

控制与决策

Control and Decision

基于特征和类别对齐的领域适应算法

赵小强, 蒋红梅

引用本文:

赵小强, 蒋红梅. 基于特征和类别对齐的领域适应算法[J]. 控制与决策, 2022, 37(5): 1203–1210.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2020.1241>

您可能感兴趣的其他文章

Articles you may be interested in

基于深度时序特征迁移的轴承剩余寿命预测方法

Remaining useful life prediction of bearing based on deep temporal feature transfer
控制与决策. 2021, 36(7): 1699–1706 <https://doi.org/10.13195/j.kzyjc.2019.1809>

基于生成对抗网络的大规模路网交通流预测算法

Traffic flow forecasting algorithm for large-scale road network based on GAN
控制与决策. 2021, 36(12): 2937–2945 <https://doi.org/10.13195/j.kzyjc.2020.0333>

基于改进卷积神经网络的动力下肢假肢运动意图识别

Intent recognition of power lower-limb prosthesis based on improved convolutional neural network
控制与决策. 2021, 36(12): 3031–3038 <https://doi.org/10.13195/j.kzyjc.2020.0326>

基于聚类簇结构特性的自适应综合采样法在入侵检测中的应用

Toward intrusion detection via cluster structure-based adaptive synthetic sampling approach
控制与决策. 2021, 36(8): 1920–1928 <https://doi.org/10.13195/j.kzyjc.2019.1672>

基于Hessian正则的自适应损失半监督特征选择

Adaptive loss semi-supervised feature selection based on Hessian regularization
控制与决策. 2021, 36(8): 1862–1870 <https://doi.org/10.13195/j.kzyjc.2019.1510>

基于特征和类别对齐的领域适应算法

赵小强[†], 蒋红梅

(1. 兰州理工大学 电气工程与信息工程学院, 兰州 730050; 2. 兰州理工大学 甘肃省工业过程先进控制
重点实验室, 兰州 730050; 3. 兰州理工大学 国家级电气与控制工程实验教学中心, 兰州 730050)

摘要: 针对现有的基于对抗学习的领域适应算法未能充分挖掘样本的可转移特征导致泛化能力较差和分类精度较低的问题, 提出基于特征和类别对齐的领域适应(FCDA)算法. 首先, 针对最大均值差异(MMD)度量准则存在的不足进行改进, 得到一种新的MID(maximizes the intra-domain density)度量函数, 分别度量具有相同标签的源域样本特征间的分布散度和相同标签的目标域样本特征间的分布散度, 实现最大化域内同类样本的类密度, 从而降低类的错分率; 其次, 为了能更深层次地学习目标样本的抽象的、可转移的特征, 从而减小域间差异, 在特征提取网络后加入残差校正块, 深化基础网络, 提高其特征的可迁移性; 最后, 将获取的特征经过联合判别网络, 通过对抗损失函数同时实现在类级和域级的对齐. 所提出的算法在数据集Office-31上平均准确率为88.6%, 在数据集Office-Home上平均准确率为67.7%, 并与其他算法相比, 验证了所提算法具备良好的泛化能力, 可以实现较高的分类性能.

关键词: 领域适应; 联合判别网络; 残差校正块; MID度量函数

中图分类号: TP391.41

文献标志码: A

DOI: 10.13195/j.kzyjc.2020.1241

开放科学(资源服务)标识码(OSID):



引用格式: 赵小强, 蒋红梅. 基于特征和类别对齐的领域适应算法[J]. 控制与决策, 2022, 37(5): 1203-1210.

Domain adaptation based on feature-level and class-level alignment

ZHAO Xiao-qiang[†], JIANG Hong-mei

(1. College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China; 2. Gansu Key Laboratory of Advanced Control for Industrial Processes, Lanzhou University of Technology, Lanzhou 730050, China; 3. National Experimental Teaching Center of Electrical and Control Engineering, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: Aiming at the problems of existing domain adaptation algorithms based on adversarial learning that they cannot effectively learn transferable features and have poor generalization ability, a domain adaptation algorithm based on feature and category alignment(FCDA) is proposed in this paper. First of all, in view of the shortcomings of the maximum mean discrepancy(MMD) measurement criteria, a new improved maximizes the intra-domain density(MID) measurement function is obtained, which measures the distribution divergence between the source domain sample features with the same label, and the distribution divergence between the target domain sample features with the same label, so as to maximize the class density of similar samples in the domain, thereby the class error rate is reduced. Then, in order to learn the abstract and transferable features of the target sample at a deeper level, and reduce the difference between domains, a residual correction block is added after the feature extraction network to deepen the basic network, and the transferability of its features is improved. Finally, the acquired features are passed through the joint discriminant network, and the alignments at the class-level and the domain-level are achieved with the adversarial loss function. The proposed algorithm has an average accuracy of 88.6% for the dataset Office-31 and an average accuracy of 67.7% for the dataset Office-Home. Compared with other algorithms, the proposed algorithm has better generalization ability and higher classification performance.

Keywords: domain adaptation; joint discriminant network; residual correction block; MID measurement function

收稿日期: 2020-09-07; 录用日期: 2021-03-16.

基金项目: 国家自然科学基金项目(61763029, 61763027); 国防基础科研项目(JCKY2018427C002); 甘肃省高等学校产业支撑引导项目(2019C-05); 甘肃省工业过程先进控制重点实验室开放基金项目(2019KFJJ01).

