大量带标签的数据转为仅有少

量标签或没有标签的数据, 研究方法也从监督学习变成半监督学习、弱监督学习和无监督学习, 而

基金项目: 河北省科技重点研发项目 (19210310D), 河北大学高层次创新人才科研启动经费项目

收稿日期: 2020-06-24

* 通讯联系人, E-mail: chenjunfen2010@126.com

无监督学习的图像任务有更重要的现实意义. 无监督学习在各个方向有各种不同的任务, 例如图像分类^[1-3]、图像聚类^[4-9]、图像分割^[10-11]、目标检测^[12]等, 虽然任务不同但目的都在于提取更有用的特征、简化后续的相关操作进而更好地完成任务. 而在图像聚类任务中, 如何提取出使同类内间距减小、异类间距增大的特征尤为重要.

无监督聚类最早的思路是关注于算法本身的 K-means, DBSCAN (Density-Based Spatial Clustering of Applications with Noise), 而后的聚类研究使用受限玻尔兹曼机^[13]、自编码器^[14]提取特征. LeNet-5^[15]的卷积、无监督特征学习卷积神经网络^[16]验证了卷积神经网络在无监督学习方面的可行性. 深度卷积神经网络拥有强大的图像特征提取能力^[17], 结合自编码器^[14]和深度网络的卷积自编码器^[18]提供了全新的图像特征提取方式. 生成对抗网络^[19]也用于特征提取. 深度卷积对抗网络的无监督表示学习^[9]证明了其可行性. 无监督聚类方法^[4]综合使用了聚类和分类来学习视觉特征, 在杂乱数据上进行无监督的图像特征预训练方法^[20]以及无监督图像分类和分割的不变信息聚类方法^[21]上等变换的思路. 至此, 无监督聚类学习从特征与聚类分步进行发展到联合训练的方向.

尽管这些优秀的深度聚类工作取得了很好的进展, 但是随着图像数据类型的多元化, 特征表示能力和聚类性能仍有提升空间. 因此本文继续探索新的网络和合适的损失函数, 提出一种基于新颖卷积自编码器的深度 Softmax 聚类算法 (Asymmetric Convolutional Auto-encoder Based Softmax Clustering, ASCAE-Softmax). 本文的主要贡献如下:

(1) 采用新的卷积形式, 通过首个卷积核与步长相等实现图像的分割式特征提取, 进一步提高了特征的鲁棒性.

(2) 提出一种非对称的卷积自编码器结构, 通过编码器和译码器的信息不对等使得所学特征更专注于特征的分解和生成.

(3) 利用 Softmax 分类的思路解决聚类问题, 将所学到的特征通过高斯方式映射成条件概率, 并获得所有数据在同一类别上的概率分布. 减小

同类间隔, 增大异类之间的间隔. 不断优化该网络的特征表示, 提高了聚类精度.

1 相关工作

以往的无监督聚类算法在特征的提取与聚类操作上相互独立, 如何将深度卷积网络提取特征的训练与聚类过程联合在一起成为无监督聚类学习的重要方向. 2016 年, Xie et al^[8]以卷积自编码器为基础, 在生成特征的层面上做了新的改进, 首次找到联合解决特征空间学习与聚类关系判别的方案 DEC (Unsupervised Deep Embedding for Clustering Analysis) 算法. 通过深度自编码器初始化数据到特征的映射, 使用 K-means 生成特征的聚类中心点. 使用 t -分布^[22]衡量特征点与特征聚类中心点的相似性, 通过软分配方式生成每个数据的聚类概率分布, 构造辅助目标分布, 使用 KL (Kullback-Leibler) 散度损失进行优化以细化聚类中心点, 同时对生成特征的网络参数进行不断调整来优化聚类结果, 直至网络收敛. 其最终的聚类结果的精度超越了以往的聚类方法, 更是无监督学习的一大进步, 因为特征的学习与聚类的判别不再是相互独立的, 而是可以联合在一起进行训练和优化. 但其使用的深度自编码器在特征的提取上仍有局限性.

