第46卷 第10期
2023年10月

计 算 机 学 报
CHINESE JOURNAL OF COMPUTERS

Vol. 46 No. 10
Oct. 2023

基于张量奇异值分解的视觉域自适应方法

李国瑞[1)] 许鹏飞[1)] 彭三城[2)3)] 阳爱民[4)]

[1)](东北大学计算机科学与工程学院 沈阳 110169)

[2)](广东外语外贸大学外国语言学及应用语言学研究中心 广州 510006)

[3)](广东外语外贸大学语言工程与计算实验室 广州 510006)

[4)](岭南师范学院计算机科学与智能教育学院 湛江 广东 524048)

摘 要 近年来,机器学习在计算机视觉中取得了许多突破性的研究进展.然而,已训练好的学习模型难以直接应用于相似但具有不同数据分布特征的其它学习任务中.域自适应技术通过抽取源域与目标域数据之间的公共特征,来实现把源域中学习到的知识迁移至目标域,从而避免针对目标域的训练数据收集和模型训练代价.但是,现有的视觉域自适应方法大都无法处理高阶的特征数据,一般都是通过简单的向量化操作将高阶张量特征转换成高维一阶向量特征.这不仅会破坏高阶特征数据内部的结构信息,而且还会增加算法的计算复杂度.为了解决上述问题,本文在保持原有张量特征结构不变的条件下,利用张量乘操作,将视觉域自适应问题抽象为求解源域和目标域的共同张量子空间以及源域和目标域特征在该共同张量子空间上投影的多变量优化问题.然后,利用张量奇异值分解和交替方向乘子法,提出一种基于张量奇异值分解的视觉域自适应方法(Visual domain Adaptation method based on TEnsor Singular value decomposition, VATES),以实现上述多变量优化问题的迭代求解.文中证明了正交张量子空间约束条件下源域与目标域表征误差最小化问题的可解性问题,并求得了相应的解析解.在公开数据集Office-Caltech-10、Office31、ImageNet-VOC2007上与17个基线模型进行对比实验.结果表明本文所提出的方法与经典的机器学习方法、非深度域自适应方法、深度域自适应方法以及张量域自适应方法相比,在无标签目标域上的图像分类精度分别提高了10.6%～43.9%、0.7%～31.1%、0.7%～24.8%以及5.7%～34.9.同时,算法的运行效率也提高了40.5%～74.3%,显著优于所对比的基线方法.实验分析也表明,VATES方法的目标域分类精度会随着所选用神经网络特征抽取能力的增强而逐渐提升.

关键词 机器学习;域自适应;张量;奇异值分解;子空间

中图法分类号 TP181 **DOI号** 10.11897/SP.J.1016.2023.02084

Visual Domain Adaptation Method Based on Tensor Singular Value Decomposition

LI Guo-Rui[1)] XU Peng-Fei[1)] PENG San-Cheng[2)3)] YANG Ai-Min[4)]

[1)](*School of Computer Science and Engineering, Northeastern University, Shenyang* 110169)

[2)](*Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies, Guangzhou* 510006)

[3)](*Laboratory of Language Engineering and Computing, Guangdong University of Foreign Studies, Guangzhou* 510006)

[4)](*School of Computer Science and Intelligence Education, Lingnan Normal University, Zhanjiang, Guangdong* 524048)

Abstract In recent years, machine learning has made a series of breakthroughs in computer vision applications. However, the well-trained machine learning model cannot be directly applied to other related machine learning tasks which have similar but different data distribution features. Thus, a large

收稿日期:2022-09-15;在线发布日期:2023-07-11. 本课题得到国家自然科学基金(61876205)、河北省自然科学基金(F2020501034)、河北省高等学校科学研究项目(ZD2021403)、教育部人文社科一般项目(20YJAZH118、19YJAZH128)、广东外语外贸大学外国语言学及应用语言学研究中心语言与人工智能重点实验室招标课题(LAI202306)的资助. **李国瑞**,博士,副教授,中国计算机学会(CCF)会员,主要研究领域为机器学习、优化理论. E-mail: lgr@neuq.edu.cn. **许鹏飞**,硕士研究生,主要研究领域为迁移学习、计算机视觉. **彭三城**(通信作者),博士,教授,中国计算机学会(CCF)高级会员,主要研究领域为领域自适应、情绪计算. E-mail: psc346@aliyun.com. **阳爱民**,博士,教授,中国计算机学会(CCF)高级会员,主要研究领域为机器学习、工业软件、文本情感分析.

number of domain adaptation methods have been proposed in the literature to take full advantage of the well-trained machine learning models as well as to avoid the expensive data collection prices and the time-consuming model training costs. By extracting the common features between the source domain and the target domain simultaneously, these domain adaptation methods can transfer the well-trained model from a source domain to a different but correlated target domain efficiently. However, most existing visual domain adaptation methods cannot deal with high-order tensorial features directly. They can only convert a high-order feature tensor into a one-order high-dimensional feature vector naively by carrying out a simple vectorization operation, which not only destroys the internal structure information within the original high-order feature tensor, but also increases its computational complexity in the subsequent operations. To overcome the above challenges, in this paper, the visual domain adaptation problem is formulated as a multiple variables optimization problem by utilizing the tensor singular value decomposition (t-SVD) and the tensor product (t-product) operator. Specifically, a common tensorial subspace between the source domain and the target domain is built constructively, and the projections of the source features and the target features on this common tensorial subspace are also computed correspondingly. In order to solve the multiple variables optimization problem, a visual domain adaptation method based on tensor singular value decomposition (VATES) is proposed in this paper by adopting the tensor singular value decomposition and the alternating direction method of multipliers (ADMM). Moreover, we prove the solvability of the multiple variables optimization problem that minimizes the representation errors in both the source domain and the target domain under the constraint of a common orthogonal tensorial subspace. The corresponding analytical solution is also provided at the same time. Extensive experiments were carried out by comparing with 17 baseline visual domain adaptation methods on three popular visual datasets like the Office-Caltech-10 dataset, the Office31 dataset, and the ImageNet-VOC2007 dataset. The experiment results show that compared with the classical machine learning methods, the non-deep domain adaptation methods, the deep domain adaptation methods, and the tensor-based domain adaptation methods, the image classification accuracy on the unlabeled target domain of our proposed VATES method is improved by 10.6%～43.9%, 0.7%～31.1%, 0.7%～24.8%, and 5.7%～34.9%, respectively. Meanwhile, the running efficiency of the VATES method is also improved by 40.5%～74.3%, which is remarkably superior to the compared visual domain adaptation methods. The excellent high-order data representation ability of the tensor singular value decomposition renders the VATES method's superior model transfer performance. Experiment analysis also shows that the classification accuracy on the unlabeled target domain of our proposed VATES method improves gradually with the enhancement of the feature extraction ability of the adopted underlying feature extractor.

