

基于双分支融合与分段域适应迁移学习的疲劳驾驶检测

李泽群, 丁飞

引用本文

李泽群, 丁飞. 基于双分支融合与分段域适应迁移学习的疲劳驾驶检测[J]. 计算机科学, 2026, 53(3): 78-87.

LI Zequn, DING Fei. [Fatigue Driving Detection Based on Dual-branch Fusion and Segmented Domain Adaptation Transfer Learning](#) [J]. Computer Science, 2026, 53(3): 78-87.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于深度学习的GIFT-128与ASCON算法神经差分区分器研究](#)

Deep Learning-based Neural Differential Distinguishers for GIFT-128 and ASCON
计算机科学, 2026, 53(3): 453-458. <https://doi.org/10.11896/jsjcx.250600176>

[基于Transformer的域自适应物联网流量入侵检测方法](#)

Transformer-based Domain Adaptation Method for IoT Traffic Intrusion Detection
计算机科学, 2026, 53(3): 443-452. <https://doi.org/10.11896/jsjcx.241200167>

[基于多任务学习的眼科视频特征融合与多维画像](#)

Multi-task Learning-based Ophthalmic Video Feature Fusion and Multi-dimensional Profiling
计算机科学, 2026, 53(3): 383-391. <https://doi.org/10.11896/jsjcx.260200058>

[基于Transformer架构的RNA二级结构预测方法](#)

Prediction Method of RNA Secondary Structure Based on Transformer Architecture
计算机科学, 2026, 53(3): 375-382. <https://doi.org/10.11896/jsjcx.250100005>

[基于少量目标数据和深度学习的行人重识别方法](#)

Pedestrian Re-identification Methods Based on Limited Target Data and Deep Learning
计算机科学, 2026, 53(3): 287-294. <https://doi.org/10.11896/jsjcx.260100073>

基于双分支融合与分段域适应迁移学习的疲劳驾驶检测

李泽群¹ 丁飞^{1,2,3}

1 南京邮电大学物联网学院 南京 210003

2 南京邮电大学现代邮政学院 南京 210003

3 南京邮电大学智慧交通学院 南京 210003

(18896709562@163.com)

摘要 疲劳驾驶是引发交通事故的重要原因之一。针对实际场景中摄像头角度、环境光等因素导致的特征提取不足和不同数据下模型适应性差的问题,提出了一种新型基于迁移学习的疲劳驾驶检测框架。该框架通过设计卷积神经网络与 Transformer 双分支特征提取与融合结构,实现 CNN 与 Transformer 的优势互补,增强了模型的特征表征能力,充分提取了驾驶员的局部与全局面部特征。为提高模型在源域与目标域之间的自适应能力,框架采取分段域适应策略,在特征提取阶段采用对抗域适应和多核最大均值差异(MK-MMD)策略,并在特征融合阶段进一步引入 MK-MMD 和最小类别混淆损失(MCC),使模型充分适应不同数据。在两个具有显著特征差异的数据集上的实验结果表明,该框架在目标域上的检测准确率达到 93.3%(A 为源域,B 为目标域)和 75.1%(B 为源域,A 为目标域),显著提升了模型的适应性与鲁棒性。

关键词:深度学习;域适应;迁移学习;疲劳检测;卷积网络;Transformer

中图分类号 TP391.9

Fatigue Driving Detection Based on Dual-branch Fusion and Segmented Domain Adaptation Transfer Learning

LI Zequn¹ and DING Fei^{1,2,3}

1 School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

2 School of Modern Posts, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

3 School of Intelligent Transportation, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Abstract Fatigue driving is one of the leading causes of traffic accidents. To address the issues of insufficient feature extraction caused by factors such as camera angles and environmental lighting, as well as poor model adaptability across different datasets, this paper proposes a novel transfer learning-based fatigue driving detection framework. The framework employs a dual-branch feature extraction and fusion architecture combining CNN and Transformer, leveraging their complementary strengths to enhance feature representation and comprehensively capture both local and global facial features of drivers. To improve the model's adaptive capability between source and target domains, a segmented domain adaptation strategy is adopted. Adversarial domain adaptation and multi-kernel maximum mean discrepancy(MK-MMD) are applied during the feature extraction stage, while MK-MMD and minimum class confusion(MCC) loss are further introduced during the feature fusion stage to enhance cross-domain adaptability. Experimental results on two datasets with significant feature disparities demonstrate that the proposed framework achieves 6 detection accuracies of 93.3%(A→B) and 75.1%(B→A) on the target domain, significantly improving the model's adaptability and robustness.

Keywords Deep learning, Domain adaptation, Transfer learning, Fatigue detection, Convolutional network, Transformer

到稿日期:2025-05-09 返修日期:2025-09-10

基金项目:国家自然科学基金(62471241);江苏省高等学校基础科学(自然科学)研究重大项目(25KJA510003);江苏省“六大人才高峰”高层次人才资助项目(DZXX-008)

This work was supported by the National Natural Science Foundation of China(62471241), Major Project of the Natural Science Foundation of the Jiangsu Higher Education Institutions of China(25KJA510003) and “Six Talent Peaks” High Level Talent Funding Project of Jiangsu Province (DZXX-008).

通信作者:丁飞(dingfei@njupt.edu.cn)

1 引言

汽车在提高出行效率和生活质量的同时,也带来了安全隐患,其中疲劳驾驶是引发交通事故的主要原因之一^[1]。据统计,疲劳驾驶导致的交通事故数量与酒驾、超速相当,造成了严重的人员伤亡和经济损失。因此,准确识别驾驶员疲劳状态并及时预警,对降低事故率、保障道路安全至关重要。

目前,驾驶员疲劳检测方法主要分为基于生理信号、车辆行为和计算机视觉3种方法。生理信号检测精度高但成本昂贵且会干扰驾驶^[2-4],车辆行为分析简便但受个体差异影响较大^[5-6]。相比之下,基于计算机视觉的非接触式检测方法凭借其较低成本和较高准确性得到了广泛应用^[7-11]。然而,上述方法多局限于同源数据集,未能充分考虑实际场景中存在的跨域分布差异问题,导致在面对真实驾驶环境复杂多变的光照条件和个体差异时,检测性能显著下降。因此,开发泛化能力的跨域自适应算法是提升疲劳检测系统实用性的关键。

为提升驾驶员疲劳检测模型在复杂环境下的适应性,迁移学习与域适应技术提供了有效解决方案。迁移学习可将实验室环境(源域)训练的模型知识迁移至真实驾驶场景(目标域),显著降低对标注数据量的需求。域适应技术则通过特征分布对齐使模型适应不同场景,增强泛化能力。尽管这些方法已展现出显著效果,但现有研究多集中于简单场景迁移,对动态多变的真实驾驶环境适应性仍显不足。因此,深入探索迁移学习和域适应技术在视觉疲劳检测中的应用,特别是针对复杂驾驶场景的优化,将成为未来研究的重要方向。

为弥补图像中全局面部特征的不足,疲劳检测通常采用多粒度特征学习方法^[12-14]。这类方法先通过全脸图像学习全局特征,再提取预定义局部区域特征,最后融合进行识别。但这种方法存在以下问题:多分支并行处理导致内存占用高、计算量大;局部与全局特征分布不一致可能引入噪声,简单的特征融合效果有限。

近年来,Transformer 凭借其全局建模能力在计算机视觉领域表现出色,而卷积神经网络(Convolutional Neural Network, CNN)则擅长局部特征提取。因此,采用 CNN-Transformer 混合架构能同时捕捉局部细节和全局特征,有效提升特征表达能力。这种混合策略在降低计算复杂度的同时,避免了传统多粒度方法存在的特征融合噪声问题。