责任编辑: 曹进德.

[†]通讯作者. E-mail: xqzhao@lut.cn.

0 引言

近年来,神经网络的研究在计算机视觉领域中取得了重大突破,这主要归功于监督学习的大量有标签样本.然而,每个新任务或新数据集的采集和标注过程费时费力、成本高,这使得基于有监督的神经网络算法难以训练出一个泛化能力较强的模型.另外,在真实场景中,由于环境的变化,训练数据集和真实数据集之间存在分布差异,导致有监督神经网络体系结构的优势也会消失.而领域适应^[1-2]就是解决将分类器从源域(训练场景)适配到目标域(测试场景)的问题^[3],因此可以为标签数据短缺的实验提供高效可行的解决方案.目前,领域适应已成功应用于图像分类^[4]、人脸识别^[5]、目标检测^[6]及语义分割^[7]等多种机器学习任务当中.

一般而言,解决领域适应问题的基本着眼点是通过减少域间的数据集偏置^[8-10],达到使源域的标签预测器适用于目标域样本的目的.随着生成对抗网络(generative adversarial nets, GAN)^[11]的不断发展,基于对抗学习(adversarial learning, AL)的算法得到广泛研究.领域对抗神经网络(domain adversarial neural network, DANN)^[12]是将对抗学习引入领域适应的算法,其通过特征提取网络与判别网络之间的对抗训练获取域不变特征,但由于梯度反转层的引入,使其存在梯度消失的问题;对抗判别领域适应(adversarial discriminative domain adaptation, ADDA)^[13]针对DANN存在的问题,设置一种基于对抗学习的领域适应算法框架,使得一般的基于对抗学习的领域适应方法都可以看作是ADDA的一种特例,但其未考虑样本的迁移性,而对齐难以迁移的源域样本将扰乱目标域样本的分布,从而影响最终的结果;受条件生成对抗网络(conditional generative adversarial nets, CGAN)^[14]的启发,条件对抗领域适应(conditional domain adversarial Network, CDAN)^[15]使用源域和目标域样本类别预测向量的熵作为条件信息来设置迁移性,并取得了较好的结果,但对于样本迁移性的思考并不充分.基于正则化条件对齐的无监督领域适应(unsupervised domain adaptation via regularized conditional alignment, RCA)^[16]设置一种基于 $2K$ 维输出的联合判别器,第一 K 维输出为源域样本的类别预测概率,第二 K 维输出为目标域样本的类别预测概率,该判别器可以同时识别类信息和域信息,从而保留更多复杂的多模态信息和较高的语义信

息,使得其在目标样本上具有更好的预测精确度.

为了获取更多的目标样本的抽象信息及增强特征的迁移性,提出基于特征和类别对齐的领域适应算法(domain adaptation based on feature-level and class-level alignment, FCDA).由于深层网络架构可以自动学习更抽象的、可转移的、期望的特征,该算法在特征提取器之后加入深度残差校正网络,使目标域样本通过该网络而源域样本不经过该网络;其次,分别度量深度残差校正网络提取的具有相同标签信息的目标域样本特征的分布散度、特征提取网络提取的具有相同标签信息的源域样本特征间的分布散度,使得源域样本和目标域样本在特征层面最小化域间差异,最大化域内密度.同时,添加的深度残差校正网络能增加现有的网络深度,使得源域样本和目标域样本在特征层面实现更精确的对齐.

1 理论基础

1.1 无监督领域适应

在无监督领域适应^[17]中,通常包含两个领域的数据集:一个包含 n_s 个已标记样本的数据集(合成数据集,计算机自动标注),称为源域 $X_s = \{(x_s, y_s)\}_{i=1}^{n_s}$, y_i 为 X_s 中的一个样本 x_i 对应的标签信息;一个包含 n_t 个未标记样本的数据集(真实数据集),但却包含实际需要预测的样本,称为目标域 $X_t = \{x_t\}_{i=1}^{n_t}$. X_s 和 X_t 分别服从于总体分布 P 和 Q , P 和 Q 存在一定的分布差异.无监督领域适应的目的是设计一种算法减小 P 与 Q 之间的分布差异,使得使用 P 训练出的模型仍然能有效地应用在 Q 上.

1.2 RCA算法

RCA算法的网络结构示意图如图1所示,其学习过程由分类部分和对抗部分并行完成:1)分类部分使用带有标签信息的源域样本依次通过特征提取网络 $f_s = F(x)$ 和预测网络 $g_s = C(f_s)$,预测网络 C 用于对源域样本的预测结果 g_s 进行分类,并对各个类别空间的梯度信息反馈特征提取网络 F ;2)对抗部分通过训练特征提取网络 F “欺骗”联合判别网络 D ,从而减小域间差异,其中 D 的维度为 $2K$,第一个 K 维为源域样本的类别信息,第二个 K 维为目标域样本的类别信息.通过训练特征提取网络 F 使其在每个域的同类样本中对抗联合判别网络 D ,使得该联合判别网络可以同时学习类别信息和域信息的分布.最后,将训练好的预测网络 C 和特征提取网络 F 应运到目标域样本分类中.