Keywords　machine learning; domain adaptation; tensor; singular value decomposition; subspace

1　引　言

近年来，随着机器学习技术的飞速发展，其在计算机视觉[1-2]、自然语言处理[3]、推荐系统[4]、智能驾驶[5]、医学诊断[6]等众多领域均取得了一系列突破性的研究成果．然而，这些算法往往都需要有大量带标签数据的支撑[7]．而且，传统机器学习方法只能针对特定学习任务进行训练，无法将已训练好的学习模型快速应用于相似但具有不同数据分布特性的其它学习任务中[8]．

针对数据标记成本高、模型应用场景缺乏扩展性等问题，域自适应技术应运而生[9]．域自适应作为迁移学习的一个重要分支，通过充分挖掘源域与目标域数据之间的公共特征，复用源域中已有的学习模型以解决目标域中的相关问题[10]．它放宽了机器学习中训练集和测试集需满足独立同分布的假设，借助源域与目标域之间的数据相关性实现了源域中

已有模型向目标域的迁移．域自适应技术在计算机视觉[11]、自然语言处理[12]、工业故障诊断[13]、医学影像分析、计算生物学等应用领域具有广泛的应用前景．

然而，现有的视觉域自适应方法大都只能处理一阶向量形式的特征数据，无法针对高阶张量特征数据进行处理．简单的向量化操作不仅会破坏高阶张量特征数据的内部结构信息，而且还会增加模型计算的复杂度[14]．针对上述问题，本文提出一种基于张量奇异值分解的视觉域自适应方法（Visual domain Adaptation method based on TEnsor Singular value decomposition，VATES）．与现有视觉域自适应方法不同，VATES方法从张量数据的内部结构出发，在保持原有张量特征结构不变的条件下，利用张量奇异值分解和张量乘运算，分别构造源域和目标域的共同张量子空间以及源域特征和目标域特征在该共同张量子空间上的投影，从而实现将源域学习到的知识向目标域进行高效迁移．

本文的主要贡献总结如下：

（1）首次将张量奇异值分解技术应用于视觉域自适应建模．通过构造源域和目标域的共同张量子空间并计算相应的特征投影，从而保持原有高阶张量特征结构的不变性．

（2）利用交替方向乘子法，设计了基于张量奇异值分解的视觉域自适应方法．通过将原始高阶多变量优化问题分解为多个高阶单变量子优化问题，从而实现了复杂优化问题的迭代求解．

（3）给出了在正交张量子空间约束条件下源域与目标域表征误差最小化问题的可解性的理论推导，并求得了相应的解析解，从而简化了高阶单变量子优化问题的求解过程．

本文第2节概述视觉域自适应领域的相关工作；第3节介绍张量基础知识；第4节描述基于张量奇异值分解的视觉域自适应问题定义及其求解方法；第5节通过对比实验分析所提方法的分类精度和运算效率；最后对本文的工作进行总结，并展望下一步的研究工作．

2 相关工作

视觉域自适应技术根据学习方法的差异可以分为基于样本的视觉域自适应方法、基于特征的视觉域自适应方法、基于模型的视觉域自适应方法和基于关系的视觉域自适应方法．其中，基于特征的视觉域自适应方法是最受关注的研究方向，其利用特征变换将源域特征与目标域特征建立关联，从而实现源域中已有知识向目标域的迁移[11]．

在基于特征的视觉域自适应方法中，通过特征变换将源域特征映射至目标域特征是最直接的一类视觉域适应方法．其中，迁移成分分析方法（Transfer Component Analysis，TCA）采用最大均值差异（Maximum Mean Difference，MMD）度量准则，最小化源域与目标域数据特征的边缘分布差异[15]．联合分布适配方法（Joint Distribution Adaptation，JDA）则同时考虑数据特征的边缘分布和条件分布，实现数据特征的联合对齐[16]．迁移联合适配方法（Transfer Joint Matching，TJM）在基于特征的域自适应方法基础上，进一步引入了基于样本的域自适应方法[17]．跨领域典型相关性分析方法（Canonical Correlation Analysis across Different Domains，CCADD）通过选择最优的特征组合，从而确保相关特征在领域间具有相似的判别性质[18]．平衡分布适配方法（Balanced Distribution Adaptation，BDA）则通过动态调节边缘分布和条件分布的权重，实现了更优的迁移效果[19]．为了更加充分地利用源域与目标域之间的相关信息，联合Wasserstein自编码器方法（Joint Wasserstein Auto Encoders，JWAE）通过利用最优传输理论构造并求解了联合特征分布之间Wasserstein距离的最小化问题[20]．

相较于直接的特征变换，子空间学习法假设源域特征和目标域特征在某个公共子空间中具有相似的分布，通过构造公共子空间并计算相应投影，从而实现源域知识向目标域的迁移[21-22]．子空间对齐方法（Subspace Adaptation，SA）通过求解一个线性变换矩阵，以实现源域特征空间和目标域特征空间的对齐[23]．子空间分布对齐方法（Subspace Distribution Alignment，SDA）通过引入概率分布自适应变换，改进了子空间对齐方法的迁移效果[24]．关联对齐方法（CORrelation ALignment，CORAL）则进一步引入一个二阶对齐矩阵，从而缩小源域特征和目标域特征之间的距离[25]．几何和统计特征联合对齐方法（Joint Geometrical and Statistical Alignment，JGSA）在子空间投影过程中同时缩小几何偏移和统计偏移，从而提升了迁移效果[26]．张量不变子空间学习方法（Tensor-Aligned Invariant Subspace Learning，TAISL）是首个将域自适应与张量结构相结合的方法，通过利用张量的Tucker分解得到公共子空间，较好地保留了高维特征结构信息，取得了不错的迁

移效果[27].

由于在流形空间中的特征通常具有良好的几何性质，可以有效地避免特征扭曲，目前也被广泛应用于基于特征的视觉域自适应方法中. 代表性的方法包括采样测地线流方法（Sampling Geodesic Flow，SGF）[28]、测地线流式核方法（Geodesic Flow Kernel，GFK）[29]以及统计流形方法（Statistical Manifold，SM）[30]等. 最近，流形嵌入分布对齐方法（Manifold Embedded Distribution Alignment，MEDA）基于流形特征变换，分别通过减小域间数据漂移和自适应分布适配来学习目标域中的分类器，进一步提升了迁移效果[31].

随着深度学习技术的飞速发展，在上述非深度域自适应方法之外，还有大量基于深度学习的域自适应方法被提出. 代表性的深度域自适应方法包括深度适配网络（Deep Adaptation Network，DAN）[32]、领域对抗迁移网络（Domain Adversarial training of Neural Networks，DANN）[33]、对抗判别领域适应网络（Adversarial Discriminative Domain Adaptation，ADDA）[34]、联合适配网络（Joint Adaptation Networks，JAN）[35]、生成适应网络（Generate To Adapt，GTA）[36]、合作对抗网络（Collaborative and Adversarial Network，CAN）[37]、联合领域对齐和判别特征学习网络（Joint domain alignment and Discriminative feature learning for unsupervised deep Domain Adaptation，JDDA）[38]、增强传输距离网络（Enhanced Transport Distance，ETD）[39]等. 由于采用了较为复杂的多层神经网络架构，这类方法均具有较好的模型迁移效果，但所需的训练样本较多，训练时间较长.

传统的视觉域自适应方法在处理输入数据特征时，简单地将高阶张量特征转换成一阶高维向量特征，不但破坏了原有特征数据的空间结构信息，而且会因数据维度的增加导致维度灾难[40-41]. 虽然TAISL方法利用张量Tucker分解在一定程度上保持了数据特征的空间结构信息，但是Tucker分解的空间复杂度随着张量维度呈指数增长，当张量维度过高时存储空间消耗巨大[42]. 近年来，涌现了张量奇异值分解[42]、张量链式分解[43]、张量环式分解[44]等张量分解方法. 它们主要是通过充分挖掘并利用高阶张量数据内部的结构信息，来保持原有高阶数据的丰富结构特征，并避免维度灾难. 目前，已被广泛应用于机器学习、信号处理、通信、神经科学和计量化学等应用领域[45].

本文针对现有研究方案的不足，提出了一种基于张量奇异值分解的视觉域自适应方法，通过利用张量奇异值分解技术构造源域和目标域的共同张量子空间以保持原有高阶张量特征结构的不变性. 进而，利用交替方向乘子法求解高阶多变量优化问题，并证明了正交张量子空间约束条件下源域与目标域表征误差最小化问题的可解性.

