

具有特征选择的多源自适应分类框架

黄学雨¹, 徐浩特^{1*}, 陶剑文²

(1. 江西理工大学 信息工程学院, 江西 赣州 341000; 2. 宁波职业技术学院 电子信息工程学院, 浙江 宁波 315100)

(* 通信作者电子邮箱 haote_xu@aliyun.com)

摘要: 对于现有的多源自适应学习方案无法有效区分多个源域中的有用信息并迁移至目标域的问题, 提出一种具有特征选择的多源自适应分类框架(MACFFS), 并将特征选择和共享特征子空间学习整合到统一框架中进行联合特征学习。具体来说, MACFFS将来自多个源域的特征数据投影至不同的潜在空间中来学习得到多个源域分类模型, 实现目标域的分类。然后, 将得到的多个分类结果进行整合用于目标域分类模型的学习。此外, 框架还利用 $L_{2,1}$ 范数稀疏回归代替传统的基于 L_2 范数的最小二乘回归来提高鲁棒性。最后, 把多种现有方法在两项任务中与MACFFS进行实验比较分析。实验结果表明, 与现有方法中表现最好的DSM相比, MACFFS节省了接近1/4的计算时间, 并且提升了大约2%的识别率。总的来说, MACFFS结合了机器学习、统计学习等相关知识, 为多源自适应方法提供了一个新的思路, 且该方法在现实场景下的识别应用中比现有方法具有更好的性能。

关键词: 多源自适应; 目标识别; 共享特征子空间; 特征选择; 迁移学习

中图分类号: TP181 **文献标志码:** A

Multi-source adaptation classification framework with feature selection

HUANG Xueyu¹, XU Haote^{1*}, TAO Jianwen²

(1. School of Information Engineering, Jiangxi University of Science and Technology, Ganzhou Jiangxi 341000, China;

2. School of Electronic Information Engineering, Ningbo Polytechnic, Ningbo Zhejiang 315100, China)

Abstract: For the problem that the existing multi-source adaptation learning schemes cannot effectively distinguish the useful information in multi-source domains and transfer the information to the target domain, a Multi-source Adaptation Classification Framework with Feature Selection (MACFFS) was proposed. Feature selection and shared feature subspace learning were integrated into a unified framework by MACFFS for joint feature learning. Specifically, multiple source domain classification models were learned and obtained by MACFFS through mapping feature data from multiple source domains into different latent spaces, so as to realize the classification of target domains. Then, the obtained multiple classification results were integrated for the learning of the target domain classification model. In addition, $L_{2,1}$ norm sparse regression was used to replace the traditional least squares regression based on L_2 norm by the framework to improve the robustness. Finally, a variety of existing methods were used to perform experimental comparison and analysis with MACFFS in two tasks. Experimental results show that, compared with the best performing Domain Selection Machine (DSM) in the existing methods, MACFFS has nearly 1/4 of the calculation time saved, and the recognition rate of about 2% improved. In general, with machine learning, statistical learning and other related knowledge combined, MACFFS provides a new idea for the multi-source adaptation method. Furthermore, this method has better performance than the existing methods in recognition applications in real scenes, which had been experimentally proven.

Key words: multi-source adaptation; object recognition; shared feature subspace; feature selection; transfer learning

0 引言

传统机器学习已经在当今的信息社会取得了瞩目的成就。而这些成就主要归功于大量的标记训练数据, 并且基于以下假设: 训练和测试数据独立于一个相同但未知的分布 (Independent and Identically Distributed, IID)^[1]。然而在实际各种领域应用中大量的数据都是属于非IID数据。并且在这些领域中已标注的数据较少, 需要人工准备充足的已标记训练数据, 这是一个需要耗费大量人力物力的过程^[2]。此外, 收

集样本的过程容易产生数据集偏差, 从而导致由训练数据集中学习得到的模型在测试数据集下呈现并非十分理想的测试结果。因此, 标记数据稀缺和数据集偏差是目前机器学习亟须解决的问题。由此领域适应学习 (Domain Adaptation Learning, DAL) 技术在现阶段得以迅速发展。在DAL任务中学习领域通常分为两个相关但不同的类型, 即源域和目标域^[3]。旨在利用源域中的训练数据来解决目标域中的学习问题, 两个域的数据分布可以相同或不同。但是强行将某些与目标域分布差异极大的源域自适应于目标域会导致DAL中

收稿日期: 2020-02-04; 修回日期: 2020-04-21; 录用日期: 2020-04-28。

基金项目: 浙江省自然科学基金资助项目 (LY19F020012); 赣州科技重点研发项目 ([2018]50-8-3)。

作者简介: 黄学雨 (1970—), 男, 江西赣州人, 副教授, 硕士, 主要研究方向: 视觉工程技术、企业信息系统; 徐浩特 (1995—), 男, 浙江温州人, 硕士研究生, 主要研究方向: 机器学习、模式识别; 陶剑文 (1973—), 男, 湖北武汉人, 教授, 博士, 主要研究方向: 模式识别、机器学习。

的“负迁移”问题^[4]。为此,多源自适应学习被提出到DAL的行列中,其旨在利用多个相关的源域,通过最小化源域之间的差异来辅助目标域的学习。多源自适应学习在现实中不同领域的应用上取得了出色的成绩。例如,文献[5]提出的A-SVM (Adaptive Support Vector Machine),该方法首先在多个源域上训练得到相应的分类模型,然后通过这些不同源域的分类模型来辅助目标域分类模型的学习。另外还有基于最小化均值偏差的学习的多源自适应方法如FastDAM (Fast Domain Adaptation Machine)^[6]等。但目前的多源自适应方法仍存在一些至今没有得到妥善的解决:

1) 如何处理域中训练数据可能含有部分无关或冗余的特征信息(例如,噪声和异常值)的问题。

由于源域的视觉训练数据可以从各种网站中随机获得,因此训练数据中的噪声和异常值比比皆是。现有方法盲目地将包含噪声和异常值的所有训练数据转换到共享子空间中,这可能导致学习得到的模型最终分类效果显著弱化。

2) 如何有效挖掘多个源域中包含的相关信息的问题。

大多数多源自适应方法通常分别处理源样本而不考虑几个源之间的相关性,这将导致多源信息无法得到充分的利用使得最终的分类型效果并非十分理想。

3) 如何充分挖掘并利用目标域未标记数据内含有的潜在信息来提高模型性能。

近年来,虽然已有的一些方法能够应对上述的部分问题,但是目前还是没有一个统一的框架能解决上述的问题。文献[7]指出,通过 $L_{2,1}$ 范数对模型矩阵行稀疏化使其拥有稀疏特征选择功能,能够筛选出最具判别性的数据特征。除了剔除特征中存在的噪声信息以外, $L_{2,1}$ 范数还可以通过消除冗余特征信息减小分类模型矩阵的特征维度提升算法计算效率。

具体来说,为了解决现有DAL方法在视觉分类任务中存在的上述问题,本文提出一种具有特征选择的多源自适应分类框架(Multi-source Adaptation Classification Framework with Feature Selection, MACFFS)。所提方法主要创新点在于:

1) 引入迹范数正则化来探索多个源之间的共享信息,并通过优化得到的权重值整合多源信息,利用 $L_{2,1}$ 范数损失函数来减轻噪声或离群值的影响,提出了一种具有特征选择的鲁棒多源自适应分类框架。

