

基于迁移学习的跨对象手语手势识别方法

王天然, 王琦, 王青山

引用本文

王天然, 王琦, 王青山. 基于迁移学习的跨对象手语手势识别方法[J]. 计算机科学, 2023, 50(6A): 220300232-5.

WANG Tianran, WANG Qi, WANG Qingshan. [Transfer Learning Based Cross-object Sign Language Gesture Recognition Method](#) [J]. Computer Science, 2023, 50(6A): 220300232-5.

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)

[面向目标识别的特征融合模糊模型及其应用](#)

Fusion Multi-feature Fuzzy Model for Target Recognition and Its Application

计算机科学, 2023, 50(6A): 220100138-7. <https://doi.org/10.11896/jsjcx.220100138>

[基于机器视觉的驾驶员视野盲区安全预警方法研究](#)

Study on Safety Warning Method of Driver's Blind Area Based on Machine Vision

计算机科学, 2023, 50(6A): 220700141-7. <https://doi.org/10.11896/jsjcx.220700141>

[基于双门控-残差特征融合的跨模态图文检索](#)

Dual Gating-Residual Feature Fusion for Image-Text Cross-modal Retrieval

计算机科学, 2023, 50(6A): 220700030-7. <https://doi.org/10.11896/jsjcx.220700030>

[基于改进Yolov4-tiny的轻量型目标检测算法](#)

Lightweight Target Detection Algorithm Based on Improved Yolov4-tiny

计算机科学, 2023, 50(6A): 220700006-7. <https://doi.org/10.11896/jsjcx.220700006>

[注意力特征融合的双生网络目标跟踪方法](#)

Attentional Feature Fusion Approach for Siamese Network Based Object Tracking

计算机科学, 2023, 50(6A): 220300237-9. <https://doi.org/10.11896/jsjcx.220300237>

基于迁移学习的跨对象手语手势识别方法

王天然 王 琦 王青山
合肥工业大学数学学院 合肥 230009
(wtrmath@mail.hfut.edu.cn)

摘 要 手语是听障人士重要的交流工具,准确识别手语可以减少健全人和听障人士之间的交流障碍。一般深度学习识别模型的性能高度依赖于所采集的数据,这导致模型跨对象泛化能力较差。因此,通过迁移学习的方法设计一种具有跨对象泛化能力的手语手势识别模型。首先,使用特征提取器融合表面肌电流(Surface Electromyography,sEMG)信号和惯性传感器(Inertial Measurement Unit,IMU)信号。然后,提出一种域对抗训练方法,其可以仅依靠源域数据完成特征提取器和域分类器的对抗训练,实现特征提取从源域到目标域的迁移。最后,在手势分类器中利用域不变特征实现手语手势跨对象识别,提高了模型的泛化能力。实验表明,在包含 200 种手语手势共 60 000 条手语样本数据集上,所提模型可将手语跨对象识别准确率提高到 85.1%。

关键词:手语手势识别;特征融合;域对抗;迁移学习;特征迁移
中图法分类号 TP391

Transfer Learning Based Cross-object Sign Language Gesture Recognition Method

WANG Tianran,WANG Qi and WANG Qingshan
School of Mathematics,Hefei University of Technology,Hefei 230009,China

Abstract Sign language is an important communication tool for hearing impaired people,and accurate recognition of sign language can reduce the communication barrier between able-bodied and hearing impaired people. The performance of general deep learning recognition models is highly dependent on the collected data,which leads to poor cross-object generalization ability of the models. Therefore, this paper designs a sign language gesture recognition model with cross-object generalization capability through a transfer learning approach. Firstly,a feature extractor is used to fuse the surface electromyography (sEMG) signal and the inertial measurement unit (IMU) signal. Then,a domain adversarial training method is proposed,which can complete the adversarial training of the feature extractor and domain classifier by relying on the source domain data only,and realize the migration of feature extraction from the source domain to the target domain. Finally,domain-invariant features are used in the gesture classifier to achieve sign language gesture cross-object recognition,which improves the generalization ability of the model in this paper. Experiments show that the proposed model can improve the accuracy of sign language cross-object recognition to 85.1% on a dataset containing 200 sign language gestures with a total of 60 000 sign language samples.