3 张量基础

张量又称为多维数组，是描述高阶数据时所使用的一种数据结构. 一般地，张量的维度数量称为张量的阶. 例如，0阶张量为标量，1阶张量为向量，2阶张量为矩阵，一般意义上的张量其阶数均大于等于3. 假设$\boldsymbol{X}\in\mathbb{R}^{I_1\times I_2\times I_3}$为一个3阶张量，则其由索引$i_1,i_2,i_3$确定的元素表示为$\boldsymbol{X}(i_1,i_2,i_3)$，其中索引需满足条件$1\leqslant i_j\leqslant I_j, j=1,\cdots,3$. 相应的，张量$\boldsymbol{X}$的第$k$个前切片表示为$\boldsymbol{X}^{(k)}=\boldsymbol{X}(:,:,k)$，即第3维取$k$时唯一确定的矩阵，其示意图如图1所示.

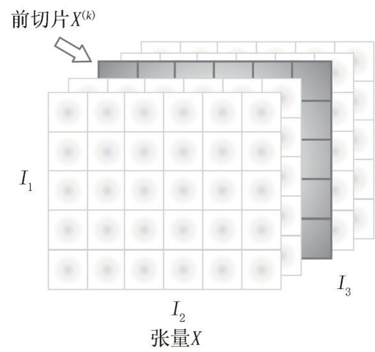

图1 张量前切片示意图

假设$\boldsymbol{Y}\in\mathbb{R}^{I_2\times I_4\times I_3}$为另一个3阶张量，则$\boldsymbol{X}$和$\boldsymbol{Y}$的张量乘运算可表示为

$$\boldsymbol{X}*\boldsymbol{Y}=fold(bcirc(\boldsymbol{X})\cdot unfold(\boldsymbol{Y})) \tag{1}$$

其中$bcirc(\ )$表示块循环操作，即：

$$bcirc(\boldsymbol{X})=\begin{bmatrix}\boldsymbol{X}^{(1)} & \boldsymbol{X}^{(I_3)} & \cdots & \boldsymbol{X}^{(2)}\\ \boldsymbol{X}^{(2)} & \boldsymbol{X}^{(1)} & \cdots & \boldsymbol{X}^{(3)}\\ \vdots & \vdots & \ddots & \vdots\\ \boldsymbol{X}^{(I_3)} & \boldsymbol{X}^{(I_3-1)} & \cdots & \boldsymbol{X}^{(1)}\end{bmatrix} \tag{2}$$

$unfold(\ )$和$fold(\ )$分别表示张量展开和折叠操作，即：

$$unfold(\boldsymbol{Y})=\begin{bmatrix}\boldsymbol{Y}^{(1)}\\ \boldsymbol{Y}^{(2)}\\ \vdots\\ \boldsymbol{Y}^{(I_3)}\end{bmatrix} \tag{3}$$

$$fold(unfold(Y))=Y \tag{4}$$

当$X*Y=I$时，称张量X和Y正交，其中I为单位张量，其第1个前切片为单位矩阵，其余前切片均为全零矩阵.

张量乘运算通过将矩阵乘法中的元素相乘替换为循环卷积，可视为矩阵乘法的高阶推广. 同时，张量乘运算也等价于傅里叶域中张量转换矩阵间的矩阵乘运算[42].

4 VATES方法

4.1 问题定义

将源域图片样本和目标域图片样本按序号排列，分别构造源域张量数据X_S和目标域张量数据X_T. 经过特征提取后，分别获得源域张量特征$\tilde{X}_S$和目标域张量特征$\tilde{X}_T$. 假设源域特征$\tilde{X}_S$和目标域特征$\tilde{X}_T$之间存在一个共同张量子空间A. 相应的，张量S和T分别表示源域特征和目标域特征在共同张量子空间A上的特征投影.

因此，源域与目标域之间的张量特征对齐问题可转化为张量子空间表征误差最小化问题，具体表示如下：

$$\min_{A,S,T}\left\|\tilde{X}_S-A*S\right\|_F^2+\left\|\tilde{X}_T-A*T\right\|_F^2 \quad \text{s.t.}\ A^T*A=I \tag{5}$$

式(5)所表示的是一个高阶多变量优化问题，其优化目标函数中的两项分别表示源域和目标域中的张量子空间表征误差，约束条件表示张量子空间A需满足正交张量约束. 通过利用张量分解，优化问题(5)避免了简单的向量化操作对高阶张量特征数据内部结构的破坏，从而保持原有高阶张量特征结构的不变性. VATES方法示意图如图2所示.

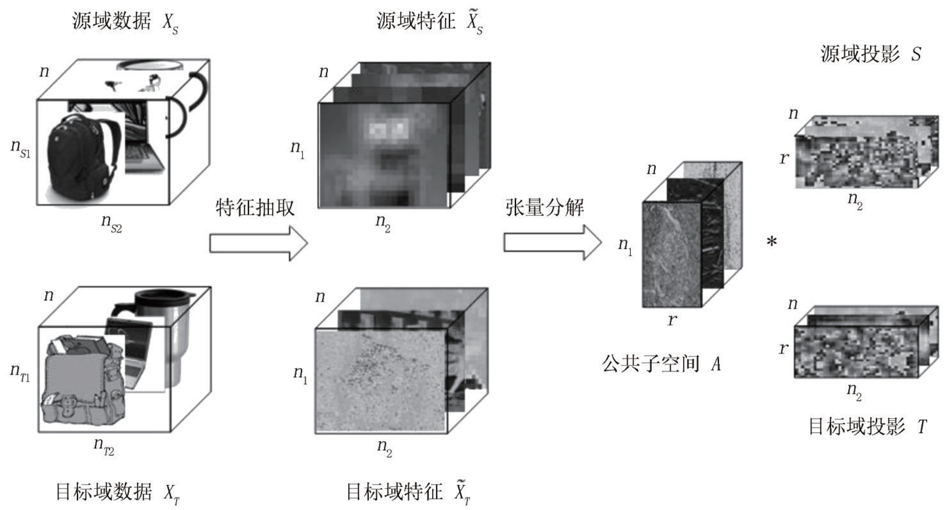

图2 VATES方法示意图

4.2 问题求解

针对正交约束的高阶多变量优化问题(5)，可利用交替方向乘子法进行求解. 即固定其余变量，循环求解公共张量子空间A、源域特征投影S和目标域特征投影T，将高阶多变量优化问题分解为多个高阶单变量子优化问题，从而对式(5)进行求解.

4.2.1 固定S和T，求解A.

此时，借助临近点思想，式(5)所示的优化问题可表示为

$$\min_{A}\left\|\tilde{X}_S-A*S\right\|_F^2+\left\|\tilde{X}_T-A*T\right\|_F^2+\rho\left\|A-A^k\right\|_F^2 \quad \text{s.t. } A^T*A=I \tag{6}$$

其中ρ表示惩罚因子，A^k表示上一次迭代获得的子空间. 为求解优化问题(6)，引入下述定理.

定理1. 优化问题(6)的最优解为$A=V*U^T$，其中U和V为张量奇异值分解得到的正交张量，即$[U,\boldsymbol{\Sigma},V]=\text{t-SVD}\left(S*\tilde{X}_S^T+T*\tilde{X}_T^T+\rho\left(A^k\right)^T\right)$.

证明. 由文献[42]可知，张量乘运算等价于傅里叶域中张量转换矩阵间的矩阵乘运算. 因此，令$\widehat{X}_S=bdiag(fft(\tilde{X}_S))$和$\widehat{X}_T=bdiag(fft(\tilde{X}_T))$分别表示源域特征张量$\tilde{X}_S$和目标域特征张量$\tilde{X}_T$在傅里叶域中的张量转换矩阵，其中$fft(\ )$表示快速傅里叶变换. 相应的，令$\widehat{A}=bdiag(fft(A))$，$\widehat{S}=bdiag(fft(S))$和$\widehat{T}=bdiag(fft(T))$分别表示子空间张量$A$、源域投影张量$S$和目标域投影张量$T$所

对应的张量转换矩阵.