为了优化这个问题, 2017 年, 联合卷积自动编码器嵌入和相对熵进行深度聚类方法 (Deep Clustering via Joint Convolutional Autoencoder Embedding and Relative Entropy Minimization, DEPICT)^[5]在对原始的无监督深度嵌入聚类算法整体思路不变的情况下使用多项逻辑回归函数计算特征点与聚类中心点的相似性; 同时将网络修改为深度卷积去噪自编码器; 输入数据方使用不包含噪声的输入数据生成辅助目标分布来优化带有噪声的输入数据的聚类概率分布. 2018 年全卷积自编码器的判别增强图像聚类方法 (Discriminatively boosted clustering, DBC)^[6]将深度自编码器优化为深度卷积自编码器, 通过判别式增强分配的方式提高聚类分配的纯度. 在其他方面不变的情况下, 得到了更优的结果.

现有的深度聚类算法大部分采用经典 K-means 算法对所提取的特征进行聚类, 并且得益

于聚类结果对优化网络权值的指导作用. 在对文献[5—8]的实验复现中发现, 深度聚类结果很大程度上取决于深度卷积自编码器初始生成特征的 K-means 聚类结果. 为此, 本文不仅在深度卷积自编码器的网络结构、卷积方式等方面做了改进, 而且在无监督深度嵌入聚类算法上采用了适用性更强 Softmax 方式.

2 深度嵌入 Softmax 聚类算法

2.1 非对称卷积特征(ASCAE)提取器 与传统的深度卷积自编码器网络相比, 本文的网络添加多个不同数量神经元的全连接层, 将整个网络修改为非对称的结构, 整个网络结构如图1所示.

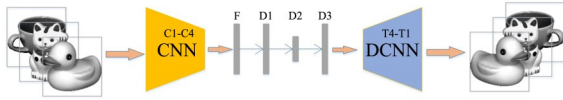


图1 非对称的卷积自编码网络(ASCAE)框架图

Fig. 1 A framework of the Asymmetric Convolutional Auto-encoder (ASCAE) network

其中, 编码器包括卷积层 C1-C4 和全连接层 F 层、D1 和 D2, 而且 F 层为 C4 的特征图变成的特征向量. 解码器包括全连接层 D2 和 D3 和反卷积层 T4-T1, 而且 D3 层的神经元个数与 F 层的相同. F 层、D1 层、D3 层为 50 个神经元, D2 层的神经元个数为类别个数左右. 由于 D2 层包含的神经元个数较少, 信息量被大大地压缩, 所以本文均使用 F 层的特征为输入图像的特征表示.

传统的卷积神经网络是模仿生物的视觉形成过程发展起来的, 通常卷积与池化操作并存且交替进行. 但有研究表明池化操作会丢失相邻数据的部分信息^[23], 因此以较大步长的卷积代替卷积和池化这一组合操作. 本文 ASCAE 网络中 C1 和 T1 层的步长与卷积核大小相同, 设定为 3. 设定步长等于卷积核大小, 相当于将原始图片划分为不相交的若干部分, 然后对各部分提取特征. 该设定降低了卷积网络的计算量, 增加了所提取特征的整体上的鲁棒性. 具体网络参数设置如表 1 所示.

所有待聚类的图片均变换为 28×28 的灰度图片且不做归一化. 在 28×28 图像上进行 3×3 卷

表1 ASCAE 网络参数设置

Table 1 Parameters settings of ASCAE network

Layers	Kernel	Strides
C1	$25 \times 3 \times 3$	3
C2	$50 \times 3 \times 3$	2
C3	$50 \times 3 \times 3$	2
C4	$50 \times 2 \times 2$	1
F	50	—
D1	50	—
D2	K	—
D3	50	—

积操作时舍弃边缘不重要的像素点. 激活函数采用无上限的 ReLU 函数. 不仅保证图片被顺利还原为 $[0, 255]$ 的范围, 还加快了网络的收敛速度. 损失目标是输入与输出之间的误差平方和函数:

$$J(\theta; X) = \frac{1}{N} \sum_{i=1}^N \|x_i - f_{\theta}(x_i)\|^2 \quad (1)$$

其中, $\theta = \{\theta_e, \theta_d\}$ 是网络参数.

在训练好的 ASCAE 网络上, 用 F 层的输出作为特征信息, 且保留 C1-F 层的网络结构和网络权重进行无监督深度嵌入聚类算法.