2) 将框架的全局优化解转换为一个广义特征分解问题,并对整个过程进行了详细的理论证明。同时给出了基于该框架的简单有效的算法步骤。

3) 分别在几个不同应用场景所对应数据集上进行了全面的实验,以验证所提出的框架的高效性与鲁棒性。

1 多源适应无监督分类框架

为便于描述,此处先提前介绍文中符号的意义。本文用 $\mathbf{A} \in \mathbf{R}^{d \times n}$ 表示大小 $d \times n$ 的矩阵, $A_{i,j}$ 对应于矩阵中的 (i,j) 元素。此外,分别用 $\|\mathbf{A}\|_{2,1} = \sum_{i=1}^n \sqrt{\sum_{j=1}^d A_{i,j}^2}$ 和 $\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^d A_{i,j}^2}$ 表示矩阵的 $L_{2,1}$ 范数和Frobenius范数。矩阵 $\mathbf{A}$ 的迹表示为 $\text{tr}(\mathbf{A})$ 。用 $\mathbf{I}_n$ 定义大小为 $n \times n$ 的单位矩阵。用 $\mathbf{1}_n \in \mathbf{R}^n$ 表示元素都是1的列向量。

1.1 问题描述

为解决现有方法存在的问题,MACFFS框架需具备以下

两个主要特性:1) 有区别地利用多个标签丰富的源域来协助标签稀缺的目标域中的学习任务;2) 通过目标领域潜在信息增强模型的性能。本文所提方法将联合共享子空间学习与多源模型迁移,并利用图流形正则化^[8]有效提升目标学习性能,最终形成一个统一的框架。

1.2 框架的提出

给定包含 n 个 d 维向量的目标数据集 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbf{R}^{d \times n}$,其对应的标签集为 $\mathbf{Y}_t \in \{0, 1\}^{n \times c}$, c 是类的数量。在多标签分类问题的情况下,对于每个输入向量 $\mathbf{x}_i \in \mathbf{X} \in \mathbf{R}^{d \times n}$ ($1 \leq i \leq n$),假设 $\mathbf{y}_i \in \mathbf{Y}_t$ 是与之相关的输出标签,如果 $\mathbf{x}_i$ 被标记属于 j 类别,则 $y_{i,j} = 1$;否则 $y_{i,j} = 0$ 。本文提出的无监督框架假设目标数据集不含有任何标签,即目标域的标签集为一个标签预测矩阵 $\mathbf{F} \in \{0, 1\}^{n \times c}$ (初始为一个全零矩阵)。将样本大小为 n_v 的第 v ($v = 1, 2, \dots, M$)个源域数据集表示为 $\mathbf{A}_v = [\mathbf{x}_1^v, \mathbf{x}_2^v, \dots, \mathbf{x}_{n_v}^v] \in \mathbf{R}^{d \times n_v}$,其对应的标签集为 $\mathbf{Y}_v = [\mathbf{y}_1^v, \mathbf{y}_2^v, \dots, \mathbf{y}_{n_v}^v] \in \mathbf{R}^{n_v \times c}$ 。

传统的有监督学习算法通过将原始数据 $\mathbf{x}$ 映射为某个类别标签值 $\mathbf{y}$ 从而获得预测函数 $f(\mathbf{x})$ ^[9]。即最小化如下正则化经验误差函数:

$$\min_j \text{loss}(f(\mathbf{x}), \mathbf{y}) + \mu \Omega(f) \quad (1)$$

其中: $\text{loss}(\cdot, \cdot)$ 为某个损失函数, $\Omega(\cdot)$ 为正则化函数, $\mu \geq 0$ 为正则化参数。视觉数据的特征之间存在某些共享属性,通过这些相关信息可以强化分类模型的学习。基于该思想,本文通过线性变换矩阵 $\mathbf{P}_v \in \mathbf{R}^{d \times r}$ 将第 v 个源域的特征投影至共享子空间中,其中 r 是特征子空间的维数。因此,可在(1)的基础上将第 v 个源域的分类模型学习形式化为以下优化问题:

$$\min_j \sum_{i=1}^{n_v} \text{loss} \left(\left((\mathbf{x}_i^v)^T \mathbf{P}_v \mathbf{Q}_v \right), \mathbf{y}_i^v \right) + \mu \Omega(f_v) \quad (2)$$

其中: $\mathbf{Q}_v \in \mathbf{R}^{r \times c}$ 为共享子空间内的权重矩阵, μ 是正则化参数。通常可以使用经典的最小二乘损失函数来学习(2)中的预测函数,但是极易受异常值和噪声的影响。因此,在本文的框架中使用 $L_{2,1}$ 范数损失函数增强其鲁棒性减轻噪声/异常值对目标数据的干扰,并利用正交约束使得在新空间中的各个特征相对独立。则该源域的共享特征子空间学习目标函数(2)可描述为:

$$\min_{\mathbf{P}_v, \mathbf{Q}_v} \left\| \mathbf{Y}_v - \mathbf{A}_v^T \mathbf{P}_v \mathbf{Q}_v \right\|_{2,1} + \mu \left\| \mathbf{P}_v \mathbf{Q}_v \right\|_{2,1} \quad (3)$$

$$\text{s. t. } \mathbf{P}_v^T \mathbf{P}_v = \mathbf{I}_r$$

$$\text{正则化项} \left\| \mathbf{P}_v \mathbf{Q}_v \right\|_{2,1} \text{使} \mathbf{P}_v \mathbf{Q}_v \text{达到行稀疏的效果,以实现在}$$

特征共享子空间中的特征选择。换句话说,通过这种方法可以对数据的所有特征信息进行联合评估,并利用不同特征之间的相关性达到信息共享的目的。

与源域类似目标域的学习函数为:

$$\min_{\mathbf{Q}_0, \mathbf{F}} \left\| \mathbf{X}^T \mathbf{Q}_0 - \mathbf{F} \right\|_F^2 + \text{tr} \mathbf{F}^T \mathbf{L} \mathbf{F} + \omega \left\| \mathbf{Q}_0 \right\|_{2,1} \quad (4)$$

其中: ω 为正则化系数; $\mathbf{Q}_0$ 为是针对目标域数据的权重矩阵,同时利用 $\|\cdot\|_{2,1}$ 保证 $\mathbf{Q}_0$ 的行稀疏。为了充分利用目标域的信息用于目标分类器的学习,加入了另一个正则化项:流形正则

化项^[8]。其中 L 为定义于目标领域数据邻接图 $\mathcal{G} = \langle \mathbf{I}, \mathbf{X} \rangle$ 上的Laplacian矩阵且 $L = \mathbf{A} - \mathbf{I}$ 。其中 $\mathbf{A}$ 为一个对角矩阵,第 i 个元素为 $\Delta_{i,i} = \sum_j \Gamma_{i,j}$; $\mathbf{I}$ 为图权值矩阵,当样本 $\mathbf{x}_i$ 和 $\mathbf{x}_j$ 为 k 近邻时,矩阵元素 $\Gamma_{i,j} = 1$,否则 $\Gamma_{i,j} = 0$ 。