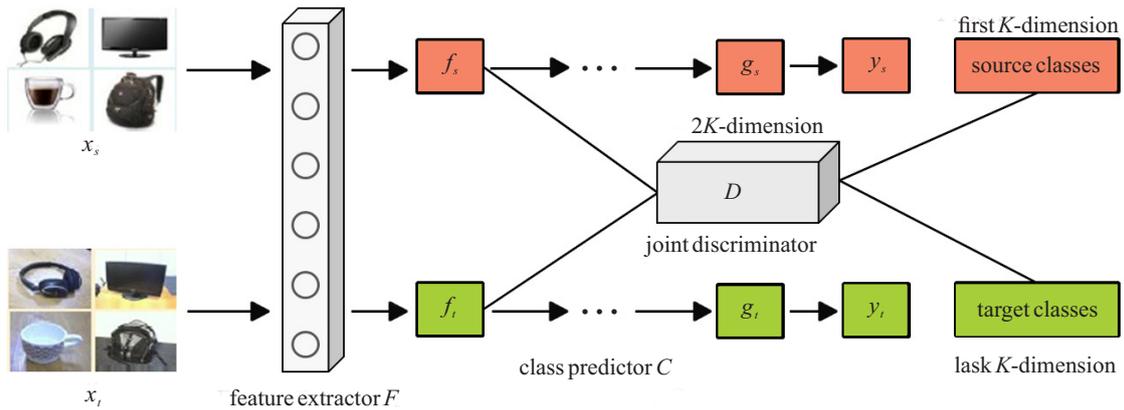


图1 RCA算法的网络结构

2 基于特征和类别对齐的对抗领域适应算法

2.1 FCDA算法的网络结构

通过对RCA算法分析发现存在以下问题:一是仅使用RCA算法的领域适应不足以最小化源域分布

和目标域分布的差异,缺少一种差异度量准则,从而提高目标域预测精确度;二是对于目标样本仅使用特征提取网络不能够提取更多抽象的、可转移的特征表示.针对上述不足,本文提出FCDA算法,其网络结构如图2所示.

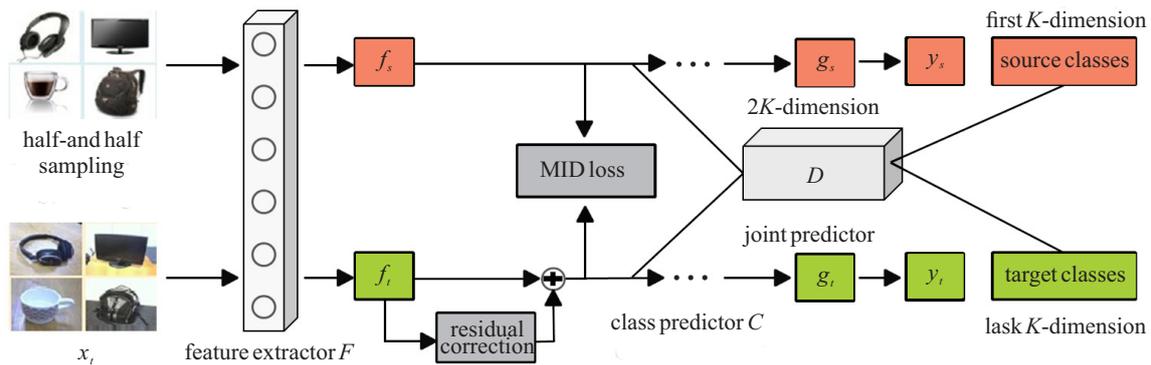


图2 FCDA算法的网络结构

由图2可知,FCDA算法的网络结构由特征提取网络F、类预测网络C、残差校正网络及联合判别网络D四部分组成,下面分别对FCDA算法的4个网络结构的特点和作用进行介绍.

2.1.1 特征提取网络

设n为源域和目标域输入样本的大小,特征提取网络F使用卷积神经网络结构,并将源域样本x_{sn}和目标域样本x_{tn}作为其输入.在训练过程中,F逐渐学习两个领域的域不变特征,并产生输出f_s = F(x_{sn})和f_t = F(x_{tn}).

2.1.2 残差校正网络

相关研究表明深层网络架构可以自动学习更抽象、可转移的特征表示^[18-19],对最终的目标样本的分类至关重要.然而,在深度网络中学习到的特征表示

只能减少,且连续的非线性变换可能会放大域之间的差异.为了在加深网络架构的同时减小域间差异,在特征提取网络加入适用于本文算法的残差校正块,其网络结构如图3所示.

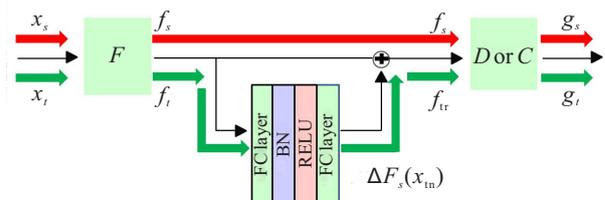


图3 残差校正块

由图3可知:只有目标域样本的特征表示f_t经过残差校正块,以学习域间差异ΔF_s(X_{tn});通过训练学习使得校正后的目标样本的特征表示f_{tr} = f_t + ΔF_s(X_{tn})更类似于源域样本的特征表示f_s,从

而达到在学习目标样本抽象的、可转移特征表示的同时在特征层面减小域间差异. 此外, 这种弱共享结构使得目标域能够有效地对其与源域的差异进行建模.