2.2 深度聚类 Softmax 算法(ASCAE-Softmax)

已知含有 N 个图像的集合 $X = [x_1, x_2, \dots, x_N]$, 通过 ASCAE 取得特征 $Z = [z_1, z_2, \dots, z_N]$, 获得关系映射为 $\varphi_{\theta}: X \rightarrow Z$. 经典 K-means 算法对 Z 聚类, 得到聚类中心集合 $C = \{c_1, c_2, \dots, c_K\}$ 和类别硬指派 $Y = \{y_1, y_2, \dots, y_N\}$, 即每个 z_i (即 x_i) 所属的 $y_i \in \{1, 2, \dots, K\}, i = 1, 2, \dots, N$.

在 ASCAE 网络后接一个带参数的 Softmax 分类器, 对特征集 $\{z_i, y_i\}_{i=1}^N$ 进行分类, 计算特征点属于某一类的概率如式(2)所示:

$$p_{ik} = p(y_i = k | z_i, W) = \frac{\exp(w_k^T z_i)}{\sum_{l=1}^K \exp(w_l^T z_i)} \quad (2)$$

其中, $W = [w_1, w_2, \dots, w_K]$ 为 Softmax 分类器的未知参数.

为了进一步减小同类的间隔, 增大异类的间隔, 先用同一类中数据的概率除以该类所有数据概率的总和即软聚类频率, 接着重新计算每个数据属于不同类别的概率分布. 即对分类概率矩阵

$P = [p_{ik}]_{N \times K}$, 其中每个特征的分类概率为 $p_i = [p_{i1}, p_{i2}, \dots, p_{iK}]$, 先对 P 每一个列向量归一化:

$$\bar{q}_{ik} = \frac{p_{ik}}{\sum_{i=1}^N p_{ik}} \quad (3)$$

此时, 特征属于类别的概率之和不为 1, 继续对其进行行向量归一化:

$$q_{ik} = \frac{\bar{q}_{ik}}{\sum_{i=1}^K \bar{q}_{il}} \quad (4)$$

得到辅助概率分布 Q .

该深度嵌入 Softmax 聚类方法的目标函数为:

$$L(\theta, W; Z) = \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K q_{ik} \lg \frac{q_{ik}}{p_{ik}} + \frac{\lambda}{2} \sum \|\theta\|_2^2 \quad (5)$$

其中, 第一项是 KL 散度描述这两个概率分布的差异性. 第二项是网络参数的 L2 正则项能稀疏化网络, 提高网络的泛化能力, 避免网络权值的爆炸式变化.

用随机梯度下降法更新网络参数和 Softmax 分类参数为:

$$\frac{\partial L}{\partial z_i} = \sum_{j=1}^K (q_{ik} - p_{ik}) w_k^T \exp(2w_k^T z_i) + \lambda \sum \theta \quad (6)$$

$$\frac{\partial L}{\partial w_k} = \sum_{j=1}^K (q_{ik} - p_{ik}) z_i \exp(2w_k^T z_i) \quad (7)$$

根据式(6)和链导法则更新 ASCAE 网络权值, 再根据式(7)更新 P 值, 然后计算 Q 值, 用 Q 值继续更新网络权重和分类参数循环往复, 直到 Q 值变化逐渐收敛至某一阈值. 流程如图 2 所示.

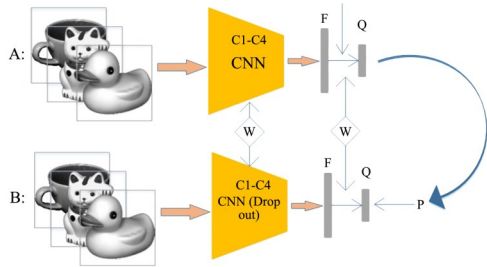


图 2 特征优化

Fig. 2 Feature optimization

3 实验与分析

本文在六个经典的图像数据集上进行实验, 评价聚类性能的数值指标包括聚类精度 (Accura-

cy, ACC) 和标准互信息 (Normalized Mutual Information, NMI). ACC 表示聚类结果的正确率, 而 NMI 衡量了预测分布与真实分布的相关程度. 两个指标值越接近 1, 说明聚类准确度越高. 实验环境为 Intel Core i5-6300HQ 处理器, NVIDIA 2.0 GB 显存, 8.0 GB RAM 内存; 基于 TensorFlow 2.0 搭建网络.

3.1 MNIST 数据集 在经典的手写数字集 MNIST 上, 设 D2 层的神经元个数等于类别数 10, 用算法 1 端到端训练 ASCAE 直至收敛. 对 F 层的特征进行 K-means 聚类分析 (重复 50 次, 取最好的聚类性能), 取编码器 F 层及其左半部分, 生成 Softmax 的聚类网络结构 (如图 2 所示), 用算法 2 优化 ASCAE 权值和 Softmax 参数 W , 直至满足停止条件.