此时，优化问题(6)等价于如下正交矩阵约束条件下的表征误差最小化问题：

$$\min_{\widehat{A}} \left\| \widehat{X}_S - \widehat{A}\,\widehat{S} \right\|_F^2 + \left\| \widehat{X}_T - \widehat{A}\,\widehat{T} \right\|_F^2 + \rho \left\| \widehat{A} - \widehat{A}^k \right\|_F^2$$

$$\text{s.t. } \widehat{A}^T \widehat{A} = \boldsymbol{I} \tag{7}$$

将目标函数展开并带入约束条件$\widehat{A}^T\widehat{A} = \boldsymbol{I}$，可将优化问题(7)转换为如下无约束优化问题：

$$\max_{\widehat{A}} Tr\left(\widehat{A}\left(\widehat{S}\,\widehat{X}_S^T + \widehat{T}\,\widehat{X}_T^T + \rho\left(\widehat{A}^k\right)^T\right)\right) \tag{8}$$

其中，$Tr(\ )$表示矩阵的迹运算.

由冯•诺依曼迹不等式[46]及$\widehat{A}$的正交性可知，当$\widehat{A} = \widehat{V}\,\widehat{U}^T$时优化问题(8)的目标函数取得极大值，其中$[\widehat{U}, \widehat{\boldsymbol{\Sigma}}, \widehat{V}] = \mathrm{SVD}\left(\widehat{S}\,\widehat{X}_S^T + \widehat{T}\,\widehat{X}_T^T + \rho\left(\widehat{A}^k\right)^T\right)$为矩阵奇异值分解所得到的矩阵. 相应的，将张量转换矩阵逆变换回高阶张量空间，可得$\boldsymbol{A} = \boldsymbol{V} * \boldsymbol{U}^T$，其中$[\boldsymbol{U}, \boldsymbol{\Sigma}, \boldsymbol{V}] = \text{t-SVD}\left(\boldsymbol{S} * \tilde{\boldsymbol{X}}_S^T + \boldsymbol{T} * \tilde{\boldsymbol{X}}_T^T + \rho\left(\boldsymbol{A}^k\right)^T\right)$为由张量奇异值分解(tensor SVD，t-SVD)所获得的正交张量[47]，同时满足张量转换关系$\widehat{U} = bdiag(fft(\boldsymbol{U}))$，$\widehat{\boldsymbol{\Sigma}} = bdiag(fft(\boldsymbol{\Sigma}))$以及$\widehat{V} = bdiag(fft(\boldsymbol{V}))$. 证毕.

根据定理1，式(6)所示的子优化问题的解为

$$\boldsymbol{A}^{k+1} = \boldsymbol{V} * \boldsymbol{U}^T \tag{9}$$

4.2.2 固定$\boldsymbol{A}$和$\boldsymbol{T}$，求解$\boldsymbol{S}$.

此时，依据临近点思想，原始优化问题(5)可表示为如下子优化问题：

$$\begin{aligned} &\min_{\boldsymbol{S}} \left\| \tilde{\boldsymbol{X}}_S - \boldsymbol{A} * \boldsymbol{S} \right\|_F^2 + \rho \left\| \boldsymbol{S} - \boldsymbol{S}^k \right\|_F^2 \\ &= \min_{\boldsymbol{S}} \left\| \frac{\boldsymbol{A}^T * \tilde{\boldsymbol{X}}_S + \rho \boldsymbol{S}^k}{1+\rho} - \boldsymbol{S} \right\|_F^2 \end{aligned} \tag{10}$$

其中，$\boldsymbol{S}^k$表示上一次迭代时获得的源域投影张量. 显然，子优化问题(10)的解为

$$\boldsymbol{S}^{k+1} = \frac{\boldsymbol{A}^T * \tilde{\boldsymbol{X}}_S + \rho \boldsymbol{S}^k}{1+\rho} \tag{11}$$

4.2.3 固定$\boldsymbol{A}$和$\boldsymbol{S}$，求解$\boldsymbol{T}$.

求解目标域特征投影$\boldsymbol{T}$与求解源域特征投影$\boldsymbol{S}$类似，依据临近点思想将优化问题(5)表示为如下子优化问题：

$$\begin{aligned} &\min_{\boldsymbol{T}} \left\| \tilde{\boldsymbol{X}}_T - \boldsymbol{A} * \boldsymbol{T} \right\|_F^2 + \rho \left\| \boldsymbol{T} - \boldsymbol{T}^k \right\|_F^2 \\ &= \min_{\boldsymbol{T}} \left\| \frac{\boldsymbol{A}^T * \tilde{\boldsymbol{X}}_T + \rho \boldsymbol{T}^k}{1+\rho} - \boldsymbol{T} \right\|_F^2 \end{aligned} \tag{12}$$

进而可得：

$$\boldsymbol{T}^{k+1} = \frac{\boldsymbol{A}^T * \tilde{\boldsymbol{X}}_T + \rho \boldsymbol{T}^k}{1+\rho} \tag{13}$$

其中，$\boldsymbol{T}^k$表示上一次迭代获得的目标域特征投影.

通过迭代求解公共子空间$\boldsymbol{A}$、源域投影$\boldsymbol{S}$和目标域投影$\boldsymbol{T}$，直至最大迭代次数K. 基于张量奇异值分解的视觉域自适应方法的伪代码如算法1所示.

算法1. 基于张量奇异值分解的视觉域自适应方法.

输入：源域张量$\boldsymbol{X}_S$，目标域张量$\boldsymbol{X}_T$，迭代阈值K，惩罚因子ρ.

输出：公共子空间$\boldsymbol{A}$，源域投影$\boldsymbol{S}$，目标域投影$\boldsymbol{T}$.

1. $\tilde{\boldsymbol{X}}_S = fe(\boldsymbol{X}_S)$，$\tilde{\boldsymbol{X}}_T = fe(\boldsymbol{X}_T)$；
2. $k=0$；
3. WHILE $k<K$：
4. $[\boldsymbol{U}, \boldsymbol{\Sigma}, \boldsymbol{V}] = \text{t-SVD}\left(\boldsymbol{S} * \tilde{\boldsymbol{X}}_S^T + \boldsymbol{T} * \tilde{\boldsymbol{X}}_T^T + \rho\left(\boldsymbol{A}^k\right)^T\right)$；
5. 更新$\boldsymbol{A}^{k+1} = \boldsymbol{V} * \boldsymbol{U}^T$；
6. 更新$\boldsymbol{S}^{k+1} = \left(\boldsymbol{A}^T * \tilde{\boldsymbol{X}}_S + \rho \boldsymbol{S}^k\right) / (1+\rho)$；
7. 更新$\boldsymbol{T}^{k+1} = \left(\boldsymbol{A}^T * \tilde{\boldsymbol{X}}_T + \rho \boldsymbol{T}^k\right) / (1+\rho)$；
8. $k=k+1$
9. END WHILE

在算法1中，$fe(\ )$表示特征抽取函数，既可以采用传统的基于特征工程的特征抽取器，如SURF(Speeded Up Robust Features)特征[48]，也可以采用基于深度学习训练得到的特征抽取器，如DeCAF6特征[49]或ResNet特征[50]等.

在算法1的迭代求解过程中，计算量最大的两类操作分别是张量奇异值分解和张量乘运算. 由文献[42]可知，张量奇异值分解的计算复杂度为$\mathrm{O}(\lceil n/2 \rceil min(n_1^2 n_2, n_1 n_2^2))$，张量乘运算的计算复杂度为$\mathrm{O}(n_1^2 n_2^2 n)$，其中$n$为样本数量，$n_1 \times n_2$为特征维度，如图2所示. 因此，算法1的计算复杂度为$\mathrm{O}(n_1^2 n_2^2 n K)$.