残差校正块只是在特征层面减小了域间差异, 为了进一步提高最终的目标样本的预测精确度, 本文提出一种新的度量函数MID来最大化源域和目标域的域内密度, 这使域内具有相同类别的样本的特征表示更加紧凑. MID的一般形式为

$$E_{(X_s, X_s^r) \sim P} |X_s - X_s^r|_l + E_{(X_t, X_t^r) \sim Q} |X_t - X_t^r|_l, \quad (1)$$

其中 $|\cdot|_l$ 为不同的范数. 在本文中, n_s 和 n_t 分别为源域样本数和目标域样本数, 则源域样本为 $\{x_{s1}, x_{s2}, \dots, x_{sn_s}\}$, 目标域样本为 $\{x_{t1}, x_{t2}, \dots, x_{tn_t}\}$, 对应的标签为 y_s 和 y_t , 其中目标域样本无标签信息, 因此 y_t 为伪标签, 则式(1)可修改为

$$\frac{1}{n_s^2} \sum_{i,j}^{n_s, n_s} \|x_{si} - x_{sj}^r\|_2^2 + \frac{1}{n_t^2} \sum_{i,j}^{n_t, n_t} \|x_{ti} - x_{tj}^r\|_2^2. \quad (2)$$

其中 x_{sj}^r 和 x_{tj}^r 分别为与 x_{si} 和 x_{ti} 具有相同分布的样本. 由于深度网络训练是批量进行的, 而式(2)中考虑了所有样本间距离, 且 x_{sj}^r 和 x_{tj}^r 的选择也不够明确. 针对上述问题, 本文对式(2)进一步改进, 只计算相对位置的样本间的距离, x_{sj}^r 和 x_{tj}^r 分别为与 x_{si} 和 x_{ti} 有相同标签的样本, 并将源域样本与目标域样本在特征空间对齐, 则可将式(2)修改为

$$l_{\text{MID}} = \frac{1}{n_s} \sum_{y_{si}=y_{sj}^r} \|f_{si} - f_{sj}^r\| + \frac{1}{n_t} \sum_{y_{ti}=y_{tj}^r} \|f_{ti} - f_{tj}^r\|, \quad (3)$$

其中 n_s 和 n_t 分别为具有相同标签信息的源域样本成对数和目标域样本成对数.

2.1.3 类预测网络

首先, 类预测网络 C 将 f_s 作为输入, 并使用源域样本的标签信息对其训练, 从而使其能够正确地分类源域样本及预测目标样本的伪标签; 然后, 将 f_s 和 f_{tr} 作为输入, 指导 C 同时对源域和目标域的数据进行分类, 同时为了使目标域的预测结果更精确, 在训练过程中, 对目标域样本在 C 上使用平方熵损失函数进行熵最小化, 则 C 对应的损失函数表示如下:

$$l_{\text{sc}} = E_{(x_s, y_s) \sim P} l_{\text{CE}}(g_s, y_s), \quad (4)$$

$$l_{\text{te}} = E_{(x_t, y_t) \sim Q} l_E(g_t). \quad (5)$$

其中

$$l_{\text{CE}}(g_s, y_s) = -\langle y_s, \log(g_s) \rangle,$$

$$l_E(g_t) = -\sum_{i=1}^n g_t \cdot g_t,$$

l_{sc} 为类预测损失函数, l_{te} 为平方熵损失函数.

2.1.4 联合判别网络

联合判别网络 D 接受 f_s 和 f_{tr} 作为输入, D 的输出为 $2K$ 维, 第一个 K 维是源域的分类分布, 第二个 K 维是目标域的分类分布, 即: 当 D 的输入为 f_s 时, 对应的输出为 $[d_s \ 0]$, 当 D 的输入为 f_{tr} 时, 对应的输出为 $[0 \ d_t]$, 其中 d_s 和 d_t 为源域样本和目标域样本的类预测概率分布. D 的功能包含两部分, 分别为: 1) 训练 D 使其能够分辨出 F 提取出的特征来自源域还是目标域; 2) 通过与 F 之间的对抗训练, 使 D 不能辨别出 F 中提取出的特征来自源域还是目标域. 通过学习域和类变量的分布在 D 中同时实现域级别和类别级别的对齐, 对应的损失函数表示如下:

$$l_{\text{dsc}} = E_{(x_s, y_s) \sim P} l_{\text{CE}}(D(f_s), [y_s, 0]), \quad (6)$$

$$l_{\text{dte}} = E_{x_t \sim Q} l_{\text{CE}}(D(f_{tr}), [0, y_t]), \quad (7)$$

$$l_{\text{dsa}} = E_{(x_s, y_s) \sim P} l_{\text{CE}}(D(f_s), [0, y_s]), \quad (8)$$

$$l_{\text{dta}} = E_{(x_s, y_s) \sim Q} l_{\text{CE}}(D(f_{tr}), [y_t, 0]). \quad (9)$$

2.2 算法步骤

FCDA算法步骤如下.

input: 源域数据集 $X_s = \{(x_s, y_s)\}_{i=1}^{n_s}$, 目标域数据集 $X_t = \{x_t\}_{i=1}^{n_t}$, 训练次数分别为 K_1, K_2 , batch size的大小为 n .

step 1: 采用ImageNet^[20]预训练模型的参数初始化网络层参数.