Keywords Sign language gesture recognition,Feature fusion,Domain adversarial,Transfer learning,Feature transfer

1 引言

截至 2021 年 4 月 1 日,全世界有 15 亿人患有不同程度的听力损失^[1],手语是他们非常重要的交流方式。然而,手语通常不能被健全人所理解,这成为听障人士融入社会的巨大障碍之一。近年来人工智能技术的快速发展,给解决这一问题带来机会。

最初,研究人员从原始信号中手动构建特征。Qi 等^[2]用主成分分析法从表面肌电流信号中提取手势动作的关键信息,减少了肌电信号中的冗余信息,并用广义回归神经网络(General Regression Neural Network,GRNN)识别手势,提高了手势识别的精度,减少了识别时间。Kim 等^[3]用动态时间规整(Dynamic Time Warping,DTW)实现惯性传感器信号的匹配,并改进了受限库伦能量(Restricted Coulomb Energy,RCE)网络,将 RCE 网络的度量替换为 DTW 距离,解决了

动态时间规整计算时间过长的问题。Joao 等^[4]使用滑动窗口从时序 sEMG 和 IMU 信号中抽取手势特征,并用遗传算法识别动态手势和静态手势的差别。Jian 等^[5]使用基于信息增益的特征选择方法,选择特定通道的 sEMG 和 IMU 信号,评估多种分类算法后,用支持向量机器做分类器,搭建了手语识别模型,可以识别 80 个美国手语词。手动提取特征的识别方法虽然可以识别手势动作,但是特征提取的不灵活性导致模型泛化性较差。

近年来,研究人员直接搭建端到端的深度学习模型,可以自适应提取特征来识别手势。Añazco 等^[6]用六轴 IMU 传感器采集手臂运动数据,并搭建端到端的循环神经网络模型,在 5 个用户的 3 种手势数据集上达到 95.7% 的识别准确率。Hu 等^[7]提出了一种混合卷积循环网络架构,其能够用深度学习方法从多通道稀疏 sEMG 信号中识别通道之间的相关性,较机器学习方法将识别准确率提高了 9.2%。他们的

通信作者:王琦(wangq@hfut.edu.cn)

工作表明了深度学习模型比手动提取特征模型有更好的泛化性能。

但是,深度学习在手势识别领域的应用高度依赖样本的质量和数量,而收集大量手语手势样本较为困难,这些导致手势识别模型缺乏足够的跨使用对象泛化能力。目前有研究者开始用迁移学习(Transfer Learning, TL)^[8]的方法来提高模型泛化能力。Shin 等^[9]使用基于样本的迁移方法实现韩国手语手势识别系统,先用聚类算法将所有用户聚类,然后对每一类用户分别训练各自的识别模型。Chen 等^[10]用基于参数的迁移学习方法在源域 sEMG 数据集上进行模型预训练,然后将其部分参数迁移到目标模型中,最后用少量目标域数据微调,实现手势识别从源域到目标域的迁移。

基于样本和基于模型的迁移方法对目标域数据要求较高,当目标域样本分布与源域有较大差别时并不能顺利完成迁移。为了探索适应性更强的迁移方法,Kang 等^[11]利用基于特征的迁移成分分析(Transfer Component Analysis, TCA)方法去除手势信息与人行走状态之间的耦合,提取不受环境影响的手势特征,完成手势识别跨行走状态迁移。Yang 等^[12]通过孪生网络(Siamese Network)得到样本对的特征映射,并利用最大均值差异(Maximum Mean Discrepancy, MMD)去除手势时空特征中的环境噪声,实现手势跨对象识别。基于特征的迁移方法提取通用性更强的域不变特征,使得后续分类任务可以不受域变化的影响,能适应域差异较大的迁移场景。因此,本文用基于特征的迁移学习方法来提高手语手势识别模型的泛化能力。

本文搭建了基于特征迁移的手语手势识别模型,用域对抗的方法实现 sEMG 和 IMU 信号的特征迁移。该模型能自适应完成两种有差异信号的特征融合,并能从原始特征中去除环境干扰,提取既有手势识别能力又有域不变性的潜在特征,实现特征从源域到目标域的迁移。实验结果表明,本文提出的模型可以提高手语手势识别跨对象的泛化能力。