算法 1 ASCAE 聚类算法 (初始化 ASCAE-Softmax 算法)

输入: 图像数据集、类簇个数 K

输出: 网络权值、特征表示; 类中心和聚类指派

① 使用端到端策略训练 ASCAE, 对目标函数 (1) 用随机梯度下降的算法最小化迭代直到收敛;

② 输出网络参数值和 F 层的特征表示;

③ K-means 算法对特征进行聚类, 得到聚类中心和聚类指派.

算法 2 ASCAE-Softmax 聚类算法

输入: 编码器的参数值和特征表示

输出: 类标向量和类簇中心

联合训练过程:

① 根据式 (2) 在 A 和 B 网络计算真实概率分布;

② 根据式 (3) 和式 (4) 将 A 网络的真实分布生成辅助概率分布;

③ 根据式 (6) 和式 (7) 在 B 网络上更新网络参数和聚类中心, 不满足停止条件时回到 ①;

④ 对于每一个数据取概率分布最大值位置为所属类簇.

ASCAE 算法的聚类指标和 ASCAE-Softmax 算法获取 P 值计算最终聚类性能指标列于表 2. 同时, 本文还在相同的实验环境下复现了 DEC^[8], DBC^[6] 和 DEPICT^[5] 三个深度聚类算法在 MNIST 上的聚类性能, 并列于表 2. 最后, 将 Li et al^[6] 的三个以往算法的聚类性能也列于表 2, 其中黑体字表示最优的性能.

表2 在 MNIST 数据集上比较 ASCAE, ASCAE-Softmax 和六个聚类算法的聚类性能

Table 2 Clustering performances of ASCAE, ASCAE-Softmax and other six clustering algorithms on MNIST dataset

Algorithms	ACC	NMI
KMS ^[6]	0.535	0.531
AEC ^[6]	0.760	0.669
IEC ^[6]	0.609	0.542
DEC	0.889	0.856
DBC	0.766	0.759
DEPICT	0.924	0.850
ASCAE	0.925	0.854
ASCAE-Softmax	0.960	0.910

其中, KMS 算法是经典 K-means 算法直接对原始图像进行聚类; AEC (Auto-encoder Based Data Clustering) 算法是先用深度自编码器提取特征, 然后用 K-means 聚类, 而训练深度自编码器的目标函数由数据重构误差和紧致性两部分组成; IEC (Infinite Ensemble For Image Clustering) 算法包含深度表示学习和聚类融合; DEC, DBC, DEPICT 算法均为同时学习特征表示和聚类中心, 采用深度自编码器学习特征表示和软 K-means 聚类, 其中 DEC, DEPICT 为参考源代码的复现, 得到与原论文相近的实验结果, DBC 在没有源代码可参考的情况下尽可能地复现所有提到的细节, 但结果不尽人意, 可能是没抓到某些训练小细节, 导致实验出现了一些偏差。

由表2可以看出, 本文的 ASCAE-Softmax 算法聚类性能最高 (与 DBC 算法^[6]最高的聚类精度 0.964 相当), 说明 Softmax 分类结果有效地引导了网络更新, 获得了更有辨识力的特征表示。而 ASCAE 算法 (直接对编码器所提取的特征进行 K-means 聚类) 稍差, 但都优于另外六个算法。且在与其它深度聚类算法相比较时, 其他算法注重联合训练的优化, 本文不仅对联合训练进行了优化, 还对自编码器进行了大量的优化, 进一步提高了联合训练时的下限, 在初次联合训练时就能得到更好的概率分布。此外, KMS 的聚类性能最差, 说明以像素为特征时, 很难区分同类图像的相似和异类之间的差异。

图3和图4是 ASCAE 和 ASCAE-Softmax 的 F 层的二维特征可视化。图3中虽然各个团簇间有明显界限, 但仍有一些数据散乱分布在整个特征空间。而图4中, 散落在整个特征空间中的点少了很多, 且较杂乱的点分布的位置相对集中, 对于后续的数据处理有很大的优势。可视化图清晰地印证了 ASCAE-Softmax 优于 ASCAE 的原因。



