



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象学报 中国图形学报

2022
07
VOL.27

ISSN1006-8961
CN11-3758/TB



舰船目标检测 P2078



第27卷第7期（总第315期）
2022年7月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院
主办单位 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 吴一戎
编辑出版 《中国图象图形学报》编辑出版委员会
通信地址 北京市海淀区北四环西路19号
邮 编 100190
电子信箱 jig@aircas.ac.cn
电 话 010-58887035
网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号
总 发 行 北京报刊发行局
订 购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
(邮政信箱: 北京399信箱 邮编: 100048)
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian | Monthly, Started in 1996

Superintended by Chinese Academy of Sciences
Sponsored by Aerospace Information Research Institute, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
Address No. 19, North 4th Ring Road West, Haidian District, Beijing, P. R. China
Zip code 100190
E-mail jig@aircas.ac.cn
Telephone 010-58887035
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ

国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



深度学习背景下视觉显著性
物体检测综述(第2112页)



提取全局语义信息的场景图
生成算法(第2214页)



融合弱监督目标定位的细粒
度小样本学习(第2226页)

学者观点

网络监督数据下的细粒度图像识别综述

魏秀参, 许玉燕, 杨健 2057

综述

海事监控视频舰船目标检测研究现状与展望

叶晨, 谔天洋, 肖灏灏, 陆海, 杨群慧 2078

深度学习行人检测方法综述

罗艳, 张重阳, 田永鸿, 郭捷, 孙军 2094

深度学习背景下视觉显著性物体检测综述

王自全, 张永生, 于英, 闵杰, 田浩 2112

图像分析和识别

图像增强对显著性目标检测的影响研究

郭继昌, 岳惠惠, 张怡, 刘迪, 刘晓雯, 郑司达 2129

混合高斯变分自编码器的聚类网络

陈华华, 陈哲, 郭春生, 应娜, 叶学义 2148

基于半监督对抗学习的图像语义分割

李志欣, 张佳, 吴璟莉, 马慧芳 2157

面向大姿态人脸识别的正面化形变场学习

胡蓝青, 阚美娜, 山世光, 陈熙霖 2171

融合时空域特征的人脸表情识别

陈拓, 邢帅, 杨文武, 金剑秋 2185

非局部注意力双分支网络的跨模态赤足足迹检索

鲍文霞, 茅丽丽, 王年, 唐俊, 杨先军, 张艳 2199

提取全局语义信息的场景图生成算法

段静雯, 闵卫东, 杨子元, 张煜, 陈鑫浩, 杨升宝 2214

融合弱监督目标定位的细粒度小样本学习

贺小箭, 林金福 2226

结合双注意力机制的道路裂缝检测

张志华, 温亚楠, 慕号伟, 杜小平 2240

融合神经网络的布料碰撞检测算法

靳雁霞, 马博, 贾瑶, 陈治旭, 芦烨 2251

双分支特征融合网络的步态识别算法

徐硕, 郑锋, 唐俊, 鲍文霞 2263

图像理解和计算机视觉

问题引导的空间关系图推理视觉问答模型

兰红, 张蒲芬 2274

结合扰动约束的低感知性对抗样本生成方法

王杨, 曹铁勇, 杨吉斌, 郑云飞, 方正, 邓小桐 2287

计算机图形学

动态书法墨迹的可回溯感评测

律睿悠, 梅莉琳, 昝跃峰, 晏涛 2300

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Review of deep learning based salient object detection(P2112)



Global semantic information extraction based scene graph generation algorithm(P2214)



Weakly-supervised object localization based fine-grained few-shot learning(P2226)

Scholar View

Review of webly-supervised fine-grained image recognition

Wei Xiushen, Xu Yuyan, Yang Jian 2057

Review

Maritime surveillance videos based ships detection algorithms: a survey

Ye Chen, Lu Tianyang, Xiao Yuhao, Lu Hai, Yang Qunhui 2078

An overview of deep learning based pedestrian detection algorithms

Luo Yan, Zhang Chongyang, Tian Yonghong, Guo Jie, Sun Jun 2094

Review of deep learning based salient object detection

Wang Ziquan, Zhang Yongsheng, Yu Ying, Min Jie, Tian Hao 2112

Image Analysis and Recognition

The analysis of image enhancement on salient object detection

Guo Jichang, Yue Huihui, Zhang Yi, Liu Di, Liu Xiaowen, Zheng Sida 2129

A Gaussian mixture variational autoencoder based clustering network

Chen Huahua, Chen Zhe, Guo Chunsheng, Ying Na, Ye Xueyi 2148

Semi-supervised adversarial learning based semantic image segmentation

Li Zhixin, Zhang Jia, Wu Jingli, Ma Huifang 2157

Large pose face recognition with morphing field learning

Hu Lanqing, Kan Meina, Shan Shiguang, Chen Xilin 2171

Spatio-temporal features based human facial expression recognition

Chen Tuo, Xing Shuai, Yang Wenwu, Jin Jianqiu 2185

Non-local attention dual-branch network based cross-modal barefoot footprint retrieval

Bao Wenxia, Mao Lili, Wang Nian, Tang Jun, Yang Xianjun, Zhang Yan 2199

Global semantic information extraction based scene graph generation algorithm

Duan Jingwen, Min Weidong, Yang Ziyuan, Zhang Yu, Chen Xinhao, Yang Shengbao 2214

Weakly-supervised object localization based fine-grained few-shot learning

He Xiaojian, Lin Jinfu 2226

Dual attention mechanism based pavement crack detection

Zhang Zhihua, Wen Yanan, Mu Haowei, Du Xiaoping 2240

Neural network fusion based cloth collision detection algorithm

Jin Yanxia, Ma Bo, Jia Yao, Chen Zhixu, Lu Ye 2251

Dual branch feature fusion network based gait recognition algorithm

Xu Shuo, Zheng Feng, Tang Jun, Bao Wenxia 2263

Image Understanding and Computer Vision

Question-guided spatial relation graph reasoning model for visual question answering

Lan Hong, Zhang Pufen 2274

A perturbation constraint related weak perceptual adversarial example generation method

Wang Yang, Cao Tieyong, Yang Jibin, Zheng Yunfei, Fang Zheng, Deng Xiaotong 2287

Computer Graphics

Traceability evaluation of flowing calligraphy strokes

Lyu Ruimin, Mei Lilin, Ze Yuefeng, Yan Tao 2300

中图法分类号: TN911 文献标识码: A 文章编号: 1006-8961(2022)07-2148-09

论文引用格式: Chen H H, Chen Z, Guo C S, Ying N and Ye X Y. 2022. A Gaussian mixture variational autoencoder based clustering network. Journal of Image and Graphics, 27(07): 2148-2156 [陈华华, 陈哲, 郭春生, 应娜, 叶学义. 2022. 混合高斯变分自编码器的聚类网络. 中国图象图形学报, 27(07): 2148-2156] [DOI:10. 11834/jig. 200467]