本文提出一种基于双分支融合与分段域适应迁移学习的疲劳驾驶检测方法。通过车载摄像头获取驾驶员面部图像,构建双分支 CNN-Transformer 网络提取局部与全局特征,并设计双向融合模块增强特征表达能力。针对模型适应性问题,提出分段域适应策略,有效提升模型在不同驾驶环境下的检测性能。本文的主要贡献如下:

1) 提出一种双分支融合与分段域适应的迁移学习框架,采用双分支特征提取与融合结构增强特征表征,并结合域适应技术提升模型准确性和鲁棒性,实现稳定可靠的状态监测。

2) 设计了双向特征融合的模块。动态融合 SwinTransformer 的全局上下文特征和 ResNeXt 的局部细节特征,提高了模型的多层次信息表征能力。

3) 提出分段域适应策略。在特征提取阶段, CNN 分支使

用对抗域适应, Transformer 分支采用 MK-MMD 方法;在特征融合阶段,结合 MK-MMD 和最小类别混淆损失(Minimum Class Confusion, MCC),有效增强了模型的跨域泛化能力。

2 相关工作

基于计算机视觉的疲劳检测方法通过分析图像和视频中的驾驶员面部特征及行为模式进行疲劳检测,已成为该领域的重要研究方向。本章将对基于计算机视觉的驾驶员疲劳检测方法进行系统总结。

传统的疲劳检测方法通过提取空间特征来检测疲劳状态。Mandal 等^[15]通过测量眼睑间距和眼睛闭合百分比来识别疲劳症状。Catalbas 等^[16]以眼球运动的监测结果为指标,对疲劳状态进行评估。Hariri 等^[17]用嘴部特征和打哈欠动作作为衡量疲劳的指标,并应用改进版 Viola-Jones 目标检测算法,针对面部及嘴部特征进行疲劳监测。Nakamura 等^[18]通过面部对齐技术,结合 KNN 算法来评估疲劳水平。然而,这些方法受限于手工特征提取的局限性,未能充分适应驾驶环境中的多变性。在实际驾驶场景中,更理想的解决方案是使模型能够自主学习并提取面部特征,以此提升任务执行的效率与准确性。

随着深度学习的快速发展,基于 CNN 的深度学习模型在计算机视觉任务中取得了突破性进展^[19]。在疲劳检测中, CNN 广泛用于空间域特征提取。它通过多层神经网络结构自动学习和分析图像数据,精准提取有效特征,无需人工手动设计特征提取算法。Li 等^[20]提出了一种轻量级 CNN 疲劳驾驶检测系统,结合眼部、嘴部及头部姿态多模态特征分析,实现边缘设备上的实时精准监测。Li 等^[8]以关键点坐标为输入建立面部倒立摆模型,以倒立摆模型计算各疲劳特征的信息熵值,基于疲劳特征信息熵值来检测疲劳驾驶。Zhao 等^[21]提出了一种基于眼睛和嘴巴状态的疲劳检测算法,该算法使用单发多框检测器算法(SSD)提取眼睛和嘴巴特征区域,并使用 VGG-16 网络来学习疲劳特征。Zhao 等^[22]提出了一种基于残差的深度三维卷积神经网络(3D-CNN),用于从驾驶员面部图像序列中提取空间和时间特征;通过循环神经网络对生成的概率向量进行分类,实现疲劳检测。Alameen 等^[23]提出一种集成了 3D-CNN 和长短期记忆的时空网络检测模型,该模型可以从相邻帧中提取时空特征,分析长序列,进行时间建模。相较于手工特征提取,深度学习模型的特征提取方法更具有灵活性。

为了充分捕获人脸的特征信息进行疲劳检测,传统方法采用多粒度网络进行特征学习^[12-14]。Celona 等^[12]提出了一种基于视觉的多任务驾驶员监控框架(MTDMF),该框架可同时分析驾驶员的头部姿势、眼睛和嘴巴状态,从而检测睡眠程度。Huang 等^[13]提出了一种多粒度深度卷积模型(RF-DCM),创新性地融合了局部和全局人脸特征,通过特征重校准机制有效克服了姿态变化的影响,显著提升了疲劳检测的鲁棒性。Lyu 等^[14]提出了一种基于面部多特征融合的驾驶员疲劳检测方法,通过多粒度面部特征提取、空间特征融合网络和 LSTM 时间特征提取构建时空特征模型。在多粒度网络中,虽然不同粒度的特征提供了多层次的判别信息,但现有模

型未对这些特征赋予差异化权重,限制了特征的深入学习和优化。此外,来自多流网络的特征在分布和表示上的一致性,可能会影响特征融合的效果。

近年来,Transformer 在计算机视觉领域取得重要进展。Dosovitskiy^[24]提出的 visual Transformer(ViT)通过图像分块和自注意力机制,有效捕捉全局上下文信息;Peng 等^[25]进一步提出 Conformer 混合架构,结合 CNN 的局部特征提取和 Transformer 的全局建模优势。受此启发,本文采用 CNN-Transformer 混合架构,协同提取面部局部细节和全局特征,提升疲劳检测的特征表示能力。

实际驾驶场景中,摄像头角度和驾驶员姿态变化导致的面部特征差异,会引发训练数据与实际数据间的特征偏移,从而影响模型性能。虽然现有域适应方法能缓解域间分布差

异^[26-27],但普遍存在两个局限:一是未能有效结合局部细节与全局特征,二是对面部特征的特异性利用不足。为此,需要开发更精细的特征融合策略,通过优化跨域特征对齐方式来提升疲劳检测的准确性。