为了将 M 个源域的知识共同作用在目标域中,进而统一学习目标域分类模型提出具有特征选择的多源自适应分类框架:

$$\begin{aligned} \arg \min_{P_v, Q_v, T_v, F_v, F, Q_0, \theta_v} & \sum_{v=1}^M \left[\left\| Y_v - A_v^T P_v Q_v \right\|_{2,1} + \right. \\ & \left. \mu \left\| P_v Q_v \right\|_{2,1} + \left\| F_v - X^T P_v Q_v \right\|_{2,1} \right] + \\ & \sum_{v=1}^M \theta_v \left\| F_v - F \right\|_F^2 + \left\| X^T Q_0 - F \right\|_F^2 + \\ & \text{tr} F^T L F + \omega \left\| Q_0 \right\|_{2,1} \end{aligned} \quad (5)$$

s. t. $\sum_{v=1}^M \theta_v = 1, \theta_v \in [0, 1], P_v^T P_v = I_r$

其中 F_v 为目标域实例通过每个源分类器所得到的目标分类标签。而 $\sum_{v=1}^M \theta_v \left\| F_v - F \right\|_F^2$ 是一个全局正则项,它要求目标域特征数据在每个源分类模型中得到的分类结果与一个统一的分类结果 F 对齐。本质上其目的为使 F_v 在不同的源分类标签矩阵之间建立桥梁,使来自一个源的信息可以被利用到另一个源。其中参数 θ_v 表示不同源对目标预测的不同贡献度。

图1为MACFFS的示意图。结合图1可以看出框架学习的具体过程如下:将多个源域的视觉数据投影至不同的隐空间并利用 $L_{2,1}$ 范数的稀疏特性选择有效特征剔除冗余信息,最后学习得到 M 个源域分类器。将无标签的目标域数据通过 M 个源域分类器得到 M 个标签矩阵,并根据不同的权重整合所有的标签矩阵最终得到一个统一的目标域标签矩阵用以目标分类器的学习。

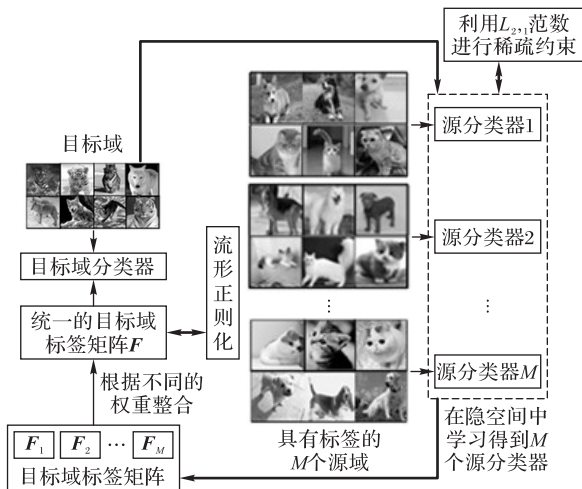


图1 MACFFS示意图

Fig. 1 Schematic diagram of MACFFS

2 优化算法

本文将采取交替优化的策略来对目标函数(5)进行求解。其主要原因为避免式中关于 $L_{2,1}$ 范数的非平滑变量导致整体目标函数不能形成一个闭合的形式。为便于优化,定义 $T_v =$

$P_v Q_v$ 。另外在优化过程中,将采取先设 θ_v 为已知常量从而优化其他变量的策略。因此,得到目标函数的优化求解表示为如下形式:

$$\begin{aligned} \Omega(P_v, Q_v, T_v, F_v, F, Q_0) = & \min \sum_{v=1}^M \left[\alpha \left\| T_v - P_v Q_v \right\|_F^2 + \mu \text{tr} T_v^T O_v T_v + \right. \\ & \text{tr} (Y_v - A_v^T T_v)^T M_v (Y_v - A_v^T T_v) + \\ & \left. \text{tr} (F_v - X^T T_v)^T V_v (F_v - X^T T_v) \right] + \\ & \sum_{v=1}^M \theta_v \left\| F_v - F \right\|_F^2 + \left\| X^T Q_0 - F \right\|_F^2 + \\ & \text{tr} F^T L F + \omega \text{tr} Q_0^T U Q_0 \end{aligned} \quad (6)$$

s. t. $P_v^T P_v = I_r$

其中 O_v, M_v, V_v, U 为对角矩阵,其第 i 个对角元素分别定义为

$$O_{i,i} = \frac{1}{2 \left\| (T_v)_i \right\|_2}, M_{ii} = \frac{1}{2 \left\| (Y_v - A_v^T T_v)_i \right\|_2}, V_{ii} = \frac{1}{2 \left\| (F_v - X^T T_v)_i \right\|_2}$$

和 $U_{ii} = \frac{1}{2 \left\| (Q_0)_i \right\|_2}$ 。其中 $(T_v)_i$ 为矩阵 T_v 的第 i 个行向量,以此类推。

对式(6)根据变量 Q_0 求导并令结果为零可得:

$$\frac{\partial \Omega}{\partial Q_0} = 2XX^T Q_0 - 2XF + 2\omega U Q_0 \Rightarrow$$

$$Q_0 = (XX^T + \omega U)^{-1} XF$$

令 $W = XX^T + \omega U$,得到 $Q_0 = W^{-1} XF$ 。将 Q_0 代入式(6)

可得:

$$\begin{aligned} \arg \min_{P_v, T_v, F_v, F, Q_v} & \sum_{v=1}^M \left[\alpha \left\| T_v - P_v Q_v \right\|_F^2 + \mu \left\| T_v \right\|_{2,1} + \right. \\ & \left. \left\| Y_v - A_v^T T_v \right\|_{2,1} + \left\| F_v - X^T T_v \right\|_{2,1} \right] + \\ & \sum_{v=1}^M \theta_v \left\| F_v - F \right\|_F^2 + \text{tr} F^T L F + \\ & \text{tr} (X^T (W^{-1} XF) - F)^T (X^T (W^{-1} XF) - F) + \\ & \text{tr} (W^{-1} XF)^T U (W^{-1} XF) \end{aligned} \quad (7)$$

s. t. $P_v^T P_v = I_r$

对上式按变量 F_v 求偏导并令其为0得到:

$$\frac{\partial \Omega}{\partial F_v} = 2V_v F_v - 2V_v X^T T_v + 2\theta_v' F_v - 2\theta_v F = 0 \Rightarrow$$

$$(V_v + \theta_v') F_v = \theta_v' F + V_v X^T T_v$$

令 $S_v = V_v + \theta_v'$,得到 $F_v = S_v^{-1} (\theta_v' F + V_v X^T T_v)$ 。将 F_v 代入到式(7)中得到新的函数表达式,对其按变量 F 求偏导并令最终的等式为0得到: $D_v T_v = G_v F$ 。其中:

$$D_v = 2\theta_v' S_v^{-1} V_v X^T - \theta_v' S_v^{-1} V_v S_v^{-1} V_v X^T - \theta_v' S_v^{-1} S_v^{-1} V_v X^T \quad (8)$$

$$G_v = \theta_v' S_v^{-1} V_v S_v^{-1} + \theta_v' I_n + I_n - \theta_v' 2S_v^{-1} + X^T W^{-1} X X^T W^{-1} X + \theta_v' S_v^{-1} S_v^{-1} - 2X^T W^{-1} X + \omega X^T W^{-1} U W^{-1} X + L \quad (9)$$