step 2: for k in $1:K_1$ do

step 2.1: 从 X_s 中采样 n 个有标记数据 $\{x_{si}, y_i\}_{i=1}^n$, 记为 x_{sn} ;

step 2.2: x_{sn} 通过网络 F 得到 f_s , 再分别通过 C 和 D 网络计算得到 g_s 和 $D(f_s)$;

step 2.3: 根据式(4)和(6)中对应的损失函数 l_{sc} 、 l_{dsc} 的计算公式, 训练 C 和 D 对源域样本进行正确分类, 目标函数为 $\min_{C, D} l_{\text{sc}} + \lambda_{\text{dsc}} l_{\text{dsc}}$.

step 3: end for

step 4: for k in $1:K_2$ do

step 4.1: 从 X_s 中采样 n 个有标记数据 $\{x_{si}, y_i\}_{i=1}^n$, 记为 x_{sn} , 从 X_t 中采样 n 个无标记数据 $\{x_{ti}\}_{i=1}^n$, 记为 x_{tn} ;

step 4.2: x_{sn} 通过网络 F 得到 $f_s = F(x_{sn})$, x_{tn} 通过网络 F 得到 $f_t = F(x_{tn})$, 再分别通过残差校正块、 C 、 D 网络计算得到 f_{tr} 、 g_s 、 g_t 、 $D(f_s)$ 、 $D(f_{tr})$;

step 4.3: 根据式(4)~(7)中对应的损失函数 l_{sc} 、 l_{te} 、 l_{dsc} 、 l_{dtc} 的计算公式, 训练 C 和 D 样本并进行正确分类, 目标函数为

$$\min_{C,D} l_{sc} + \lambda_{te} l_{te} + \lambda_{dsc} l_{dsc} + \lambda_{dtc} l_{dtc};$$

step 4.4: 根据式(3)、(8)和(9)中对应的损失函数 l_{MID} 、 l_{dsa} 、 l_{dta} 的计算公式, 训练 F 使 D 不能分辨出样本来自于源域还是目标域, 目标函数为

$$\min_F \lambda_{MID} \cdot l_{MID} + \lambda_{dsa} l_{dsa} + \lambda_{dta} l_{dta}.$$

step 5: end for

3 实验结果及分析

为了验证本文算法的可行性和有效性, 将本文算法与残差神经网络(ResNet)、DANN、ADDA、CDAN及RCA等主流算法在目前流行的公开数据集Office-31^[21]、ImageCLEF-DA^[22]及Office-Home^[23]上进行比较.

3.1 实验数据集

Office-31是一个基于领域适应且应用较为广泛的数据集, 一共包含4652张图片, 分为31个类别, 图片源于3个不同的领域, 分别为在线电商图片Amazon(A)、摄像头拍摄的低分辨率图片Webcam(W)和数字单反相机拍摄的高分辨率图片DSLR(D). 实验中, 对这3个领域的数据集设置6种迁移任务.

ImageCLEF-DA包含12个类别, 图片取自3个不同的领域, 分别为加州理工学院的图像数据库Caltech-256(C)、用于图像识别最大的数据库ImageNet中的ILSVRC 2012(I)及用于分类和检测视觉对象的一个基准测试集Pascal VOC 2012(P), 实验中, 对这3个领域的数据集设置6种迁移任务.

Office-Home是一个更具迁移难度的数据集, 共包含15500张图片, 分为65个类别, 图片取自4个领域, 分别为艺术画图片Art(Ar)、插画图片Clipart(CI)、商品展示图片Product(Pr)和实际生活中所拍摄的图片Real World(Rw). 在实验中, 对这4个领域的图片设置12种迁移任务. Office-Home的迁移难度主要体现在: 首先, Office-Home数据集中各领域差异大、类别多、相似度低; 其次, Office-Home数据集包含的类别数较多, 模型需要学习较为复杂的模态.

3.2 实验设置

本文实验所用的深度学习框架为Pytorch^[24], 在搭载GPU为RTX2080Ti的服务器实验环境下使用Python3.6, 在网络框架中, 使用ResNet-50^[25]作为

基础的特征提取网络, 其初始学习率为0.001, 其中ResNet的初始参数为使用ImageNet预训练的模型参数. C 和 D 由两个全连接层构成, 其初始学习率为0.001. 对于优化器的设置, 采用动量为0.9的SGD更新参数, 同时采用与文献[8]相同的优化策略, 学习率 η_p 由公式 $\eta_p = \eta_0 / (1 + \alpha p)^\beta$ 计算所得, 其中 p 指模型训练完成程度, 范围为 $0 \sim 1$, 并设置 η_0 、 α 和 β 的优化器参数组合. 在实验过程中, 参数 λ_{dsc} 和参数 λ_{dtc} 的取值范围为 $\{0.1, 0.5, 1.0\}$, 参数 λ_{MID} 和参数 λ_{te} 的取值范围为 $\{0.1, 1.0\}$, 参数 λ_{dsa} 和参数 λ_{dta} 的取值范围为 $\{0.05, 0.1, 0.5, 1.0\}$, 而这些参数在数据集Office-31及数据集ImageCLEF-DA的最优设置分别如表1、表2所示.