3 基于双分支融合与分段域适应迁移学习的疲劳驾驶检测

本章将给出基于双分支融合与分段域适应迁移学习的疲劳驾驶检测方法的详细设计过程。

3.1 框架概述

本文提出了一种基于迁移学习的疲劳驾驶检测框架,旨在解决传统方法在特征提取和域适应方面的不足,提升疲劳驾驶监测的精度与鲁棒性。本文方法总体流程如图 1 所示。

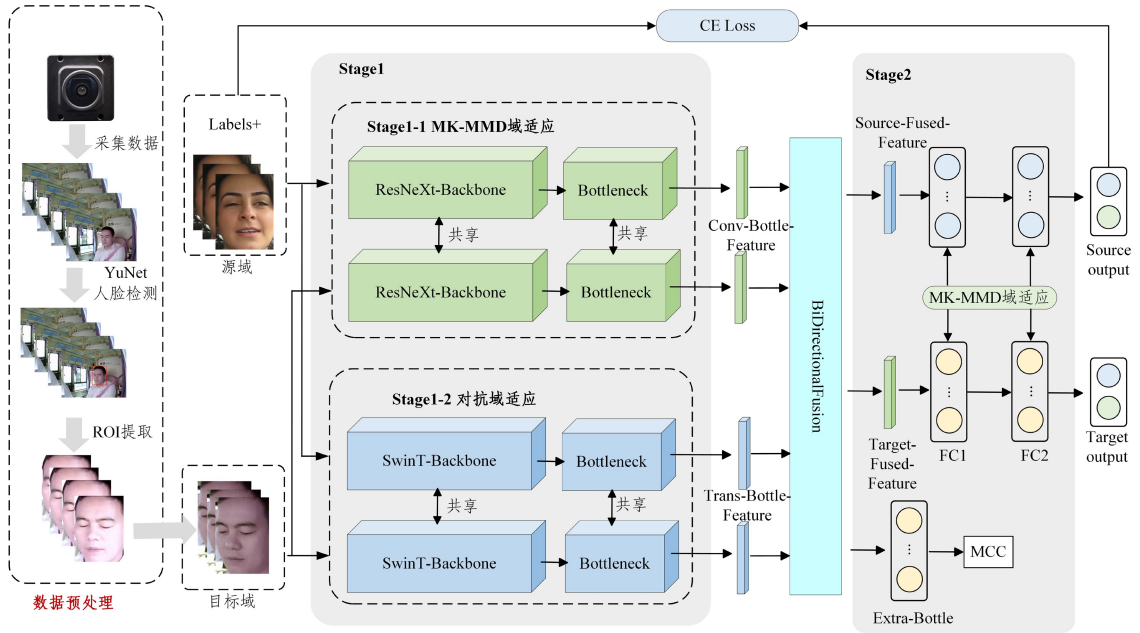


图 1 基于双分支融合与分段域适应迁移学习的疲劳驾驶检测的模型架构

Fig. 1 Model architecture of fatigue driving detection based on dual-branch fusion and segmented domain adaptation transfer learning

在数据预处理阶段,使用 YuNet 模型对车载终端采集的驾驶员视频数据进行快速人脸区域检测,高效定位脸部区域并提取特征,为后续模型分析和疲劳驾驶状态检测提供输入数据。

本文提出了一种基于双分支融合和分段域适应的迁移学习架构(Swin-Rest Adaptation Network, SRAN),用于疲劳驾驶检测。该方法通过双分支特征提取与融合架构以及分段域适应策略,提升目标域数据集上的疲劳驾驶检测准确度。

在双分支架构设计中,通过 CNN 和 Transformer 分别提取局部与全局特征,并采用 Bidirectional 模块进行特征融合。针对跨域差异问题,创新性地引入分段域适应策略:在特征提取阶段,CNN 分支使用 MK-MMD 方法对齐域分布,Transformer 分支采用对抗训练增强适应性;在特征融合阶段,结合 MK-MMD 和 MCC 损失优化特征映射。通过这些创新架构和策略,SRAN 能够在复杂驾驶环境中实现高精度的疲劳驾驶检测,显著提升了模型的准确性和稳定性。

3.2 基于 CNN 与 Transformer 的双分支特征提取与融合

SRAN 架构通过并行部署 CNN 和 Transformer 两种深

度学习结构(见图 1 中间部分),旨在同时捕获图像中的局部空间特征和全局空间特征,以此提升模型对驾驶员疲劳状态监测的准确性。在主干网络设计中,CNN 分支和 Transformer 分支分别采用 ResNeXt 网络和 Swin Transformer(SwinT)网络,这两种网络分别负责提取局部和全局特征。随后,通过双向融合模块对两个分支提取的特征进行信息融合,增强特征表示能力。此外,为了增强模型对多样化驾驶环境的适应能力,引入域适应策略提升跨域泛化性能。域适应策略的具体细节将在 3.3 节中详细阐述。

CNN 结构擅长捕捉局部空间特征,特别是在静态图像中的物体识别和细节分析方面具有显著优势。在本文中,CNN 特征提取分支采用了 ResNeXt 与 Bottleneck 结合的架构。ResNeXt 通过引入分组卷积操作,增强了模型的表达能力和计算效率。其结构在处理图像时能够有效减少参数量,同时保持较高的准确性,适合用于提取驾驶员面部区域和其他静态特征。Bottleneck 通过压缩和扩展操作,有效减少了计算量,同时保持了较强的特征表达能力。这一设计使得模型在进行大规模图像数据处理时,能够高效提取到关键的空间信

息,从而为驾驶疲劳状态的判断提供重要依据。

Transformer 结构在捕捉长程依赖和全局信息方面具有无可比拟的优势,尤其在时序数据的处理上展现出强大的能力。本文采用了 Swin Transformer(SwinT)与 Bottleneck 作为特征提取分支。SwinT 是近年来在视觉任务中表现优异的 Transformer 变种^[28],它通过分层的局部窗口和跨窗口自注意力机制,使其在处理大规模图像时不仅能够捕捉全局信息,还能有效捕捉图像局部的细节,从而提高对驾驶员疲劳状态的识别能力。Bottleneck 不再赘述。

为了进一步提升模型的表现,本文设计了双向融合模块(Bidirectional Fusion Module, BFM),用于动态融合 Swin-Transformer 的全局特征和 ResNeXt 的局部特征。双分支网络分别提取了局部特征和全局特征,而如何有效融合这些特征是接下来的关键。BFM 模块接收 ResNeXt 和 SwinT 分支的特征,并在多个维度上进行加权融合,确保特征的互补性和协同效应。图 2 展示了该模块的结构。其融合过程可表示为:

$$\text{Concat}(\sigma(\text{Trans-Feature} * \text{Kernel}), \sigma(\text{Conv-Feature} * \text{Kernel}))$$

其中,Concat 表示通道维度拼接操作, σ 表示 ReLU 激活函数,* 表示卷积运算。通过拼接操作,将两种不同卷积路径提

取的特征进行融合,从而获得更丰富的特征表示。这种融合策略不仅保留了原始特征的多样性,而且通过结合两种特征提取器的优势,提高了特征的表达能力。

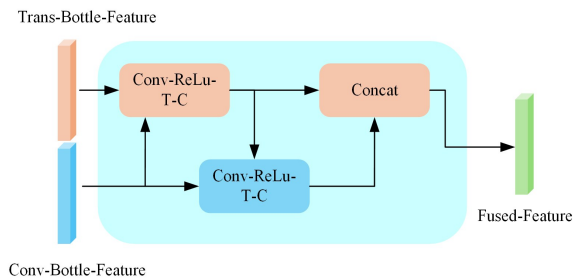


图 2 双向融合模块

Fig. 2 Bidirectional fusion module

3.3 基于迁移学习的疲劳驾驶检测模型分段域适应训练

本文设计了一种分段域适应训练策略,分别在低层特征和高层特征上采取不同的适应方法,旨在充分发挥分支网络中 ResNeXt 和 SwinT 各自的优势以及融合特征的优势,以实现更有效的目标域适应。

