研究生课程考试成绩单
（试卷封面）

<table>
<tr><td>院　　系</td><td colspan="2">网络空间安全学院</td><td>专业</td><td colspan="3">网络空间安全</td></tr>
<tr><td>学生姓名</td><td colspan="2">陈根文</td><td>学号</td><td colspan="3">220235183</td></tr>
<tr><td>课程名称</td><td colspan="6">人工智能与机器学习</td></tr>
<tr><td>授课时间</td><td colspan="2">年　月至　年　月</td><td>周学时</td><td></td><td>学分</td><td></td></tr>
<tr><td>简
要
评
语</td><td colspan="6"></td></tr>
<tr><td>考核论题</td><td colspan="6"></td></tr>
<tr><td>总评成绩
（含平时成绩）</td><td colspan="6"></td></tr>
<tr><td>备注</td><td colspan="6"></td></tr>
</table>

任课教师签名：_____

日期：

注：1. 以论文或大作业为考核方式的课程必须填此表，综合考试可不填。“简要评语”栏缺填无效。

2. 任课教师填写后与试卷一起送院系研究生秘书处。

3. 学位课总评成绩以百分制计分。

东 南 大 学 试 题 纸

课程 人工智能与机器学习 2023—2024 学年第一学期

学 号 220235183 姓 名 陈根文 得 分

（本试卷共 2 页）

一、 结合兴趣、可能的毕业设计课题或读书报告，撰写一篇能够体现应用人工智能与机器学习思想或技术的论文；或设计与实现一个智能系统并写出相应技术报告（70%）。

要求：1）论文或技术报告内容应包括：题目、作者、作者联系地址、中英文摘要、主题词、正文、主要参考文献。

2）论文或技术报告正文部分字数不少于 3000 字，主要包括：问题、问题的研究背景、问题的已有解决方法及其分析、本文所提出的解决方案及比较、结论和待继续的工作等。

3）论文题目自拟，或从下列题目中选取：

a) 融合规则和深度学习的入侵检测
b) 基于粒子群算法的网络负载均衡优化
c) 基于决策树方法的网络安全风险分析
d) 基于生成对抗网络的网络对抗策略生成
e) 基于数据挖掘技术的社交网络舆情分析
f) 基于 LSTM 和注意力机制的网络态势感知
g) 引入强化学习机制的网络路由规划模型
h) 基于回答集编程的中间人攻击检测模型
i) 基于知识图谱和主动学习的智能问答模型
j) 结合知识图谱和深度学习的智能推荐
k) 知识融合中语义不一致性检测
l) 基于概率搜索的旅行商问题求解策略分析

二、 面试（30%）

附：论文框架要求

网格环境下基于信任模型的动态级调度

袁禄来　曾国荪　姜黎立　蒋昌俊
（同济大学计算机科学与工程系　上海　201804）

摘　要　网格用户、资源和服务的不确定性潜在地影响网格应用任务的正常执行，这样使得设计既能减小应用任务执行时间又能减小欺骗可能性的调度算法十分困难.参考社会学的人际关系信任模型，建立网格节点信任推荐机制，并利用 D-S 理论对推荐证据进行综合分析，从而定义出基于不确定性推理理论的信任度计算函数.将该函数并入 DLS 算法得到“可信”动态级调度算法（TDLS），从而在计算调度级别时考虑网格节点的可信程度.仿真结果证实，提出的 TDLS 算法以小的时间花费为代价，能有效提高任务在信任方面的服务质量需求.

关键词　网格计算；可信调度；信任模型；不确定性推理；DLS 算法

Dynamic Level Scheduling Based on Trust Model in Grid Computing

YUAN Lu-Lai　ZENG Guo-Sun　JIANG Li-Li　JIANG Chang-Jun
(*Department of Computer Science and Technology, Tongji University, Shanghai*　201804)

Abstract　The uncertainty of Grid users, resources and services may play a negative affect on the execution of Grid tasks, which makes it difficult to design a scheduling algorithm to minimize execution time and cheat probability of Grid tasks. Referring to the social trust relationship, the authors introduce a trust model that is based on the uncertainty reasoning theory (D-S theory). In addition, by combining the trust model and Dynamic Level Scheduling (DLS) algorithm, the authors propose a novel scheduling algorithm that is called Trustworthy and Dynamic Level Scheduling (TDLS). The algorithm takes the Grid nodes' trust degree into account when calculating the scheduling-level of task-node pairs. Simulations prove that the algorithm can efficiently satisfy the QoS requirement in trust, with costing a few more time.

Keywords　Grid computing; trustworthy scheduling; trust model; uncertainty reasoning; DLS algorithm

面向 WIFI 行为跨域识别的机器学习方法

陈根文

(东南大学网络空间安全学院江苏 211189)

摘要 近年来,WIFI 信号以其灵活性和良好的传播特性等优势,已在无线感知人体行为中得到广泛的应用.然而,目前 WIFI 信号对人体行为的感知效果受人体位置、面部朝向等领域因素影响,当这些因素改变时识别效果下降明显.如何设计出与领域因素无关的可跨域的 WIFI 行为识别方法已经成为了当前 WIFI 感知中的挑战.针对当前 WIFI 感知无法实现跨域的问题,本文提出了一种面向 WIFI 行为跨域识别的机器学习模型 WCDR.首先提取出与领域因素无关的行为特征,然后使用卷积神经网络进行学习和预测.本文使用 Widar3.0 数据集进行实验,结果表明,WCDR 模型可以实现 95%的训练识别准确率和在人体位置、面部朝向两种领域因素上进行跨域的 87.62%、71.17%的跨域识别准确率.与现有跨域模型相比较,WCDR 实现了较高的跨域识别准确率,具有较稳健的跨域识别能力.

关键词 WIFI 感知;跨域识别;机器学习;卷积神经网络

Machine Learning Methods for Cross-Domain Recognition of WIFI Behavior

CHENGen-Wen

(*SchoolofCyberScienceandEngineering,SoutheastUniversity,Jiangsu*211189)

Abstract In recent years, WIFI signals have been widely used in wireless perception of human behavior due to their flexibility and good propagation characteristics. However, the perception effect of WIFI signals on human behavior is currently influenced by factors such as human position and facial orientation, When these factors change, the recognition performance decreases significantly. How to design cross domain WIFI behavior recognition methods that are independent of domain factors has become a challenge in current WIFI perception. To address the problem that current WIFI perception cannot achieve cross domain recognition, this paper proposes a machine learning model WCDR for cross domain WIFI behavior recognition. Firstly, behavior features that are independent of domain factors are extracted, Then, convolutional neural networks were used for learning and prediction. In this paper, the Widar3.0 dataset was used for experiments, and the results showed that the WCDR model can achieve a training recognition accuracy of 95% and a cross domain recognition accuracy of 87.62% and 71.17% on the two domain factors of human position and facial orientation. Compared with existing cross domain models, WCDR achieved higher cross domain recognition accuracy and has more robust cross domain recognition abilities.

Keywords WIFI perception; cross-domain recognition; machine learning; convolutional neural networks

目录

1 引言

1.1 选题背景

随着科技的不断发展，行为识别技术已成为人工智能领域的关键研究方向，可分为基于计算机视觉和基于无线信号的两大感知技术。基于计算机视觉的方法主要依赖于视频流中的动作检测，例如摔倒检测和手势识别，但存在受光线影响、跨域问题和隐私侵犯等挑战。

基于无线信号的感知技术，尤其是基于 WIFI 信号的方法，通过分析信号传播来实现对人体行为的感知。相较于计算机视觉，这种方法不受光照条件的限制，且具备强大的抗干扰性和隐私保护能力。在识别人体行为时，WIFI 信号的干扰程度和特性与不同行为呈现相关性，通过采集并分析 WIFI 信号，机器学习模型能够实现不同干扰图像与人体行为的精准映射，从而达到对人体行为进行识别分类的目的。尤其随着深度学习的迅速发展，越来越多基于深度学习的 WIFI 信号识别模型涌现，其中神经网络模型占据核心地位。设计健壮且准确的神经网络模型成为实现 WIFI 行为跨域识别的关键要素，能够提高模型的鲁棒性和识别准确率。因此，为实现对人体行为的跨域识别，开发强大的神经网络模型至关重要。

1.2 国内外研究现状

WIFI 感知受到多种因素影响，包括工作频段、多径效应、通信带宽、天线数目、节点数目等领域因素，而这些问题均会增大 WIFI 感知模型的拟合度并降低 WIFI 感知的准确性。为了解决上述问题，国内外学者们提出了不同的模型。

在使用 WIFI 信号进行手势识别的现有模型中，ChenningLi 等人提出了 WiHF[1]模型，该模型主要解决基于 WIFI 信号的手势识别问题，是一个应用跨域进行手势识别的模型。该模型提出了一个有效的解决实时提取动作改变的算法，并设计了一个深度神经网络进行手势识别。该模型获取基于用户身份跨域的手臂手势和 WIFI 运动特征进行特征提取。系统设计包括数据获取模块、模式分析模块和协作式双任务模块。系统中包括获取信道状态信息数据，运动改变模式提取和双任务模块。其中双任务为数据适应和特征提取。具体实验中，该模型使用 Widar3.0 数据集，实现了 97.65%的总体准确率并实现了对手势位置、手势朝向和手势环境的 90.30%，79.14%和 89.67%的跨域准确率。虽然模型可以通过一组 WIFI 信号发射器和 WIFI 信号接收器来跨域识别人体的手势，但是也存在一些局限性。首先该模型仍然无法完全克服背景干扰，即当背景干扰增