表1 基于Office-31数据集的最优参数设置

参数	A→W	D→W	A→D	W→D	D→A	W→A
λ_{dsc}	1.0	1.0	1.0	1.0	0.5	1.0
λ_{dtc}	1.0	1.0	1.0	1.0	0.5	1.0
λ_{MID}	0.1	0.1	0.1	0.1	0.1	0.1
λ_{te}	0.1	0.1	0.1	0.1	1.0	1.0
λ_{dsa}	0.05	0.05	0.05	0.05	0.1	0.1
λ_{dta}	0.05	0.05	0.05	0.05	0.1	0.1

表2 基于ImageCLEF-DA数据集的最优参数设置

参数	C→I	P→I	C→P	P→C	I→C	I→P
λ_{dsc}	1.0	1.0	0.5	1.0	1.0	1.0
λ_{dtc}	1.0	1.0	0.5	1.0	1.0	1.0
λ_{MID}	0.1	0.1	0.1	0.1	0.1	0.1
λ_{te}	0.1	0.1	1.0	0.1	0.1	1.0
λ_{dsa}	0.05	0.05	0.1	0.05	0.05	0.1
λ_{dta}	0.05	0.05	0.1	0.05	0.05	0.1

3.3 实验结果及分析

Office-31数据集是领域适应中使用最为广泛的数据集之一, 其对应的实验结果如表3所示.

由表3可知: 所提出的FCDA算法与ResNet^[25]、DANN、ADDA、基于联合适配的领域适应算法(JAN)^[22]、CDAN及RCA算法进行对比, 其在6个迁移任务上都具有更好的性能. 与仅使用二分类判别网络执行域级别对齐的DANN及ADDA算法相比, FCDA算法同时对齐了域级别信息和类级别信息; 与使用类别预测向量的熵作为条件信息来设置迁移性的CDAN算法相比, FCDA算法在增强特征迁移性的同时获取了更多目标样本的抽象信息. 由于本文提出的FCDA算法和RCA算法共享相似的对抗网络,

表3 基于Office-31数据集的实验结果

方法	A→W	D→W	A→D	W→D	D→A	W→A	Avg
ResNet-50	68.4	96.7	99.3	68.9	62.5	60.7	76.1
DANN	82.0	96.9	99.1	79.7	68.2	67.4	82.2
ADDA	86.2	96.2	98.4	77.8	69.5	68.9	82.9
JAN	85.4	97.4	99.8	84.7	68.6	70.0	84.3
CDAN	93.1	98.6	100.0	92.9	71.0	69.3	87.5
RCA	92.8	97.9	100.0	91.3	68.0	69.5	86.6
FCDA	95.0	98.9	100.0	93.8	71.2	72.5	88.6

RCA算法与FCDA算法之间的比较可以反映本文所提MID损失函数及残差校正块在解决特征迁移性、减小源域和目标域分布时的有效性.对于最具挑战性的两个迁移任务D→A和W→A,本文算法FCDA比RCA算法的精确度分别高出3.2%和3.0%.就6个迁移任务的平均精确度而言,FCDA算法比RCA算法高2.0%,原因是本文算法提出的MID损失函数及残差校正块提高了RCA算法的分类性能.

Office-31数据集的图片只来源于办公场景,场景较为单一,而ImageCLEF-DA数据集的图片包含的场景较为丰富,包括人、动物、车辆等不同的场景,因

此,测试本文提出的FCDA算法在ImageCLEF-DA数据集各迁移任务上的精确度可验证FCDA算法在不同场景中的有效性.基于ImageCLEF-DA数据集的实验结果如表4所示.

由表4结果可知:当Pascal作为目标域时,迁移任务变得具有挑战性,其他迁移任务的精确度都高于90%,与RCA算法相比,平均精度值提高了2.0%,与CDAN算法相比,平均精度值提高了1.1%,这说明本文提出的FCDA算法不仅在应用广泛的办公场景中具有良好的分类性能,而且在其他不同的场景中也具有较高的分类效果.

表4 基于ImageCLEF-DA数据集的实验结果

方法	C→I	P→I	C→P	P→C	I→C	I→P	Avg
ResNet-50	74.8	83.9	91.5	78.0	65.5	91.2	80.7
DAN	74.5	82.2	92.8	86.3	69.2	89.8	82.5
DANN	75.0	86.0	96.2	87.0	74.3	91.2	85.0
MADA	75.0	87.9	96.0	88.8	75.2	92.2	85.8
CDAN	76.7	90.6	97.0	90.5	74.5	93.5	87.1
RCA	78.5	92.3	97.7	92.0	76.9	94.8	88.7
FCDA	79.0	94.3	98.5	93.0	78.3	95.2	89.7

与Office-31数据集及ImageCLEF-DA数据集相比,Office-Home数据集中图片的类别较多,各领域之间的差异较大,迁移任务更多.因此,测试本文提出的FCDA算法在Office-Home数据集各迁移任务上的精确度可验证FCDA算法在复杂场景中的有效性.基于Office-Home数据集的实验结果如表5所示.与RCA算法相比,本文算法在12个迁移任务的精确度上都得到了提升,平均精确度提高了1.7%,验证了本文提出的图像分类算法在复杂场景上的有效性和先进性.

为了进一步验证本文算法的有效性,对Office-31数据集中的任务A→W进行可视化,如图4所示,红点

表示源域数据,蓝点表示目标域数据.