3.3.1 模型训练与分段域适应策略

在第一阶段自适应训练中,CNN 分支网络使用 MK-MMD 域适应方法进行模型训练,如图 3 上部分所示。

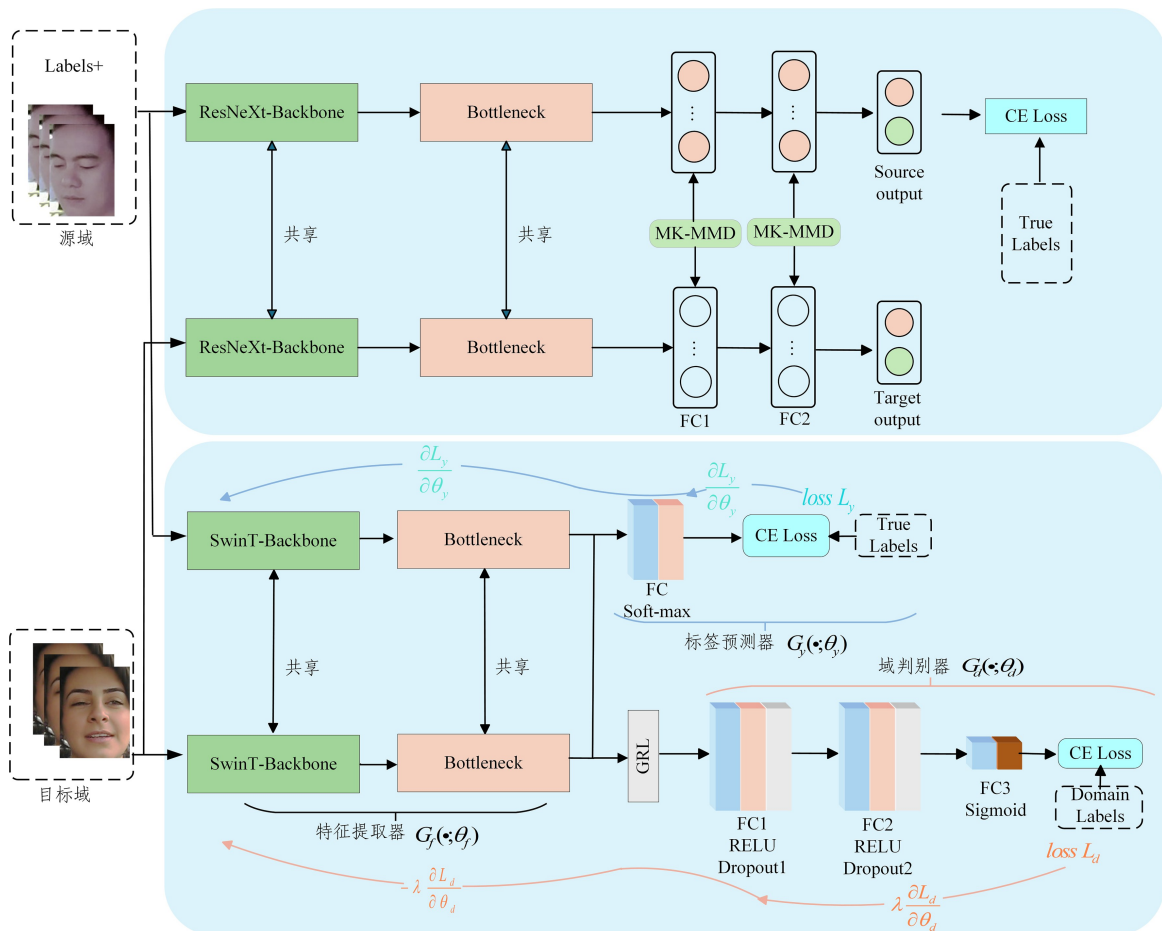


图 3 第一阶段域适应训练

Fig. 3 First stage of domain adaptation training

模型的输入包括源域和目标域输入,两者通过共享的 ResNeXt 骨干网络进行特征提取。为实现跨域局部特征的

分布匹配,采用多核最大均值差异(MK-MMD)域适应方法(详见 3.3.2 小节)。MK-MMD 通过核空间映射来计算源域

与目标域特征分布的差异,在目标函数中引入正则项:

$$MK-MMD^2 \triangleq \left\| \frac{1}{n_s} \sum_{i=1}^{n_s} \phi(x_s^i) - \frac{1}{n_t} \sum_{j=1}^{n_t} \phi(x_t^j) \right\|_{\mathcal{H}_k}^2 \quad (1)$$

其中, $\phi(\cdot)$ 表示复合核函数映射, n_s 和 n_t 分别是源域和目标域的样本数量, x_s^i 和 x_t^j 代表源域和目标域中的样本, \mathcal{H}_k 为具有特征核 k 的再生核希尔伯特空间(RKHS)。通过最小化 MK-MMD, 可以实现源域和目标域特征分布的对齐。Transformer 分支使用对抗域适应方法进行模型训练, 如图 3 下部分所示, 模型的输入同样分为源域和目标域输入。在特征提取的基础上, 引入域对抗域适应(详见 3.3.3 节)。对抗域适应利用梯度反转层(Gradient Reversal Layer, GRL)实现跨域特征对齐。在训练过程中, GRL 会反转域判别器的梯度, 促使特征提取器学习域不变特征。通过这种对抗机制, 模型能有效缩小域间差异, 提升目标域上的泛化性能。

第二阶段, 通过双向特征融合模块实现特征融合, 之后进一步采用 MK-MMD, 并引入最小类别混淆损失(Minimum Category Confusion, MCC), 以提升分类置信度。MCC 通过抑制目标域预测矩阵的非对角线元素来实现。

$$\mathcal{L}_{MCC} = \frac{1}{C^2 - C} \sum_{r \neq z} \left(\frac{1}{n_r} \sum p_r(x_t) p_z(x_t) \right)^2 \quad (2)$$

其中, C 是标签类别(疲劳和非疲劳)数量, 即 $C=2$; $p_r(x_t)$ 是模型对第 r 类的预测概率。通过最小化 MCC, 可以减少类别间的混淆, 提高模型分类准确性。

3.3.2 MK-MMD 域适应

多核最大均值差异(MK-MMD)^[29]通过显式最小化源域和目标域的特征分布差异来增强特征迁移能力。该方法在适配层计算两域间的 MK-MMD 距离, 并将其作为损失函数进行优化。当损失收敛至预设阈值时, 即可实现有效的跨域特征对齐, 从而提升目标域上的预测性能。设 \mathcal{H}_k 为具有特征核 k 的再生核希尔伯特空间(RKHS), 分布 P 在 \mathcal{H}_k 中的均值嵌入 $\mu_k(P)$ 是一个唯一的元素, 使得对于所有 $f \in \mathcal{H}_k$, 有 $E_{x \sim P} f(x) = \langle f(x), \mu_k(P) \rangle_{\mathcal{H}_k}$ 。概率分布 P 和 Q 之间的 MK-MMD $d_k(P, Q)$ 被定义为 P 和 Q 的均值嵌入之间的 RKHS 距离。MK-MMD 的平方形式定义为:

$$d_k^2(P, Q) \triangleq \| E_P[\phi(x_s)] - E_Q[\phi(x_t)] \|_{\mathcal{H}_k}^2 \quad (3)$$

其中, x_s 和 x_t 分别代表源域和目标域抽取的样本。最重要的性质是, 当且仅当 $d_k^2(P, Q) = 0, P = Q$ 。在实际计算中, $\phi(\cdot)$ 为核函数 \mathcal{K} 的隐式映射, $\| \cdot \|_{\mathcal{H}_k}^2$ 可通过核矩阵计算。

$$d_k^2 = \frac{1}{n_s^2} \sum_{i,j=1}^{n_s} k(x_s^i, x_s^j) + \frac{1}{n_t^2} \sum_{i,j=1}^{n_t} k(x_t^i, x_t^j) - \frac{2}{n_s n_t} \sum_{i=1}^{n_s} \sum_{j=1}^{n_t} k(x_s^i, x_t^j) \quad (4)$$

与特征映射 ϕ 相关联的特征核 $k(x_s, x_t) = \langle \phi(x_s), \phi(x_t) \rangle$ 被定义为 m 个正定(PSD)核 $\{k_u\}$ 的凸组合, 即:

$$\mathcal{K} \triangleq \{k = \sum_{u=1}^m \beta_u k_u : \sum_{u=1}^m \beta_u = 1, \beta_u \geq 0, \forall u\} \quad (5)$$

优化后的目标函数表示为:

$$\min_{\theta} \frac{1}{n_s} \sum_{i=1}^{n_s} J(\theta(x_s^i), y_s^i) + \lambda_1 \sum_{l=1}^{l_2} d_k^2(\mathcal{D}_s^l, \mathcal{D}_t^l) \quad (6)$$