2 模型结构

本节给出跨对象识别问题的数学描述,并提出基于迁移学习的手语手势识别模型。该模型可以自适应地解决跨对象识别分布不匹配导致的泛化性问题。最后,提出对抗训练算法,该算法可以在不依赖目标域数据的情况下,提高手语手势识别的泛化能力。

2.1 问题描述

首先给出手语跨使用对象识别的问题描述。令 $X \in R^k$ 为手语信号输入空间, k 表示输入特征维数, $Y = \{1, 2, 3, \dots, N\}$ 为手语手势标签集合, $\tilde{C} = \{1, 2, \dots, M\}$ 为多源域标签集合,其中 N 和 M 分别表示手语手势的个数和手语使用对象的个数。事实上,在实时手语识别过程中,目标域数据的域信息异于已知的源域信息。为了区分目标域和源域,令目标域的域标签为 0。因此,域标签集合为 $C = \tilde{C} \cup \{0\}$,进而 $X \times Y \times C$ 构成了手语识别数据集。令 D_S 和 D_T 是分别定义在源域集合 $\{(x, y, i) | x \in X, y \in Y, i \in \tilde{C}\}$ 和目标域集合 $\{(x, y, 0) | x \in X, y \in Y\}$ 上的分布。跨使用对象手语识别的目标就是在含有限对象的源域上训练,能够得到在有新对象的目标域上具有良好泛化性能的手语识别模型。跨域目标损失描述为:

$$R_T(\eta) = \frac{P}{(x, y) \sim T} (\eta(x) \neq y) \quad (1)$$

其中, $\eta: X \rightarrow Y$ 是目标分类器, $\frac{P}{(x, y) \sim T} (\eta(x) \neq y)$ 表示目标分类器结果 $\eta(x)$ 和手势标签 y 不一致的概率。

Ben-David^[13]证明了当源域数据 $S \sim (D_S)^n$ 和目标域数据 $T \sim (D_T)^n$ 以比例 $\delta: 1-\delta$ 分布时,式(1)中跨域目标损失的上界为:

$$R_T(\eta) \leq R_S(\eta) + \sqrt{\frac{4}{n} \left(d \log \frac{2n}{d} + \log \frac{4}{\delta} \right)} + \lambda + \hat{d}_H(S, T) + 4 \sqrt{\frac{1}{n} \left(d \log \frac{2n}{d} + \log \frac{4}{\delta} \right)} \quad (2)$$

其中, $\lambda \geq \inf_{\eta^* \in \mathcal{H}} [R_S(\eta^*) + R_T(\eta^*)]$; $\hat{d}_H(S, T)$ 为两个数据集 S 和 T 上的 H-散度; $\mathcal{H}$ 为 VC 维为 d 的潜在函数空间;

$$R_S(\eta) = \frac{1}{n} \sum_{i=1}^n I(\eta(x_i) \neq y_i) \quad (3)$$

是源域的经验损失; $I(\eta(x_i) \neq y_i)$ 为示性函数。

从式(2)可以看出,分类器 $\eta^* \in \mathcal{H}$ 在源域和目标域的风险都需要处于较低水平时,跨域目标损失较小。从式(3)还可以看出,只有当任务分类器既在源域有较高的识别能力,又能使得 H-散度足够低时,才能获得比较低的目标损失。而找到一种不能区分源域和目标域的特征表示可以降低两个域之间的 H-散度。本文提出的训练方法就受此启发,通过减少多源域特征表示中的域信息,可以使特征表示对域没有区分能力,进而降低跨对象手语手势识别任务的损失。

2.2 域对抗手语识别网络

本节提出域对抗手语手势识别模型,如图 1 所示,包括特征提取器、手势分类器和域分类器。其中,特征提取器和手势分类器组成了一个完整的前馈网络用于手势识别,而对抗训练部分由特征提取器和域分类器组成。