图3 MNIST 在 ASCAE 训练后的 F 层特征

Fig. 3 F-layer features of ASCAE on MNIST dataset

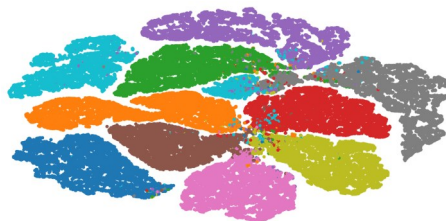


图4 MNIST 上 ASCAE-Softmax 的 F 层特征

Fig. 4 F-layer features of ASCAE-Softmax on MNIST dataset

3.2 COIL-20 物品数据集 COIL-20 数据集是 20 个外形简单旋转角度不同的物品图像, 每个物品有 72 张 128×128 灰度图片, 部分图片展示于图5中。实验前先将图像缩小为 28×28 , 且 D2 层的神经元个数设为 20。本文还在相同的实验环境下复现了 DEC^[8], DBC^[6] 和 DEPICT^[5] 三个深度聚类算法在 COIL-20 上的聚类性能, 并列于表3。四个聚类算法和以往常用的两个聚类算法的聚类性能也列于表3, 其中黑体字表示最优的性能。

根据表3可知, 联合训练的 ASCAE-Softmax 和 ASCAE 和 DEN 相比, 聚类精度提高了 1.5% 和 3.0%, KMS 仍是最差。在相同的实验环境下, 与 DEC, DBC, DEPICT 相比亦在聚类精度上获得了些许提升, 很大程度上说明了初始聚类概率分布对于最终结果的影响。虽然 ACC 精度优于其他六个算法, 但 NMI 不及 DEN 算法。而且在实验中发现, 较小的数据量不利于获取辅助目标的

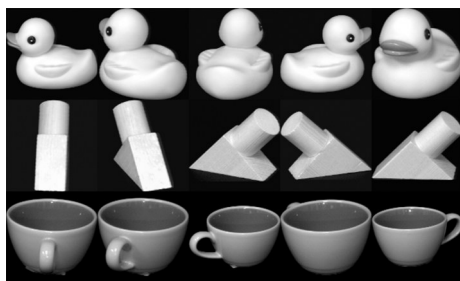


图 5 COIL-20数据集中的部分图片

Fig. 5 Some images of COIL-20 dataset

表 3 ASCAE, ASCAE-Softmax 和六个聚类算法在 COIL-20 上的聚类性能

Table 3 Clustering performances of ASCAE, ASCAE-Softmax and other six clustering algorithms on COIL-20 dataset

Algorithms	ACC	NMI
KMS ^[6]	0.592	0.767
DEN ^[6]	0.725	0.870
DEC	0.731	0.813
DBC	0.724	0.822
DEPICT	0.749	0.825
ASCAE	0.740	0.823
ASCAE-Softmax	0.755	0.833

概率分布,使优化聚类特征较为困难.当自编码器提取出优异的特征,提供了较高的聚类精度,后续的联合优化能有限地提升聚类精度.

图6展示了COIL-20数据集在深度嵌入聚类算法的过程可视化.可以看出不同类别特征之间的距离逐渐增大,相同类别的特征逐渐变得紧凑.特征空间是逐渐向同类紧凑、异类稀疏的方向发展.

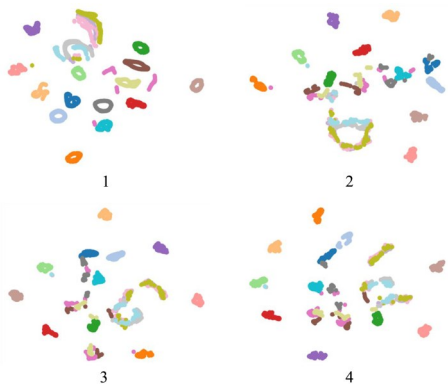


图 6 COIL-20 上 ASCAE-Softmax 聚类过程的可视化

Fig. 6 Visualization of ASCAE-Softmax clustering process on COIL-20 dataset

3.3 人脸数据集 在四个形态、角度各异的人脸图像集(CAS-PEAL-R1(<http://www.jdl.ac.cn/peal/index.html>), BioID-Face(<https://www.bioid.com/About/BioID-Face-Database>), IMM-Face(<http://www.imm.dtu.dk/~aam/aamexplorer/>), UMIST(<https://blog.csdn.net/garfielder007/article/details/51480770>)上进行了第三组实验,数据集的差异性如图7所示.先将人脸图像缩小为 28×28 的灰度图像.其次,通过镜像、旋转、调整亮度、饱和度等操作,将每个人的图片量扩充50倍.此组实验中,D2层的神经元个数为20.聚类性能如表4所示,其中黑体字表示最优的性能.