即 $F = G_v^{-1} D_v T_v$ 。将得到的 F 与 F_v 代入到式(7)中得到式(10):

$$\arg \min_{P_v, T_v, Q_v} [\alpha \operatorname{tr} (T_v - P_v Q_v)^T (T_v - P_v Q_v) + \mu \operatorname{tr} T_v^T O_v T_v +$$

$$\operatorname{tr} (Y_v - A_v^T T_v)^T M_v (Y_v - A_v^T T_v)] - \operatorname{tr} T_v^T D_v^T G_v^{-1} D_v T_v +$$

$$\operatorname{tr} T_v^T (\theta_v^r X V_v S_v^{-1} S_v^{-1} V_v X^T + X V_v S_v^{-1} V_v S_v^{-1} V_v X^T -$$

$$2 X V_v S_v^{-1} V_v X^T + X V_v X^T) T_v$$

$$\text{s. t. } P_v^T P_v = I_r \quad (10)$$

对上式按变量 T_v 求偏导并令其为 0 得到: $T_v = H_v^{-1} (A_v M_v Y_v + \alpha P_v Q_v)$, 其中:

$$H_v = \theta_v^r X V_v S_v^{-1} S_v^{-1} V_v X^T + X V_v S_v^{-1} V_v S_v^{-1} V_v X^T + \alpha I_d +$$

$$\mu O_v + A_v M_v A_v^T + X V_v X^T - D_v^T G_v^{-1} D_v - 2 X V_v S_v^{-1} V_v X^T$$

将 $T_v = H_v^{-1} (A_v M_v Y_v + \alpha P_v Q_v)$ 代入到 (10) 中得到新的函数表达式。对新的到的表达式按变量 Q_v 求偏导并令其为零, 可得:

$$Q_v = K_v^{-1} (\alpha P_v^T H_v^{-1} A_v M_v Y_v) \quad (11)$$

其中 $K_v = \alpha I_r - \alpha^2 P_v^T H_v^{-1} P_v$ 。最后得到目标函数:

$$\max_{P_v} \operatorname{tr} (Y_v^T M_v^T A_v^T H_v^{-1} P_v) K_v^{-1} (P_v^T H_v^{-1} A_v M_v Y_v)$$

$$\text{s. t. } P_v^T P_v = I_r \quad (12)$$

问题 (12) 的最大化问题可转化为一个广义特征分解问题, 具体形式由如下定理给出。

定理 1 目标函数 (12) 的全局最优解可通过求解如下迹最大化问题获得:

$$\max_{P_v^T P_v = I_r} \operatorname{tr} ((P_v^T Z_v P_v)^{-1} (P_v^T J_v P_v)) \quad (13)$$

其中: $J_v = (H_v^{-1} A_v M_v Y_v Y_v^T M_v^T A_v^T H_v^{-1})$, $Z_v = (I_d - \alpha H_v^{-1})$ 。

证明 根据矩阵的特性 $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, 优化问题 (12) 等价于:

$$\max_{P_v^T P_v = I_r} \operatorname{tr} (I_r - \alpha P_v^T H_v^{-1} P_v)^{-1} (P_v^T J_v P_v) \quad (14)$$

其中 $J_v = (H_v^{-1} A_v M_v Y_v Y_v^T M_v^T A_v^T H_v^{-1})$ 。令 $Z_v = (I_d - \alpha H_v^{-1})$, 由此定理得证。而优化问题 (12) 可由矩阵 $Z_v^{-1} J_v$ 的特征分解获得。

下面对变量 θ_v 进行优化, 由于除变量 θ_v 以外其他变量都可视为常量因此可得到如下目标函数:

$$\min_{\theta_v} \sum_{v=1}^M \theta_v^r \|F_v - F\|_F^2$$

$$\text{s. t. } \sum_{v=1}^M \theta_v = 1, \theta_v \in [0, 1] \quad (15)$$

根据经验设定 θ_v 的指数 $r = 2$ 并利用拉格朗日乘子 δ , 将上述问题转化为如下拉格朗日函数:

$$\tau(\theta_v, \delta) = \sum_{v=1}^M \theta_v^2 \|F_v - F\|_F^2 - \delta \left(\sum_{v=1}^M \theta_v - 1 \right) \quad (16)$$

对上式子按变量 θ_v 求偏导并令其为零, 可得:

$$\frac{\partial \tau}{\partial \theta_v} = 2 \theta_v \left(\|F_v - F\|_F^2 \right) - \delta \Rightarrow \theta_v = \frac{\delta}{2} \left(\|F_v - F\|_F^2 \right)^{-1} \quad (17)$$

由于 $\sum_{v=1}^M \theta_v = 1$, 可得:

$$\theta_v = \frac{\theta_v}{\sum_{v=1}^M \theta_v} = \frac{1 / \left(\|F_v - F\|_F^2 \right)}{\sum_{v=1}^M 1 / \left(\|F_v - F\|_F^2 \right)} \quad (18)$$

3 半监督式 MACFFS

3.1 半监督框架与算法的提出

在实际情况中目标域都会含有极少量的标签数据, 为了充分利用这些标签, 本文提出具有特征选择的多源自适应分

类框架的半监督形式: 将来自目标域已标记与未标记的实例集合分别表示为 $X_l = \{\mathbf{x}_i\}_{i=1}^{n_l}$ 与 $X_u = \mathbf{x}_u \Big|_{u=n_l+1}^{n_l+n_u}$ ($n = n_l + n_u, n_l \ll n_u$)。目标训练样本对应的标签集为 $Y_l = [Y_l^T, Y_u^T]^T \in \{0, 1\}^{n \times c}$, 其中 $Y_l \in \{0, 1\}^{n_l \times c}$, 而 $Y_u \in \{0, 1\}^{n_u \times c}$ 是一个全零矩阵, c 是类的数量。目标预测标签矩阵 F 应该与实际的目标标签矩阵 Y_l 相一致。因此半监督形式的多源自适应分类框架如下:

$$\arg \min_{P_v, Q_v, T_v, F_v, F, Q_0, \theta_v} \left[\|Y_v - A_v^T P_v Q_v\|_{2,1} + \mu \|P_v Q_v\|_{2,1} + \right.$$

$$\left. \|F_v - X^T P_v Q_v\|_{2,1} \right] + \sum_{v=1}^M \theta_v^r \|F_v - F\|_F^2 + \|X^T Q_0 - F\|_F^2 +$$

$$\operatorname{tr} F^T L F + \omega \|Q_0\|_{2,1} + \|F - Y_l\|_F^2$$

$$\text{s. t. } \sum_{v=1}^M \theta_v = 1, \theta_v \in [0, 1], P_v^T P_v = I_r \quad (19)$$