由图4可知:在只使用源域样本训练算法时,目标域数据散乱地分布,未能观察到任何关于目标域样本间的类别信息及源域与目标域之间的适应信息,这说明源域数据与目标域数据之间存在较大的差异;在使用RCA算法对其进行领域适应分类后,域之间的类间距离变小;而在使用FCDA算法对其进行领域适应分类后,源域数据与目标域数据的类间距离变得更小,具有相同类别的源域样本与目标域样本较好地拟合在一起.相较于RCA算法,FCDA算法的错分样本的数量更少,适应性能更强.

表5 基于Office-Home数据集的实验结果

方法	Ar→Cl	Ar→Pr	Ar→Rw	Cl→Ar	Cl→Pr	Cl→Rw	Pr→Ar	Pr→Cl	Pr→Rw	Rw→Ar	Rw→Cl	Rw→Cl	Avg
ResNet-50	34.9	50.0	58.0	37.4	41.9	46.2	38.5	31.2	60.4	53.9	41.2	59.9	46.1
DAN	43.6	57.0	67.9	45.8	56.5	60.4	44.0	43.6	67.7	63.1	51.5	74.3	56.3
DANN	45.6	59.3	70.1	47.0	58.5	60.9	46.1	43.7	68.5	63.2	51.8	76.8	57.6
JAN	45.9	61.2	68.9	50.4	59.7	61.0	45.8	43.4	70.3	63.9	52.4	76.8	58.3
CDAN	50.7	70.6	76.0	57.6	70.0	70.0	57.4	50.9	77.3	70.9	56.7	81.6	65.8
RCA	50.0	71.0	76.9	60.7	69.3	70.5	59.8	46.5	77.7	71.1	55.8	82.1	66.0
FCDA	53.6	72.3	78.1	61.5	70.9	71.5	62.3	50.5	79.8	72.3	56.7	83.4	67.7

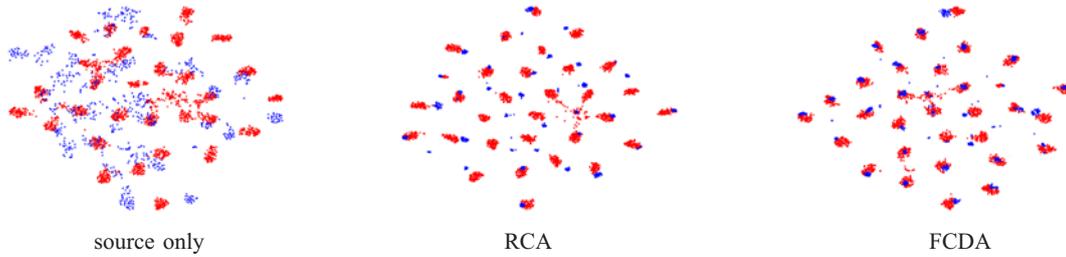


图4 T-SNE可视化

为了验证本文算法的稳定性及收敛性,对 Office-31 数据集上的任务 A→W 在 RCA 和 FCDA 算法上的训练曲线进行了可视化(如图5所示).

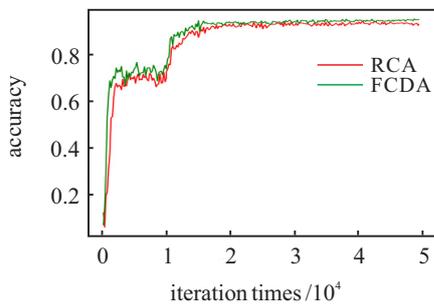


图5 算法收敛性对比实验

在图5中:前10000次迭代是无领域适应算法收敛性对比实验时目标样本的预测平均精确度.

由图5可以看到,未进行领域适应时,目标样本的预测精确度较低且处于震荡状态,加入领域适应后,目标样本的精确度快速上升,并最终趋于稳定,且 RCA 和 FCDA 算法都有着较为稳定的结果. FCDA 算法相对于 RCA 算法性能更好,结果也更具鲁棒性.

图6为不同迭代次数对应的 MID 度量函数.

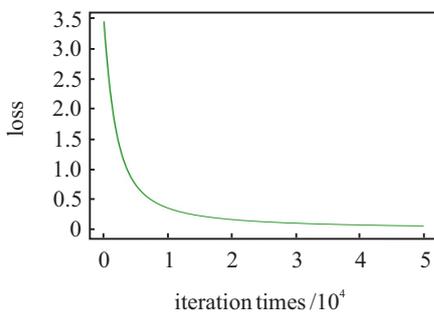


图6 MID度量函数

由图6可知, MID 度量函数随着迭代次数的增加稳定地减小,直至收敛,进一步验证了 FCDA 算法的稳定性.

4 结论

针对现有基于对抗学习的领域适应算法获取目标样本的抽象的、可转移的特征能力较弱从而导致

领域适应的分类精度不高的问题,本文提出了一种基于特征和类别对齐的对抗领域适应算法 FCDA,该算法学习源域数据分布和目标域数据分布的共享特征嵌入,并联合残差校正块和 MID 度量函数进一步减小域间差异,增大同一域内的相同类别的样本的类密度.同时,该算法还利用联合判别网络学习多模态信息,在两个领域适应的公开数据集上的实验结果显示了 FCDA 算法在无监督领域适应分类中的有效性和优越性.如何解决在源域数据集较小的两个领域中引起的迁移能力弱化的问题需要进一步研究.