强时，识别准确率迅速降低。其次该模型无法准确分析行为动态，即不定期的动态行为会为多普勒频移谱分析(DFS)带来一定困难。再次模型的分布式开销较大。并且具有阴影效应，其中交叉方向的识别准确率将比别的受影响严重。总结起来是由于特征提取算法不够成熟等种种原因，其鲁棒性和识别准确率仍有待提高。而在模型 Widar3.0[2]实现的跨域识别中，该模型通过学习无线信号的动作代表来解决具有挑战性的人类手势识别问题。模型指出，基于 WIFI 信号的手势识别可以在特殊数据域内达到较好的效果，但是不加域适应时很难用于新的目标域。为了解决上述问题，提出了零工作量的跨域手势识别模型--Widar3.0。该模型主要分为两部分：体坐标运动曲线(BVP)模块和手势识别模块。在 BVP 模块中，首先构造信道状态信息(CSI)的多普勒表示，再从多普勒频谱(DFS)中提取体坐标运动曲线(BVP)，再进行 BVP 估算，动态链路选取。识别机制包括 BVP 正则化，空间特征提取，时态建模，异常值检测。从结果上看，该模型可实现 92.7%的总体准确率，并在总体上实现了域独立，相比之下 CARM 模型不可实现跨域识别。该模型指出通过比较去噪声 CSI、DFS 和 BVP 作为输入特征，并使用 CNN-GRU 做实验，BVP 获得较高准确率。尽管如此，该模型也具有局限性，LOS 需求，即模型需要保持一条 WIFI 发射器到人体再到 WIFI 接收器的一条直接路径，否则模型可能不会正常工作。KaixuanChen[3]等人提出了“个体差异性-任务一致性”模型。这个模型考虑到人类产生的数据在半监督学习中由于人类不同的生物条件和行为模式，不断受着分布变化的影响，进而提出了一个稳健的分布式半监督学习的模型。该模型考虑标签数据和非标签数据之间的差异性和不变性，学习了减少个人特殊差异性和保留任务特殊不变性的潜在特征。该模型通过将对抗性消耗转化为参数优化来实现减少个体差异性的目的，同样的解决方式可应用于保留任务不变性的解决上。在训练环节，该模型使用一个非常强的分类器来最小化特征分布差异，但是存在分散编码器的注意力的问题。所以在优化上，该模型采用阈值来实现差异性判别性之间的最大值和最小值的平衡，并改用强大的解码器。实验中，该模型使用四种不同的数据集，与顶尖水平比较看，可实现 95%、60%、70%的准确率。总的来说，该模型提出了一种健壮的半监督来处理标签和没有标记的数据，进一步提出了减少个体差异性和保留任务一致性的解决办法。最后实验结果证明，该模型是可以用于实际应用并解决实际问题的。JianweiLiu[4]等人提出了基于梯度的无设备独立于域的手势识别模型 DI。该模型的提出背景是随着 VR 游戏和智能家居等的发展，传统的方法受域间隙的影响较大，即在某一个域下训练出的模型应用在另一个域时会具有较低的测试效果。然而通过对抗性学习、迁移学习和体坐标速度曲线(BVP)实现跨域手势识别会或多或少具有些缺陷。这个模型考虑到传统解决方案的工作原理是利用手势间隙作为手势之间的边界，但域差距会导

致边界模糊，进一步降低识别度准确性。模型定义了领域差距的概念，然后提出一个可以消除领域差距，进一步实现领域独立的手势识别的方法，可以达到更高的准确率。模型提出使用渐变映射的符号映射产生域间隙消除器用来提高识别精度。总体上 DI 模型包括两个模块，数据获取和 AH-Net 处理。从结果上看，DI 模型可实现 94.45%的识别准确率。与四种顶尖水平方法：EI、跨域感知、WiAG 和 Widar3.0 相比，DI 实现了最高的识别准确率和最少的识别限制。从总体思想来说，DI 是通过减小域差距来实现跨域手势识别，达到了比较高的识别准确率。YuanrunFang[5]等人提出一种面向 WIFI 的跨场景迁移活动识别系统。该模型的提出背景是商业 WIFI 被应用于室内定位，人类活动识别和 WIFI 想象等，然而在这些场景中，都无法实现跨域识别。为了解决这一问题，YuanrunFang 等人提出了 WiTransfer 模型，该模型是一个基于商业 WIFI 的跨域迁移感知模型。该模型试图获取信道状态信息(CSI)数据并建立一个卷积神经网络模型进行训练，最后使用这种已经训练好的模型做迁移学习，即直接跨域使用模型识别。具体方法首先是 CSI 数据预处理。通过异常值删除和过滤，数据规范化，统一输入维度。其次训练分类器，最后迁移至新领域。在具体实验中，该模型使用五种环境下的三种训练集，其中原始的 CSI 数据集可达 78.6GB 大小，并使用 MATLAB 进行 CSI 原始数据的预处理，并使用两块 NVIDIAGeForceRTX2080 显卡进行计算。从实验结果上看，WiTransfer 模型可实现三个场景下的 99%的训练准确率，在迁移学习中，可以实现 91.56%的测试准确率。整体上看，当迁移到目标域中，模型的准确率有所降低。同时，WiTransfer 模型同样也存在一些问题，该模型没有搞清究竟是什么因素在影响着 CSI 数据在跨域识别中的效果，以及如何避免识别准确率的负向变化等等问题。YueZheng[6]等人提出基于 WIFI 的零工作量跨域手势识别模型。该模型提出的背景是在 WIFI 已经普遍部署的情况下，仍然没有一种明确的算法来适应跨域识别。虽然已经提出了许多不同的算法，当出现新目标域的数据时，仍然需要大量工作量来实现在新的目标域中的再训练，这无疑限制了模型的鲁棒性。为了提高域适应能力并实现零工作量跨域识别，他们对 Widar3.0 进行了设计。该模型考虑到没有跨域功能的原始影响因素，实现了粗略跨域功能的运动跟踪，并获取跨域识别的潜在特征，吸取了以往 Widar3.0 的经验和教训。该模型大体上是一个分析 CSI 数据，提取 BVP 并构建 DFS 进行预测的过程。大体上与模型[2]一致，此处不再赘述。总的来说，该模型实现了基于 WIFI 的零工作量跨域手势识别系统来推进跨域感知并达到完全零工作量感知的效果。但是模型存在不同分布变化下准确率分布不均衡的问题。PetteriNurmi[7]等人提出了交叉感知模型，即面向跨站点和大规模的无线传感识别。该模型提出使用漫游让 WIFI 在新环境中实现感知，并且应用于更大规模的问题。为了减少感知模型训练数

据收集的开销，交叉感知模型采用在每个目标域中使用漫游模型，并使用机器学习进行离线训练。模型采用“专家混合”的方法来捕捉 WIFI 模型的输入。该模型的中心思想包括两点，高效率进行跨站点工作和解决大规模问题。其中漫游 WIFI 测量使用了一个人工神经网络(ANN)进行跨环境 WIFI 翻译。首先 ANN 捕获线性层和非线性层之间的关系，然后在训练域训练这个 ANN 以致在新环境中减少训练开销。这个 ANN 是一个多层神经网络，以 WIFI 测量作为输入，产生一个集成的目标域样本作为输出。包括一个噪声控制和训练示例配对的过程。最后进行迁移学习。在解决大规模问题方面，先获取无线信号特征，然后构建专家模型、专家选择器，最后控制计算开销。从实验结果上看，该模型可实现超过 80%的平均准确率，而专家选择器可实现 94.5%的步态识别准确率和 98.5%的手势识别准确率。该模型同样存在局限性，单链路无线设置和单个平台随环境变化微小而不是显著性的，即没有一套较好的复杂设置。SiamakYousefi[8]等人提出了基于长短期记忆网络来进行识别人体行为的模型。该模型采用 CSV 文件格式的信道状态信息(CSI)数据集，包括 7 中不同的人体行为。该模型实现行为识别分为两部分。第一部分为对原始 CSI 数据集进行 CARM 模型去噪声和特征提取，并仍以 CSV 格式存入文件。第二部分为行为识别，该模型构建了长短期记忆网络，输入层神经元 90，隐层神经元 200，输出层神经元 7，是一个将形状为时间-特征值的 CSI 数据流经过 LSTM 网络处理后分为 7 个不同的类别的过程。从实验结果上看，该模型实现了对所有种类行为的 75%的识别准确率。ZhangMenghuan[9]等人提出了一个基于 CNN 和超声传感的手势识别模型，该模型是一种非视距下的超声传感系统。该模型首先通过多普勒频移(DFS)和信道脉冲来实现对人体行为的感知，接着模型将会从得到的数据中提取特征，输入到卷积神经网络中实现识别人体发出的动作，该模型还实现了身份认证。从该模型的实验结果上看，模型实现了 96%的识别准确率。YuGu[10]等人提出了一种基于 CNN 和 LSTM 融合的信道注意力机制的手势识别模型。同时具有 CNN 和 LSTM 让模型具有了既可以在时间和空间两个维度中同时提取数据特征，这使得该模型在 Widar3.0 数据集上取得了良好的效果。XianjiaMeng[11]等人提出了面向信道状态信息(CSI)的模型 FaSee。模型的主要算法是 k-means 算法和动态时间规整(DTW)的结合，分层实现对 CSI 数据的特征识别。该模型在 9 类不同的手势上可以达到 94.75%的识别准确率。YulingYan[12]等人提出了使用自适应惯性权重的粒子群来优化神经网络中的反向传播来实现识别手势识别。从实验结果上看，该算法可以实现在自建数据集上 81.26%的识别准确率。ZhengjieWang[13]等人提出了面向 CSI 数据的深度学习手势识别模型 WiDG。该模型可以识别人体在空间中手写的 0-9 数字。WiDG 模型采用 CNN 实现了视距下 97.2%和非视距下 95.7%的识别准确率，实验证明了 CNN 可以用来处理 CSI

数据流。上述五种手势识别模型均未实现跨域的任务，都是基于同一目标域下的手势进行识别。JianchunGeng[14]等人提出了基于注意力机制的双流卷积神经网络，可以识别复杂背景下的人体手势。该模型分为两部分，从背景下提取手势，然后使用 CNN 识别手势。从实验结果上看，模型在提取手势和手势识别上相对于其他模型达到了较高的 F-score。该模型对本实验有了一定的指引，通过从复杂背景下提取手势可以完成本实验的跨域识别任务。JiangDehao[15]等人提出了使用 GAN 进行手势识别的模型 WiGAN。模型使用 GAN 提取手势特征并融合特征，然后使用支持向量机(SVM)进行分类。实验结果显示，WiGAN 可以在两个数据集上达到 98%和 95.6%的识别准确率。ZhanjunHao[16]等人提出了利用 CSI 数据的 SVM 模型 Wi-SL。模型还采用 K-means 算法结合 Begging 算法对 SVM 进行优化。实验结果显示模型可以在三种不同的场景下达到 95.8%平均识别准确率。ColliGuillermo[17]等人提出了基于肌电图(EMG)的传感器融合的手势分类模型。模型使用 MYO 臂环采集动作发出者的肌肉活动和运动数据，在 3 类不同的手势上可以实现最高 84.6%的识别准确率，使用最小二乘向量机可实现 92.9%的识别准确率。

WIFI 感知技术是一种重要的基于无线信号的感知技术，具有很大的潜在价值。目前现存模型只是在同域下实现了较高的识别准确率，在进行跨域识别时的模型效果相比同域下均相差较大。这说明了模型方法对于所有领域因素并不是公平的稳健的。因此，要研究出一套稳健的公平的有效的解决 WIFI 受领域因素变化的影响的算法是当务之急。