对其优化需对原框架进行相应的替换, 具体内容如下:

$$G_v \rightarrow \tilde{G}_v = (X^T W^{-1} X X^T W^{-1} X - 2 X^T W^{-1} X +$$

$$\omega X^T W^{-1} U W^{-1} X + (2 + \theta_v^r) I_n + L +$$

$$\theta_v^{3r} S_v^{-1} S_v^{-1} - 2 \theta_v^{2r} S_v^{-1} + \theta_v^{2r} S_v^{-1} V_v S_v^{-1}) \quad (20)$$

$$F \rightarrow F = \tilde{G}_v^{-1} (D_v T_v + Y_l) \quad (21)$$

$$H_v \rightarrow \tilde{H}_v = (\theta_v^r X V_v S_v^{-1} S_v^{-1} V_v X^T + D_v^T \tilde{G}_v^{-1} D_v -$$

$$2 X V_v S_v^{-1} V_v X^T + \alpha I_d + A_v M_v A_v^T +$$

$$X V_v X^T + X V_v S_v^{-1} V_v S_v^{-1} V_v X^T + \mu O_v)$$

$$T_v \rightarrow T_v = \tilde{H}_v^{-1} (\alpha P_v Q_v + \xi_v) \quad (23)$$

其中 $\xi_v = A_v M_v Y_v + D_v^T \tilde{G}_v^{-1} Y_l$ 。

$$Q_v \rightarrow Q_v = (I_r - \alpha P_v^T \tilde{H}_v^{-1} P_v)^{-1} P_v^T \tilde{H}_v^{-1} \xi_v \quad (24)$$

在除 P_v 以外其他变量已知的情況下, 对 (16) 的全局最优解可通过求解如下迹最大化问题:

$$\max_{P_v^T P_v = I_r} \operatorname{tr} ((P_v^T \tilde{Z}_v P_v)^{-1} (P_v^T \tilde{J}_v P_v)) \quad (25)$$

其中 $\tilde{J}_v = \tilde{H}_v^{-1} \xi_v \xi_v^T \tilde{H}_v^{-1}$, $\tilde{Z}_v = (I_d - \alpha (\tilde{H}_v^{-1})^{-1})$ 。

综上, 可通过交替优化各变量的方式对半监督型的 MACFFS 框架进行迭代优化, 具体算法如下。

算法 1 半监督式 MACFFS。

输入: 多个源领域训练数据集与其所对应的标签 $\{(\mathbf{x}_i^s, \mathbf{y}_i^s), i = 1, 2, \dots, n, v = 1, 2, \dots, M\}$; 目标域数据集 $\{(\mathbf{x}_i, \mathbf{y}_i), i = 1, 2, \dots, n\}$ (其中标签未知样本对应的 $\mathbf{y}_i$ 为元素全为 0 的且维度为 c 的向量); KNN 的近邻数量 k , 参数 α, μ, ω 。

输出: 收敛完毕的 F, Q_0, F_v, P_v, Q_v 和 θ_v 。

初始化: 设置 $t = 0$, 并随机初始化 $\theta_v^0, Q_v^0, Q_0^0, P_v^0, F_v^0, F^0$ 并将 O_v^0, M_v^0, V_v^0, U^0 设置为单位矩阵。

步骤:

构造 K 近邻图并计算 L ;

$t = 0$;

while ($|\operatorname{Max} \Omega_t - \operatorname{Min} \Omega_t| / \operatorname{Max} \Omega_t > 10^{-4}$)

{For ($v = 1$; $v \leq M$; $v++$)

{计算 $W^t = X X^T + \omega U^t$;

计算 $S_v^t = V_v^t + \theta_v^t I_n$;

根据式 (8) 计算 D_v^t ;

根据式 (20) 计算 $\tilde{G}_v^t$;

根据式 (22) 计算 $\tilde{H}_v^t$;

计算 $\xi_v^t = \lambda A_v M_v^t Y_v + (D_v^t)^T (\tilde{G}_v^t)^{-1} Y_l$

计算 $\tilde{Z}_v^t = (I_d - \alpha (\tilde{H}_v^t)^{-1})$;

计算 $\tilde{J}_v^t = (\tilde{H}_v^t)^{-1} \xi_v^t (\tilde{H}_v^t)^{-1}$;

对矩阵 $(\tilde{Z}_v^t)^{-1} \tilde{J}_v^t$ 进行特征分解,求得 P_v^{t+1} ;

根据(24)更新 Q_v^{t+1} ;

根据(23)更新 T_v^{t+1} ;

更新 $F^{t+1} = (\tilde{G}_v^t)^{-1} (D_v^t T_v^{t+1} + Y_i)$;

更新 $F_v^{t+1} = (S_v^t)^{-1} ((\theta_v^t)^T F^{t+1} + V_v^t X^T T_v^{t+1})$

更新 $Q_0^{t+1} = (W_v^t)^{-1} X F^{t+1}$;

更新 $O_v^{t+1}, M_v^{t+1}, V_v^{t+1}, U^{t+1}$;

根据式(18)更新 θ_v^{t+1} ;

}

更新 $t = t + 1$;

}

算法1由里外两个循环组成,外循环为一个迭代循环,而里循环依次对每个源域的相关变量进行优化更新。当遍历完所有源域后,将所有更新后的变量代入到式(19)得到一个值 Ω_i ,并继续进行下一轮的迭代。当达到条件 $|\text{Max } \Omega_i - \text{Min } \Omega_i| / \text{Max } \Omega_i < 10^{-4}$ 时,预示着算法收敛迭代完成。最后输出所有优化完成的变量。

3.2 目标域样本标签计算方法

本文将这些多个源域分类模型得到的决策值通过不同源域的贡献度 θ_v 线性融合作为最终决策值。换句话说,来自目标域的样本 x_i 对应的标签决策值 y_i 由下式给出:

$$y_i = \sum_{v=1}^M \theta_v x_i P_v Q_v \quad (26)$$

4 实验与结果

4.1 实验设置

分别对本文提出的 MACFFS 与相关的多源自适应方法在 Caltech-256+Office 数据集和 TRECVID 2005 数据集上进行实验分析,以证明所提方法对各种任务的广泛适用性和相较于其他方法性能方面的优异性。

根据实际的视觉分类任务,将所提方法与几个多源自适应方法进行比较。具体方法如下:

- 1) TCA(Transfer Component Analysis)^[10];
- 2) A-SVM^[5];
- 3) DSM(Domain Selection Machine)^[11];
- 4) FastDAM^[6];
- 5) Multi-KT(Multi Model Knowledge Transfer)^[12];
- 6) MACFFS 变体方法:MACFFS_1 和 MACFFS_2。

其中加入 MACFFS 变体方法的目的是验证本文所提框架各个组件的必要性,其具体形式与验证目的如下:

1) 设置 $\theta_v = 1/M$ 的 MACFFS,简称为 MACFFS_1;通过此设置使框架在相同权重多源适应的情况下进行性能评估。

2) 设置 $\mu = 0$ 和 $\omega = 0$ 的 MACFFS,简称 MACFFS_2;此设置的目的是评估没有特征选择项的 MACFFS 的性能。

关于 TCA、DSM、FastDAM、Multi-KT 和 A-SVM 中的参数按照各自文献中的最优设置来选择。其中高斯核参数 γ 通过在源集上交叉验证得到一个公共值。

在本文的方法中主要有3个模型参数,即 α, μ 和 ω ,这些参数在 $\{10^{-6}, 10^{-5}, \dots, 10^6\}$ 范围内进行调整,根据经验设置最近邻数 $k=5$ 。本研究将使用五重交叉验证法针对每个参数选取最优值,每次参数的设置都将重复5次实验,并取平均值作为最终的实验结果。