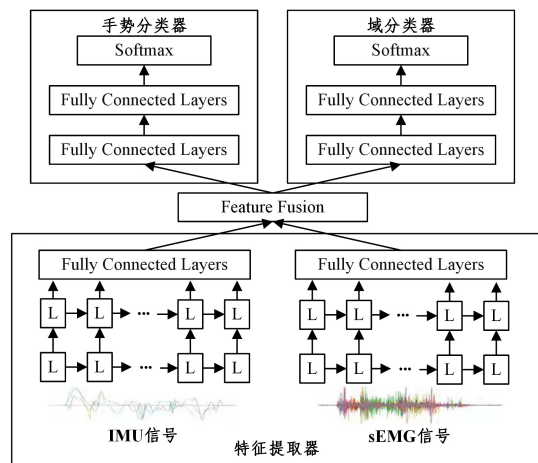


图 1 迁移学习手语手势识别模型结构

Fig. 1 Structure of transfer learning sign language gesture recognition model

首先,特征提取器将接收手势信号作为输入,然后输出相应的潜在特征表达。经过对抗训练后,输出的特征表达将对域变化有不变性,同时又保留对手语手势的识别能力。

中国手语的表达与时序相关,如果要从手语手势的动作中准确识别手势含义,手势信号的时间顺序至关重要。另外,手势动作长短不同又要求特征提取能解决循环网络长期依赖的问题。因此,本文选择长短期记忆^[14](Long Short-Term Memory, LSTM)神经网络从手语手势中提取信息。它既可以综合考虑整段序列的所有信息来提取特征,加入遗忘门也能自适应解决长期依赖的问题。

特征提取器使用两层 LSTM 网络学习时序信号中的关联信息,并用两层全连接降维,最终得到 m 维潜在特征表示。这个过程可以表示为:

$$f = F(x) \in \mathbb{R}^m$$

其中, f 表示潜在特征, $F(\cdot)$ 表示特征提取器, $\mathbb{R}^m$ 为 m 维特征空间。

其次,为了度量目标分布与特征表示条件分布之间的差异,用 KL 散度^[15]来度量两个分布之间的统计距离。可以证明^[16]:

$$D_{KL}(Q \parallel P) = -H(Q, Q) + H(Q, P) \quad (4)$$

其中, $D_{KL}(\cdot)$ 表示 KL 散度, $H(\cdot)$ 表示交叉熵, Q 和 P 表示两个不同的分布。 $H(Q, Q)$ 只与分布 Q 本身有关, KL 散度 $D_{KL}(Q \parallel P)$ 将退化为交叉熵 $H(Q, P)$, 因此可以直接用交叉熵来度量两个分布的差异。

最后,令 $P_d(c|F(x))$ 和 $P_l(y|F(x))$ 分别表示样本关于域和手势的条件分布:

$$P_d(c|F(x)) = \text{softmax}(DC(F(x)))$$

$$P_l(y|F(x)) = \text{softmax}(SC(F(x)))$$

其中, $DC(\cdot)$ 表示域分类器, $SC(\cdot)$ 表示手势分类器, $\text{softmax}(\cdot)$ 为 softmax 激活函数。

2.3 域对抗训练方法

为了提高任务分类器在目标域的泛化能力,希望潜在特征表示能具有域不变性。因此,在对抗训练过程中,首先进行正向训练,最小化域分类器输出和域标签之间的统计距离,以使域分类器有识别域的能力。然后进行反向训练,最大化域分类器输出和域标签之间的差异,使特征表示不含域信息,见图 2。在对抗训练过程中,正向训练和反向训练交替进行共同优化,它们之间的比例对模型效果的影响会在后文进一步研究。