1.3 研究意义

在感知领域中，基于 WIFI 的感知具有非常重要的作用。基于 WIFI 的感知是无线感知技术中的一项重要的技术。相对于大多视距感知技术，基于 WIFI 的无线信号的感知技术可以突破视距感知技术下受领域因素、光线等因素影响严重的限制，保护隐私权，在感知技术中已经被广泛应用。然而，即使使用 WIFI 信号作为感知载体，也仍然存在一些问题，典型的问题为无法实现跨域识别，也即典型的例子：鱼可以在海洋背景下被识别，但在沙漠背景下却无法被识别。也就是说，无论是步态识别还是手势识别等等，受领域因素影响仍比较严重，无法实现跨域识别，这也是当前感知技术中比较困难的挑战之一。当前的实现跨域识别的模型有：WiHF 模型、Widar3.0 模型、个体差异性-任务一致性模型、基于梯度的无设备独立于域的手势识别模型、WiTransfer 模型、WiAG 模型和交叉感知模型等等。这些模型都已经取得了一些不错的可观的效果。但是由于对跨域识别影响的因素太多，这些模型或多或少都会达不到一些想要的效果。随着机器学习的不断发展，特别是深度学习在这些模型中在跨域感知方

面多多少少取得成功后，使用深度学习的跨域识别模型也是不断涌现。除了神经网络的设计，使用什么样的数据，怎样处理原始数据更是重中之重。是使用 CSI 原始数据作为输入，还是说使用从 CSI 原始数据中提取出的体坐标速度曲线(BVP)作为输入，还是说再从 BVP 基础上构造多普勒频谱(DFS)作为输入，应该通过具体实验来探究哪种输入会使深度神经网络模型能达到更好的跨域识别准确率，换句话说，哪种输入作为特征受领域因素变化影响最小，即跨域识别影响程度最小的，可以更好地保持任务一致性，减小域间差异性的。

跨域识别已经应用在现实生活中，比如跨域行人重识别(REID)、跨领域活动识别等多个领域。比如，在跨域行人重识别中，当训练集和测试集的属性有很大差别，如背景、视角、光线等有较大的差别，即源域和目标域之间有较大差别时，实现跨域识别。通过跨域行人重识别，可以实现跨域跟踪。此外，跨域识别还会对步态识别、手势识别等当前识别应用有重大贡献，跨域识别可以增强识别模型的鲁棒性，可以提高识别模型的测试准确率。而良好的公平的稳健的神经网络模型是实现较好的跨域识别模型的重中之重。总之，使用稳健的神经网络模型来实现跨域识别，对当前识别模型具有重要的意义。

1.4 本文贡献和结构

本文主要的贡献如下：

(1)本文提出了一个面向 WIFI 行为跨域识别的机器学习模型 WCDR 来实现跨域识别人体行为。

(2)本文设计了一个公平的数据预处理算法并结合卷积神经网络保证了跨域识别的鲁棒性。

(3)本文通过实验证明了 WCDR 模型具有更高的识别准确率和更强的跨域能力。

本文结构为：

第一章：介绍了跨域识别的相关背景，重点介绍了跨域识别的发展现状尤其是当前比较顶尖的跨域识别模型，同时介绍了跨域识别的研究意义和应用领域，最后对本文结构进行了说明。

第二章：主要介绍本文所使用的相关技术，主要是现存的特征提取算法和深度学习相关技术，对各个模型所使用的特征提取算法进行了相关总结。总结了机器学习和深度学习相关的基础知识，对卷积神经网络、循环神经网络和分类器等相关技术模型进行了详细的介绍。

第三章：提出了面向 WIFI 行为跨域识别的机器学习模型 WCDR。从整体上介绍了模型的整体设计，然后分别介绍了数据预处理模块和跨域识别模块的设计。其中数据预处理模块详细介绍了如何进行特征提取、数据切片，重点介绍了公平良好的特征提取算法。跨域识别模块提出了实现跨域识别的卷积神经

网络模型：WCDR。

第四章：通过具体实验证明 WCDR 模型的跨域识别性能。首先介绍了 WCDR 模型的前端展示，接着展示模型 WCDR 的实验结果，采用 Widar3.0 数据集对模型 WCDR 进行训练，实现跨域识别，并对模型的参数进行了必要的说明。通过实验证明 WCDR 模型的训练和验证效果，然后通过实验测试 WCDR 模型在 CSI 数据集上进行人体位置和面部朝向的跨域识别效果，接着简要分析了模型的性能，最后，对实验结果进行了总结分析，并与现有模型作对比。

结论：对本文进行整体总结，总结 WCDR 模型仍存在的局限性，并提出接下来的改进方向。

2 相关技术

2.1 CARM 模型中数据预处理算法

在基于 WIFI 信号进行感知的识别任务模型中，最经常使用的便是 CSI 数据。典型的做法就是直接将原始 CSI 数据直接输入到分类器中进行训练，然后在将训练好的模型用来测试数据集。从实验结果上看，直接使用原始 CSI 数据作为输入的模型在识别准确率上不是很高，也就是说，原始 CSI 数据集并不合适直接用来做分类任务。本文也对原始 CSI 数据进行了实验，实验结果如图 2.1 所示。

```
Epoch 0/4
----------
train Loss: 0.0001 Acc: 0.0645
valid Loss: 0.0001 Acc: 0.0648

Epoch 1/4
----------
train Loss: 0.0001 Acc: 0.0648
valid Loss: 0.0001 Acc: 0.0649

Epoch 2/4
----------
train Loss: 0.0001 Acc: 0.0649
valid Loss: 0.0001 Acc: 0.0649

Epoch 3/4
----------
train Loss: 0.0001 Acc: 0.0649
valid Loss: 0.0001 Acc: 0.0657

Epoch 4/4
----------
train Loss: 0.0001 Acc: 0.0657
valid Loss: 0.0001 Acc: 0.0664

Training complete in 0m 1s
Best val Acc: 0.0664
```

图 2.1　ANN 对原始 CSI 数据分类效果

针对分类准确率较低的问题，实验者们分析原始 CSI 数据是“噪声性的”，所以不会显示出不同活动的独特特征[18]。所以为机器学习模型过滤噪声并提取出一定特征成为必要过程。下面介绍典型的 CARM 模型的去噪声和特征提取方法[18]。

2.1.1 去噪声

CSI 数据流的噪声[19]主要来自 WIFI 接收器和 WIFI 发射器的 WIFINIC 中的内部状态转换。例如，当 WIFI 传输的功率改变时，会使 CSI 数据流中出现高幅度脉冲和突发噪声，它们同时影响 CSI 数据流中的所有样本。有关实验表明，

传统的滤波器对这种 CSI 数据噪声过滤后的效果不佳。而 CARM 模型利用人体的活动对 CSI 流中产生的影响是相关的这一特性，进而对原始 CSI 数据流进行降噪。

CARM 模型使用主成分分析(PCA)对 CSI 数据进行降噪[20]。首先从 CSI 流中减去 CSI 样本数据的振幅平均值来标准化。接下来进行相关估计，CARM 计算相关矩阵 HTxH，相关矩阵的维度为 NxN，其中 N 是 CSI 流的数量。第三步是特征分解，CARM 对相关矩阵进行特征分解，得到特征向量。最后进行运动信号重建，CARM 使用式 2.1 进行计算主成分，最后 CARM 舍弃了第一个主成分而保留了接下来的五个主成分来用于特征提取。

$$h_i = H \times q_i \tag{2.1}$$

其中q_i和h_i分别是第 i 个特征向量和第 i 个主要成分。

2.1.2 特征提取

人体行为在 CSI 数据流上具有两个重要的属性，分别是时间和频率[19]。CARM 使用离散小波变换(DWT)从多种分辨率中提取频率。DWT 给信道频率响应(CFR)信号计算不同级别的能量，其中不同能量级别对应不同的频率范围。具体表现为 DWT 级别的能量越高，人体行为的速度就越有可能存在于与该级别相关的频率范围内。DWT 的优点表现为在时间和频率上的分辨率权重平均，可以轻易地区分开高速行为运动和低速行为运动。另外，DWT 减少了数据的大小。

在 CARM 模型具体进行 DWT 特征提取时，将 PCA 分解为 12 个级别，从 0.15Hz 到 300Hz。CARM 对五个主要成分的 DWT 结果进行取平均值，来获取不同主成分中的动作信息。最后，CARM 提取出一个 27 维特征向量，其中包含 3 个主要特征，分别是级别能量，即运动能量强度、速度变化率、身体速度。接下来 CARM 将用提取出的特征构建隐马尔可夫(HMM)模型。

2.2 常见分类模型

线性模型(LinearModel)[21]]是机器学习中的一个经典模型，它是一个通过学习分类样本的线性组合的特征来进行分类的模型。在一个 N 维样本 $x=[x_1,x_2,...,x_N]^T$ 中，它的线性函数为式 2.2[21]：

$$f(x;w) = w_1x_1 + w_2x_2 + \cdots + w_Nx_N + b = w^Tx + b \tag{2.2}$$

其中 $w=[w_1,w_2,...,w_N]^T$ 是 N 维权重向量，b 为偏置。即直接用 $y=f(x;w)$来预测目标。在线性分类问题中，y 是预测目标的标签(Label)，由于函数$f(x;w)$的对应的值域是实数[21]，所以没有办法直接使用 f 函数来进行分类预测，于是引入一个非线性的决策函数(DecisionFunction)$g(f)$来预测目标。对于二分类问题时，g 函数可以说是一个符号函数 $sgn()$。典型的线性分类模型有逻辑回归和支持向量机等等。

对于二分类(BinaryClassification)问题，模型的标签只有两种取值，一般为0-1 标签。典型的二分类问题有计算机视觉中的猫狗分类，自然语言处理中的情感二分类等等。在二分类问题中[21]，一般需要一个线性判别函数 $f(x;w)=w^T+b$。样本空间中满足 $f(x;w)=0$ 的坐标点构成一个用于分割的超平面(Hyperplane)，也就是决策边界(DecisionBoundary)。在二分类问题中超平面将样本空间一分为二，每个部分对应一个分类。二分类问题主要做的就是通过学习样本特征，寻求类别间的决策边界，对样本进行二分类，然后在数据集上进行测试，目标就是实现比较准确的分类模型。