4.2 数据集描述

4.2.1 目标识别任务数据集

Caltech-256+Office 数据集包含来自4个域的图像: Amazon(A)、DSLR(Digital Singular Lens Reflex)(D)、Webcam(W)和 Caltech-256(C)。Caltech-256 与 Office 数据集之间共有10个通用类,总共包含2533张图像^[13]。在实验中,所有图像均按保留比例调整大小为 150×150 并对所有图像使用6种特征表示:CH(Color Histogram)、LSS(Local Self-Similarity)、PHOG(Pyramid Histogram of Oriented Gradients)、SIFT(Scale-Invariant Feature Transform)、CSIFT(Color Scale-Invariant Feature Transform)^[14] 和 SURF(Speeded Up Robust Feature)^[15] 特征。将每幅图像的6种特征表示拼接组合为一个新的特征向量。当作为源域时,Amazon/Caltech 每类使用20个训练样本,而 DSLR/Webcam 每类使用8个训练样本。当其作为目标域,都仅使用3个已标记的样本作为训练数据,目标域中的其余数据用于测试。另外,本文从 Caltech-256 中随机抽取10类和20类组成两个集合分别进行目标识别。第二组是在第一组的基础上再随机添加10类获得。在基于 Caltech-256 数据集的多源适应任务中,将依次把其中一个类别视为目标域,其余的视为源域,最终得到的平均识别率用于评估几个方法在 Caltech-256 数据集中多源情况下的识别性能,该数据集包含256个对象类别的图像以及杂乱类别的图像,这些杂乱类别用作对象与背景问题的负类别。

4.2.2 视频概念识别任务数据集

TRECVID 2005^[16] 是最大的视频语料库之一。该数据集由从6个不同广播频道收集的108个小时视频节目的61901个关键帧组成,包括2个汉语频道 CCTV 与 NTDTV(New Tang Dynasty Television),3个英语频道 CNN(Cable News Network)、MSNBC(Microsoft National Broadcasting Corporation)和 NBC(National Broadcasting Company),以及一个阿拉伯语频道 LBC(Lebanese Broadcasting Corporation Television)。表1列出了每个频道中关键帧的总数(除 MSNBC 以外)。为所有关键帧提取6个视觉特征:SIFT、SURF、GiST(Generalized Search Tree)^[17]、LBP(Local Binary Pattern)、PHOG 和 WT(Wavelet Texture)。每个样本的多个特征经过与目标识别任务中相同的方法组合得到新视觉特征向量。从 LSCOM-lite 词典中选择了36个视频概念,该词典涵盖了广播新闻视频中存在的36个主要视觉概念,包括对象、位置、人物、事件和节目。通过手动注释这36个概念,以描述 TRECVID 2005 数据集中关键帧的视觉内容。在实验中,本文将两个英语频道(CNN 和 NBC)和两个汉语频道用作源域,选择阿拉伯语频道作为目标域,并在其中随机采样10个样本以进行标记加入到训练数据中。另外本文还通过设置 CNN_ENG 为目标域,其余的频道为源域测试不同目标域标记样本的数量对性能的影响。在所有设置中除训练数据外,来自目标域的其余样本用作测试数据集。

表1 TRECVID 2005 数据集说明

Tab. 1 Description of TRECVID 2005 dataset

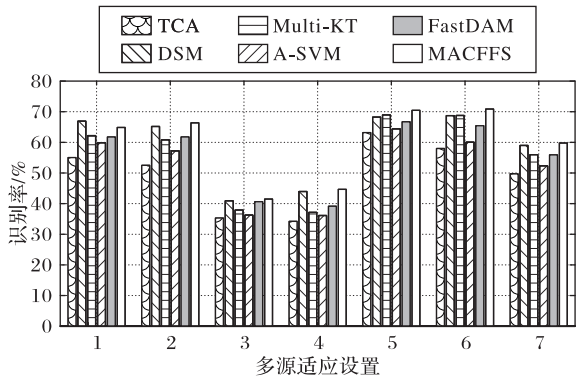
频道	关键帧数	频道	关键帧数
CNN_ENG	11 025	NTDTV_CHN	6 481
NBC_ENG	9 322	LBC_ARB	15 272
CCTV_CHN	10 896		

4.3 实验结果与分析

实验中分别比较了本文的方法与其他基准方法在多个源适应设置中对上述两个跨域学习任务的识别性能,并在图2~5

中报告了实验结果。其中,图4和图5(b)分别显示了所有比较方法在不同标记目标样品数量下的识别结果。根据不同的算法在不同实验设置中的识别率,可以分析得到如下几个结论:

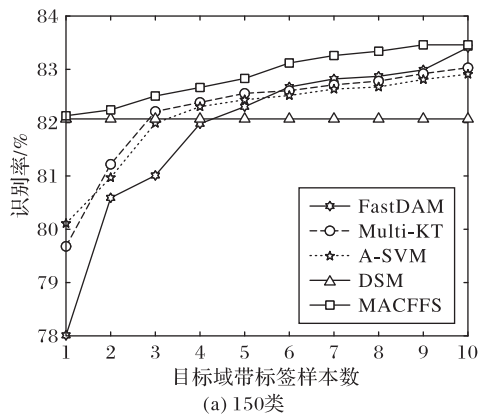
1)如图2、3与图5(a)所示,Multi-KT和FastDAM的性能不稳定。这可能是因为所选的内核函数并不适合所有情况。从图5(a)可以观察到,Fast-DAM和Multi-KT实现了相似的性能。但是,在大多数情况下,Multi-KT比Fast-DAM稍差一些。这是由于在多源自适应在视频识别这样的复杂场景中,Multi-KT几乎无法有效地估计要自适应的源域的权重。而另一个有趣的观察结果是,在大多数情况下,TCA与A-SVM通常比其他DAL方法差,这可能是由于TCA在多个源域的情况该方法只能平均分配给每个源相同的权重,而A-SVM表现较差的原因是由于目标域中的有价值的未标记数据在A-SVM中未被完全利用。在大多数情况下,本文所提出的方法的性能都要优于其他比较方法。例如,在图2中,在大多数情况下所提方法的结果始终优于其他算法(除了在D, A→W目标识别任务中识别率略低于DSM方法)。这证明了本文的模型在视觉识别中具有较高的有效性。