其中, θ 表示模型参数集合; $\frac{1}{n_s} \sum_{i=1}^{n_s} J(\theta(x_s^i), y_s^i)$ 是训练误差项, 它计算了模型在源数据集 $\{x_s^i, y_s^i\}$ 上的平均损失; n_s 是源域样本数量; J 是交叉熵损失函数, 它测量了模型预测 $\theta(x_s^i)$ 与真实标签 y_s^i 之间的差异; λ_1 是正则化参数, 本文参考了文献 [29] 中的参数设置方案, 将 λ_1 权衡系数设为 0.3; $\sum_{l=1}^{l_2} d_k^2(\mathcal{D}_s^l, \mathcal{D}_t^l)$ 是正则化项, 它测量了不同层次 l 上的源域分布 \mathcal{D}_s^l 和目标域分布 \mathcal{D}_t^l 之间的距离, l_1 和 l_2 是应用 MK-MMD 正则化的层索引, 以图 1 为例, 即为 FC1 和 FC2 的层索引。

3.3.3 对抗迁移域适应

域对抗迁移网络由特征提取器、标签预测器和域判别器组成。特征提取器将输入数据映射到一个共享特征空间, 该空间既要便于标签预测器对源域数据分类, 又要使域判别器难以区分数据来自源域还是目标域。标签预测器负责对源域数据进行分类, 而域判别器则尝试判断特征的来源域。特征提取器和标签预测器构成一个前馈神经网络, 域判别器通过梯度反转层(GRL)与特征提取器相连。训练时, 网络同时优化两个目标: 1) 最小化源域数据的分类误差; 2) 通过 GRL 使两个域的特征分布尽可能相似。这种对抗训练实现了源域向目标域的知识迁移, 从而提升了目标域的分类性能。

特征提取器为 $G_f(\cdot; \theta_f)$, 其参数为 θ_f ; 预测器为 $G_y(\cdot; \theta_y)$, 其参数为 θ_y 。给定一个源样本 (x_i, y_i) (输入样本 x_i , 对应的标签 y_i), 负对数似然作为损失函数, 其标签预测器的损失为:

$$\mathcal{L}_y(G_y(G_f(x_i; \theta_f); \theta_y), y_i) = \log \frac{1}{G_y(G_f(x_i; \theta_f); \theta_y)_{y_i}} \quad (7)$$

因此在源域上的优化目标为:

$$\min_{\theta_f, \theta_y} \left[\frac{1}{n_s} \sum_{i=1}^{n_s} \mathcal{L}_y^i(\theta_f, \theta_y) + \lambda_2 \cdot R(\theta_f) \right] \quad (8)$$

其中, $\mathcal{L}_y^i(\theta_f, \theta_y) = \mathcal{L}_y(G_y(G_f(x_i; \theta_f); \theta_y), y_i)$ 表示第 i 个样本的标签预测损失表示为 $R(\theta_f)$ 是一个可选的正则化器; λ_2 是人为设置的正则化参数, $\lambda_2 \cdot R(\theta_f)$ 用于防止神经网络过拟合。

域判别器 $G_d(\cdot; \theta_d)$ 的损失为(负对数似然作为损失函数):

$$\mathcal{L}_d(G_d(G_f(x_i; \theta_f); \theta_d), d_i) = d_i \log \frac{1}{G_d(G_f(x_i; \theta_f); \theta_d)} + (1-d_i) \log \frac{1}{1-G_d(G_f(x_i; \theta_f); \theta_d)} \quad (9)$$

其中, d_i 表示第 i 个样本的域标签, 用于标识样本属于源数据集还是目标数据集。

源域数据带标签, 目标域数据无标签。模型训练时利用源域标签学习, 最终预测目标域标签。为此, 可在目标函数(式(9))中引入领域适应项, 得到以下正则化项:

$$R(\theta_f) = \max_{\theta_d} \left[-\frac{1}{n_s} \sum_{i=1}^{n_s} \mathcal{L}_d^i(\theta_f, \theta_d) - \frac{1}{n_t} \sum_{i=n_s+1}^{n_s+n_t} \mathcal{L}_d^i(\theta_f, \theta_d) \right] \quad (10)$$

其中, $\mathcal{L}_d^i(\theta_f, \theta_d) = \mathcal{L}_d(G_d(G_f(x_i; \theta_f); \theta_d), d_i)$ 。

优化后的总目标函数为:

$$E(\theta_f, \theta_y, \theta_d) = \frac{1}{n_s} \sum_{i=1}^{n_s} \mathcal{L}_y^i(\theta_f, \theta_y) - \lambda_2 \left(\frac{1}{n_s} \sum_{i=1}^{n_s} \mathcal{L}_d^i(\theta_f, \theta_d) + \frac{1}{n_t} \sum_{i=n_s+1}^{n_s+n_t} \mathcal{L}_d^i(\theta_f, \theta_d) \right) \quad (11)$$

本文参考对抗域适应领域经典研究^[30]的建议范围 $[0.1, 1.0]$,将权衡系数 λ_2 设为0.4。针对疲劳检测任务中常见的显著域差异(如实验室环境与真实驾驶场景在光照、姿态等方面的差异),适度提高 λ_2 值有助于增强域不变特征的学习。 λ_2 值稍高于基准值0.3,既保证了足够的域对齐效果,又避免了对源域分类性能的过度影响。

4 实验与结果分析

4.1 数据集

为了验证本文提出的基于双分支融合与分段域适应迁移学习的疲劳驾驶检测方法的有效性,在两个差异显著的图像数据集(A数据集和B数据集)上进行对比实验。这两个数据集在摄像头角度、驾驶员姿态和脸部特征等方面存在明显差异(见图4),为验证双分支融合网络的特征提取与融合能力以及分段域适应策略的有效性提供了理想的测试环境。此外,为了测试方法在复杂光照条件下的鲁棒性,还构建了一个低照度环境下的疲劳检测数据集(C数据集,见图5)。

A数据集^[31]为国外数据集,样本采集于小轿车内,以西方和中东地区的人脸为主。该数据集共包含1448个样本,涵盖了多样化的驾驶场景和个体特征。由于其采集环境和样本来源的多样性,A数据集在数据分布上表现出较高的异质性。

B数据集^[32]为国内数据集,样本采集于公交车内,实验对象均为黄种人面孔。该数据集专注于公交车驾驶员的图像数据,仅包含3名公交车司机,样本数量为2931个。B数据集在数据分布上表现出较高的同质性。

C数据集为自建数据集,样本为低照度环境下的人脸图像,共132张图像。

实验过程在3个迁移方向上进行。A→B方向:源域为A数据集、目标域为B数据集。同理,还有B→A和A→C方向。这3个方向上的迁移学习旨在评估模型在不同数据分布下的适应性和泛化能力。