对于多分类(Multi-classclassification)问题[21]，模型的标签的取值种类数大于2，即有多种分类。根据可以将多分类问题向多个二分类问题转化的思想，可以使用多个线性判别函数来解决多分类问题。即将 N 分类问题(N>2)，通过二分类转化为(N-1)分类问题，再通过二分类转化为(N-2)分类问题，以此类推，直到转化为二分类问题，再得到分类结果。多分类问题的目标与二分类问题相同，即找到多组合适的权重，使得最后的多分类模型达到一个良好的分类效果。下面将介绍四种典型的分类模型。

2.2.1 贝叶斯分类器

首先介绍贝叶斯决策理论[22]贝叶斯理论是一种概率统计的理论。该理论提供了一种更新概率估计的方法，通过考虑新的观测数据，可以调整先前的概率估计。以下是贝叶斯理论的基本概念和原理：

1.先验概率(PriorProbability)：在考虑任何观测数据之前，我们对事件的概率进行初始猜测，称为先验概率。通常用$P(A)$表示，其中 A 是某个事件。

2.似然函数(Likelihood)：在得到新的观测数据后，我们想要评估事件的概率有多大。似然函数描述了给定某个事件的情况下，观测到这些数据的概率。通常用$P(D|A)$表示，其中 D 是观测数据。

3.后验概率(PosteriorProbability)：贝叶斯理论的核心是计算后验概率，即在考虑新的观测数据后，我们对事件的新估计概率。后验概率通过先验概率和似然函数的乘积得到，然后进行归一化。通常用$P(A|D)$表示。如式 2.3 所示

$$P(A|D)=\frac{P(D|A)\cdot P(A)}{P(D)} \tag{2.3}$$

其中，$P(D)$是归一化因子，确保后验概率的总和为 1。

4.证据或边缘似然(EvidenceorMarginalLikelihood)：在分母$P(D)$中，$\Sigma P(D|A_i)\cdot P(A_i)$，其中$A_i$是所有可能的假设或事件。这部分被称为证据，也称为边缘似然。

接下来介绍朴素贝叶斯分类器[23]。朴素贝叶斯分类器就是用贝叶斯决策理论构成的监督学习分类器。朴素贝叶斯分类器使用 GaussianNB 和 plot_classifier

两个程序包。首先构造输入的数据。构建一个二维的 numpy 数组 X 和一个一维的 numpy 数组 y，其中 X 是需要进行分类的样本，y 是进行监督学习时使用的标签。这时，已经把需要输入的数据和标签分别输入到变量 X 和 y 中了。下面需要建立一个朴素贝叶斯分类器模型。模型可以通过初始化函数，并在样本(X,y)上训练，然后再在测试集 X 上进行预测，即可达到朴素贝叶斯分类器对样本的预测。若想分析一下所构建的分类器的准确性，可以通过计算准确率来看贝叶斯模型的预测效果。此外，还可以将数据可视化，即通过逻辑回归的 plot_classifier 函数将 X 和 y 画在坐标系内并给出分类区域。在优化方面，可以通过将数据集分割成训练集和测试集并用交叉验证结合来优化贝叶斯分类器。另一种数据可视化是混淆矩阵可视化。混淆矩阵(confusionmatrix)[24]计算把一个 X 类预测成结果 y 的样本数。当然，在理想情况下，希望混淆矩阵中对角线上的元素值越高越好，因为对角线上的元素即为将第 i 类元素预测成第 i 类的样本数，即预测准确的样本数。

对于朴素贝叶斯分类器，下面将通过对正态分布朴素贝叶斯模型(GaussianNaiveBayesmodel)来探究朴素贝叶斯分类器对 CSI 数据流处理的效果。将 CSI 数据流经 PCA 方法去噪声和 CARM 特征提取，并保留后五个主要成分后的特征向量作为 X 样本，每行向量所归属的动作类别作为 y 标签(其中 CSI 数据流包含 9 中基本动作)，再将(X,y)作为朴素贝叶斯分类器的输入，然后对样本 X 进行预测，最后实验结果和混淆矩阵分别如图 2.2 和图 2.3 所示。

从实验结果上看，使用朴素贝叶斯分类器进行基于 WIFI 信号的手势行为跨域识别时，识别准确率只有 13%，在其他指标如召回率等也表现为很低。通过实验分析，贝叶斯决策模型不适用于对 CSI 数据流的分类。接下来将介绍另外一种典型的分类模型。

	precision	recall	f1-score	support
class_1	0.00	0.00	0.00	4
class_2	0.33	1.00	0.50	1
class_3	0.00	0.00	0.00	1
class_4	0.00	0.00	0.00	3
class_5	0.00	0.00	0.00	2
class_6	0.20	1.00	0.33	1
class_7	0.00	0.00	0.00	0
class_8	0.00	0.00	0.00	2
class_9	0.00	0.00	0.00	1
accuracy			0.13	15
macro avg	0.06	0.22	0.09	15
weighted avg	0.04	0.13	0.06	15

图 2.2　朴素贝叶斯分类器对 CSI 数据流分类结果

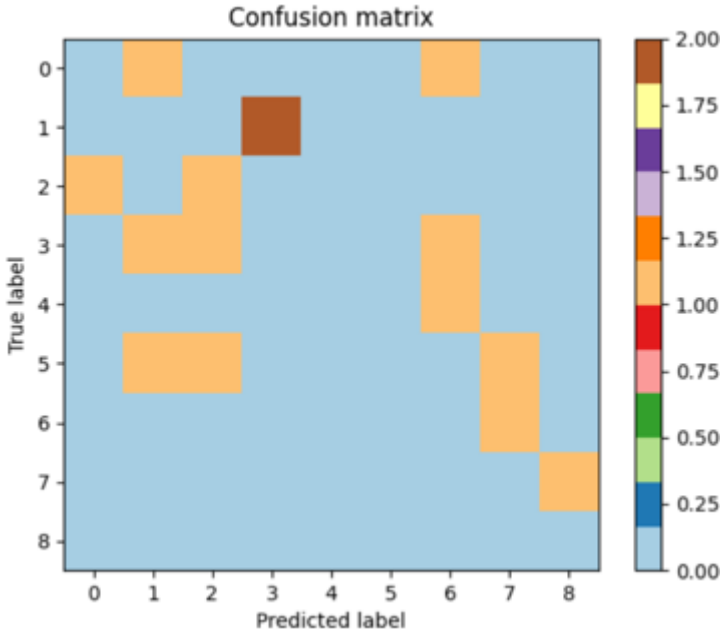


图 2.3　朴素贝叶斯分类器对 CSI 数据流分类混淆矩阵

2.2.2 全连接神经网络

人工神经网络[21](ArtificialNeuralNetwork，ANN)，是计算机科学家们模拟神经科学构建的数学模型。ANN 通过对人们大脑的神经元构成的网络进行模拟，构建人工神经元，并根据一定的拓扑排序来建立人工神经网络，也称为神经网络(NeuralNetwork，NN)。神经网络是一种大规模的并行分布式处理器[21]，具有学习经验知识的能力。从网络结构上看，神经网络具有多种网络结构，例如卷积神经网络和循环神经网络，还有本节将要介绍的全连接神经网络。全连接神经网络(DenseNeuralNetwork，DNN)是一种前馈网络。顾名思义，前馈网络就是整个网络中的信息是朝着一个方向进行传播，不进行反向传播信息。DNN 和卷积神经网络(CNN)都属于前馈网络。前馈神经网络[21](FeedforwardNeuralNetwork，FNN)是最早最简单的 ANN，由于其由多层构成的特点，FNN 也被称为多层感知器(Multi-LayerPerceptron，MLP)，即由多层逻辑回归模型和非线性激活函数构成。DNN 通常由多层神经元构成，总体上分为三类，输入层，中间层和输出层，其中中间层也叫做隐藏层，DNN 结构如图 2.4 所示。

DNN 在输入层输入，通过每一层的信息传递，最后在输出层形成输出。在每一层中的神经元都有自身的权重组，一层神经元形成一个权重矩阵，输入向量对每一层的权重矩阵进行矩阵乘法，依次作乘法到最后一层，最后得出输出矩阵。上述算法即 DNN 的前向计算，但是怎么对 DNN 的行为进行限制呢？通过损失函数(LossFunction)作为限制 DNN 的指标。最开始时 DNN 中的参数都是随机的[25]，需要让 DNN 知道自己的输出结果和真实值相差多少，就是通过损

失函数来完成这一任务。典型的损失函数有 0-1 损失函数、绝对值损失函数和均方误差损失函数等等。有了损失函数，需要使用损失值来对网络中原来的随机参数进行有序的调整，使其可以输出有意义的结果。这一步称为梯度下降[25]，从根本意义上说，目的就是优化网络中的参数，最后使得损失值达到最小。借鉴数学中的梯度一概念，在高等数学中，梯度是沿着 f_t 上升最快的方向。若想要让损失值达到最小，可以沿着与梯度相反的方向，让损失值下降最快，最后就可以达到可观的损失值。梯度下降法的迭代公式如式 2.4 和式 2.5。

$$W_{ij} = W_{ij} - \alpha \times \frac{\partial}{\partial W_{ij}} L(\omega, b) \tag{2.4}$$

$$b_i = b_i - \alpha \times \frac{\partial}{\partial b_i} L(\omega, b) \tag{2.5}$$