注: 横轴的序号1~7分别代表D, A→W; D, A, C→W; W, D→A; W, D, C→A; A, W→D; A, C, W→D和平均结果

图2 Caltech-256+Office数据集的识别率

Fig. 2 Recognition rate on Caltech-256+Office dataset



(a) 150类

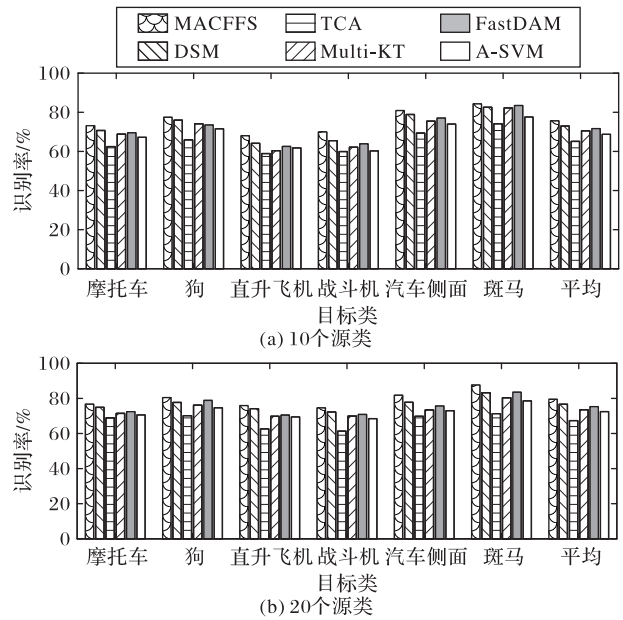
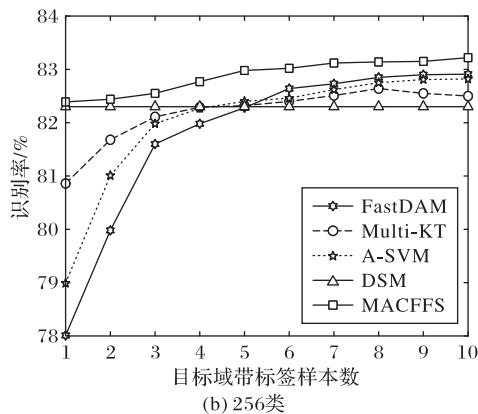


图3 Caltech-256数据集的识别率

Fig. 3 Recognition rate on Caltech-256 dataset

2)为了进一步评估在具有不同先验信息的情况下算法的性能,本文在实验中通过更改标记目标样本的数量,以研究所提方法的性能。从图4和图5(b)可以看出除DSM以外,所有方法在所有情况下都可以通过标记更多的目标样本实现更高的识别精度,这表明利用标记的目标数据来改善学习性能是有益的。当带标签的目标样本数量逐渐增加时,FastDAM方法获得最为显著的性能提升。此外,可以观察到本文的方法在所有多个源设置中性能都可以平滑地提升。这说明即使仅使用较小数量的标记目标数据,MACFFS仍可以获得较高的识别准确率。但是,其他DAL方法(尤其是FastDAM)只能在标记的目标样品数量相对较大时,才能得到令人满意的性能。



(b) 256类

图4 多源扩展性

Fig. 4 Multi-source scalability

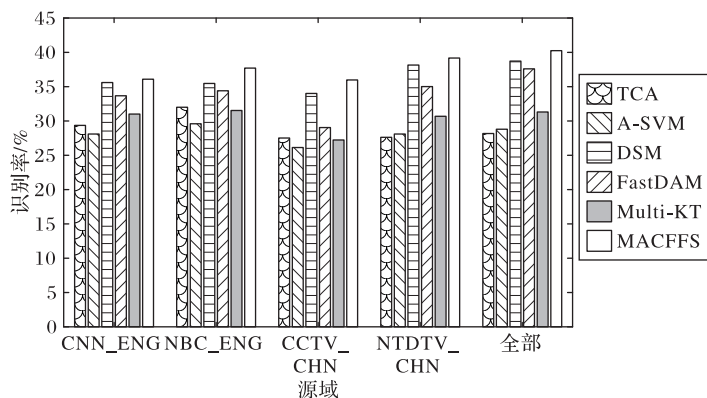
3)在图3中,20类的识别率在某种程度上明显高于10类的识别率。在其他任务(例如在图5(a)的视频概念识别任务)上也可以观察到相同的结果,即当源域数量逐渐增加时,本文的方法在性能上获得了明显的改善。此外,可以在图5中观察到,图5(b)中各方法的性能比5(a)要更好,可以由A-SVM与Multi-KT两个方法在两图中的比较观察得到。具体的,在图5(a)中A-SVM与Multi-KT识别率大概为30%(A-SVM甚至

在30%以下),而在图5(b)中,两个算法的识别率都有提升(在标记样本数达到10时,两个方法的识别率都高于35%)。这是由于在图5(b)任务中源域中有与目标域类似的英文频道因此源域可迁移至的目标域的有用信息会更多。在进行实验之前,本研究期望通过增加相关源域的数量,从而增加找到适用于目标域的有用先验知识的可能性。而这些实验结果验证了该方法的有效性:结合从更多来源领域中获得的判别信

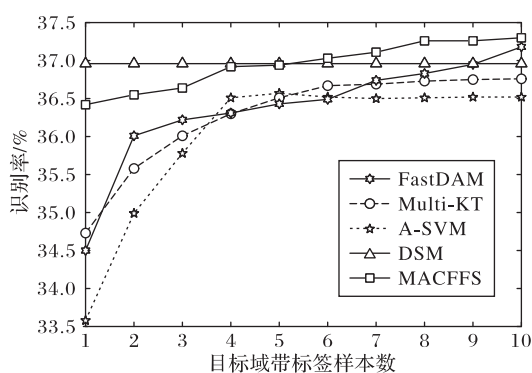
息可以进一步提高适应性能。

4)对于任何开放式学习系统,已知对象类别的数量会随着时间的增长而增加。由于需要针对新任务检查每个已知模型的可靠性,随着源域的大量增加可能会在领域适应学习中存在能否扩展的问题。具体而言,对于100数级以上的源域个数,上述领域适应方法在计算方面变得极为昂贵且对识别性能会有所影响。因此本文对来自Caltech-256数据集的150个和256个

对象类进行了实验,分别在图4中报告了MACFFS、Multi-KT、DSM、FastDAM和A-SVM的实验结果。在这两种情况下,可以看出对于极少的已标记目标训练样本,正确选择每个源域的权重相对于所有来源的平均值是更有效的方式:在少于三个目标训练样本时,MACFFS和Multi-KT优于FastDAM和A-SVM。同时这也表明了当拥有足够的训练样本和丰富的域集时,最有效的方法是不要忽略任何源域的信息。



(a) 不同源域设置下识别率



(b) 不同目标域带标签样本数下识别率

图5 视频概念识别任务中不同设置下的识别率

Fig. 5 Recognition rates under different settings in video concept recognition task

除此之外,本文还在两个视觉识别任务中分别测试了MACFFS与MACFFS变体方法(MACFFS_1和MACFFS_2)的实验性能,其结果显示在图6中。通过分析可得到如下结论:

1) MACFFS性能明显优于MACFFS_1。这些效益归因于每个源域在信息迁移至目标域时具有不同的权重可以达到减少冗余信息的效果,从而有利于提高性能。

2) 通过MACFFS与MACFFS_2的性能比较可知,如果去除特征选择项,MACFFS的性能将在一定程度上退化。原因是由低级视觉特征表示的图像可能会引入过多的噪声信息,导致其降低性能。这同时也证明了协同回归中特征选择的必要性。