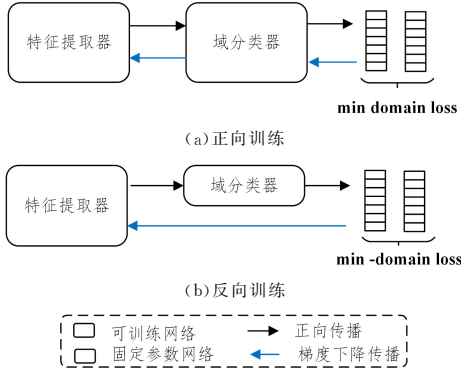


图 2 对抗训练流程图

Fig. 2 Adversarial training process

(1) 正向训练

更新特征提取器和域分类器的参数,最小化模型输出和域标签的交叉熵,使特征提取器和域分类器组成的域识别网络有对域的识别能力,见图 2(a)。令 X_s 是源域数据, C_s 是相应的域标签,这个过程可以表示为:

$$\min_{\theta_F, \theta_D} L_d(X_s, C_s; \theta_F, \theta_D)$$

$$L_d(X_s, C_s; \theta_F, \theta_D)$$

$$= E_{(x, c) \sim (X_s, C_s)} \sum_{k=1}^M c_k \times \log P_d(\hat{c} | F(x; \theta_F); \theta_D)$$

其中, $c = (c_1, c_2, \dots, c_M)$ 表示样本的域标签向量, $L_d(\cdot)$ 表示域分类器的交叉熵损失, θ_F 和 θ_D 分别表示特征提取器和域分类器的网络参数。

(2) 反向训练

反向训练固定域分类器的参数,只更新特征提取器 F , 见图 2(b)。这时域分类器已经有了一定识别域的能力,最大化特征条件分布和样本域标签分布之间的交叉熵,可以增大两个分布之间的统计距离。优化过程可表示为:

$$\max_{\theta_F} L_d(X_s; \theta_F)$$

$$L_d(X_s; \theta_F) = E_{(x, c) \sim (X_s, C_s)} \sum_{k=1}^M c_k \times \log P_d(\hat{c} | F(x; \theta_F))$$

2.4 手势识别训练

域对抗迁移训练可以使得网络提取的特征表示 f 不能区分域,降低跨域识别目标损失。使用不含域信息的特征表示完成手语手势识别,最小化手势分类器网络输出和真实手势分布的交叉熵。令 Y 为源域手语手势标签集合,优化目标可以表示为:

$$\min_{\theta_F, \theta_L} L_l(X_s, Y; \theta_F, \theta_L)$$

$$L_l(X_s, Y; \theta_F, \theta_L)$$

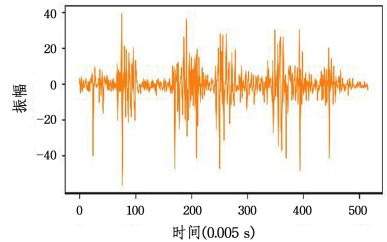
$$= E_{(x, y) \sim (X_s, Y)} \sum_{k=1}^N y_k \times \log P_l(\hat{y} | F(x; \theta_F); \theta_L)$$

其中, $y = (y_1, y_2, \dots, y_N)$ 表示样本手势标签, θ_L 表示手势分类器网络参数。

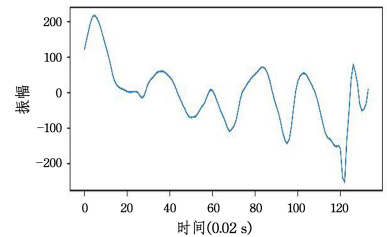
3 实验

本节在采集的数据集上评估所提出的手语识别模型的识别准确率和泛化性能。使用带有表面肌电流传感器和惯性传感器的手环作为手语手势信号采集装置,它可以同时采集手腕表面肌电流、手部运动加速度和旋转角速度等序列信号。其中,加速度和旋转角速度为采集频率 50 Hz 的 10 维数据,反映手臂的三维运动和旋转。表面肌电信号为采集频率 200 Hz 的 8 维数据,反映手腕上 8 个位置的表面肌电流的变化。采集的数据可表示为 $[x_{imu}, x_{emg}] = x \in X$, 其中 $x_{imu} \in \mathbb{R}^{10 \times T_1}$ 为物理运动信号, $x_{emg} \in \mathbb{R}^{8 \times T_2}$ 为肌电流信号。 T_1, T_2 分别为两种信号的帧数。

从听障人士日常使用需要出发,本文选取了 200 个常用手语词作为实验的识别目标,15 位志愿者作为手语使用对象采集数据。每位志愿者重复每个手语词 20 遍,共采集到 60000 条手语手势数据。手语词“名片”的 sEMG 信号和 IMU 信号如图 3 所示。