(a) A数据集



(b) B数据集

图4 A和B数据集部分样本

Fig. 4 Partial samples of A and B datasets

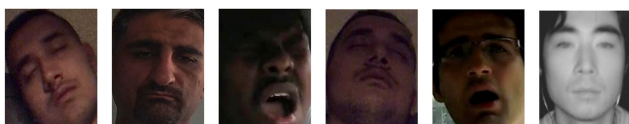


图5 C数据集部分样本

Fig. 5 Partial samples of C dataset

4.2 实验设置

本文实验平台配置如下:搭载 Intel® Core™ i7-10700K CPU,主频为3.80 GHz,配备32 GB内存,并配置了 NVIDIA GeForce RTX 3060显卡,显存为12 GB。操作系统为 Windows 11。利用 PyTorch 框架构建和训练模型,并采用了自适应矩估计优化器 Adam 进行优化。初始学习率设定为0.001,权重衰减为0.0005。评价指标准确率(ACC)的描述如下:

$$ACC = \frac{TP + TN}{TP + FP + FN + TN} \quad (12)$$

其中,TP为真正例,FP为假正例,FN为假负例,TN为真负例。

4.3 实验结果分析

4.3.1 双分支特征提取与融合可视化分析

本文实验使用 Grad-CAM 生成了 ResNeXt 分支、SwinT 分支和双分支融合网络的脸部热力激活图(见图6)。结果显示,ResNeXt 分支关注局部区域,主要聚焦脸部小范围;SwinT 分支关注区域更广,能捕捉全局信息。融合后的热力图显示,模型对关键区域的关注显著增强,红色区域增多且更集中于眼部和嘴部等重要区域。这表明双分支融合整合了局部和全局特征,生成了更全面、精确的特征表示,提升了关键面部特征的捕获能力,为疲劳检测提供了更具鲁棒性的特征支持。

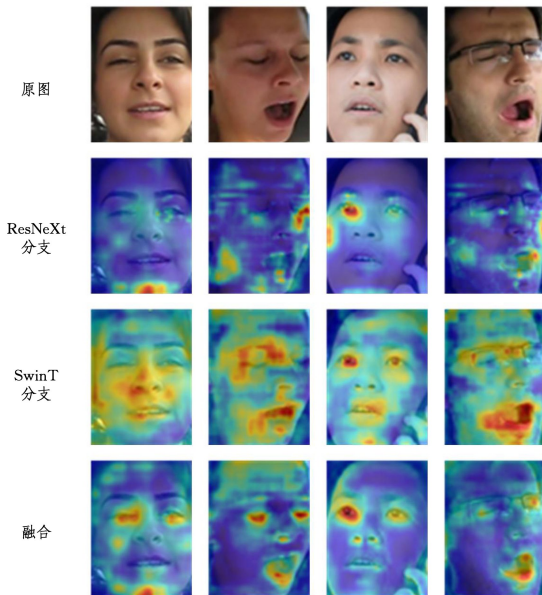


图6 ResNeXt分支、SwinT分支和融合后的热力图对比
(电子版为彩图)

Fig. 6 Comparison of heat maps of ResNeXt branch, SwinT branch and the fused one

本文采用 t-SNE 对分支网络和融合网络进行可视化分析,如图7(a)和图7(b)所示。相比于图7(a)和图7(b)的特征分布,图7(c)中疲劳和非疲劳两类特征的分布更加清晰,聚类更加紧密,且两类特征之间的重叠较少。SwinT-Rest 双通道融合模型在特征区分上表现出更优的性能。这可能是由于单一模型在特征提取上的局限性,而本文融合模型通过结合两种模型的优势,实现了特征的互补和增强,从而获得了更好的特征区分能力。

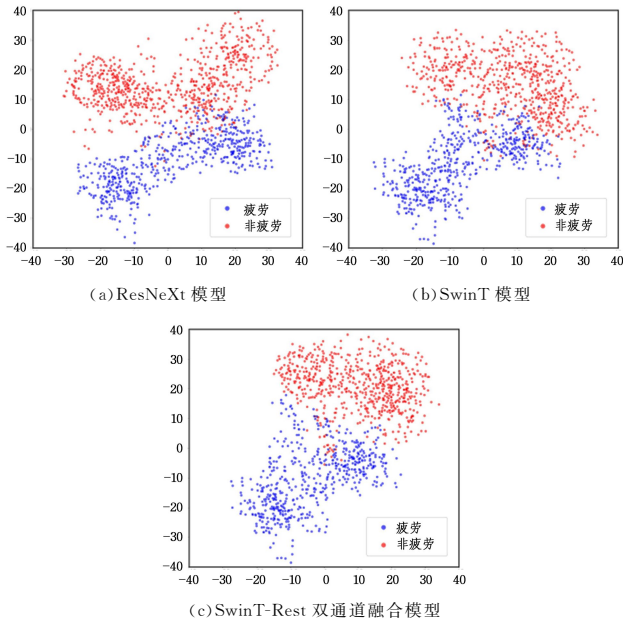


图7 使用 t-SNE 对分支网络和融合网络可视化

Fig. 7 Visualize branching and fusion networks using t-SNE

4.3.2 对比实验

将本文方法与其他域适应方法进行对比,对比方法包括 DAN^[29], DANN^[30], CDAN^[33], MADA^[34], SENTRY^[35]。DAN 通过最小化源域和目标域在再生核希尔伯特空间 (RKHS) 中的特征分布差异来实现跨域迁移。DANN 通过引入领域判别器和梯度反转层 (GRL) 来最小化源域和目标域特征分布差异。CDAN 通过引入条件对抗训练,同时考虑特征和预测标签的分布差异,进一步提高模型在目标域上的适应性和分类性能。MADA 通过多个领域判别器对不同类别特征进行细粒度对齐,有效解决跨域数据中的类别不匹配问题。SENTRY 通过随机变换检查一致性,选择性优化目标域实例的预测熵,从而提升目标域分类性能。

图 8 对比了各域适应模型的性能表现。在 A→B 和 B→A 两个任务方向上,SRAN 分别取得 93.3% 和 75.1% 的准确率,显著优于其他模型。特别是在 A→B 方向上,SRAN 的领先优势明显。此外,在低照度条件下(绿色线),本文方法展现出更强的适应能力,性能明显优于现有主流方法,体现了其优异的泛化能力。

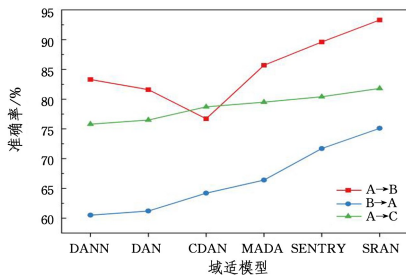


图8 域适应模型在疲劳检测数据集上的性能对比

(电子版为彩图)

Fig. 8 Performance comparison of domain adaptation models on fatigue detection datasets

学习方向(A→B 和 B→A 方向)中的准确率表现。纵轴表示准确率,横轴为模型类型和采用的域适应方法,括号中为所使用的域适应方法,包括 ResNet18 (MK-MMD),Dinov2 (MK-MMD),DeiT (MK-MMD),ResNeXt (MK-MMD),ResNeXt (Adversarial),SwinT (MK-MMD),SwinT (Adversarial),以及本文提出的 SRAN。

在 A→B 方向中,本文方法的准确率为 93.3%,高于其他方法的结果。SwinT (Adversarial) 和 ResNeXt (MK-MMD) 紧随其后,分别达到 90.7% 和 91.6%。其他方法的准确率则为 66.6%~88.3%。在 B→A 方向中,各方法的准确率整体有所下降,但 SRAN 仍以 75.1% 的准确率领先,SwinT (Adversarial) 以 72.3% 紧随其后,其他模型的准确率为 64.5%~72.1%。

SRAN 在两个方向中均表现出色,尤其在 A→B 方向中显著领先。这表明 SRAN 能够有效应对源域与目标域之间的特征偏移,展现了其在复杂迁移学习方向中的强大适应能力。使用对抗方法的 SwinT (Adversarial) 在两个方向中表现稳定,尤其在 B→A 方向中效果明显;而 ResNeXt 在使用对抗方法时,在 A→B 方向上的效果欠佳。尽管 MK-MMD 方法在 ResNeXt, Dinov2 等模型中表现不错,尤其是 ResNeXt (MK-MMD) 方法在两个方向上的效果都很优异。但在 B→A 方向上,SwinT (MK-MMD) 方法的效果不佳。这可能与 MK-MMD 方法在处理某些特定模型时的适配性不足有关。