5 实验分析

本文实验运行环境为MATLAB R2020b，运行平台为一台配置为AMD R7-5800U CPU、16 GB内存的个人计算机.

5.1 数据集

本文对比实验中使用的数据集均为视觉域自适应领域中常见的公共数据集，包括：

（1）Office-Caltech-10数据集[31]. 该数据集包含在不同角度、背景、分辨率条件下拍摄或手工绘制的图像. 数据集中共有2533个样本，10个类，包含Amazon、Caltech、DSLR、Webcam四个域，可构成12组域自适应任务.

（2）Office31数据集[32]. 该数据集包含在不同角度、背景、分辨率条件下拍摄的日常办公物品图像. 数据集中共有4652个样本，31个类，包含Amazon、DSLR、Webcam三个域，可构成6组域自适应任务.

（3）ImageNet-VOC2007数据集[31]. 该数据集分别从ImageNet和PASCAL VOC2007数据集中抽取同一类图像，用于测试简单背景和复杂背景间的视觉域自适应任务. 数据集中共有10717个样本，5个类，包含ImageNet和VOC2007两个域，可构成2组域自适应任务.

实验数据集及相应域中的示例图像如表1所示. 为了方便描述，下文采用域名首字母代替相应的域名. 实验评价指标采用无标签目标域上的图像分类精度. 分类精度值越高，表明域自适应算法的迁移效果越好.

表1 数据集和相应的域

数据集	域			
Office-Caltech-10				
	(A)mazon	(C)altech	(D)SLR	(W)ebcam
Office31				
	(A)mazon	(D)SLR	(W)ebcam	
ImageNet-VOC2007				
	(I)mageNet	(V)OC2007		

5.2 实验设置

本文实验将VATES方法与以下四类17种基线方法进行对比：

(1) 经典机器学习方法

① SVM[51]：利用核技巧在特征空间中寻找间隔最大化的非线性分类器.

② PCA[52]：利用线性变换提取数据主要特征分量以实现低维空间上的分类.

(2) 非深度域自适应方法

① TCA[15]：采用最大均值差异度量针对不同领域数据特征的边缘分布差异实现最小化.

② GFK[29]：利用核技巧在源域与目标域之间构造无穷多个子空间实现知识迁移.

③ JDA[16]：同时考虑数据特征的边缘分布和条件分布实现数据特征的联合对齐.

④ TJM[17]：联合基于特征和基于样本的域自适应方法实现跨域知识迁移.

⑤ CORAL[25]：同时对齐源域和目标域的一阶和二阶统计特征实现跨域知识迁移.

⑥ JGSA[26]：在子空间投影过程中同时缩小几何偏移和统计偏移以提升知识迁移效果.

⑦ MDTL[53]：在迁移时利用流形子空间学习同时缩小类内分布差异并扩大类间分布差异.

(3) 深度域自适应方法

① DAN[32]：基于深度神经网络在再生核希尔伯特空间中利用多核最大均值差异和多层特征实现迁移.

② DANN[33]：基于深度神经网络采用单核最大均值差异和分类误差进行特征提取和知识迁移.

③ DCORAL[54]：基于深度神经网络学习非线性变换以对齐激活层的统计分布特征.

④ ADDA[34]：基于深度神经网络对抗训练分类器损失和判别器损失以实现跨域知识迁移.

⑤ ProCA[55]：通过构造目标域内类别检测策略和域间原型导引对齐方法，实现跨越知识迁移.

⑥ CKB[56]：在跨域迁移时利用条件核Bures度量刻画不同域间的条件分布差异，同时构建神经网络对其进行估计.

(4) 张量域自适应方法

① NTSL[27]：利用张量Tucker分解构造公共子空间和源域、目标域投影以实现跨域知识迁移.

② TAISL[27]：在NTSL方法基础上同时利用张量Tucker分解和矩阵对齐方法构造公共子空间和投影以实现跨域知识迁移.

5.3 迁移效果分析

(1) Office-Caltech-10数据集

不同算法在Office-Caltech-10数据集上的分类精度如表2所示. 从表2中可以看出，本文提出的VATES方法在众多基线对比方法中取得了最优的分类效果. 其中，传统机器学习算法SVM和PCA

表2　Office-Caltech-10数据集上分类精度对比

算法	C-A	C-W	C-D	A-C	A-W	A-D	W-C	W-A	W-D	D-C	D-A	D-W	均值
SVM	91.6	80.7	86.0	82.2	71.9	80.9	67.9	73.4	100	72.8	78.7	98.3	82.0
PCA	88.1	83.4	84.1	79.3	70.9	82.2	70.3	73.5	99.0	71.7	79.2	98.0	81.7
TCA	89.8	78.3	85.4	82.6	74.2	81.5	80.4	84.1	100	82.3	89.1	99.7	85.6
GFK	88.2	77.6	86.6	79.2	70.9	82.2	69.8	76.8	100	71.4	76.3	99.3	81.5
JDA	89.6	85.1	89.8	83.6	78.3	80.3	84.8	90.3	100	85.5	91.7	99.7	88.2
TJM	88.8	81.4	84.7	84.3	71.9	76.4	83.0	87.6	100	83.8	90.3	99.3	86.0
CORAL	92.0	80.0	84.7	83.2	74.6	84.1	75.5	81.2	100	76.8	85.5	99.3	84.7
JGSA	91.4	86.8	93.6	84.9	81.0	88.5	85.0	90.7	100	86.2	92.0	99.7	90.0
MDTL	93.4	92.2	88.5	88.8	90.5	90.4	88.1	92.3	100	87.2	92.4	98.9	91.9
DAN	93.1	91.5	91.2	88.6	91.5	90.5	88.6	92.2	99.9	87.8	90.1	98.6	91.9
DANN	93.2	89.4	91.1	87.8	77.8	82.4	81.3	82.9	99.5	82.1	84.7	98.5	87.6
DCORAL	89.8	97.3	91.0	91.9	100	90.5	83.7	81.5	90.1	88.6	80.1	92.3	89.7
ADDA	93.6	90.5	90.1	87.9	86.2	89.7	86.6	88.9	98.4	86.1	89.5	96.2	90.3
ProCA	92.9	91.1	90.3	88.2	87.2	90.7	85.9	89.3	98.1	87.2	90.2	97.7	90.7
CKB	93.3	92.9	91.7	87.5	89.8	93.0	85.8	92.8	100	83.4	92.3	99.7	91.9
NTSL	89.6	80.4	87.7	78.5	77.3	83.1	80.0	85.8	97.8	79.8	87.8	95.4	85.3
TAISL	90.0	85.3	90.6	80.1	77.9	85.1	82.6	85.6	97.7	84.0	87.6	95.9	86.9
VATES	96.3	93.1	90.2	87.6	90.9	92.4	85.2	96.5	98.7	85.4	97.7	97.6	**92.6**

因缺少有效的迁移机制，取到了最低的分类精度．以TCA为代表的非深度域自适应方法因采用特征分布对齐等技术在不同程度上提升了无标签目标域上的分类精度．而深度域自适应方法因采用了更加强大的深度学习技术，在目标域上的分类精度获得了大幅度的提升．本文提出的VATES方法借助特征张量分解技术，在保持原有特征空间结构不变的同时，充分挖掘源域与目标域之间的共性，从而获得了最高的目标域分类精度．与传统机器学习算法相比，分类精度提升了10%以上，与同类型的张量域自适应方法NTSL和TAISL相比，分类精度也分别提高了7.3%和5.7%.