(a) sEMG 信号



(b) IMU 信号

图 3 “名片”的信号图

Fig. 3 Signals of “business card”

3.1 有对象依赖的手语手势识别实验

本部分对每个志愿者的每个手语词数据都保留 80% 作为训练集, 剩余 20% 为测试集, 来验证所提模型对手语手势的识别能力。同时, 排除手语使用对象对实验的影响, 不考虑模型跨对象的泛化性能, 保证每个志愿者的数据都同时出现在训练集和测试集中。为了证明本文模型特征提取部分的有效性, 将该部分的 LSTM 网络用多层感知机 (Multilayer Perceptron, MLP) 和 CNN 两种网络结构替换做了对比, 实验结果如表 1 所列。

表 1 对象依赖实验准确率统计指标表

Table 1 Statistical indicators of accuracy of object-dependent experiments

(单位: %)				
模型	均值	标准差	最大值	最小值
MLP	73.8	4.3	81.5	67.5
CNN	80.4	5.0	88.7	76.6
LSTM	90.1	3.5	95.1	85.8

本文提出的基于 LSTM 的手语翻译模型识别准确率均值达到 90.1%, 且比另外两个模型更稳定。原因是 LSTM 不但可以保留前序信号中的有用信息, 其遗忘机制又能很好地处理长期依赖导致的梯度爆炸和梯度消失问题。

3.2 双通道特征融合实验

表面肌电流和惯性运动信号对手势动作的描述重点不同, 本实验研究两种信号各自描述手势动作的差别以及本文模型特征融合的效果。选择日常问候、交通、就餐购物、医疗和休闲旅游等 5 个常见的对话场景, 研究两种信号在各种生活场景中的识别效果, 每个场景选取 20 个手语词, 部分手语词如表 2 所列, 实验结果如图 4 所示。从图 4 可以看出, 不同场景下两种通道的识别效果不同, 在医疗场景单惯性信号的识别效果好于肌电流信号, 在其他场景肌电流信号优势更大。值得注意的是, 在所有场景中, 特征融合模型都取得了比单通道模型更高的识别准确率, 说明本文模型所用的特征融合方式可以有效结合两种通道的信息, 有从多个维度获取手势动作信息的能力。

表 2 分场景手语词表

Table 2 Scene sign language words

场景	手语词条
日常问候	名片, 天气, 保暖, 平安, 顺利
交通	飞机, 轮船, 汽车, 身份证
就餐购物	菜谱, 烤鸭, 帽子, 大衣
医疗	挂号, 外科, 头疼, 咳嗽, 血压
休闲旅游	空调, 热水, 登记表, 前台

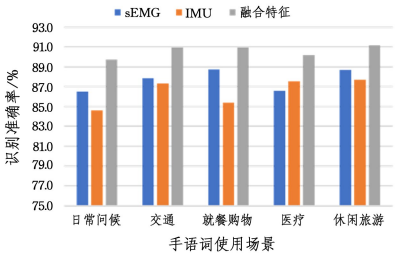


图 4 各信号识别效果对比

Fig. 4 Recognition effect comparison of each signal

3.3 对象独立的手语手势识别实验

本节研究提出的迁移学习方法的有效性。为了研究本文

模型手语识别跨人的泛化能力, 将采用留出交叉验证^[17]的策略评估模型的跨域泛化能力。从 15 个志愿者中随机选取 5 个作为目标域的迁移对象, 其他志愿者为源域对象。选择源域对象的数据为训练集, 目标域对象数据为测试集进行实验, 保证源域数据和目标域数据来自不同的手语使用对象。将本文模型与两个深度学习的迁移方法和两个机器学习的迁移方法进行对比: 域对抗神经网络^[18] (Domain Adaptive Neural Network, DANN) 是基于特征的迁移学习方法, 通过直接在网络中加入梯度反转层来消除特征的域信息, 实现特征从源域到目标域的迁移; 参数迁移学习^[10] (Parameter Transfer Learning, PTL) 是基于参数的迁移学习方法, 通过迁移全部或部分源域网络参数, 并在目标域微调, 实现域间的迁移。迁移成分分析^[19] (Transfer Component Analysis, TCA) 将源域和目标域数据映射到低维空间降维, 并拉近特征之间的距离; 联合分布适应^[20] (Joint Distribution Adaptation, JDA) 映射源域和目标域特征, 并同时适配两者的边缘分布和条件分布。