其中W_{ij}和b_i是 DNN 中的参数，L是损失函数，α是学习率。

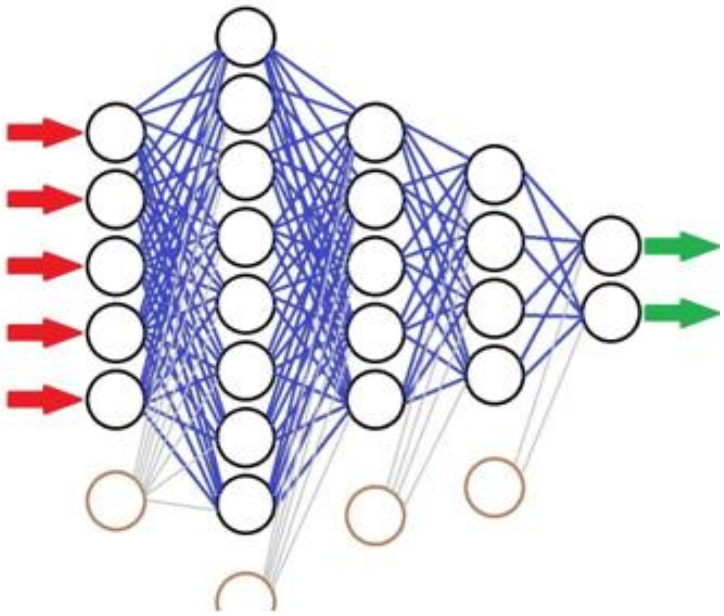


图 2.4　DNN 结构示例

在计算损失函数对权重和偏置的偏导数的过程中，使用反向传播算法[25]。具体过程是，对于每一层神经元，计算出残差 $\delta_i^{(l)}$，即该神经元的值对最终输出的偏差的影响程度，从顶层开始，逐层向下计算每一层神经元的残差值，再根据残差值计算损失函数L对权重 W 和偏置b的偏导数，进而更新 W 和b。其中残差计算偏导数的过程如式 2.6 和式 2.7。

$$\frac{\partial L(w,b)}{\partial w_{ij}^{(l)}} = \delta_i^{(l)} \times \alpha_j^{(l-1)} \tag{2.6}$$

$$\frac{\partial L(w,b)}{\partial b_i^{(l)}} = \delta_i^{(l)} \tag{2.7}$$

对于 DNN，下面将探究三层 DNN 对 CSI 数据流的处理效果。将与朴素贝叶

斯分类器输入相同的 CSI 数据，并进行相同的预处理，输入到 DNN 中，留出验证集和测试集，再进行五折交叉验证，最后手势识别的效果如图 2.5 所示。

从实验结果上看，三层 DNN 对处理后的 CSI 数据流的识别准确率为 10.42%，大于对原始 CSI 数据流的 6.64%，与朴素贝叶斯分类器分类效果相近，但是同样仍然较低。通过具体实践得出，对原始 CSI 数据进行去噪声和特征提取确实可以提高分类准确率，但是全连接神经网络和贝叶斯决策理论却不适用于处理数据流类的数据。

```
Epoch 0/4
----------
train Loss: 0.0493 Acc: 0.1042
valid Loss: 0.0493 Acc: 0.1042

Epoch 1/4
----------
train Loss: 0.0493 Acc: 0.1042
valid Loss: 0.0491 Acc: 0.1042

Epoch 2/4
----------
train Loss: 0.0491 Acc: 0.1042
valid Loss: 0.0490 Acc: 0.1042

Epoch 3/4
----------
train Loss: 0.0490 Acc: 0.1042
valid Loss: 0.0488 Acc: 0.1042

Epoch 4/4
----------
train Loss: 0.0488 Acc: 0.1042
valid Loss: 0.0486 Acc: 0.1042

Training complete in 0m 1s
Best val Acc: 0.1042
```

图 2.5　三层 DNN 对 CSI 数据流分类结果

2.2.3 普通循环神经网络

一般的线性分类模型不适用于处理 CSI 数据流类的数据，可以观察到，将原始 CSI 数据经过 CARM 去噪声和特征提取进行数据预处理后，再分别作为朴素贝叶斯分类器和全连接神经网络的输入。在朴素贝叶斯中，将五个主成分输入到分类器中，由于 CSI 数据是数据流类的数据，输入到分类器后，丢失了数据流的一连串连续的时间特征，将每个时间点离散式地进行学习，此时无法学习到前后相连的信息。在 DNN 中具有同样的问题，DNN 是一种前馈网络，也将 CSI 数据流“离散”地处理，必然也会丢失 CSI 数据流的连续特征。

通过分析，机器学习方法识别人体行为的关键是在于捕捉每个动作在一段

时间内具有独特的 CSI 变化。而处理连续数据流是前馈网络所不能完成的。具有处理数据流能力的两种神经网络模型分别是循环神经网络和一维卷积神经网络。下面将介绍两种循环神经网络，其中包括普通 RNN 和 LSTM 模型。

DNN 无法预测 CSI 数据流是因为 DNN 是一种单纯的前馈网络，在 DNN 中，信息的传递是单向的。这种特征的好处是模型简单、易学习，但是也降低了 DNN 的适应能力。重新考虑要实现的目标，很容易知道是要一种能够处理序列化数据的模型，也即能处理前后相关数据的模型，这时来考虑一下循环神经网络(RNN)。

循环神经网络(RecurrentNeuralNetwork，RNN)[21]是一类具有短期记忆功能的神经网络。RNN 的特点是，每一层的神经元不仅仅可以接收其他神经元的信息，也能接收自身传来的信息。RNN 处理时序数据的原理是使用自反馈的神经元。RNN 的结构如图 2.6 所示[21]。设模型的输入序列 $x=(x_1,x_2,...,x_N)$，RNN 通过式 2.8 计算带反馈边的隐藏层的活性值 h_t[21]。其中 $h_0=0$，f 函数是一个非线性函数。

$$h_t = f(h_{t-1}, x_t) \tag{2.8}$$

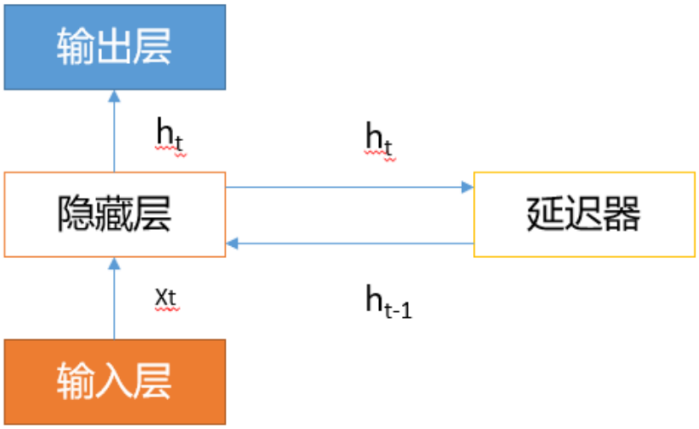


图 2.6　循环神经网络

典型的 RNN 有简单循环网络[21]。简单循环网络(SimpleRecurrentNetwork，SRN)是最简单的 RNN。SRN 由输入层、1 个隐藏层、输出层构成。SRN 与前馈网络的一个重要的区别是，在前馈网络中，隐藏层之间是无连接的，而 SRN 增加了隐藏层之间的反馈连接。一个 SRN 可表示为式 2.9[21]。

$$h_t = f(Uh_{t-1} + Wx_t + b)$$
$$y_t = Vh_t \tag{2.9}$$

其中 h 是隐状态，f 是非线性激活函数，x_t 是输入，y_t 是输出。

在具体的机器学习应用中，RNN 可以应用于多种任务，如序列到类别、序

列到序列[21]。序列到类别典型的例子就是文本分类，序列到序列典型的例子有序列标注、编码解码。在 RNN 中的参数学习中，RNN 采用随时间反向传播算法[21]。随时间反向传播算法(BackPropagationThroughTime，BPTT)，用于 RNN 计算梯度。损失函数 L 对参数 u 的偏导数可通过式 2.10 计算。

$$\frac{\partial L_t}{\partial u_{ij}} = \sum_{k=1}^{t} \frac{\partial^+ z_k}{\partial u_{ij}} \frac{\partial L_t}{\partial z_k} \tag{2.10}$$

进而，RNN 计算损失函数 L 对参数 U 的梯度，如式 2.11 所示。

$$\frac{\partial L}{\partial U} = \sum_{t=1}^{T} \sum_{k=1}^{t} \delta_{t,k} h_{k-1}^{T} \tag{2.11}$$

同理可得，损失函数 L 对权重 W 和偏置 b 的梯度如式 2.12 和式 2.13 所示。

$$\frac{\partial L}{\partial W} = \sum_{t=1}^{T} \sum_{k=1}^{t} \delta_{t,k} x_{k}^{T} \tag{2.12}$$

$$\frac{\partial L}{\partial b} = \sum_{t=1}^{T} \sum_{k=1}^{t} \delta_{t,k} \tag{2.13}$$

BPTT 算法得到参数的梯度后经过前向计算和反向计算就可以对参数进行更新[21]。据此，RNN 可以通过“反复”来获得比前馈网络更强的处理计算能力。然而，强如 RNN 也会存在一些问题。SRN 中主要存在的问题是梯度消失和梯度爆炸问题。梯度消失问题是当间隔较大时，梯度会变得非常小。梯度爆炸问题是当间隔较大时，梯度会变得很大。这两种问题统称长期依赖问题。下面将通过具体实验来探究 SRN 对 CSI 数据流的处理能力。

```
loss:  2.217214345932007
loss:  2.319049119949341
loss:  2.3118057250976562
loss:  2.3288462162017822
loss:  2.327744722366333
loss:  2.3017821311950684
loss:  2.4834752082824707
loss:  2.4346094131469727
loss:  2.4578676223754883
loss:  2.4783482551574707
loss:  2.519265651702881
loss:  2.518771171569824
loss:  2.593949317932129
loss:  2.6168766021728516
loss:  2.2843704223632812
loss:  2.368624210357666
loss:  2.3750874996185303
loss:  2.324338912963867
loss:  2.3399784564971924
loss:  2.2620720863342285
loss:  2.3269083499908447
loss:  2.35239839553833
loss:  2.3726468086242676
loss:  2.3719019889831543
loss:  2.366888999938965
```

图 2.7　SRN 对 CSI 数据流训练损失

本文构造一个简单的 SRN，共包括三层，1 层输入层，1 层隐藏层和 1 层输

出层。其中隐层的神经元个数为 90，输出层的神经元个数为 9。在每层的前向计算中，该层自动获取输入数据维度，通过 GRU 后再经过一个密集连接层(90 输入神经元、9 输出神经元)。SRN 的输入数据与 DNN 的不同，保留了时间维度的信息，构造出“样本数-时间特征”的输入数据格式。最后在样本上进行迭代，输出损失值。实验结果如图 2.7 所示。

从实验结果可以看出，起始损失率为 2.403，训练到最后的损失率为 2.366，损失率并没有降到较低的水平。究其原因，本文认为可能是 CARM 对原始 CSI 数据流的特征提取使其丢失了循环神经网络所需要的时间维度特征，致使 SRN 对 CSI 数据流的训练损失仍然较高。

2.2.4 长短期记忆网络

考虑到 SRN 中存在长期依赖问题，为了避免梯度爆炸和梯度消失问题，可以通过权重衰减和梯度截断[21]来解决。式 2.14 可以有效地改进梯度消失问题。

$$h_t = h_{t-1} + g(x_t, h_{t-1}, \theta) \tag{2.14}$$

然而 SRN 仍然存在梯度爆炸和记忆容量问题[21]。对于梯度爆炸问题，在计算误差项时，梯度可能会过大，从而导致梯度爆炸。对于记忆容量问题，随着 ht 不断存储信息，会出现饱和现象。为了解决上述两个问题，可以引入门控机制[21]，即基于门控的循环神经网络(GatedRNN)，长短期记忆网络就属于一种 GatedRNN。