(2) Office31数据集

本文方法与其它17种基线对比方法在Office31数据集上的分类精度如表3所示．从表3中可以看出，由于类别数量的增加，所有方法的分类精度与Office-Caltech-10数据集上的分类结果相比都有一定程度的降低．显然，分类类别越多，模型迁移难度越大．但是，本文提出的VATES方法仍然具有显著的分类精度优势，大幅度领先于传统的机器学习方法．与JGSA等非深度域自适应方法和ADDA等深度域自适应方法相比，VATES方法虽然未采用复杂的统计特征计算和深度神经网络，却仍然具有较高的目标域分类精度．与同为张量域自适应方法的NTSL和TAISL相比，VATES方法借助具有更强高维数据结构表征能力的张量奇异值分解，分类精度提升了将近25%.

表3　Office31数据集上分类精度对比

算法	A-W	D-W	W-D	A-D	D-A	W-A	均值
SVM	47.4	77.4	81.3	53.8	39.3	36.3	56.0
PCA	36.5	68.7	70.5	40.5	38.2	37.8	48.7
TCA	59.0	90.2	88.2	57.8	51.6	47.9	65.8
GFK	58.4	93.6	91.0	58.6	52.4	46.1	66.7
JDA	73.6	96.5	98.6	80.7	64.7	63.1	79.5
TJM	79.1	94.6	96.6	81.1	71.2	72.9	82.6
CORAL	79.3	94.3	99.4	74.8	56.4	63.4	78.0
JGSA	81.1	97.2	99.0	84.3	75.8	76.5	85.7
MDTL	82.1	97.3	99.2	84.4	74.8	77.3	85.8
DAN	80.5	97.1	99.6	78.6	63.6	62.8	80.4
DANN	82.0	96.9	99.1	79.7	68.2	67.4	82.2
DCORAL	66.4	95.7	99.2	66.8	52.8	51.5	72.1
ADDA	86.2	96.2	98.4	77.8	69.5	68.9	82.9
ProCA	82.5	99.1	99.6	81.8	65.2	64.1	82.1
CKB	86.5	97.1	98.2	82.3	70.6	70.5	84.2
NTSL	50.8	84.4	88.2	56.1	45.7	42.6	61.3
TAISL	50.7	84.5	88.5	56.4	45.9	43.2	61.5
VATES	85.8	97.2	98.4	83.5	71.6	79.7	**86.1**

(3) ImageNet-VOC2007数据集

所有域自适应方法在ImageNet-VOC2007数据集上的分类精度如表4所示．由于该数据集中的图像尺寸有所增加，图像背景信息也由简单的纯色背景或单调背景换成了较为复杂的多元素背景，因此

分类任务难度也有所增加.相应的,从表4中的分类精度结果可以发现,各种算法的目标域分类精度均低于前两个数据集中获得的计算结果.

从表4中的实验结果可以看出,本文提出的VATES方法仍然具有最高的目标域分类精度,不但远高于经典的机器学习方法,也高于实验中对比的其它基线非深度域自适应方法和深度域自适应方法.VATES方法具有高分类精度的潜在原因在于其通过张量子空间表征下的域特征投影能够有效地保留源域和目标域内的数据特征,进而可获得较优的迁移效果.

表4 VOC2007-ImageNet数据集上分类精度对比

算法	I-V	V-I	均值
SVM	52.4	42.7	47.6
PCA	58.4	65.1	61.8
TCA	63.7	64.9	64.3
GFK	59.5	73.8	66.7
JDA	63.4	70.2	66.8
TJM	63.7	73.0	68.4
CORAL	59.6	70.3	65.0
JGSA	52.3	70.6	61.5
MDTL	66.7	71.2	68.9
DAN	65.9	69.7	67.8
DANN	66.5	71.3	68.9
DCORAL	67.3	72.6	69.9
ADDA	67.1	71.3	69.2
ProCA	66.9	72.1	69.5
CKB	67.3	71.5	69.4
NTSL	55.9	59.9	57.7
TAISL	56.1	60.1	58.1
VATES	67.5	72.7	**70.1**

5.4 迁移效率分析

为了比较各个算法的运行效率,表5分别展示了JGSA、MDTL、TAISL和VATES四种方法在Office-Caltech-10、Office31和ImageNet-VOC2007三个数据集上的运行时间.其中,Office-Caltech-10数据集采用了SURF特征,特征维度为800;Office31数据集采用ResNet50网络来抽取特征,其维度为4096;ImageNet-VOC2007数据集采用了DeCaf6模型来抽取特征,其维度为4096.

从表5中可以看出,本文方法具有最短的运行时间.与非深度域自适应方法JGSA和MDTL相比,其运行时间开销分别仅占38.4%~59.5%和25.7%~39.7%.与同为张量域自适应方法的TAISL相比,其运行时间开销仅占7.2%.VATES方法的高效性主要得益于简单的模型结构和高效的张量奇异值分解实现.因此,VATES方法在保持较高的分类精度条件下具有最高的运行效率.

表5 运行时间对比(秒)

数据集	JGSA	MDTL	TAISL	VATES
Office-Caltech-10	4.2	6.3	34.7	**2.5**
Office31	16.2	26.1	N/A	**8.6**
ImageNet-VOC2007	409.6	611.9	N/A	**157.3**

5.5 模型参数分析

在VATES方法中,参数ρ表示临近点惩罚因子,其对算法精度的影响如图3所示.图3中显示的是Office-Caltech-10数据集中的C-A迁移任务,其它数据集中的迁移任务与之相似.从图3中可以看出,VATES方法的分类精度随着ρ值增加呈现先上升后下降的趋势,并最终趋于平稳.当ρ值为1时,VATES方法的分类精度能获得最优值.

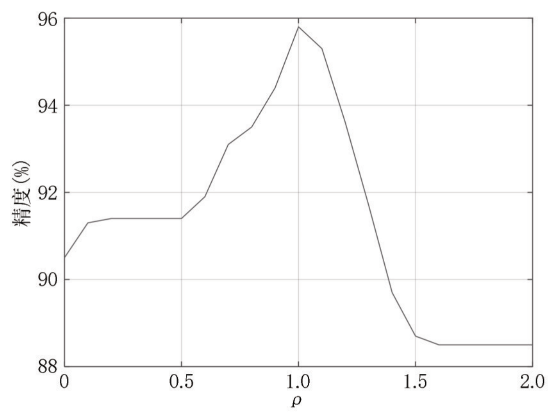

图3 参数ρ对算法精度的影响

图4展示了利用不同的卷积神经网络所抽取的特征对VATES方法迁移效果的影响,其中横轴代表Office-Caltech-10数据集上的不同域自适应任务,纵轴代表目标域上的图像分类精度.实验中分别使用了AlexNet、VGGNet16、VGGNet19、ResNet50四种经典的卷积神经网络,其层数及特征抽取数量如表6所示.

从图4中可以看出,随着网络特征抽取能力的增强,本文方法的目标域分类精度逐渐增加.因此,VATES方法的域自适应性能会随着特征抽取能力的强化而逐渐提升.同时,利用不同网络所抽取的特征在不同实验组别上保持相同的变化趋势.因此,VATES方法中的特征抽取对域自适应效果的影响具有稳定性.