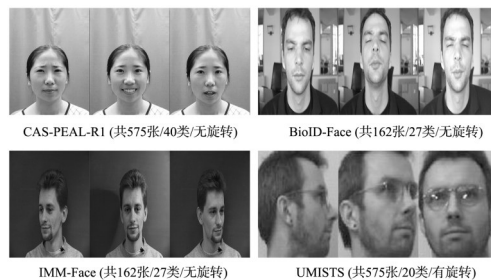


图 7 四个人脸图像集上的部分图片示例

Fig. 7 Image examples from the four human face datasets

表 4 ASCAE 和 ASCAE-Softmax 算法在四个人脸图像集上的聚类性能

Table 4 Clustering performances of ASCAE and ASCAE-Softmax on four face datasets

	ASCAE		ASCAE-Softmax	
	ACC	NMI	ACC	NMI
CAS-PEAL-R1	0.890	0.958	0.900	0.959
BioID-Face	0.852	0.949	0.895	0.950
IMM-Face	0.550	0.760	0.571	0.763
UMISTS	0.442	0.647	0.447	0.654

从表4可以看出,在无旋转的人脸数据集CAS-PEAL-R1和BioID-Face上可以获得较好的实验结果,但是在有旋转角度的IMM-Face, UMISTS上,得到与MNIST和COIL-20上相同的结论.即有旋转角度时,聚类精度较差;无角度的数据上有很高的聚类精度.结合了Softmax的联合训练虽然对于精度的提升不是很大,但在特

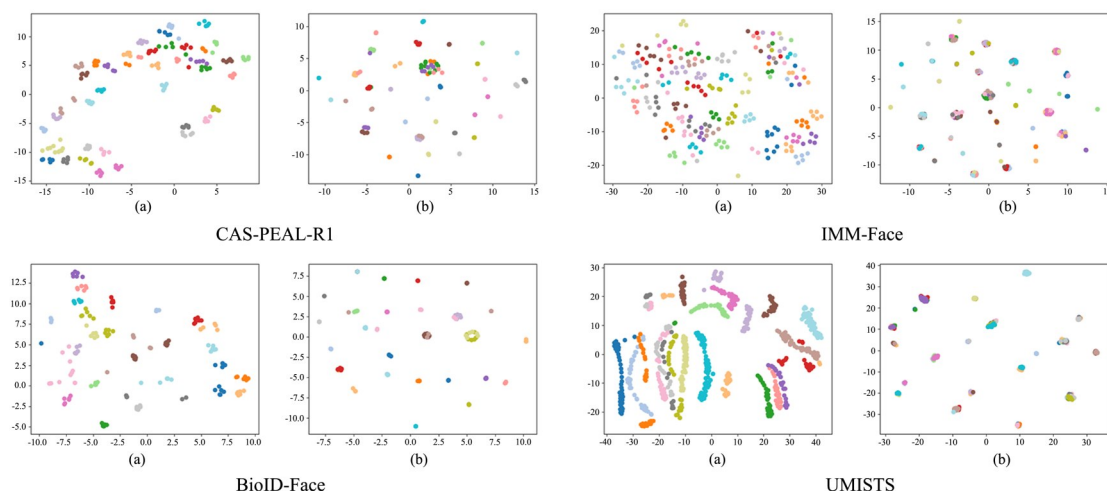


图8 四个人脸数据集上 ASCAE 和 ASCAE-Softmax 的 F 特征可视化

Fig. 8 Visualization of F features: (a)ASCAE,(b) ASCAE-Softmax on the four face datasets

征的可视化方面贡献突出,表现为特征在同类上更为紧凑,在异类上更为稀疏,从而类别间更易划分。

4 结 论

本文提出一种基于非对称卷积自编码器的特征提取器. 用卷积核大小与步长相同的卷积方式在传统的卷积编码器上加了多个全连接层,且全连接层的神经元个数不一,整个网络呈非对称结构. 在获取聚类特征后,采用深度 Softmax 聚类方法. 用 Softmax 函数计算每个数据点的聚类概率分布,使用不含 Dropout 的网络获取构造辅助目标分布,优化有 Dropout 层的网络. 不断优化特征得到了更优的聚类概率分布和特征生成空间。