长短期记忆网络(LongShort-TermMemoryNetwork，LSTM)[21]是 RNN 中的一种变体，可以有效解决 SRN 的梯度消失和梯度爆炸问题。LSTM 网络对 SRN 的改进主要包括内部状态和门控机制。在 LSTM 网络中，内部状态c_t计算公式[21]如式 2.15 所示。

$$\begin{aligned} c_t &= f_t \cdot c_{t-1} + \imath_t \cdot \mathcal{C}_t \\ h_t &= o_t \cdot \tanh(c_t) \end{aligned} \tag{2.15}$$

其中f_t、$\imath_t$、o_t三个门来控制信息传递，候选状态如式 2.16 所示。

$$\mathcal{C}_t = \tanh(W_x x_t + U_C h_{t-1} + b_c) \tag{2.16}$$

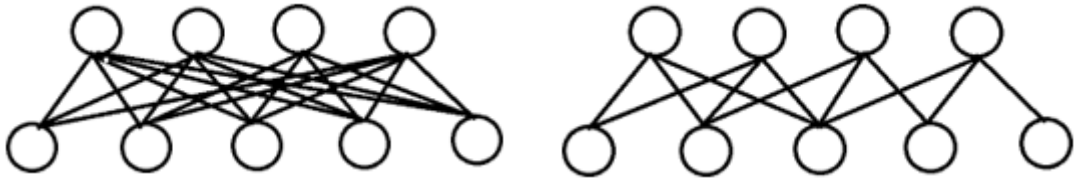
在 LSTM 网络[21]中，$\imath_t$为输入门，f_t为遗忘门，o_t为输出门。引入内部状态门控机制，LSTM 网络通过记忆单元 c 长期存储历史信息，具体来说，在网络中，LSTM 网络使用参数 weight 和 bias 收集前一层的输出数据并向下一层和本层传递信息[26]。长期记忆也是一种参数，它代表着从训练集中学到的经验，记忆单元 c 学习某个关键信息，并保存一定的时间间隔。根据 LSTM 网络的长期记忆的特征，对 CSI 数据流的分析可以完全使用 LSTM 网络来驾驭，并可以达到可观的识别效果。

2.3 卷积神经网络

2.3.1 卷积神经网络概念

卷积神经网络一般由卷积层、汇聚层和全连接层构成[21]。局部连接层的构成是在卷积层中的每一个神经元都只和前一层的局部窗口内的神经元相连，形成了一个局部连接网络。权重共享指的是同一层中所有的神经元的卷积核是一样的，共享一个卷积核。

CNN 用卷积层代替了全连接层。全连接层和卷积层的对比如图 2.8 所示。



(a)全连接层　　(b)卷积层

图2.8　全连接层和卷积层对比

卷积层由输入特征映射组、输出特征映射组和卷积核构成。输入 X 是一个包含二维张量切片组的三维张量。输出 Y 也是一个包含二维张量切片组的三维张量。卷积核 W 是一个四维张量(U,V,P,D)，每个切片矩阵(U,V)是一个二维卷积核。

汇聚层(PoolingLayer)[21]也即子采样层(SubsamplingLayer)，用来对特征数量降维。卷积层可以减少神经网络中神经元连接的维度，而汇聚层可以减少特征数量的神经元数，即减少特征数量维度。在 CNN 中，汇聚也被称为池化操作。常用的池化函数有最大池化和平均池化。最大池化(MaxPooling)选择某个区域内所有神经元的最大活性值来作为该区域的表示，从而达到降维效果。平均池化选取某个区域内所有神经元的平均活性值作为该区域的表示，同样达到降维效果。

典型的 CNN 均是由卷积层、池化层和全连接层共同构成。目前卷积核趋于更小，不断取代池化层，CNN 正在趋于全卷积网络。

2.3.2 几种经典的卷积神经网络

目前广泛使用的深层卷积神经网络[21]包括 LeNet-5、AlexNet、Inception 网络、残差网络和 VGG16 等等。

LeNet-5[21]于上世纪 90 年代在美国广泛应用于银行识别支票的手写数字识别。网络共有 7 层，输入形状为(32,32)，输出即 10 类阿拉伯数字 0-9 的各类得分。其中 7 层网络依次为卷积层、池化层、卷积层、池化层、卷积层、全连接层和输出层构成。

AlexNet[21]是第一个现代深度卷积神经网络模型，采用 ReLU 非线性激活函数，使用 Dropout 防止过拟合，使用数据增强来提高准确率，并在 2012 年 ImageNet 图像分类竞赛中获得冠军。网络结构为 5 个卷积层、3 个池化层和 3 个全连接层(softmax 激活输出)。

Inception 网络[21]有很多版本，v1 版本在 2014 年 ImageNet 图像分类竞赛获得冠军。

残差网络(ResidualNetwork，ResNet)[21]向非线性的卷积层增加线性直连边(残差连接)来提高信息传播效率。

VGG16[27]模型是一个训练好的用于分类任务的模型。模型含有的权重的大小会达到 500MB 左右。模型包含了冻结层，冻结层可以阻止更新卷积块的权重值。模型可以被训练来对大到 1000 个类别进行分类，小到 2 个类别进行分类(如猫狗分类)。

2.3.3 其他卷积方式

卷积用来实现高维特征向低维特征的转换，表示为 $z=Wx$。转置卷积为卷积的反向变化，即实现低维特征向高维特征的转换，表示为 $x=W^{T}z$。举例说明，神经网络中的前向计算和反向传播也是一种转置关系。

空洞卷积[21]是一种无需增加参数的数量就可以增加输出单元感受野的方法，也称为膨胀卷积。空洞卷积通过在卷积核中插入空洞的方法来增大卷积核的尺寸，从而增大感受野。

3 面向 WIFI 行为跨域识别的机器学习模型 WCDR

本章将详细介绍面向 WIFI 行为跨域识别的机器学习模型 WCDR。在第一小节从整体上介绍 WCDR 的模型设计；在第二小节详细介绍面向卷积神经网络输入的 CSI 数据预处理方法；在第三小节详细介绍卷积神经网络的结构设计和进行 WIFI 行为跨域识别的详细过程。

3.1 WCDR 模型整体设计

模型 WCDR 利用 WIFI 信号对行为的跨域识别对于动作的类型没有特别的要求，对于不同目标域下分类行为的泛化能力很强，需要解决的就是域独立[28]问题，即在不相关的域之间寻求一致性规则，这种一致性规则不仅可以适用于同域下识别，还适用于跨域识别。而在跨域识别中，对原始数据集的预处理尤为重要，一个好的数据预处理算法可以大幅度提高模型跨域的成功率。在跨域识别领域中，CARM 模型具有完整的数据预处理算法，分为去噪声和特征提取两部分。但是在 CSI 数据上进行 CARM 去噪声和特征提取，一方面容易导致原始 CSI 数据流丢失大量的特征，另一方面会产生不利于卷积神经网络输入的数据格式，最后由于缺乏 CARM 预处理的相关代码，难以实现人工匹配输入数据格式。模型不但要对原始 CSI 数据流进行预处理，而且还要转化 CSI 数据的存储形式。在原来使用 DNN 处理 CSI 数据的过程中，采用直接读取 dat 格式的 Widar3.0 中的 CSI 原始文件，而这种读取 dat 格式文件的形式会耗费大量的时间，平均 916ms 读取一个 dat 文件，而一个 CSI 数据用户文件夹下有 6000 多个 dat 文件，读取一个用户文件夹将要花费 1.5 个小时。所以 WCDR 模型采用将 dat 格式文件一次性读出，经数据预处理后转化成 npz 格式文件的方式重新存储原始 CSI 数据来减少后续模型读取数据的时间成本。

使用 WCDR 模型进行行为跨域识别，通过将 npz 数据进行数据预处理，再输入到卷积神经网络中进行训练、验证和测试，最终能够达到较小的训练损失和较高的识别准确率。

在 CARM 模型思路和 CNN 的基础上，本文设计了一种鲁棒性强的跨域识别模型 WCDR。模型结构图如图 3.1 所示，该模型主要分为两个阶段进行行为跨域识别，第一个阶段是数据预处理阶段，其中对输入的待处理原始 CSI 数据进行数据集筛选和特征提取，可以得到一个特征值矩阵。第二个阶段是跨域识别阶段，将不同人体位置和面部朝向下的 CSI 数据输入到 CNN 中，并不断迭代，减小验证损失，提高识别准确率。通过数据预处理和跨域识别两个阶段的运行，最终可以得到一个良好的跨域识别效果。

使用 WCDR 进行 WIFI 行为跨域识别，一方面在预处理模块使用增强模型

泛化能力的数据集筛选和特征提取算法进行数据预处理，形成易于卷积神经网络学习的数据格式，另一方面在跨域识别模块设计一个卷积神经网络模型，通过人工分类标签和监督学习训练模型，使得在减小读取数据的时间成本的同时又保证了良好的跨域识别的效果。WCDR 模型的整体结构图如图 3.1 所示。

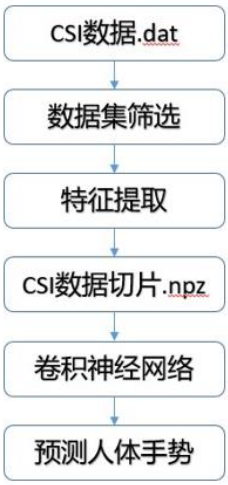


图 3.1　WCDR 模型整体结构图

3.2 预处理模块设计

预处理模块设计的目的主要有两个，第一个是去除原始 CSI 数据流中的噪声，并提取 CSI 数据流中的有用的信息，便于神经网络处理，减小神经网络的计算量；第二个是形成预处理模块和跨域识别模块的接口，生成便于输入神经网络的数据格式。对原始 CSI 数据流进行数据预处理可分为两个阶段：数据集筛选和特征提取。

3.2.1 数据集筛选

在数据集筛选过程中，首先获取数据集。在数据读取过程中，模型从原始 CSI 数据集中的 dat 文件读取数据，其中原始 CSI 数据集文件格式为用户名-手势种类-躯体位置-面部朝向-重复次数-WIFI 接收器名，这就完成了数据读取环节。通过对整体数据集的分析，本文发现用户名、手势种类、躯体位置、面部朝向、重复次数和 WIFI 接收器这 6 种领域因素均对 CSI 波形的影响较大。因此，模型拟控制躯体位置和面部朝向这两个领域因素不变，其他领域因素改变的形式来筛选原始 CSI 数据集，构成模型所需的原始 CSI 数据训练集，本文余下部分的原始 CSI 数据集未经说明均指筛选后的原始 CSI 数据集。数据集筛选的结构图如图 3.2 所示。