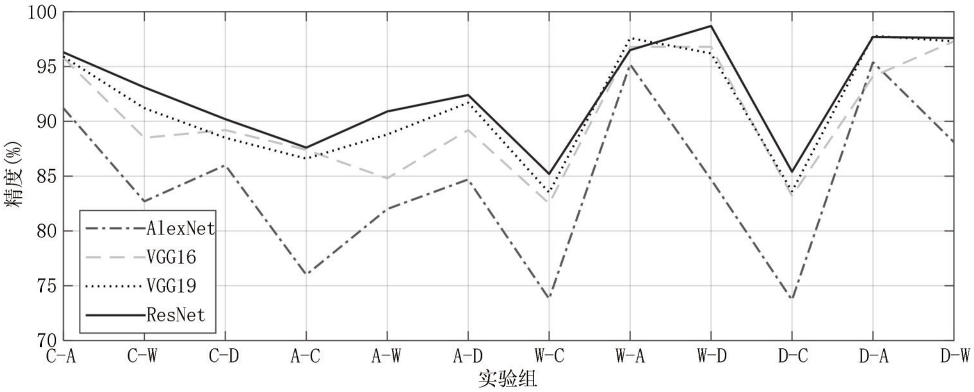

图4　不同的特征抽取方法对VATES精度的影响

表6　卷积神经网络的特征抽取参数

卷积神经网络	层数	特征数
AlexNet	8	4096
VGGNet16	16	4096
VGGNet19	19	4096
ResNet50	50	2048

6　结　论

本文针对现有视觉域自适应方法中缺乏有效处理高维特征数据的问题，提出了一种基于张量奇异值分解的视觉域自适应方法VATES. 该方法在保持原有张量特征结构不变的条件下，利用张量奇异值分解构造了源域和目标域共享的张量子空间，从而实现了源域模型向目标域的迁移. 基于多组公开数据集上的对比实验结果表明，VATES方法在保持较优迁移效果的同时，具有最高的运行效率. 未来将进一步研究更为复杂的张量链式分解结构下的视觉域自适应方法. 此外，也将针对开放集场景下的视觉域自适应方法开展研究.

致　谢　衷心感谢审稿专家对本文工作的指导！

参　考　文　献

［1］ Rempe D，Birdal T，Hertzmann A，et al. HuMoR：3D human motion for robust pose estimation//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal，Canada，2021：11488-11499

［2］ CHEN Jia，CHEN Ya-Song，LI Wei-Hao，et al. Application and prospect of deep learning in video object segmentation. Chinese Journal of Computers，2021，44(3)：609-631 (in Chinese)
(陈加，陈亚松，李伟浩等. 深度学习在视频对象分割中的应用与展望. 计算机学报，2021，44(3)：609-631)

［3］ Yang L，Shen Y，Mao Y，et al. Hybrid curriculum learning for emotion recognition in conversation//Proceedings of the 36th AAAI Conference on Artificial Intelligence. 2022：12341-12350

［4］ Qi T，Wu F，Wu C，et al. Personalized news recommendation with knowledge-aware interactive matching//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2021：61-70

［5］ Zhu Z，Zhao H. A survey of deep RL and IL for autonomous driving policy learning. IEEE Transactions on Intelligent Transportation Systems，2022，23(9)：14043-14065

［6］ Niu S，Liu M，Liu Y，et al. Distant domain transfer learning for medical imaging，IEEE Journal of Biomedical and Health Informatics，2021，25(10)：3784-3793

［7］ SUN Qi-Yu，ZHAO Chao-Qiang，TANG Yang，et al. A survey on unsupervised domain adaptation in computer vision tasks. Scientia Sinica Technologica，2022，52(1)：26-54 (in Chinese)
(张琦钰，赵超强，唐漾等. 基于无监督域自适应的计算机视觉任务研究进展. 中国科学：技术科学，2022，52(1)：26-54)

［8］ Ye Y，Lin Q，Ma L，et al. Multiple source transfer learning for dynamic multiobjective optimization. Information Sciences，2022，607：739-757

［9］ Zhuang F，Qi Z，Duan K，et al. A comprehensive survey on transfer learning. Proceedings of the IEEE，2021，109(1)：43-76

［10］ Niu S，Liu Y，Wang J，et al. A decade survey of transfer learning (2010—2020). IEEE Transactions on Artificial Intelligence，2020，1(2)：151-166

［11］ Wang M，Deng W. Deep visual domain adaptation：A survey. Neurocomputing，2018，312：135-153

［12］ Wu H，Xu K，Song L，et al. Domain-adaptive pretraining methods for dialogue understanding//Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics. Bangkok，Thailand，2021：665-669

［13］ Zhang W，Li X，Ma H，et al. Universal domain adaptation in fault diagnostics with hybrid weighted deep adversarial learning. IEEE Transactions on Industrial Informatics，2021，17(12)：7957-7967

［14］ Sidiropoulos N，Lathauwer L，Fu X，et al. Tensor decomposition for signal processing and machine learning. IEEE Transactions on Signal Processing, 2017, 65(13): 3551-3582

[15] Pan S, Tsang I, Kwok J, et al. Domain adaptation via transfer component analysis. IEEE Transactions on Neural Networks, 2010, 22(2): 199-210

[16] Long M, Wang J, Ding G, et al. Transfer feature learning with joint distribution adaptation//Proceedings of the 2013 IEEE International Conference on Computer Vision. Sydney, Australia, 2013: 2200-2207

[17] Long M, Wang J, Ding G, et al. Transfer joint matching for unsupervised domain adaptation//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA, 2014: 1410-1417

[18] ZHANG Bo, SHI Zhong-Zhi, ZHAO Xiao-Fei, et al. Canonical correlation analysis across different domains. Chinese Journal of Computers, 2015, 38(7): 1326-1336 (in Chinese)
(张博，史忠植，赵晓非等. 一种基于跨领域典型相关性分析的迁移学习方法. 计算机学报，2015, 38(7): 1326-1336)

[19] Wang J, Chen Y, Hao S, et al. Balanced distribution adaptation for transfer learning//Proceedings of the 2017 IEEE International Conference on Data Mining. New Orleans, USA, 2017: 1129-1134

[20] CAO Jie-Zhang, MO Lang-Yuan, DU Qing, et al. Joint distribution matching method and applications based on optimal transport theory. Chinese Journal of Computers, 2021, 44(6): 1233-1245 (in Chinese)
(曹杰彰，莫朗元，杜卿等. 基于最优传输理论的联合分布匹配方法及应用. 计算机学报，2021, 44(6): 1233-1245)

[21] Elhadji-Ille-Gado N, Grall-Maes E, Kharouf M. Transfer learning for large scale data using subspace alignment//Proceedings of the 16th IEEE International Conference on Machine Learning and Applications. Cancun, Mexico, 2017: 1006-1010

[22] LI Xin, LI Zhe-Min, Wei Ju-Hui, et al. Cross-domain adaptive learning model based on feature separation. Journal of Computer Research and Development, 2022, 59 (1) : 105-117 (in Chinese)
(李鑫，李哲民，魏居辉等. 基于特征分离的跨域自适应学习模型. 计算机研究与发展，2022, 59(1): 105-117)

[23] Fernando B, Habrard A, Sebban M, et al. Unsupervised visual domain adaptation using subspace alignment//Proceedings of the 2013 IEEE International Conference on Computer Vision. Sydney, Australia, 2013: 2960-2967

[24] Sun B, Saenko K. Subspace distribution alignment for unsupervised domain adaptation//Proceedings of the 26th British Machine Vision Conference. Swansea, UK, 2015: 1-10

[25] Sun B, Feng J, Saenko K. Return of frustratingly easy domain adaptation//Proceedings of the 30th AAAI Conference on Artificial Intelligence. Palo Alto, USA, 2016: 2058-2065

[26] Zhang J, Li W, Ogunbona P. Joint geometrical and statistical alignment for visual domain adaptation//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Phoenix, USA, 2017: 1859-1867

[27] Lu H, Zhang L, Cao Z, et al. When unsupervised domain adaptation meets tensor representations//Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice, Italy, 2017: 599-608

[28] Gopalan R, Li R, Chellappa R. Domain adaptation for object recognition: an unsupervised approach//Proceedings of the 2011 IEEE International Conference on Computer Vision. Barcelona, Spain, 2011: 999-1006