混合高斯变分自编码器的聚类网络

陈华华*, 陈哲, 郭春生, 应娜, 叶学义

杭州电子科技大学通信工程学院, 杭州 310018

摘要: 经典的聚类算法在处理高维数据时存在维数灾难等问题, 使得计算成本大幅增加并且效果不佳。以自编码或变分自编码网络构建的聚类网络改善了聚类效果, 但是自编码器提取的特征往往比较差, 变分自编码器存在后验崩塌等问题, 影响了聚类的结果。为此, 本文提出了一种基于混合高斯变分自编码器的聚类网络。**方法** 使用混合高斯分布作为隐变量的先验分布构建变分自编码器, 并以重建误差和隐变量先验与后验分布之间的 KL 散度 (Kullback-Leibler divergence) 构造自编码器的目标函数训练自编码网络; 以训练获得的编码器对输入数据进行特征提取, 结合聚类层构建聚类网络, 以编码器隐层特征的软分配分布与软分配概率辅助目标分布之间的 KL 散度构建目标函数并训练聚类网络; 变分自编码器采用卷积神经网络实现。**结果** 为了验证本文算法的有效性, 在基准数据集 MNIST (Modified National Institute of Standards and Technology Database) 和 Fashion-MNIST 上评估了该网络的性能, 聚类精度 (accuracy, ACC) 和标准互信息 (normalized mutual information, NMI) 指标在 MNIST 数据集上分别为 95.86% 和 91%, 在 Fashion-MNIST 数据集上分别为 61.34% 和 62.5%, 与现有方法相比性能有了不同程度的提升。**结论** 实验结果表明, 本文网络取得了较好的聚类效果, 且优于当前流行的多种聚类方法。**关键词:** 聚类; 混合高斯分布; 变分自编码器 (VAE); 软分配; KL 散度

A Gaussian mixture variational autoencoder based clustering network

Chen Huahua*, Chen Zhe, Guo Chunsheng, Ying Na, Ye Xueyi

School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China

Abstract: Objective Effective automatic grouping of data into clusters, especially clustering high-dimensional datasets, is one of the key issues in machine learning and data analysis. It is related to many aspects of signal processing applications, including computer vision, pattern recognition, speech and audio recognition, wireless communication and text classification. Current clustering algorithms are constrained of high computational complexity and poor performance in processing high-dimensional data due to the dimension disaster. Deep neural networks based clustering methods have its potential for real data clustering derived of their high representational ability. Autoencoder (AE) or variational autoencoder (VAE) clustering networks improve clustering effectiveness. But, their clustering performance is easy to be distorted intensively because of poor features extraction in distinguishing clear and unclear data or posterior collapse to clarify determining its posterior parameters of the latent variable of VAE, and they are insufficient to segment multiple classes, especially share

收稿日期: 2020-09-08; 修回日期: 2020-11-02; 预印本日期: 2020-11-09

* 通信作者: 陈华华 iseealv@hdu.edu.cn

基金项目: 浙江省“领雁”研发攻关计划项目 (2022C03065)

Supported by: the “Leading Goose” Technology Research and Development Program of Zhejiang Province (2022C03065)

very similar mean and variance in the context of clustering a multiclass dataset or two different classes. We demonstrate a clustering network based on VAE with the prior of Gaussian mixture (GM) distribution in terms of the deficiency of AE and VAE. **Method** The VAE, a maximum likelihood generative model, maximizes evidence lower bound (ELBO) via minimizing model reconstruction errors. Its difference of potential cost is through Kullback-Leibler (KL) divergence between the posterior distribution and the hypothesized prior, and then establishes maximum marginal log-likelihood (LL) of the data observed. Due to the approximate posterior distribution used VAE as a benched Gaussian distribution, it is challenged to match the ground truth posterior and have its priority of the KL term in ELBO, and the latent variable space may be arbitrarily complicated or even multimodal. To further improve the description of latent variables, a VAE is facilitated based on a latent variable prior of GM distribution. Its GM distribution prior linked data representation is approximated using the posterior distribution of the latent variable composed of a GM model, and the reconstruction error and the KL divergence based cost function between posterior and prior distribution is adopted to train the GM model based VAE. Due to the KL divergence between two GM distribution functions without a closed form solution, we use the approximate variational lower bound solution of the cost function with the aid of the fact that the KL divergence between the two single Gaussians has a closed-form solution, and implement the VAE using GM distribution priors optimization to resolve the KL divergence. A VAE based clustering network is constructed through a clustering layer combination behind the VAE. To improve the clustering performance, the STUDENT's t -distribution is used as a kernel to compute the soft assignment of the latent features of the VAE between the embedded point and the cluster centroid. Furthermore, a KL divergence cost is constructed between the soft assignment and its auxiliary target distribution. The commonly used VAE utilizes fully-connected neural networks to compute the latent variable, which generates more over fitted data parameters. Thus, the clustering network is carried out by convolutional neural networks (CNNs), which consist of three convolutional layers and two fully-connected layers, without fully-connected neural networks, and no pooling layers used in the network because it will result in loss of useful information of the data. The network is trained by optimizing the KL divergence cost using stochastic gradient descent (SGD) method with the initial network parameters from the VAE. Our clustering network was obtained by the two-step training mentioned above like acquired VAE, as the initial value to train the following clustering layer. **Result** To test the effectiveness of the proposed algorithm, our network is evaluated on the multiclass benchmark datasets MNIST (Modified National Institute of Standards and Technology Database) which contains images of 10 categories of handwritten digits, and Fashion-MNIST which consists of grayscale images associated to a 10 segmented label. Our algorithm achieves 95.86% accuracy (ACC) and 91% normalized mutual information (NMI) on MNIST, ACC 61.34% and 62.5% NMI on Fashion-MNIST. Our network demonstration has the similar performance to ClusterGAN with fewer parameters and less memory space. The experimental results illustrate that our network achieves feasible clustering performance. **Conclusion** We construct a VAE based clustering network with the prior of GM distribution. A novel framed VAE is established to improve the representation ability of the latent variable based on a latent variable prior of GM distribution. The KL divergence between posterior and prior GM distribution is optimized to achieve latent variable features of VAE and reconstruct its input well. To improve the clustering performance, the clustering network is trained by optimizing the KL divergence between the soft distribution of the latent features of the VAE and the auxiliary target distribution of the soft assignment. We focus on the issue of where the number of Gaussian components in prior and posterior is different and the ability of the representation of the model on complex texture features further.