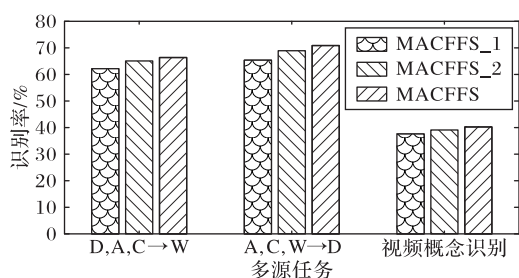


图6 MACFFS变体方法在多个识别任务中的识别率

Fig. 6 Recognition rates of MACFFS variant methods in multiple recognition tasks

4.4 算法收敛性与时间复杂度分析

由上述可知,本文所提算法1为一个交替优化过程,图7显示了在上述两个真实数据集上的算法的收敛曲线。从图7可以看出,目标值通常会在10次迭代中收敛。这是由于所需优化的函数为一个凸函数,因此函数可以在少量次数的迭代后收敛。另外可以从图8中看出本文的算法在计算时间并不占优势(计算时间与最快的A-SVM方法相差10s左右)。导致这个现象的主要原因是由于算法中对一些矩阵逆的计算所需要的时间复杂度为 $O(d^3)$,另外对矩阵 $(\tilde{Z}_v)^{-1}\tilde{J}_v$ 的特征分解

的时间复杂度为 $O(d^3)$ 。因此,算法整体的计算时间会相对较长。但是从各个实验结果中可以看出通过牺牲一定的时间效率所获得的性能的提升还是非常可观的。

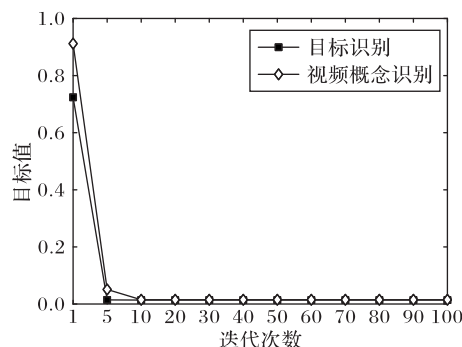


图7 算法收敛曲线

Fig. 7 Algorithm convergence curves

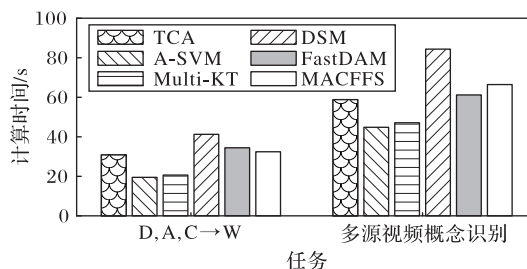


图8 算法计算时间

Fig. 8 Calculation times of algorithms

5 结语

本文提出了一种新的具有特征选择的多源自适应分类框架(MACFFS),旨在将来自多个源域的特征数据同时在不同潜在空间中学习得到分类模型,并将得到的源域分类模型用以目标数据的分类,最终对得到的分类结果进行整合并有区分性地帮助目标分类模型的学习。此外,还通过在几个

跨域视觉识别任务上进行实验和分析,证明本文的方法与其他相关的最新算法相比的优越性与框架中几个组成项的重要性。然而,在MACFFS中仍然存在两个问题需要在后续的工作中展开研究并设法解决。第一个问题是如何降低该方法在高维和大规模数据集中计算的时间复杂度。MSMFR中的最佳模型参数是数据相关的。因此,另一个问题则是如何自动确定所提方法的最优参数。

参考文献 (References)

- [1] 李茂莹,杨柳,胡清华. 同构迁移学习理论和算法研究进展[J]. 南京信息工程大学学报(自然科学版), 2019, 11(3):269-277. (LI M Y, YANG L, HU Q H. A survey on theories and algorithms about homogeneous transfer learning[J]. Journal of Nanjing University of Information Science and Technology (Natural Science Edition), 2019, 11(3):269-277.)
- [2] 韩嵩,韩秋弘. 半监督学习研究的述评[J]. 计算机工程与应用, 2020, 56(6):19-27. (HAN S, HAN Q H. Review of semi-supervised learning research [J]. Computer Engineering and Applications, 2020, 56(6):19-27.)
- [3] 贾颖霞,郎丛妍,冯松鹤. 基于类别相关的领域自适应交通图像语义分割方法[J]. 计算机研究与发展, 2020, 57(4):876-887. (JIA Y X, LANG C Y, FENG S H. A semantic segmentation method of traffic scene based on categories-aware domain adaptation [J]. Journal of Computer Research and Development, 2020, 57 (4): 876-887.)
- [4] JIA X, SUN F. Unsupervised deep domain adaptation based on weighted adversarial network [J]. IEEE Access, 2020, 8: 64020-64027.
- [5] YANG J, YAN R, HAUPTMANN A G. Cross-domain video concept detection using adaptive SVMs[C]// Proceedings of the 15th ACM International Conference on Multimedia. New York: ACM, 2007: 188-197.
- [6] DUAN L, TSANG I W, XU D. Domain transfer multiple kernel learning [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(3):465-479.
- [7] MA Z, NIE F, YANG Y, et al. Web image annotation via subspace-sparsity collaborated feature selection [J]. IEEE Transactions on Multimedia, 2012, 14(4): 1021-1030.
- [8] NIE F, XU D, TSANG I W H, et al. Flexible manifold embedding: a framework for semi-supervised and unsupervised dimension reduction [J]. IEEE Transactions on Image Processing, 2010, 19 (7):1921-1932.
- [9] 姚哲,陶剑文. 多源适应多标签分类框架[J]. 计算机工程与应用, 2017, 53 (7): 88-96, 170. (YAO Z, TAO J W. Multi-source adaptation multi-label classification framework via joint sparse feature selection and shared subspace learning [J]. Computer Engineering and Applications, 2017, 53(7): 88-96, 170.)
- [10] PAN S J, TSANG I W, KWOK J T, et al. Domain adaptation via transfer component analysis [J]. IEEE Transactions on Neural Networks, 2011, 22(2): 199-210.
- [11] DUAN L, XU D, CHANG S F. Exploiting web images for event recognition in consumer videos: a multiple source domain adaptation approach [C]// Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2012:1338-1345.
- [12] TOMMASI T, ORABONA F, CAPUTO B. Learning categories from few examples with multi model knowledge transfer [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(5): 928-941.
- [13] GHIFARY M, BALDUZZI D, KLEIJIN W B, et al. Scatter component analysis: a unified framework for domain adaptation and domain generalization [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(7):1414-1430.
- [14] GENDREAU M, HERTZ A, LAPORTE G. A tabu search heuristic for the vehicle routing problem [J]. Management Science, 1994, 40 (10): 1276-1289.
- [15] BAY H, ESS A, TUYTELAARS T, et al. Speeded-Up Robust Features (SURF) [J]. Computer Vision and Image Understanding, 2008, 110(3): 346-359.
- [16] SMEATON A, OVER P, DOHERTY A R, et al. Video shot boundary detection: Seven years of TRECVID activity [J]. Computer Vision and Image Understanding, 2010, 114 (4): 411-418.
- [17] OLIVA A, TORRALBA A. Modeling the shape of the scene: a holistic representation of the spatial envelope [J]. International Journal of Computer Version, 2001, 42(3): 145-175.

This work is partially supported by the Natural Science Foundation of Zhejiang Province (LY19F020012), the Key Research and Development Project of Ganzhou Science and Technology ([2018]50-8-3).

HUANG Xueyu, born in 1970, M. S., associate professor. His research interests include visual engineering technology, enterprise information system.

XU Haote, born in 1995, M. S. candidate. His research interests include machine learning, pattern recognition.

TAO Jianwen, born in 1973, Ph. D., professor. His research interests include pattern recognition, machine learning.