实验表明,本文提出的基于非对称卷积编码器的深度 Softmax 聚类算法在没有角度的图像集上有很好的结果,优于其他深度聚类算法. 在有旋转的图像上,聚类精度不是特别突出但仍能优于部分深度聚类算法. 在本文的六个图像集生成的特征空间上,拥有同类紧凑异类稀疏的独特优势. 因此本文的方法是行之有效的图像深度聚类方法之一。

显然,图像类型的多样性带来更多的不确定性,增加了图像级别知识颗粒的不可分辨性. 如何有效地进行图像聚类分析将是下一步的研究内容。

参考文献

- [1] Zhang L H, Qi G J, Wang L Q, et al. AET vs . AED: unsupervised representation learning by auto-encoding transformations rather than data//2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA: IEEE, 2019:2547—2555.
- [2] Kang G L, Jiang L, Yang Y, et al. Contrastive adaptation network for unsupervised domain adaptation//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA: IEEE, 2019:4893—4902.
- [3] Wu Z R, Xiong Y J, Yu S X, et al. Unsupervised feature learning via non - parametric instance discrimination//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018:3733—3742.
- [4] Caron M, Bojanowski P, Joulin A, et al. Deep clustering for unsupervised learning of visual features//The 15th European Conference on Computer Vision (ECCV). Springer Berlin Heidelberg, 2018:132—149.
- [5] Dizaji K G, Herandi A, Deng C, et al. Deep clustering via joint convolutional autoencoder embedding and relative entropy minimization//2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017:5736—5745.

- [6] Li F F, Qiao H, Zhang B, et al. Discriminatively boosted image clustering with fully convolutional auto - encoders. *Pattern Recognition*, 2018, 83: 161—173.
- [7] Yang J W, Parikh D, Batra D. Joint unsupervised learning of deep representations and image clusters//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 5147—5156.
- [8] Xie J Y, Girshick R, Farhadi A. Unsupervised deep embedding for clustering analysis. 2016, arXiv: 1511.06335.
- [9] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. arxiv: 1511.06434, 2015.
- [10] Wang W G, Song H M, Zhao S Y, et al. Learning unsupervised video object segmentation through visual attention//IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA: IEEE, 2019: 3064—3074.
- [11] He K M, Fan H Q, Wu Y X, et al. Momentum contrast for unsupervised visual representation learning. arXiv:1911.05722, 2020.
- [12] Vo V H, Bach F, Cho M, et al. Unsupervised image matching and object discovery as optimization//IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA: IEEE, 2019: 8287—8296.
- [13] Fischer A, Igel C. An introduction to restricted Boltzmann machines//Alvarez L, Mejail M, Gomez L, et al. Progress in pattern recognition, image analysis, computer vision, and applications (CIARP). Springer Berlin Heidelberg, 2012: 14—36.
- [14] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors. *Nature*, 1986, 323(6088): 533—536.
- [15] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, 86(11): 2278—2324.
- [16] Dosovitskiy A, Springenberg J T, Riedmiller M, et al. Discriminative unsupervised feature learning with convolutional neural networks//Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press, 2014: 766—774.
- [17] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks. *Science*, 2006, 313(5786): 504—507.
- [18] Masci J, Meier U, Cireşan D, et al. Stacked convolutional auto - encoders for hierarchical feature extraction//The 21th International Conference on Artificial Neural Networks. Springer Berlin Heidelberg, 2011: 52—59.
- [19] Goodfellow I, Pouget - Abadie J, Mirza M, et al. Generative adversarial nets//Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press, 2014: 2672—2680.
- [20] Caron M, Bojanowski P, Mairal J, et al. Unsupervised pre-training of image features on non-curated data//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South): IEEE, 2019: 2959—2968.
- [21] Ji X, Vedaldi A, Henriques J F. Invariant information clustering for unsupervised image classification and segmentation//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South): IEEE, 2019: 9865—9874.
- [22] Van Der Maaten L, Hinton G. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 2008, 9: 2579—2605.
- [23] Springenberg J T, Dosovitskiy A, Brox T, et al. Striving for simplicity: the all convolutional net. 2015, arXiv:1412.6806.

(责任编辑 杨可盛)