Key words: clustering; Gaussian mixture distribution; variational autoencoder (VAE); soft assignment; Kullback-Leibler (KL) divergence

0 引言

随着科学技术的发展,数据的规模呈指数型增长,这些海量数据往往蕴含着许多极具价值的潜在

信息,如何捕获并挖掘这些隐藏的潜在信息是一个亟需解决的问题,具有重要的现实意义。

聚类分析是数据挖掘领域的一种研究方法,也是目前用于发现数据间潜在信息的主要方法之一。聚类分析旨在发现数据间潜在的关系,并根据数据

的特征等将数据聚为不同的类别,也称为簇,使得簇内的数据具有较小的差异性,簇间的数据具有较大的差异性。聚类分析已广泛应用于用户画像(张海涛等,2018)、协同过滤推荐系统(吴湖等,2010)、基因分析(岳峰等,2008)、异常检测(成宝芝等,2017)和文本聚类(路荣等,2012)等领域,吸引了越来越多的学者加入到聚类分析的研究队伍中。

1 相关工作

MacQueen(1967)提出的 K-means 聚类算法和 Ester 等人(1996)提出的 DBSCAN(density-based spatial clustering of applications with noise)聚类算法是两个经典的聚类方法。K-means 算法的基本思想是将数据分为不同的类别,每一个类别具有一个聚类中心,根据数据到各聚类中心的距离更新优化聚类划分,反复迭代得到一个最佳的聚类结果。DBSCAN 是一种基于密度的聚类算法,根据邻域内的点数判断该区域是否属于密集区域并形成临时聚簇,将相连的临时聚簇进行合并得到更合理的聚簇分配。

然而,随着数据维度的提高,经典的聚类算法在应对高维数据时往往存在维数灾难等问题,使得计算成本大幅增加并且效果不佳。近年来,深度神经网络得到了飞速的发展。因此,越来越多的研究人员将目光转向利用深度学习进行聚类分析。相比于传统方法,深度学习能够更好地得到数据特征的低维表示,提高聚类效果。Yang 等人(2017)提出了 DCN(deep clustering network)模型,该模型训练一个自编码器 AE(autoencoder),得到数据特征的低维表示,然后结合 K-means 聚类算法,将重构损失和 K-means 聚类损失进行联合优化,该模型聚类效果高于同时期的传统算法。Xie 等人(2016)提出了 DEC(deep embedded clustering)模型,模型训练一个堆叠的自编码器得到数据特征的低维表示,然后在此基础上构建聚类网络。Opochinsky 等人(2020)和 Chazan 等人(2019)用多个自编码器建立深度聚类网络,每个自编码器表示一个簇,取得了更好的聚类效果。Duan 等人(2019)采用自编码生成一个深度嵌入网用于数据降维,学习一个 softmax 自编码器用于估计簇的数目,获得了较好的实验结果。但是自

编码器的目的主要是为了降维,网络训练目标是使解码器输出和输入尽可能逼近,当训练样本与预测样本不符合相同分布时,提取的特征往往比较差。与自编码器相关的另一种编码器是变分自编码器(variational autoencoder, VAE),Jiang 等人(2017)和 Lim 等人(2020)提出了变分自编码器用于聚类,先用 VAE 生成隐层特征,然后用混合高斯分布拟合隐层特征,聚类效果优于经典的聚类方法和一些生成式聚类方法。变分自编码器采用标准正态分布作为先验,容易导致后验塌陷(Guo 等,2020),对不同类别数据的分布不能较好地逼近,影响编码和解码结果。为此,本文引入混合高斯分布作为先验,构建混合高斯自编码器生成隐层特征,学习数据的特征分布,并以此编码器与聚类层结合形成聚类网络,通过优化编码器隐层特征的软分配分布与软分配概率辅助目标分布之间的 KL 散度(Kullback-Leibler divergence)对聚类网络进行训练。在基准数据集 MNIST(Modified National Institute of Standards and Technology Database)和 Fashion-MNIST 上的实验结果表明,本文网络取得了较好的聚类效果,且优于当前多种流行的聚类方法。

2 聚类模型

2.1 混合高斯变分自编码器

变分自编码器(Kingma 和 Welling, 2014)是深度学习领域一种近似推理的有向模型,它利用变分推断与深度学习相结合,能够学习到数据结构化的特征表示或分布,是生成式建模中的一种重要方法。标准变分自编码器的目标函数可表示为

$$L(\theta, \phi; \mathbf{x}) = -D_{\text{KL}}[q_{\phi}(\mathbf{z} | \mathbf{x}) \| p_{\theta}(\mathbf{z})] + E_{q_{\phi}(\mathbf{z} | \mathbf{x})}[\log p_{\theta}(\mathbf{x} | \mathbf{z})] \quad (1)$$

式中,当近似后验分布与假设先验分布之间的 KL 散度 $D_{\text{KL}}[q_{\phi}(\mathbf{z} | \mathbf{x}) \| p_{\theta}(\mathbf{z})]$ 最小时,目标函数 $L(\theta, \phi; \mathbf{x})$ 下限最大,从而使模型达到最优。其中, E 表示求期望, θ 和 ϕ 分别是先验分布 p 和后验分布 q 的参数, $\mathbf{x}$ 和 $\mathbf{z}$ 分别是变分自编码器的输入和隐层特征。但是,标准变分自编码器中采用标准正态分布作为先验,可能引起后验塌陷(Guo 等,2020),并且容易忽略一些潜在的变量约束,导致对不同类别数据的分布不能较好地逼近,影响自编码网络的编

码和解码结果。在这里,本文引入混合高斯分布作为先验,构建变分自编码器。混合高斯先验可表示为

$$p_{\theta}(\mathbf{z}) = \sum_{i=1}^M \pi_i N(\mathbf{z}^{(i)}; \tilde{\boldsymbol{\mu}}^{(i)}, \tilde{\boldsymbol{\sigma}}^{(i)^2} \mathbf{I}) \quad (2)$$

$$\sum_{i=1}^M \pi_i = 1 \quad (3)$$

式中, π_i 是第 i 个单高斯分布的系数, $\tilde{\boldsymbol{\mu}}^{(i)}$ 和 $\tilde{\boldsymbol{\sigma}}^{(i)^2}$ 代表第 i 个单高斯分布的均值和方差, $\mathbf{z}^{(i)}$ 表示第 i 个单高斯生成的隐层特征, $\mathbf{I}$ 是单位矩阵, M 表示高斯分量的个数, $N(\cdot; \cdot)$ 表示高斯分布。相应地,近似后验 $q_{\phi}(\mathbf{z} | \mathbf{x})$ 可表示为

$$q_{\phi}(\mathbf{z} | \mathbf{x}) = \sum_{i=1}^M \omega_i N(\mathbf{z}^{(i)}; \boldsymbol{\mu}^{(i)}, \boldsymbol{\sigma}^{(i)^2} \mathbf{I}) \quad (4)$$

$$\sum_{i=1}^M \omega_i = 1 \quad (5)$$

式中, ω_i 是第 i 个单高斯分布的系数, $\boldsymbol{\mu}^{(i)}$ 和 $\boldsymbol{\sigma}^{(i)^2}$ 代表第 i 个单高斯分布的均值和方差。式(1)中等号右侧的第 1 项是混合高斯分布近似后验与先验之间的 KL 散度,目前尚无高效的算法可以获得它的解析解。为此, Hershey 和 Olsen(2007)提出了一种近似解法,采用变分推断获得 KL 散度的上界。于是最小化 KL 散度转换为最小化其近似上界,该算法简述如下:

设数据 $\mathbf{x}$ 的混合高斯分布 $f(\mathbf{x})$ 和 $g(\mathbf{x})$ 分别可表示为

$$f(\mathbf{x}) = \sum_{i=1}^M \alpha_i N(\mathbf{x}; \bar{\boldsymbol{\mu}}^{(i)}, \bar{\boldsymbol{\sigma}}^{(i)^2} \mathbf{I}) \quad (6)$$

$$g(\mathbf{x}) = \sum_{i=1}^M \beta_i N(\mathbf{x}; \boldsymbol{\mu}'^{(i)}, \boldsymbol{\sigma}'^{(i)^2} \mathbf{I}) \quad (7)$$

式中, α_i 和 β_i 分别是 $f(\mathbf{x})$ 和 $g(\mathbf{x})$ 的第 i 个单高斯分布的系数, $\sum_{i=1}^M \alpha_i = 1$, $\sum_{i=1}^M \beta_i = 1$, $\bar{\boldsymbol{\mu}}^{(i)}$ 和 $\bar{\boldsymbol{\sigma}}^{(i)^2}$ 分别是 $f(\mathbf{x})$ 中第 i 个单高斯分布的均值和方差, $\boldsymbol{\mu}'^{(i)}$ 和 $\boldsymbol{\sigma}'^{(i)^2}$ 分别是 $g(\mathbf{x})$ 中第 i 个单高斯分布的均值和方差。于是, $f(\mathbf{x})$ 和 $g(\mathbf{x})$ 之间的 KL 散度上界可以表示为

$$D_{\text{KL}}(f(\mathbf{x}) \| g(\mathbf{x})) \leq \sum_{i=1}^M \alpha_i (\log(\alpha_i / \beta_i) + D_{\text{KL}}(f_i(\mathbf{x}) \| g_i(\mathbf{x}))) \quad (8)$$

式中, $f_i(\mathbf{x})$ 和 $g_i(\mathbf{x})$ 分别代表 $f(\mathbf{x})$ 和 $g(\mathbf{x})$ 中等号右侧第 i 个单高斯分布。

因此,式(1)中等号右侧第 1 项可以表示为

$$\begin{aligned} & -D_{\text{KL}} \left[\sum_{i=1}^M \omega_i N(\mathbf{z}^{(i)}; \boldsymbol{\mu}^{(i)}, \boldsymbol{\sigma}^{(i)^2} \mathbf{I}) \right. \\ & \quad \left. \parallel \sum_{i=1}^M \pi_i N(\mathbf{z}^{(i)}; \tilde{\boldsymbol{\mu}}^{(i)}, \tilde{\boldsymbol{\sigma}}^{(i)^2} \mathbf{I}) \right] \geq \\ & - \sum_{i=1}^M \omega_i \left(\log \frac{\omega_i}{\pi_i} + D_{\text{KL}}(N(\mathbf{z}^{(i)}; \boldsymbol{\mu}^{(i)}, \boldsymbol{\sigma}^{(i)^2} \mathbf{I}) \parallel \right. \\ & \quad \left. N(\mathbf{z}^{(i)}; \tilde{\boldsymbol{\mu}}^{(i)}, \tilde{\boldsymbol{\sigma}}^{(i)^2} \mathbf{I})) \right) = - \sum_{i=1}^M \omega_i \left(\log \frac{\omega_i}{\pi_i} + \right. \\ & \quad \left. \sum_{j=1}^J \frac{1}{2} \left(\log \frac{\tilde{\sigma}_j^{(i)^2}}{\sigma_j^{(i)^2}} + \frac{(\mu_j^{(i)} - \tilde{\mu}_j^{(i)})^2 + \sigma_j^{(i)^2}}{\tilde{\sigma}_j^{(i)^2}} - 1 \right) \right) \end{aligned} \quad (9)$$

式中, $\tilde{\mu}_j^{(i)}$ 表示 $\tilde{\boldsymbol{\mu}}^{(i)}$ 的第 j 个元素, $\tilde{\sigma}_j^{(i)^2}$ 表示 $\tilde{\boldsymbol{\sigma}}^{(i)^2}$ 的第 j 个元素, J 为 $\tilde{\boldsymbol{\mu}}^{(i)}$ 和 $\tilde{\boldsymbol{\sigma}}^{(i)^2}$ 的维数。

式(1)中等号右侧的第 2 项是重构项,它的计算方式与标准变分自编码器类似,具体为

$$E_{q_{\phi}(\mathbf{z} | \mathbf{x})} [\log p_{\theta}(\mathbf{x} | \mathbf{z})] \approx \frac{1}{L} \sum_{l=1}^L \log p_{\theta}(\mathbf{x} | \mathbf{z}_l) \quad (10)$$

式中, L 是采样的数量, $\mathbf{z}_l$ 的下标 l 表示第 l 次采样。由式(9)和式(10)可得到混合高斯分布后验与先验的变分下界,即

$$\begin{aligned} L(\theta, \phi; \mathbf{x}) \geq & - \sum_{i=1}^M \omega_i \left(\log \frac{\omega_i}{\pi_i} + \sum_{j=1}^J \frac{1}{2} \times \right. \\ & \left(\log \frac{\tilde{\sigma}_j^{(i)^2}}{\sigma_j^{(i)^2}} + \frac{(\mu_j^{(i)} - \tilde{\mu}_j^{(i)})^2 + \sigma_j^{(i)^2}}{\tilde{\sigma}_j^{(i)^2}} - 1 \right) \right) + \\ & \frac{1}{L} \sum_{l=1}^L \log p_{\theta}(\mathbf{x} | \mathbf{z}_l) \end{aligned} \quad (11)$$

式(11)即为混合高斯分布自编码器的目标函数。

2.2 聚类网络

本文在混合高斯变分自编码器的基础上,使用编码器部分作为数据空间和特征空间之间的初始映射,将编码器和聚类层组合成聚类网络,如图 1 所示。图 1 中,以学习获得的混合高斯分布自编码器部分作为聚类网络的编码器部分,结合聚类层通过最小化辅助目标分布和软分配分布之间的 KL 散度学习聚类网络。具体过程如下:

假设存在数据集 $\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(N)}\}$, N 为数据样本的数量,输入数据 $\mathbf{x}^{(i)}$ 到混合高斯变分自编码器中得到隐层特征 $\mathbf{z}^{(i)}$,使用欧氏距离计算隐层特征 $\mathbf{z}^{(i)}$ 到聚类中心 $\mathbf{c}^{(t)}$ 的距离, $\mathbf{c}^{(t)}$ 表示第 t 个聚类中心,并使用 t 分布(van der Maaten 和 Hinton, 2008)衡量隐层特征 $\mathbf{z}^{(i)}$ 到聚类中心 $\mathbf{c}^{(t)}$ 之间的相似