可以看出,SRAN 方法在不同方向上展现了良好的适应性和稳定性,证明了其在迁移学习中的有效性。对比图 8 和图 9 中 A→B 和 B→A 方向的性能曲线可以看出,源域数据多样性对目标域性能有显著影响。尽管数据集 A 的样本数量少于数据集 B,但 A 数据集因人物多样性更丰富,模型在 A→B 方向表现更好。这表明在迁移学习中,源域数据的多样性比数据规模更重要,其有助于模型学习更广泛的特征,提升对目标域的泛化能力。

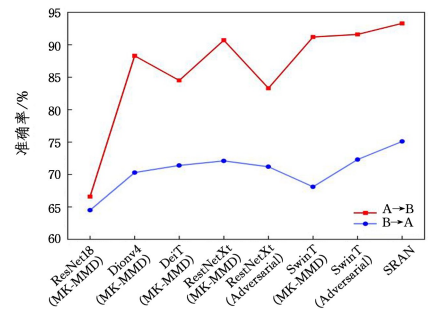


图9 不同深度模型在不同域适应方法上的性能对比

Fig. 9 Performance comparison of different deep models on different domain adaptation methods

4.3.3 域适应消融实验

为了比较不同域适应策略的有效性,进行了消融实验,结果如表 1 所列。在 A→B 任务中,stage1 的两个子模块 stage1-1 和 stage1-2 均能有效提升模型性能,而完整的 stage1 策略(stage1-1+stage1-2)在 A→B 和 B→A 任务上分别达到

92.2%和73.9%,优于单独使用任一子模块的性能,表明二者具有协同作用。此外,stage2策略单独使用时也展现出优于Baseline的性能(A→B:91.7%,B→A:73.4%)。最重要的是,stage1+stage2的联合策略在两个任务上均取得最佳性能(A→B:93.3%,B→A:75.1%),相比Baseline分别提升了2.7个百分点和4.3个百分点,充分验证了多阶段域适应策略的有效性和模块间的互补性。

表1 域适应消融实验

Table 1 Ablation experiment of domain adaptation

域适应策略	A→B	B→A
Baseline	90.6	70.8
stage1-1	91.5	72.7
stage1-2	91.8	73.1
stage1	92.2	73.9
stage2	91.7	73.4
stage1+stage2	93.3	75.1

采用 t-SNE 方法对网络中不同点的特征分布进行可视化,通过颜色标注区分源域和目标域的示例。具体而言,黄色点代表目标域(数据集 B)特征分布,紫色点代表源域(数据集 A)分布。源域和目标域未域适应与适应后的特征分布分别如图 10 和图 11 所示。图 10 显示,未经域适应的源域和目标域特征分布存在明显分离,表明存在显著域间差异。经域适应调整后,两域特征分布明显收敛,差异显著减小。这一改进证明域适应方法能有效缩小域间差异,提升模型在目标域的泛化性能,从而优化迁移学习效果。

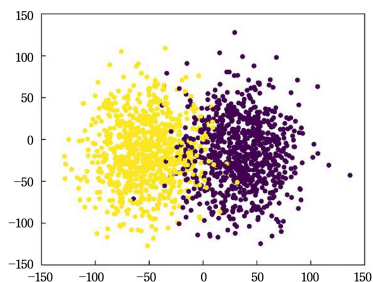


图 10 源域和目标域未域适应的特征分布(电子版为彩图)

Fig. 10 Feature distributions of the source domain and the target domain in the unsupervised domain adaptation

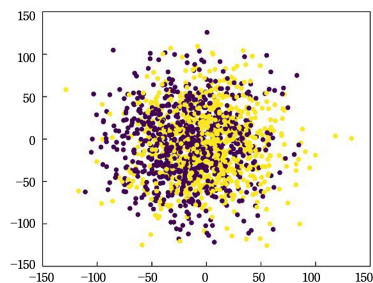


图 11 源域和目标域域适应后的特征分布(电子版为彩图)

Fig. 11 Feature distributions of the source domain and the target domain after domain adaptation

接着,分别在采用分段域适应策略和未采用域适应的情况下,对比了模型在目标域上的准确率,结果如图 12 所示。其中,红蓝线表示 A→B 方向的准确率变化,绿紫线表示 B→

A 方向的准确率变化。训练初期,采用域适应和未采用域适应的模型的准确率相似,表明域适应在初始阶段对性能的提升不明显。但随着迭代次数的增加,采用域适应的模型的准确率显著高于未采用域适应的模型。在 B→A 方向,迭代次数超过 60 后,未采用域适应的模型的准确率明显下降,可能是由于在源域过拟合,而采用域适应的模型的准确率未出现明显下降。这表明域适应策略能有效缩小源域和目标域之间的差异,提升模型在目标域的泛化能力,从而提高疲劳驾驶检测的准确率。

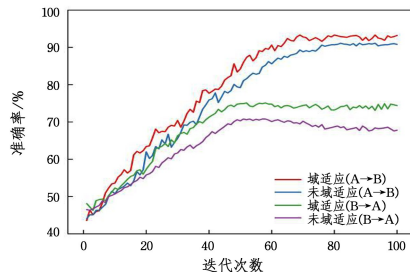


图 12 SwinT-ResNeXt 并联网路采用域适应和未采用域适应的性能对比(电子版为彩图)

Fig. 12 Performance comparison of SwinT-ResNeXt parallel network with and without domain adaptation

最后,对比了两阶段域适应策略和单阶段域适应策略在目标域上的准确率表现,结果如图 13 所示。图中上方 3 条曲线表示 A→B 方向,下方 3 条曲线表示 B→A 方向。随着迭代次数增加,所有策略的准确率均上升。在初期,两阶段和单阶段域适应策略表现相似,但随着迭代进行,两阶段域适应策略的准确率逐渐超过其他策略。总体来看,两阶段域适应策略的准确率始终高于单阶段策略。这表明两阶段域适应策略能更有效地缩小源域与目标域之间的差异,提升模型在目标域的泛化能力和性能。这一发现进一步证明了多阶段域适应策略在迁移学习中的潜力,为相关研究提供了有力支持。

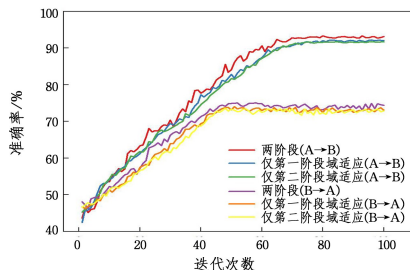


图 13 模型采用两阶段域适应与采用一阶段域适应在目标域上的准确率对比

Fig. 13 Comparison of accuracy on the target domain between the model adopting two-stage domain adaptation and that adopting one-stage domain adaptation

结束语 总体而言,本文 SRAN 模型通过双分支网络架构有效融合全局和局部特征,结合对抗域适应和 MK-MMD 方法,显著提升了跨域性能。实验表明,该模型在 A→B 等任务中表现优异,能有效缓解特征偏移问题,增强模型泛化能力。特别是在源域和目标域差异较大的场景下,SRAN 展现