参考文献(References)

- [1] Long M S. Transfer learning: Problems and methods[D]. Beijing: Tsinghua University, 2014.
- [2] Pan S J, Yang Q. A survey on transfer learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10): 1345-1359.
- [3] Bousmalis K, Silberman N, Dohan D, et al. Unsupervised pixel-level domain adaptation with generative adversarial networks[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017: 95-104.
- [4] Lunga D, Yang H L, Reith A, et al. Domain-adapted convolutional networks for satellite image classification: A large-scale interactive learning workflow[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(3): 962-977.
- [5] Sohn K, Liu S F, Zhong G Y, et al. Unsupervised domain adaptation for face recognition in unlabeled videos[C]. Proceedings of the 2017 IEEE International Conference on Computer Vision. Washington, 2017: 5917-5925.
- [6] Chen Y, Li W, Sakaridis C, et al. Domain adaptive faster R-CNN for object detection in the wild[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington, 2018: 3339-3348.
- [7] Zou Y, Yu Z, Kumar B V K V, et al. Unsupervised domain adaptation for semantic segmentation via class-balanced self-training[C]. Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2018: 297-313.
- [8] Wang M, Deng W H. Deep visual domain adaptation: A

- survey[J]. *Neurocomputing*, 2018, 312: 135-153.
- [9] 王雪松, 赵季娟, 程玉虎. 基于极限学习机的无监督领域适应分类器[J]. *控制与决策*, 2020, 35(4): 861-869. (Wang X S, Zhao J J, Cheng Y H. Unsupervised adaptive classifier based on extreme learning machines[J]. *Control and Decision*, 2020, 35(4): 861-869.)
- [10] 臧绍飞, 程玉虎, 王雪松. 基于最大分布加权均值嵌入的领域适应学习[J]. *控制与决策*, 2016, 31(11): 2083-2089. (Zang S F, Cheng Y H, Wang X S. Domain adaptation learning based on maximum distribution weighted mean discrepancy[J]. *Control and Decision*, 2016, 31(11): 2083-2089.)
- [11] Goodfellow I, Poudet-abadie J, Mirza M, et al. Generative adversarial nets[C]. *Proceedings of the 27th International Conference on Neural Information Processing Systems*. Cambridge: MIT Press, 2014: 2672-2680.
- [12] Ganin Y, Ustinova E, Ajakan H, et al. Domain-adversarial training of neural networks[J]. *The Journal of Machine Learning Research*, 2016, 17(1): 2096-2030.
- [13] Tzeng E, Hoffman J, Saenko K, et al. Adversarial discriminative domain adaptation[C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Washington, 2017: 2962-2971.
- [14] Mirza M, Osindero S. Conditional generative adversarial nets[J/OL]. 2014, Arxiv: 1411.1784.
- [15] Long M S, Cao Z J, Wang J M, et al. Conditional adversarial domain adaptation[C]. *Proceedings of the 32nd Conference on Neural Information Processing Systems*. Cambridge: MIT Press, 2018: 1647-1657.
- [16] Cicek S, Soatto S. Unsupervised domain adaptation via regularized conditional alignment[C]. *2019 IEEE Conference on Computer Vision and Pattern Recognition*. Washington, 2019: 1415-1420.
- [17] Ganin Yaroslav, Victor Lempitsky. Unsupervised domain adaptation by backpropagation[C]. *The 32nd International Conference on Machine Learning*. New York: ACM, 2015: 1180-1189.
- [18] Long M S, Cao Y, Wang J M, et al. Learning transferable features with deep adaptation networks [C]. *Proceedings of the 32nd International Conference on Machine Learning*. New York: ACM, 2015: 97-105.
- [19] Yosinski J, Clune J, Bengio Y, et al. How transferable are features in deep neural networks?[C]. *Proceedings of the 27th International Conference on Neural Information Processing Systems*. Cambridge: MIT Press, 2014: 3320-3328.
- [20] Deng J, Dong W, Socher R, et al. ImageNet: A large-scale hierarchical image database[C]. *2009 IEEE Conference on Computer Vision and Pattern Recognition*. Washington, 2009: 248-255.
- [21] Saenko K, Kulis B, Fritz M, et al. Adapting visual category models to new domains[C]. *Proceedings of the 11th European Conference on Computer Vision*. Berlin: Springer, 2010: 213-226.
- [22] Long M S, Zhu H, Wang J M, et al. Deep transfer learning with joint adaptation networks[C]. *Proceedings of the 34th International Conference on Machine Learning*. New York: ACM, 2017: 2208-2017.
- [23] Venkateswara H, Eusebio J, Chakraborty S, et al. Deep hashing network for unsupervised domain adaptation[C]. *2017 IEEE Conference on Computer Vision and Pattern Recognition*. Washington, 2017: 5385-5394.
- [24] Paszke A, Gross S, Chintala S, et al. Automatic differentiation in pyTorch[C]. *Proceedings of the 31st Conference on Neural Information Processing Systems*. Cambridge: MIT Press, 2017.
- [25] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]. *2016 IEEE Conference on Computer Vision and Pattern Recognition*. Washington, 2016: 770-778.

作者简介

赵小强(1969—), 男, 教授, 博士生导师, 从事故障检测与诊断、图像处理等研究, E-mail: xqzhao@lut.cn;

蒋红梅(1995—), 女, 硕士生, 从事图像处理的研究, E-mail: 18709467853@163.com.

(责任编辑: 闫妍)