度 s_{it} , 也即特征 $\mathbf{z}^{(i)}$ 分配到聚类类别 t 的软分配概率 (Dempster 等, 1977), s_{it} 计算为

$$s_{it} = \frac{(1 + \|\mathbf{z}^{(i)} - \mathbf{c}^{(t)}\|^2 / \chi)^{-\frac{\chi+1}{2}}}{\sum_{t=1}^T (1 + \|\mathbf{z}^{(i)} - \mathbf{c}^{(t)}\|^2 / \chi)^{-\frac{\chi+1}{2}}} \quad (12)$$

式中, χ 代表 t 分布的自由度, 由于聚类属于无监督学习, 无法交叉验证 χ 的取值, 因此本文取 $\chi = 1$ 。将数据输入到训练获得的混合高斯变分自编码器得到隐层特征, 然后在特征空间采用混合高斯模型进行聚类, 得到 T 个聚类中心 $\{\mathbf{c}^{(1)}, \mathbf{c}^{(2)}, \dots, \mathbf{c}^{(T)}\}$ 作为初始化的聚类中心。

本文采用最小化辅助目标分布和软分配之间的 KL 散度实现对模型的优化, 因此辅助目标分布的选择对聚类效果至关重要。Xie 等人 (2016) 提出辅助

目标分布应该更加注重高置信度的数据点, 以提高聚类的准确性。同时, 需要归一化代价函数对每个聚类中心的贡献, 防止出现过大的聚簇导致隐层的特征空间扭曲。因此, 将软分配概率的辅助目标分布 p_{it} 定义为

$$p_{it} = \frac{s_{it}^2 / f_t}{\sum_{i=1}^M s_{it}^2 / f_t} \quad (13)$$

式中, $f_t = \sum_i s_{it}$ 代表软聚类中类别 t 的频率。

由此, 得到聚类层的损失 L_{cluster} , 定义为软分配分布 s 与辅助目标分布 p 之间的 KL 散度, 采用随机梯度下降法优化损失函数。具体为

$$L_{\text{cluster}} = D_{\text{KL}}(p \| s) = \sum_i \sum_t p_{it} \log \frac{p_{it}}{s_{it}} \quad (14)$$

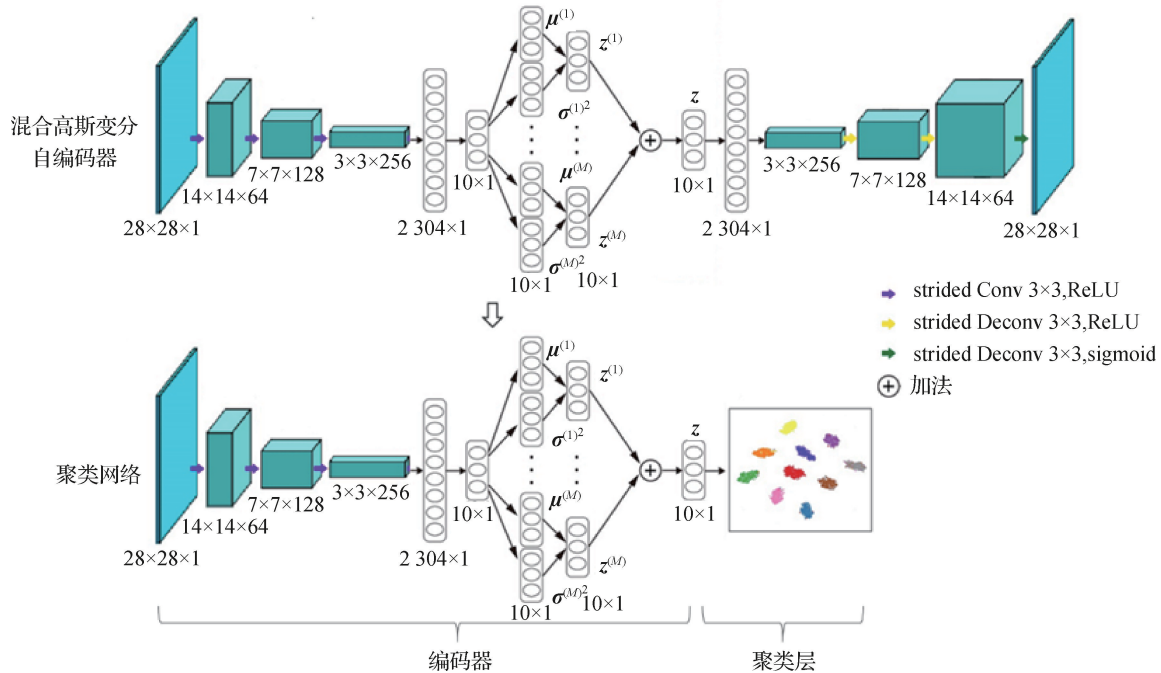


图1 聚类网络结构图

Fig. 1 Structure of the clustering network

2.3 网络结构

本文中的变分自编码器采用卷积神经网络实现, 实现的网络结构如图1所示。网络的编码器部分, 首先是核大小为 3×3 、步长为2、通道数为64的卷积层, 激活函数采用 ReLU (rectified linear unit) 函数。考虑到池化层在实现下采样时存在丢失有用信息的不足 (Sabour 等, 2017), 本文采用核大小为 3×3 、步长为2的卷积层实现下采样以保留重要信息。然后级联大小为 3×3 、步长为2、通道数为

128的卷积层, 再级联大小为 3×3 、步长为2、通道数为256的卷积层, 然后连接两个全连接层, 维数分别是2304和10。

解码器部分在结构上与编码器部分是对称的, 首先是两个级联的全连接层, 它们的维数分别是10和2304, 然后级联一个维数变形层 (reshape 层), 将数据维度从2304转换为 $3 \times 3 \times 256$, 然后级联3个卷积核大小为 3×3 的反卷积层, 它们分别具有128通道、64通道、1通道。与编码器类

似, 解码器采用步长为 2 的反卷积层实现上采样。解码器在最后一层的反卷积层使用 sigmoid 函数作为激活函数, 其余卷积层、反卷积层的激活函数都采用 ReLU 函数。

聚类网络中, 自编码器的结构直接采用混合高斯分布自编码器的编码部分, 并以学习获得的编码器参数作为初始值, 进一步按式 (14) 所示目标函数优化学习聚类网络。

3 实验结果分析

3.1 实验配置

实验采用的计算机配置如下: Intel (R) Core (TM) i5-7300HQ@2.50 GHz CPU, Windows 10, 编译器 Python3.6, 内存为 8 GB, 编程环境为 TensorFlow 和 Keras, 编程语言为 Python。网络训练参数值采用正态分布随机初始化, batch size 为 100, 优化方法为 Adam 优化器, 学习率为 0.000 1。