[29] Gong B, Shi Y, Sha F, et al. Geodesic flow kernel for unsupervised domain adaptation//Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, USA, 2012: 2066-2073

[30] Baktashmotlagh M, Harandi M, Lovell B, et al. Domain adaptation on the statistical manifold//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA, 2014: 2481-2488

[31] Wang J, Feng W, Chen Y, et al. Visual domain adaptation with manifold embedded distribution alignment//Proceedings of the 26th ACM International Conference on Multimedia. Seoul, Korea, 2018: 402-410

[32] Long M, Cao Y, Wang J, et al. Learning transferable features with deep adaptation networks//Proceedings of the 32nd International Conference on Machine Learning. Lille, France, 2015: 97-105

[33] Ganin Y, Ustinova E, Ajakan H, et al. Domain-adversarial training of neural networks. Journal of Machine Learning Research, 2011, 17: 1-35

[34] Tzeng E, Hoffman J, Saenko, et al. Adversarial discriminative domain adaptation//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017: 7167-7176

[35] Long M, Zhu H, Wang J, et al. Deep transfer learning with joint adaptation networks//Proceeding of the 34th International Conference on Machine Learning, Sydney, Australia, 2017: 2208-2217

[36] Sankaranarayanan S, Balaji Y, Castillo C D, et al. Generate to adapt: aligning domains using generative adversarial networks//Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 8503-8512

[37] Zhang W, Ouyang W, Li W, et al. Collaborative and adversarial network for unsupervised domain adaptation//Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 3801-3809

[38] Chen C, Chen Z, Jiang B, et al. Joint domain alignment and discriminative feature learning for unsupervised deep domain adaptation//Proceedings of the 33rd AAAI Conference on Artificial Intelligence. Honolulu, USA, 2019, 3296-3303

[39] Li M, Zhai Y, Luo Y, et al. Enhanced transport distance for unsupervised domain adaptation//Proceedings of the 2020 IEEE Conference on Computer Vision and Pattern Recognition. Seattle, USA, 2020: 13936-13944

[40] Ji Y, Wang Q, Li X, et al. A survey on tensor techniques and applications in machine learning. IEEE Access, 2019, 7: 162950-162990

[41] Cichocki A, Lee N, Oseledets I, et al. Tensor networks for dimensionality reduction and large-scale optimization: part 1 low-rank tensor decompositions. Foundations and Trends in Machine Learning, 2016, 9(4): 249-429

[42] Lu C, Feng J, Chen Y, et al. Tensor robust principal component analysis with a new tensor nuclear norm. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(4): 925-938

[43] Ko C, Batselier K, Daniel L, et al. Fast and accurate tensor completion with total variation regularized tensor trains. IEEE Transactions on Image Processing, 2020, 29: 6918-6931

[44] Liu J, Zhu C, Liu Y. Smooth compact tensor ring regression. IEEE Transactions on Knowledge and Data Engineering, 2022, 34(9): 4439-4452.

[45] Bi Y, Lu Y, Long Z, et al. Tensors for data processing: theory, methods, and applications. Salt Lake City: Academic Press, 2022

[46] Albert M, Ingram O, Barry A. Inequalities: theory of majorization and its applications. Second edition. New York: Springer, 2011

[47] Wang A, Jin Z, Tang G. Robust tensor decomposition via t-SVD: near-optimal statistical guarantee and scalable algorithms. Signal Processing, 2020, 167: 107319

[48] Bay H, Ess A, Tuytelaars T, et al. Speeded-up robust features (SURF). Computer Vision and Image Understanding, 2008, 110(3): 346-359

[49] Donahue J, Jia Y, Vinyals O, et al. DeCAF: a deep convolutional activation feature for generic visual recognition// Proceedings of the 31st International Conference on Machine Learning, Beijing, China, 2014: 647-655

[50] Zheng J, Wu W, Zhao Y, et al. Transresnet: transferable resnet for domain adaptation//Proceedings of the 28th IEEE International Conference on Image Processing, Anchorage, USA, 2021, 764-768

[51] Chandra M, Bedi S. Survey on SVM and their application in image classification. International Journal of Information Technology, 2021, 13: 1-11

[52] Candès E, Li X, Ma Y, et al. Robust principal component analysis? Journal of the ACM, 2011, 58(3): 1-37

[53] Quan X, Wu D, Zhu M, et al. Manifold discriminative transfer learning for unsupervised domain adaptation//Proceedings of the 28th International Conference on Neural Information Processing. Bali, Indonesia, 2021: 321-334

[54] Sun B, Saenko, K. Deep CORAL: correlation alignment for deep domain adaptation//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, Netherlands, 2016: 443-450

[55] Lin H, Zhang Y, Qiu Z, et al. Prototype-guided continual adaptation for class-incremental unsupervised domain adaptation// Proceedings of the 17th European Conference on Computer Vision. Tel Aviv, Israel, 2022: 351-368

[56] Luo Y, Ren C. Conditional bures metric for domain adaptation// Proceedings of the 2021 IEEE Conference on Computer Vision and Pattern Recognition. 2021: 13989-13998

LI Guo-Rui, Ph. D., associate professor. His research interests include machine learning and optimization theory.

XU Peng-Fei, M. S. candidate. His research interests include transfer learning and computer vision.

PENG San-Cheng, Ph. D., professor. His research interests include domain adaptation and emotion computing.

YANG Ai-Min, Ph. D., professor. His research interests include machine learning, industrial software, and sentiment analysis.

Background

As an important and widely-concerned research topic in machine learning, domain adaptation can transfer the trained model from the source domain to the target domain by extracting and utilizing the common features shared between two domains. Thus, the expensive data collection and time-consuming model training on the target domain are unnecessary and can be avoided thoroughly. Nowadays, domain adaptation has been widely used in the fields of computer vision, natural language processing, audio processing, and industrial fault diagnostics, and so on.

However, the existing visual domain adaptation methods cannot deal with high-order tensorial features directly. They can only convert high-order feature tensors into one-order high-dimensional feature vectors by simple vectorization operations. Such naive operations not only destruct the internal structure information within the original high-order tensors but also increase the computational complexity of the domain adaptation methods.

Aiming at preserving the valuable high-order structural information and increasing the computational efficiency simultaneously, we first formulated the visual domain adaptation problem as a high-order multiple variables optimization problem. It constructs a common tensor subspace shared between source and target domain, and computes the projections of source and target features simultaneously. Secondly, we proposed the visual domain adaptation method based on tensor singular value decomposition (VATES) to solve the newly formulated optimization problem. It decomposes the original high-order optimization problem into several simple univariate optimization sub-problems by following the framework of alternating direction method of multipliers. Finally, we proved the solvability and deduced the analytical solution of the optimization sub-problem which minimizes the representation errors of source and target domains under the constraint of orthogonal common tensor subspace.

To demonstrate the effectiveness and efficiency of the proposed VATES method, we carried out extensive experiments on three popular visual domain adaptation datasets like Office-Caltech-10 dataset, Office31 dataset, and ImageNet-VOC2007 dataset, and compared it with 17 baseline methods. The experiment results showed that the proposed VATES method can not only provide higher classification accuracy for image classification tasks on unlabeled target domain than other baseline methods, but also present superior efficiency at the same time. Meanwhile, the performance of the proposed VATES method under different parameters and settings was also analyzed in the paper.

This work was partially supported by the National Natural Science Foundation of China under Grant No. 61876205, the Natural Science Foundation of Hebei Province of China under Grant No. F2020501034, the Science and Technology Project of Hebei Education Department of China under Grant No. ZD2021403, the Ministry of Education of Humanities and Social Science project under Grant Nos. 19YJAZH128 and 20YJAZH118, and the Laboratory of Language and Artificial Intelligence, Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies under Grant No. LAI202306.