在数据预处理前，WCDR 模型读取原始 CSI 数据流的格式为：时间-子载波-接收天线-发射天线，CSI 数据流的格式如图 3.3 所示。

(a)在躯体位置上进行跨域的训练集筛选

(b)在面部方向上进行跨域的训练集筛选

图 3.2　数据集筛选结构图

```
.dat data shape:  (1370, 30, 3, 1)
.dat data shape:  (1373, 30, 3, 1)
.dat data shape:  (1394, 30, 3, 1)
.dat data shape:  (1391, 30, 3, 1)
.dat data shape:  (1374, 30, 3, 1)
.dat data shape:  (1374, 30, 3, 1)
.dat data shape:  (1051, 30, 3, 1)
.dat data shape:  (1045, 30, 3, 1)
.dat data shape:  (1060, 30, 3, 1)
.dat data shape:  (1053, 30, 3, 1)
.dat data shape:  (1049, 30, 3, 1)
.dat data shape:  (1057, 30, 3, 1)
.dat data shape:  (1131, 30, 3, 1)
.dat data shape:  (1135, 30, 3, 1)
.dat data shape:  (1153, 30, 3, 1)
.dat data shape:  (1143, 30, 3, 1)
.dat data shape:  (1131, 30, 3, 1)
.dat data shape:  (1140, 30, 3, 1)
```

图 3.3　原始 CSI 数据流格式

得到该形状的 numpy.array 型数组后，本文将会将原始 CSI 数据流显示成波形的形式进行可视化分析。本文通过对原始 CSI 数据 dat 文件的批量可视化后发现 WIFI 接收器会对原始 CSI 数据(未筛选前)的振幅的取值范围和变化趋势造成较大的影响，并降低模型识别准确率，因此模型对原始 CSI 数据(未筛选前)

进行了筛选处理，并选取同一 WIFI 接收器获取同一人体的 CSI 数据集作为本模型的原始 CSI 数据集。其中不同 WIFI 接收器对 CSI 数据波形的影响如图 3.4 所示。

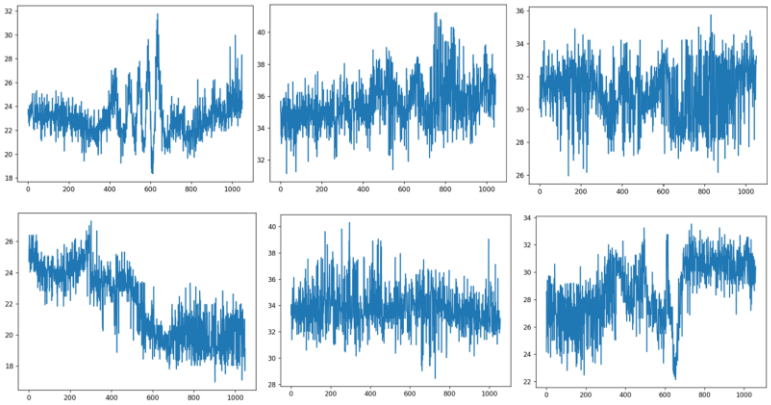


图 3.4　6 个不同 WIFI 接收器下原始 CSI 数据流振幅波形图

3.2.2 特征提取

在对 Widar3.0 数据集进行数据集筛选后，WCDR 模型对筛选后的 CSI 数据流进行特征提取环节。在特征提取环节，由于原始 CSI 数据为复数，因此模型首先对 CSI 数据的子载波进行取振幅和相位处理得到两组矩阵，接下来将振幅和相位矩阵在子载波维度上进行拼接，共同组成特征向量。由于有三根接收天线，模型将三根接收天线的特征向量进行简易的拼接处理，此时 CSI 数据变为时间-特征值这一二维数据，实现了对原始 CSI 数据的降维。最后，为了易于输入卷积神经网络进行训练，将 CSI 数据在时间维度上切片处理。以样本(1370,30,3,1)为例，特征提取的结构图如图 3.5 所示。

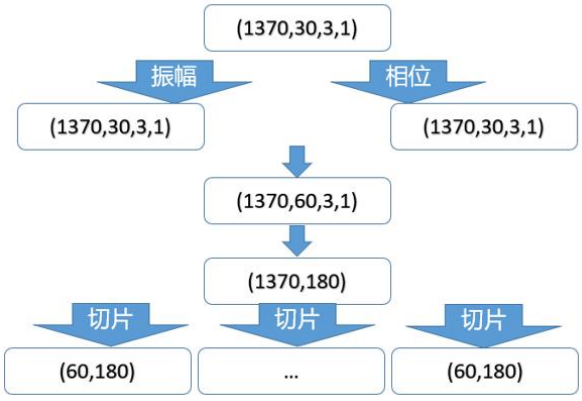


图3.5　特征提取结构图

对 CSI 数据流进行特征提取，一方面是为了减小卷积神经网络的计算量，提高神经网络的训练效率；另一方面有利于提高卷积神经网络的处理效果，通过提取输入数据中的有效信息，过滤掉一些潜在的混淆因子，可以在一定程度

上提高网络的训练准确率。通过特征提取，每一个特征向量(180 维，表示为一行数据)已经包含了人体活动的信道状态信息(CSI)的振幅值和相位信息，即包含了 CSI 数据的主要特征，换句话说，在通过特征提取后，每一种人体活动已经有了各自的振幅变化和相位变化，即在经过卷积神经网络训练之前，使用预处理算法已经对 CSI 数据流分析出了一定的信息。当然，在一些相近的手势之间只经过数据预处理模块肯定还是分辨不很清楚，不过这些细微的区别会在卷积神经网络中被学习，被放大，被区分进而被分辨。因此，数据预处理在总体上会对 WCDR 模型起到良好的促进作用。

3.3 跨域识别模块设计

跨域识别模块的设计思想主要是通过卷积神经网络对数据预处理后的 CSI 数据切片特征向量流进行学习、训练、验证，然后改变躯体位置和面部方向这两种领域因素进行跨域测试。注意 CNN 所处理的特征向量并不是线性代数中的矩阵的特征向量，而是 CSI 数据流经过数据预处理后保留的主要特征所组成的张量(tensor)。通过跨域识别阶段，可以对神经网络模型添加损失函数来监测模型对真实样本预测的偏离程度，在 CNN 训练的过程中可以反复迭代训练样本数据，使得最终的损失函数降到最低，对人体行为的识别准确率达到满意的效果。

3.3.1 跨域识别模块整体设计

跨域识别模型可以分为以下三个部分，第一部分是构建训练样本及标签和验证样本及标签，第二部分是使用卷积神经网络在训练集上学习并在测试集上进行测试，第三部分是将一些中间结果和最后的测试效果保存至文档。经过这三个部分，最终就可以得到 WCDR 模型进行跨域识别的识别准确率和损失值，了解到 WCDR 模型在对跨域识别的处理效果。

跨域识别模块的第一部分是构建数据集，目的是将数据预处理输出的 CSI 数据流格式转化为易于输入卷积神经网络的数据格式，首先定义出 WCDR 模型进行监督学习的监督指标训练集损失、训练集准确率、验证集损失和验证集准确率。然后采取留出验证集，将前 90%数据用于训练，最后 10%数据用于验证，并通过数据不断向前滚动的方式，将验证集始终留在最后 10%，便于保持代码的一致性。为了便于操作，将十种活动的训练样本和标签分别组成一个总的训练样本和标签。接着使用同样的方法留出验证集。为了保证面向对象特性，WCDR 模型重写 Dataset 类并包含样本和标签成员，并构建训练集对象和验证集对象，这样就完成了构建数据集部分。

为了能够使用 WCDR 模型可以对人体活动训练集进行学习，WCDR 模型的第二部分是设计卷积神经网络。模型将(60,180)的切片样本转换成(60,60,3)的形式，即将 CSI 数据流当作一幅(60,60)像素大小的三通道图像输入到卷积神经

网络进行处理。经过卷积层、激活函数、池化层和线性层最后分为十类输出。WCDR 模型的第三部分是模型的监督学习，WCDR 模型利用第一部分生成的训练集、验证集和测试集输入到第二部分设计的卷积神经网络来训练验证，并预测人体活动。模型实验的具体流程如下。

在数据预处理阶段完成后，每一种动作对应生成的样本和标签首先在训练、验证和测试文件夹下保存为 npz 文件。当跨域识别模块执行时，模型首先从 npz 文件中将 CSI 特征向量信息读入至模型中，每一个动作的样本数据格式为样本数-切片长度-特征值。接着模型会将每一个动作在样本维度上重排列，采取留出验证集的方式进行验证，并在测试文件夹中的测试集上进行测试，即每次从训练样本随机将验证集留出作为模型验证。接下来模型在不同动作的样本上迭代，在每种动作的样本中以一定的步长进行迭代训练，并获取每次迭代的损失和准确率。

在跨域识别的过程中，需要的是转换样本数据至卷积神经网络易于学习训练的数据格式，这样做的目的有两个：一是便于建构卷积神经网络，让 CNN 的参数具有变量可解释性；二是让经过特征提取后的 CSI 数据流变得具有可解释性，符合人类的思考和逻辑方式。输出每次迭代的损失和准确率，是为了让卷积神经网络进行学习数据时不至于是一个完全的黑盒模型，通过可视化数据，可以看到 CNN 在学习数据时的一个优化过程，具有视觉可解释性。从整体上看，跨域识别模块是一个面向对象的模块设计，对每一步的函数进行了封装，尽可能模拟了人的思维方式，是一个大数据和机器学习结合的面向对象模型。

3.3.2 卷积神经网络实现

在模型的跨域识别阶段使用的是卷积神经网络，以(60,60,3)形式的 CSI 数据切片作为输入，以分为十种人体手写数字动作为分类输出，模型设计的卷积神经网络的结构如表 3.1 所示。

在 WCDR 模型中最重要最核心的部分就是卷积神经网络，通过对数据流类数据进行切片并附标签处理，这使得卷积神经网络可以用来处理数据流类数据。卷积神经网络由卷积层、池化层和全连接层组成。从表 3.1 可以看出，WCDR 模型的卷积神经网络有 7 层卷积层，使用 LeakyReLU 激活函数进行非线性激活，4 层最大池化层，3 次归一化处理，然后接 3 层全连接层并使用 LeakyReLU 激活共同组成。