为评估本文方法的有效性, 采用了聚类分析中常用的 MNIST (LeCun 等, 1998) 数据集和 Fashion-MNIST (Xiao 等, 2017) 数据集分别进行实验评测。MNIST 数据集是 LeCun 等人 (1998) 在美国国家标准与技术研究院提供的手写数据集的基础上筛选, 并进行了尺寸标准化及数字中心化等处理的标准数据集, 由 60 000 个训练样本和 10 000 个测试样本组成, 每个样本都是 28×28 像素的灰度图像; Fashion-MNIST 是一个替代 MNIST 手写数字集的图像数据集, 由德国科技公司 Zalando 旗下的研究部门提供, 涵盖了 10 种类别共 70 000 个不同时尚商品的正面灰度图像, 图像大小为 28×28 像素, 包括 T 恤、裤子、套衫、裙子、外套、凉鞋、衬衫、运动鞋、包和靴子等 10 类商品。为简化问题求解, 实验中使用的混合高斯中高斯分量的个数 $M = 10$, 各分量的混合系数取经验值为 $1/10$, 初始化聚类中心的个数 $T = 10$ 。

3.2 评价指标

采用聚类精度 (accuracy, ACC) 和标准互信息 (normalized mutual information, NMI) 作为评估指标。

聚类精度 ACC 用于衡量算法得到的聚类标签准确性, 计算为

$$ACC = \max_m \frac{1}{N} \sum_{i=1}^N 1(l^{(i)} = m(\hat{l}^{(i)})) \quad (15)$$

式中, $l^{(i)}$ 为给定真实数据类别标签, $\hat{l}^{(i)}$ 为算法得到的聚类标签, $m(\cdot)$ 是聚类标签和真实数据类别标签之间所有可能的一对一的映射, 目的是在聚类标签与真实数据类别标签之间找到最佳映射, 这里采用匈牙利算法 (Kuhn, 2005) 得到最佳映射。 $1(\cdot)$ 表示指示函数, 当括号内的表示为真时值为 1, 否则为 0。

标准化互信息 NMI 是衡量两个随机事件之间相关性的一个重要指标, 也是常用的聚类评估指标之一, 这里用做衡量聚类标签与真实数据类别标签的契合程度。标准化互信息计算为

$$NMI(l, \hat{l}) = \frac{MI(l, \hat{l})}{\sqrt{H(l)H(\hat{l})}} \quad (16)$$

式中, $MI(l, \hat{l})$ 是真实数据类别标签 l 和网络得到的聚类标签 $\hat{l}$ 的互信息, $H(l)$ 和 $H(\hat{l})$ 分别是它们的信息熵。

3.3 实验结果

本文模型分别在 MNIST 和 Fashion-MNIST 数据集上进行了聚类实验。为了进一步验证本文方法的有效性, 与高斯混合模型 (Gaussian mixture model, GMM) (Fraley 和 Raftery, 1998)、VAE + K-means (Kingma 和 Welling, 2014)、DEC (deep embedded clustering) (Xie 等, 2016)、IDEC (improved DEC) (Guo 等, 2017)、GMVAE (Dilokthanakul 等, 2017)、KADC (K-autoencoders deep clustering) (Opochinsky 等, 2020)、VaDE (variational deep embedding) (Jiang 等, 2017)、DCVA (deep clustering with VAE) (Lim 等, 2020) 和 ClusterGAN (clustering in generative adversarial networks) (Mukherjee 等, 2019) 等算法进行对比, 对比结果如表 1 和表 2 所示。

从表 1 和表 2 可以看出, 本文模型与 GMM、VAE + K-means、DEC、IDEC、GMVAE、VaDE 和 DCVA 相比, 在 MNIST 和 Fashion-MNIST 数据集上的聚类精度 ACC 和标准互信息 NMI 均有较大提升。与 KADC 相比, 本文方法在 Fashion-MNIST 数据集上的标准互信息 NMI 低于 KADC。与 ClusterGAN 算法相比, 本文算法在 MNIST 数据集上的聚类 ACC 和标准互信息 NMI 优于 ClusterGAN, 但在 Fashion-MNIST 数据集上略逊于 ClusterGAN 算法。除了 GMM 和 ClusterGAN 方法, 其他方法都是从 AE 或 VAE 基础上发展起来的, 其中性能最好的 VaDE 方法对聚簇

表1 不同算法在 MNIST 和 Fashion-MNIST 数据集上的 ACC 比较
Table 1 Comparison of ACC among different algorithms on MNIST and Fashion-MNIST datasets

算法	MNIST	Fashion-MNIST
GMM	52.73	49.37
VAE + K-means	70.92	45.57
DEC	84.30	58.6
IDEC	88.06	58.9
GMVAE	88.54	56.35
KADC	88	60
VaDE	94.46	-
DCVA	88.75	-
ClusterGAN	95	63
本文	95.86	61.34

注:加粗字体表示各列最优结果;“-”表示原文献未做评价。

表2 不同算法在 MNIST 和 Fashion-MNIST 数据集上的 NMI 比较
Table 2 Comparison of NMI among different algorithms on MNIST and Fashion-MNIST datasets

算法	MNIST	Fashion-MNIST
GMM	49.26	48.69
VAE + K-means	70.34	46.74
DEC	83.64	57.5
IDEC	86.7	55.7
GMVAE	87.03	53.59
KADC	86	65
VaDE	-	-
DCVA	81.9	-
ClusterGAN	89	64
本文	91	62.5

注:加粗字体表示各列最优结果;“-”表示原文献未做评价。

采用单个质心表示嵌入空间中的向量,而本文方法采用一个混合高斯自动编码器网络表示嵌入空间中的向量,这样对每个聚簇可以实现更丰富的表示,同时辅助目标分布的引入提高了聚类的准确性,防止了隐层的特征空间扭曲,使得算法具有良好的聚类表现;ClusterGAN 和本文方法虽然都属于生成式聚

类方法,但是两者无论是从网络结构还是实现思路上属于两个不同的方法类别,ClusterGAN 是目前聚类性能最好的生成对抗网络,本文方法与该方法的性能相当。总体而言,本文算法具有较好的聚类结果。

此外,由表 1 和表 2 还可知,尽管 MNIST 和 Fashion-MNIST 都是 28 × 28 像素的高维图像数据,具有相同的数据维度,但是无论是 ACC 指标还是 NMI 指标,各方法除了 GMM 方法在改变数据集时指标下降比较小,其他方法都出现了大幅度下降,这是因为 MNIST 数据集由灰度变化范围小的图像组成,纹理特征信息比较单一,主要是字符的边界信息,在描述数据特征时对各模型的特征表述能力要求相对较低,而 Fashion-MNIST 数据集由灰度变化范围大的图像组成,纹理特征信息比较丰富,如衣服、鞋子和包等丰富的内部纹理和边缘,在描述数据特征时对各模型的特征表述能力要求较高,而目前的方法对 Fashion-MNIST 数据集的特征信息表达均不是很强,导致聚类性能明显劣于 MNIST 数据集。