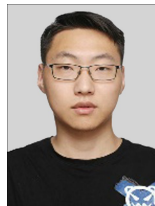
出卓越的适应性和稳定性,为域适应问题提供了可靠的解决方案。

本文方法在处理源域和目标域特征分布偏移方面表现出色,但在理论上,基于 MK-MMD 的域适应方法在非一对一映射关系下可能存在局限性,导致特征空间歧义并影响模型性能。针对此问题,后续研究将从以下方向改进:一是引入细粒度特征对齐机制(如类别感知方法)优化特征映射;二是结合注意力机制区分疲劳相关特征和域相关特征;三是建立更完善的评估体系,量化非一对一映射对性能的影响。此外,本文尚未考虑标签分布差异的挑战。未来研究将引入开域域适应技术,探讨目标域和源域在特征和标签分布上均存在差异的情况,以提升模型在复杂场景下的适应性和鲁棒性。

参考文献

- [1] AHMED S K, MOHAMMED M G, ABDULQADIR S O. Road traffic accidental injuries and deaths: A neglected global health issue[J]. *Health Science Reports*, 2023, 6(5): 1240.
- [2] AN J, CAI Q, SUN X, et al. Attention-based cross-frequency graph convolutional network for driver fatigue estimation[J]. *Cognitive Neurodynamics*, 2024, 18(5): 3181-3194.
- [3] JIA H J, XIAO Z J, JI P. End-to-end fatigue driving EEG signal detection model based on improved temporal-graph convolution network[J]. *Computers in Biology and Medicine*, 2023, 152: 106431.
- [4] PENG B, ZHANG Y, WANG M, et al. TA-MFFNet: Multi-feature fusion network for EEG analysis and driving fatigue detection based on time domain network and attention network[J]. *Computational Biology and Chemistry*, 2023, 104: 107863.
- [5] CAI S X, DU C K, ZHOU S Y, et al. Detection of Fatigue Driving State Based on Vehicle Operating Data[J]. *Transportation Systems Engineering and Information*, 2020, 20(4): 77.
- [6] CHEN L W, CHEN H M. Driver Behavior Monitoring and Warning With Dangerous Driving Detection Based on the Internet of Vehicles[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 22(11): 7232-7241.
- [7] XU H Z, HAO D S, XU X T, et al. Expressway small object detection algorithm based on deep learning[J]. *Journal of Jilin University(Engineering and Technology Edition)*, 2025, 55(6): 2003-2014.
- [8] LI T Z, ZHANG T C, LI C, et al. Driver Fatigue Detection Based on Facial Inverted Pendulum Model and Information Entropy [J]. *Transportation Systems Engineering and Information*, 2023, 23(5): 24.
- [9] CHEN L, WEI Z. Research on railway dispatcher fatigue detection method based on deep learning with multi-feature fusion [J]. *Electronics*, 2023, 12(10): 2303.
- [10] LIU M Z, XU X, HU J, et al. Real time detection of driver fatigue based on CNN-LSTM[J]. *IET Image Processing*, 2022, 16(2): 576-595.
- [11] YI Y, ZHOU Z, ZHANG W, et al. Fatigue detection algorithm based on eye multi feature fusion[J]. *IEEE Sensors Journal*, 2023, 23(7): 7949-7955.
- [12] CELONA L, MAMMANA L, BIANCO S, et al. A multi-task CNN framework for driver face monitoring[C]// 2018 IEEE 8th International Conference on Consumer Electronics-Berlin(ICCE-Berlin). IEEE, 2018: 1-4.
- [13] HUANG R, WANG Y, LI Z, et al. RF-DCM: Multi-granularity deep convolutional model based on feature recalibration and fusion for driver fatigue detection[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2020, 23(1): 630-40.
- [14] LYU J, YUAN Z, CHEN D. Long-term multi-granularity deep framework for driver drowsiness detection [J]. arXiv: 1801.02325, 2018.
- [15] MANDAL B, LI L, WANG G S, et al. Towards detection of bus driver fatigue based on robust visual analysis of eye state[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2016, 18(3): 545-557.
- [16] CATALBAS M C, CEGOVNIK T, SODNIK J, et al. Driver fatigue detection based on saccadic eye movements[C]// 2017 10th International Conference on Electrical and Electronics Engineering(ELECO). IEEE, 2017: 913-917.
- [17] HARIRI B, ABTAHI S, SHIRMOHAMMADI S, et al. A yawning measurement method to detect driver drowsiness [C] // IEEE International Instrumentation and Measurement Technology Conference. 2012.
- [18] NAKAMURA T, MAEJIMA A, MORISHIMA S. Detection of Driver's drowsy facial expression[C]// 2013 2nd IAPR Asian Conference on Pattern Recognition. IEEE, 2013: 749-753.
- [19] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 770-778.
- [20] LI X, CAO L, ZHANG G, et al. Driver fatigue detection based on convolutional neural network and face alignment for edge computing device [C] // Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering. 2021: 2699-711.
- [21] ZHAO G, HE Y, YANG H, et al. Research on fatigue detection based on visual features [J]. *IET Image Processing*. 2022, 16(4): 1044-53.
- [22] ZHAO L, WANG Z, ZHANG G, et al. Driver drowsiness recognition via transferred deep 3D convolutional network and state probability vector[J]. *Multimedia Tools and Applications*, 2020, 79(35): 26683-26701.
- [23] ALAMEEN S A, ALHOTHALI A M. A Lightweight Driver Drowsiness Detection System Using 3DCNN With LSTM[J]. *Computer Systems Science & Engineering*, 2023, 44(1): 895-912.
- [24] DOSOVITSKIY A. An image is worth 16×16 words: Transformers for image recognition at scale [J]. arXiv: 2010.11929, 2020.
- [25] PENG Z, HUANG W, GU S, et al. Conformer: Local features coupling global representations for visual recognition[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 367-376.
- [26] WANG M, DENG W. Deep visual domain adaptation: A survey [J]. *Neurocomputing*, 2018, 312: 135-153.

- [27] TENG S H, HUANG L L, ZHANG W. Dual Strategies and Confidence-based Domain Adaptation Learning[J]. Journal of Chinese Computer Systems. 2025, 46(5):1135-1146.
- [28] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 10012-10022.
- [29] LONG M, CAO Y, WANG J, et al. Learning transferable features with deep adaptation networks[C]// International Conference on Machine Learning. PMLR, 2015: 97-105.
- [30] GANIN Y, USTINOVA E, AJAKAN H, et al. Domain-adversarial training of neural networks[J]. Journal of Machine Learning Research, 2016, 17(59):1-35.
- [31] Driver-Drowsiness-Detection-Gk0ws-Jii6r Dataset[EB/OL]. <https://gitee.com/lzq320147/driver-drowsiness-detection-gk0ws-jii6r>.
- [32] Drowsy-Driving-Det2[EB/OL]. <https://gitee.com/lzq320147/drowsy-driving-det2>.
- [33] LONG M, CAO Z, WANG J, et al. Conditional adversarial domain adaptation[C]// Proceedings of the 32nd International Conference on Neural Information Processing Systems. 2018: 1647-1657.
- [34] PEI Z, CAO Z, LONG M, et al. Multi-adversarial domain adaptation[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2018.
- [35] PRABHU V, KHARE S, KARTIK D, et al. Sentry: Selective entropy optimization via committee consistency for unsupervised domain adaptation[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 8558-8567.



LI Zequn, born in 1997, postgraduate. His main research interest is intelligent assisted driving systems.



DING Fei, born in 1981, Ph.D, professor, is a senior member of CCF (No. H2242S). His main research interests include crowd and participatory sensing, cyber-physical systems.

(责任编辑:何杨)