随机选取不同志愿者划分方案进行实验, 不同划分方案中各迁移方法的对比如图 5 所示, 各迁移方法识别准确率的统计指标如表 3 所列。其中, PTL no-FT 表示没有微调的参数迁移学习, 失去迁移优势, 退化为一般深度学习。

表 3 各迁移方法识别准确率统计指标

Table 3 Statistical indicators of recognition accuracy of each migration method

(单位: %)				
模型	均值	标准差	最大值	最小值
DANN	77.3	3.1	81.34	69.3
PTL	88.2	2.1	90.56	81.45
PTL no-FT	67.5	4.3	72.3	60.0
TCA	49.7	3.8	43.8	56.9
JDA	59.0	3.4	53.6	65.3
本文模型	85.1	3.1	89.9	79.1

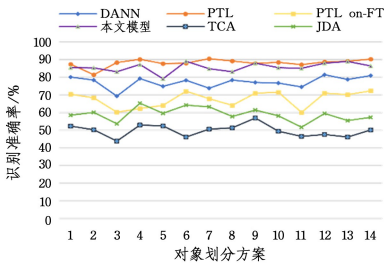


图 5 不同样本划分方案中各迁移方法的对比

Fig. 5 Comparison of migration methods in different sample division schemes

如图 5 所示, 本文提出的迁移学习模型在所有样本划分方案中都优于 DANN 和无迁移优势的 PTL no-FT 方法, 说明该模型可以有效提高手语手势跨对象识别的泛化能力。但是, 在大多数样本划分方案中, PTL 模型的识别准确率都高于本文模型。这是因为参数迁移方法将预训练模型在少量目标域数据上再次训练, 模型已经适应了目标域数据的分布, 这种要求在实时翻译场景中不容易被满足。而本文提出的模型在对抗训练过程并不需要目标域数据, 可以应用于实时识别场景, 并且在大多数样本划分方案中 PTL 模型效果相差不多, 说明对抗训练过程提高了模型跨使用对象的泛化能力。

最后, 本节还通过改变对抗训练中正向更新和反向更新

比例的实验找到两个步骤的最佳比例,结果如图 6 所示。实验结果表明,当正向训练和反向训练比例为 2:1 时模型有最好的识别效果,准确率达到 85.1%。

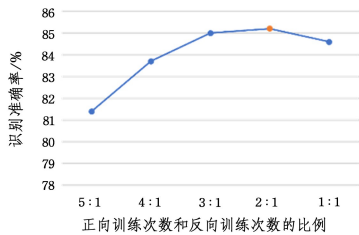


图 6 不同正反训练比例模型的准确率

Fig. 6 Accuracy of models with different positive and negative training ratios

结束语 本文提出了一种基于迁移学习的手语手势识别模型。该模型能融合表面肌电信号和运动信号的特征,提取既有手语手势识别能力,又有域不变性的潜在特征,实现了手语手势的跨对象识别,从而提高模型的泛化能力。实验证明,在多个日常对话场景中,本文模型提取的融合特征都能弥补单模态模型的缺陷,将对象依赖的手语手势平均识别准确率提高到 90.1%;本文提出的迁移学习方法,能够在缺乏目标域数据的情况下,实现源域到目标域的迁移,将跨对象手语手势识别准确率提高到 85.1%。未来工作将从手语手势动作结构出发,探究手语动作规律,进一步提高手语识别泛化能力。

参 考 文 献

[1] World Health Organization. Deafness and hearing loss [EB/OL]. (2021-04-01)[2022-03-21]. <https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss>.

[2] QI J Q,JIANG G Z,LI G ,et al. Surface EMG Hand Gesture Recognition System Based on PCA and GRNN[J]. Neural Comput Appl,2020,32(10):6343-6351.