图 2 列出了本文模型在 MNIST 和 Fashion-MNIST 数据集中每个类的 10 幅图像,其中每一行对应一个聚簇。由图 2(a)可知,本文模型对 MNIST 数据集的聚类结果较为准确,但出现了若干“4”和“9”混淆的情况,这一结果与“4”和“9”的外观特征相似有关。由图 2(b)可知,Fashion-MNIST 数据集下的每个聚类依次为凉鞋、外套、靴子、套衫、裤子、运动鞋、T 恤、衬衫、包和裙子。聚类结果中的凉鞋、靴子、裤子、运动鞋、T 恤衫、包和裙子等类别的聚类较为准确,在外套、套衫和衬衫这 3 类中出现了若干次混淆的情况,这一结果的产生与这 3 类物体的外观较为相似有关,区分这 3 类物品更依赖于内部纹理特征的差异。

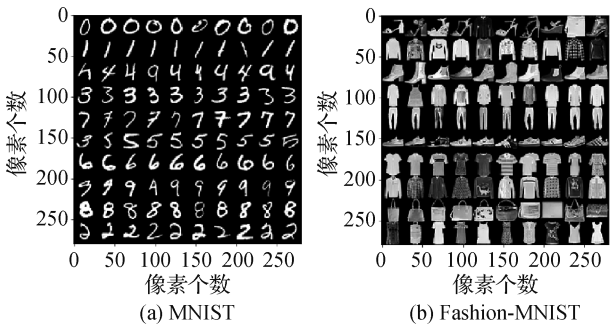


图2 MNIST 和 Fashion-MNIST 数据集的聚类结果
Fig. 2 Clustering results on MNIST and Fashion-MNIST datasets
((a) MNIST; (b) Fashion-MNIST)

图 3 列出了本文模型与 ClusterGAN 算法在 Fashion-MNIST 数据集下的重建结果。由图 3 可知, ClusterGAN 算法重建得到图像的纹理特征较本文模型清晰,其在纹理特征的提取与重建上优于本文方法。因此,ClusterGAN 的聚类效果在 Fashion-MNIST 数据集上略优于本文算法。

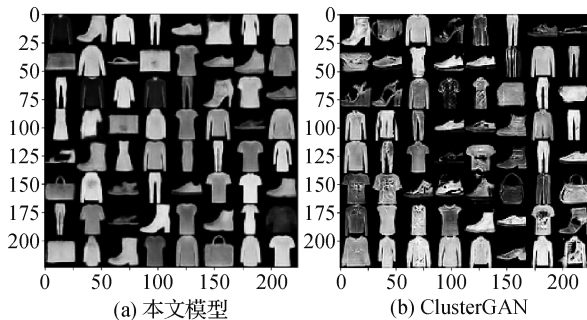


图 3 本文模型与 ClusterGAN 算法的重建结果
Fig. 3 Reconstruction of the proposed network and ClusterGAN
((a) ours; (b) ClusterGAN)

3.4 结构复杂度分析

本文模型在 Fashion-MNIST 数据集上的聚类指标略低于 ClusterGAN 算法。本文模型和 ClusterGAN 算法都属于生成式聚类方法,不同的是 ClusterGAN 采用生成对抗网络用于聚类,而本文采用混合高斯分布的变分自编码器。对这两种算法的模型参数量进行了比较分析,如表 3 所示。

表 3 本文模型与 ClusterGAN 算法的模型参数量比较
Table 3 Comparison of the number of parameters in the proposed network and ClusterGAN

算法	模型参数量
ClusterGAN	13 199 785
本文	1 225 417

注:加粗字体表示最优结果。

由表 3 可知,本文模型参数量不及 ClusterGAN 算法的 1/10,远小于 ClusterGAN 算法,是一个更轻量级的网络模型,这使得本文模型占用更小的存储空间,降低了对内存的需求,同时能够实现更快的运行速度,在 Fashion-MNIST 数据集上的性能差异小于 2%。

4 结 论

本文提出了一种基于混合高斯变分自编码器的

聚类网络模型,以混合高斯分布为先验建立变分自编码器,学习数据的特征分布,然后将编码器与聚类层结合构建聚类网络,采用编码器隐层特征的软分配分布与软分配概率辅助目标分布之间的 KL 散度作为目标函数,对网络进行训练和优化。在基准数据集 MNIST 和 Fashion-MNIST 上进行了评价和比较,对比实验结果表明,采用混合高斯自动编码器网络对每个聚簇可以实现更丰富的表示,辅助目标分布的引入使得算法具有良好的聚类表现,使得本文方法在聚类 ACC 和标准互信息 NMI 指标都优于当前的一些聚类算法,取得了较好的聚类效果。

但是本文算法也存在两个不足:1)在模型建立上,先验和后验中的高斯分量个数设为相同,虽然实验结果优于当前的一些聚类算法,但是处理方法只是对实际情况的简化处理,与实际情况存在一些差别,实际更一般的情况中,先验和后验中的高斯分量个数并不相同,如何优化求解这个问题是个难题;同时,本文中各高斯分量的混合系数不是通过数学优化求得的最佳混合系数,而是根据经验设定为等概率混合的,缺乏完善的理论支持。2)在模型对信息的表达能力上,当处理复杂纹理特征时,纹理特征的重建效果有待提高。上述问题将是下一步研究的重点。

参考文献 (References)

- Chazan S E, Gannot S and Goldberger J. 2019. Deep clustering based on a mixture of autoencoders//Proceedings of the 29th IEEE International Workshop on Machine Learning for Signal Processing. Pittsburgh, USA: IEEE: 1-6 [DOI: 10.1109/MLSP.2019.8918720]
- Cheng B Z, Zhao C H, Zhang L L and Zhang J P. 2017. Joint spatial preprocessing and spectral clustering based collaborative sparsity anomaly detection for hyperspectral images. Acta Optica Sinica, 37(4): 296-306 (成宝芝, 赵春晖, 张丽丽, 张健沛. 2017. 联合空间预处理与谱聚类的协同稀疏高光谱异常检测. 光学学报, 37(4): 296-306) [DOI: 10.3788/AOS201737.0428001]
- Dempster A P, Laird N M and Rubin D B. 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society: Series B (Methodological), 39(1): 1-22 [DOI: 10.2307/2984875]
- Dilokthanakul N, Mediano P A M, Garnelo M, Lee M C H, Salimbeni H, Arulkumaran K and Shanahan M. 2017. Deep unsupervised clustering with Gaussian mixture variational autoencoders [EB/OL]. [2020-05-06]. <https://arxiv.org/pdf/1611.02648.pdf>
- Duan L, Aggarwal C, Ma S and Sathe S. 2019. Improving spectral clustering with deep embedding and cluster estimation//Proceedings of 2019 IEEE International Conference on Data Mining. Beijing, Chi-