[3] KIM M,CHO J,LEE S ,et al. IMU Sensor-based Hand Gesture Recognition for Human-machine Interfaces[J]. Sensors,2019,19(18):3827.

[4] JOAO L,MIGUEL S,NUNO M ,et al. Hand/Arm Gesture Segmentation by Motion Using IMU and EMG Sensing[C]// International Conference on Flexible Automation and Intelligent Manufacturing. Amsterdam:Elsevier,2017:107-113.

[5] WU J,SUN L,JAFARI R. A Wearable System for Recognizing American Sign Language in Real-time Using IMU and Surface EMG Sensors[J]. IEEE J. Biomed. Health Informatics,2016,20(5):1281-1290.

[6] AÑAZCO E V,HAN S J,KIM K ,et al. Hand Gesture Recognition Using Single Patchable Six-axis Inertial Measurement Unit via Recurrent Neural Networks[J]. Sensors,2021,21(4):1404.

[7] HU Y,WONG Y K,WEI W T ,et al. A Novel Attention-based Hybrid CNN-RNN Architecture for sEMG-based Gesture Recognition[J]. Plos One,2018,10(13):<https://10.1371/journal.pone.0206049>.

[8] YOSINSKI J,CLUNE J,BENGIO Y ,et al. How Transferable are Features in Deep Neural Networks[C]// Annual Conference on Neural Information Processing Systems. Montreal: NIPS,2014:3320-3328.

[9] SHIN S,BAEK Y,LEE J H ,et al. Korean Sign Language Recognition Using EMG and IMU Sensors Based on Group-dependent NN Models[C]// IEEE Symposium Series on Computational Intelligence. New York:IEEE,2017:1770-1776.

[10] CHEN X,LI Y,HU R ,et al. Hand Gesture Recognition Based on Surface Electromyography Using Convolutional Neural Network with Transfer Learning Method [J]. IEEE J Biomed Health Informatics,2021,25(4):1292-1304.

[11] YANG J,ZOU H,ZHOU Y ,et al. Learning Gestures from WiFi: A Siamese Recurrent Convolutional Architecture [J]. IEEE Internet Things J,2019,6(6):10763-10772.

[12] KANG P,LI J,FAN B ,et al. Wrist-worn Hand Gesture Recognition While Walking via Transfer Learning[J]. IEEE J Biomed Health Informatics,2022,26(3):952-961.

[13] BEN-DAVID S,BLITZER J,CRAMMER K ,et al. Analysis of Representations for Domain Adaptation [C]// Annual Conference on Neural Information Processing Systems. Montreal: NIPS,2006:137-144.

[14] ZHOU J,XU W. End-to-end Learning of Semantic Role Labeling Using Recurrent Neural Networks[C]// Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL,2015:1127-1137.

[15] GOODFELLOW L J,BENGIO Y,COURVILLE A C. Deep Learning[M]. Massachusetts:MIT Press,2016:47-49.

[16] GOODFELLOW L J,POUGET A J,MIRZA M ,et al. Generative Adversarial Nets[C]// Annual Conference on Neural Information Processing Systems. Montreal:NIPS,2014:2672-2680.

[17] LI Y,ZHENG W M,ZONG Y ,et al. A Bi-hemisphere Domain Adversarial Neural Network Model forEEG Emotion Recognition[J]. IEEE Transactions on Affective Computing,2021,12(2):494-504.

[18] YAROSLAV G,EVGENIYA U,HANA A ,et al. Domain-Adversarial Training of Neural Networks[J]. J Mach Learn Res,2016,17(59):1-35.

[19] PAN S J,TSANG I W,KWOK J S ,et al. Domain Adaptation via Transfer Component Analysis [J]. IEEE Trans Neural Networks,2011,22(2):199-210.

[20] LONG M,WANG J,DING G ,et al. Transfer Feature Learning with Joint Distribution Adaptation [C] // IEEE International Conference on Computer Vision. New York:IEEE Press,2013:2200-2207.



WANG Tianran, born in 1994, postgraduate. His main research interests include transfer learning and gesture recognition.



WANG Qi, born in 1975, Ph.D, associate professor. Her main research interests include gesture recognition and edge computing.