- na; IEEE; 170-179 [DOI: 10.1109/ICDM.2019.00027]
- Ester M, Kriegel H P, Sander J and Xu X W. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise [EB/OL]. [2020-05-06]. <https://max.book118.com/html/2017/0725/124226331.shtm>
- Fraley C and Raftery A E. 1998. How many clusters? Which clustering method? Answers via model-based cluster analysis. *The Computer Journal*, 41(8): 578-588 [DOI: 10.1093/comjnl/41.8.578]
- Guo C S, ZHOU J L, Chen H H, Ying N, Zhang J W and Zhou D. 2020. Variational autoencoder with optimizing Gaussian mixture model priors. *IEEE Access*, 8: 43992-44005 [DOI: 10.1109/ACCESS.2020.2977671]
- Guo X F, Gao L, Liu X W and Yin J P. 2017. Improved deep embedded clustering with local structure preservation//*Proceedings of the 26th International Joint Conference on Artificial Intelligence*. Melbourne, Australia: AAAI; 1753-1759 [DOI: 10.24963/ijcai.2017/243]
- Hershey J R and Olsen P A. 2007. Approximating the Kullback Leibler divergence between Gaussian mixture models//*Proceedings of 2007 IEEE International Conference on Acoustics, Speech and Signal Processing*. Honolulu, USA: IEEE; IV-317-IV-320 [DOI: 10.1109/ICASSP.2007.366913]
- Jiang Z X, Zheng Y, Tan H C, Tang B S and Zhou H N. 2017. Variational deep embedding: an unsupervised and generative approach to clustering//*Proceedings of the 26th International Joint Conference on Artificial Intelligence*. Melbourne, Australia: AAAI; 1965-1972 [DOI: 10.24963/ijcai.2017/273]
- Kingma D P and Welling M. 2014. Auto-encoding variational Bayes [EB/OL]. [2020-05-06]. <https://arxiv.org/pdf/1312.6114.pdf>
- Kuhn H W. 2005. The Hungarian method for the assignment problem. *Naval Research Logistics*, 52(1): 7-21 [DOI: 10.1002/nav.20053]
- LeCun Y, Bottou L, Bengio Y and Haffner P. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11): 2278-2324 [DOI: 10.1109/5.726791]
- Lim K L, Jiang X D and Yi C Y. 2020. Deep clustering with variational autoencoder. *IEEE Signal Processing Letters*, 27: 231-235 [DOI: 10.1109/LSP.2020.2965328]
- Lu R, Xiang L, Liu M R and Yang Q. 2012. Discovering news topics from microblogs based on hidden topics analysis and text clustering. *Pattern Recognition and Artificial Intelligence*, 25(3): 382-387 (路荣, 项亮, 刘明荣, 杨青. 2012. 基于隐主题分析和文本聚类的微博客中新闻话题的发现. *模式识别与人工智能*, 25(3): 382-387) [DOI: 10.3969/j.issn.1003-6059.2012.03.004]
- MacQueen J. 1967. Some methods for classification and analysis of multivariate observations [EB/OL]. [2020-05-06]. <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=4172CEB4912D3E21EF68579C888BA56?doi=10.1.1.308.8619&rep=rep1&type=pdf>
- Mukherjee S, Asnani H, Lin E and Kannan S. 2019. ClusterGAN: latent space clustering in generative adversarial networks [EB/OL]. [2020-05-06]. <https://arxiv.org/pdf/1809.03627v1.pdf>
- Opochinsky Y, Chazan S E, Gannot S and Goldberger J. 2020. K-autoencoders deep clustering//*Proceedings of 2020 IEEE International Conference on Acoustics, Speech and Signal Processing*. Barcelona, Spain: IEEE; 4037-4041 [DOI: 10.1109/ICASSP40776.2020.9053109]
- Sabour S, Frosst N and Hinton G E. 2017. Dynamic routing between capsules [EB/OL]. [2020-05-06]. <https://arxiv.org/pdf/1710.09829.pdf>
- van der Maaten L and Hinton G. 2008. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9: 2579-2605
- Wu H, Wang Y J, Wang Z, Wang X L and Du S Z. 2010. Two-phase collaborative filtering algorithm based on co-clustering. *Journal of Software*, 21(5): 1042-1054 (吴湖, 王永吉, 王哲, 王秀利, 杜拴柱. 2010. 两阶段联合聚类协同过滤算法. *软件学报*, 21(5): 1042-1054) [DOI: 10.3724/SP.J.1001.2010.03758]
- Xiao H, Rasul K and Vollgraf R. 2017. Fashion-MNIST: a novel image dataset for benchmarking machine learning algorithms [EB/OL]. [2020-05-06]. <https://arxiv.org/pdf/1708.07747.pdf>
- Xie J Y, Girshick R and Farhadi A. 2016. Unsupervised deep embedding for clustering analysis [EB/OL]. [2020-05-06]. <https://arxiv.org/pdf/1511.06335.pdf>
- Yang B, Fu X, Sidiropoulos N D and Hong M Y. 2017. Towards K-means-friendly spaces: simultaneous deep learning and clustering [EB/OL]. [2020-05-06]. <https://arxiv.org/pdf/1610.04794.pdf>
- Yue F, Sun L, Wang K Q, Wang Y J and Zuo W M. 2008. State-of-the-art of cluster analysis of gene expression data. *Acta Automatica Sinica*, 34(2): 113-120 (岳峰, 孙亮, 王宽全, 王永吉, 左旺孟. 2008. 基因表达数据的聚类分析研究进展. *自动化学报*, 34(2): 113-120) [DOI: 10.3724/SP.J.1004.2008.00113]
- Zhang H T, Cui Y, Wang D and Song T. 2018. Study of online healthy community user profile based on concept lattice. *Journal of the China Society for Scientific and Technical Information*, 37(9): 912-922 (张海涛, 崔阳, 王丹, 宋拓. 2018. 基于概念格的在线健康社区用户画像研究. *情报学报*, 37(9): 912-922) [DOI: 10.3772/j.issn.1000-0135.2018.09.006]

作者简介



陈华, 1975年生, 男, 副教授, 主要研究方向为数字图像处理、计算机视觉和模式识别。

E-mail: iseealv@hdu.edu.cn

陈哲, 男, 硕士研究生, 主要研究方向为数字图像处理。

E-mail: 631950238@qq.com

郭春生, 男, 教授, 主要研究方向为视频分析与模式识别。

E-mail: gcs@hdu.edu.cn

应娜, 女, 副教授, 主要研究方向为信号处理、语音信号处理及其应用。E-mail: yingna@hdu.edu.cn

叶学义, 男, 副教授, 主要研究方向为图像处理、模式识别、信息隐藏。E-mail: xueyiye@hdu.edu.cn