从表中可以看出 CSI 数据的输入尺寸是(60,60,3)，经过两层卷积层后输出尺寸是(56,56,64)，然后通过最大池化层减小尺寸为(28,28,64)。接下来再经过两层卷积层后输出尺寸是(24,24,128)，同样再次通过最大池化层减小尺寸为(12,12,128)。下面再经过两层卷积层后输出尺寸为(8,8,256)，然后进行最大池化为(4,4,256)。再通过最后一次卷积层输出为(2,2,512)，然后最大池化为(1,1,512)。

此时 CSI 数据尺寸变为了(512)的形状，然后通过一层全连接层变为(4096)，再通过一层全连接层变为(10)。到此，一个(60,180)的 CSI 数据切片样本被分类处理完毕。

表3.1　卷积神经网络的结构表

Layer	OutputShape
initial	(60,60,3)
Conv2d(3,64,3)	(58,58,64)
Conv2d(64,64,3)	(56,56,64)
LeakyReLU	(56,56,64)
MaxPool2d(2,2)	(28,28,64)
Conv2d(64,128,3)	(26,26,128)
Conv2d(128,128,3)	(24,24,128)
BatchNorm2d(128)	(24,24,128)
LeakyReLU	(24,24,128)
MaxPool2d(2,2)	(12,12,128)
Conv2d(128,256,3)	(10,10,256)
Conv2d(256,256,3)	(8,8,256)
BatchNorm2d(256)	(8,8,256)
LeakyReLU	(8,8,256)
MaxPool2d(2,2)	(4,4,256)
Conv2d(256,512,3)	(2,2,512)
BatchNorm2d(512)	(2,2,512)
LeakyReLU	(2,2,512)
MaxPool2d(2,2)	(1,1,512)
Linear(512,4096)	(4096)
LeakyReLU	(4096)
Linear(4096,4096)	(4096)
LeakyReLU	(4096)
Linear(4096,10)	(10)

卷积层在处理 CSI 数据流的任务中使得神经网络的性能得到了提升，并且达到了良好的效果。因此 WCDR 模型使用卷积神经网络对 CSI 数据流进行训练、预测。

3.3.3 使用卷积神经网络跨域识别人体行为

本文使用 WCDR 模型进行跨域识别时，拟在领域因素上进行跨域。与领域因素有关的因素如人体特征、躯体位置、面部方向、WIFI 接收器在空间中所处位置等等。这些因素都会使得目标域发生变化。本模型拟采用 Widar3.0 数据集下在躯体位置和面部方向上与训练集不同的数据集进行模型测试，评估模型的

跨域效果。

在 WCDR 模型中使用卷积神经网络进行 CSI 数据流的监督学习，在经过 CSI 数据流的预处理后，将时间样本数-时间切片-特征值矩阵作为卷积神经网络的输入，输出对行为种类的预测，在每次迭代中输出该批次样本训练的准确率和损失值，并梯度下降减小损失值，进行反向传播更新神经元的参数，最终得到最小化损失值的参数值。通过不断更新神经元参数值，模型损失值被降至一个较低的水平，模型可以实现对人体行为尽可能准确的预测，并且由于 WIFI 信号的信道状态信息(CSI)受领域因素干扰小，所以使用卷积神经网络的 WCDR 模型可以在跨领域因素域的条件下对人体行为进行精准识别。

4 实验与分析

4.1 实验工具

4.1.1 实验工具的设计

跨域识别不同于域内识别，从概念上来说给人以抽象的感觉，因此设计一个简单易懂的可视化界面有助于对跨域识别的过程有一个宏观的理解，可视化界面的设计大抵说明使用 WCDR 模型进行跨域识别的流程，使用 WCDR 进行跨域识别的流程图如图 4.1 所示。

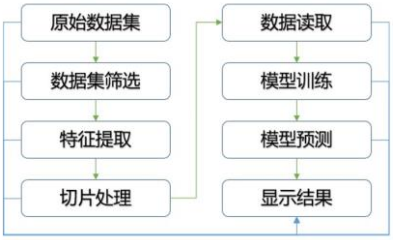


图4.1　　WCDR 模型跨域识别流程

从图 4.1 可以看出，在进行数据预处理之前，可以通过选择数据来加载数据到模型中并可以在界面中显示，然后通过数据预处理对原始数据进行数据集筛选和特征提取并将处理后的数据样本切片输出至 npz 文件中，同样也可以在界面中显示。数据预处理完毕后，此时点击训练模型，模型就会使用卷积神经网络对数据进行学习训练，数据的迭代结果会显示到界面中，并将最后的识别准确率和损失值保存至 txt 文本中。最后，点击过程可视化可以将卷积神经网络的训练过程中准确率和损失值的变化通过折线图显示出来。

在 WCDR 模型设计图的基础上，就可以得出模型界面应该包括的基本控件有：选择数据、显示数据、数据预处理、数据切片处理、训练模型、预测、显示模型效果和过程可视化，同时根据可解释性的特点可以知道界面应该有的可视化功能是：显示准确率损失值变化曲线、显示最终模型训练准确率、查看混淆矩阵等。

4.1.2 实验工具的实现

WCDR 模型界面的实现是基于 HTML5 语言，HTML[29]即超文本标记语言，是一种文本类、解释执行的标记语言，是一个在 Internet 上编写网页的语言，也可以用来开发前端界面。HTML 于 1993 在互联网发布，在 20 世纪 90 年代 HTML 语言快速发展，发布版本有 2.0 版、3.2 版、4.0 版，到 1999 年的 4.01 版，与另两种文本标记语言 XML 和 XHTML 不断竞争，2010 年 HTML 为扩大应用

范围、层次，不再用文档类型定义(DTD)，减少了原来严格的规范，从此大量浏览器厂商开始承认 HTML5，Google 也开始使用 HTML5。HTML5 增加了一些非常实用的功能，可以部分替代原来的 JavaScript，这使得用户制作网页的效率变得更高，在与 Python 交互的方面也可以得心应手。

由于 HTML5 在网页开发方面的强大功能，以及使用 JavaScript 调用 Python 可执行文件的方便性，因此使用 HTML5 进行 WCDR 模型界面的开发，主要使用到的功能模块如表 4.1 所示。

表 4.1　WCDR 模型界面部分功能模块

模块名称	功能说明
WCDR.select_data	选择数据
WCDR.pretreatment	数据预处理
WCDR.train	模型训练
WCDR.predict	模型预测
WCDR.print_result	结果显示

接下来详细地介绍如何使用 HTML5 实现 WCDR 模型界面的主要功能：首页界面、选择数据、数据预处理、模型训练、模型预测、显示结果。

1.选择数据：如图 4.2 所示，每一个功能界面由四个部分组成。第一部分是“WIFI 行为跨域识别”标题。第二部分是结果显示区域，其中可以显示加载的数据、训练的结果和可视化的过程。第三部分是操作区域，由选择数据、数据预处理、模型训练、模型预测和显示结果五个按钮组成。第四部分主要显示当前的状态，即在模型流程中的哪一步骤。

加载数据使用的是 html5 语言中的 file 弹窗，file 弹窗具有获取系统文件资源管理器的文件目录的功能，通过 file 弹窗可以选择要加载到 WCDR 模型中的 dat 文件，再点击 upload 按钮，python 后端会将 dat 文件加载到模型缓冲区中，并通过 txt 文件形式保存在 html 界面内。

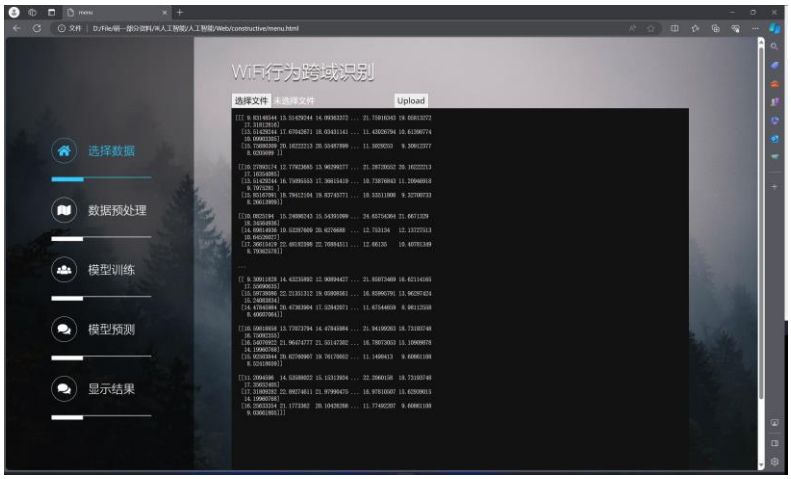


图 4.2　选择数据界面

2.数据预处理：如图 4.3 所示，点击数据预处理按钮会调用后台 Python 可执行文件，Python 后端对加载的数据进行数据集筛选和特征提取，并将结果输出成 npz 文件的形式。再次点击显示结果，可以将 npz 文件中的数据预处理完毕的 CSI 数据显示在界面中的数据显示区。

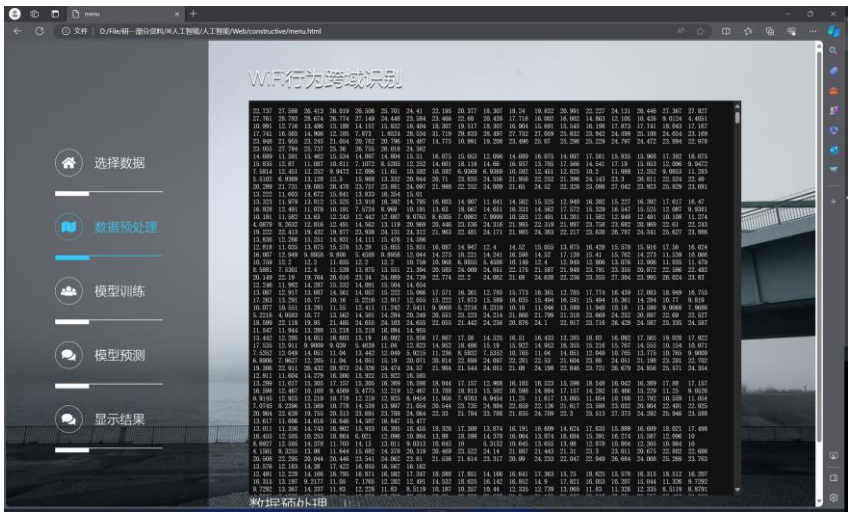


图 4.3　数据预处理界面

3.模型训练：如图 4.4 所示，模型训练就是将数据预处理后的 CSI 数据切片流输入到卷积神经网络中进行学习。Python 后端会先对 CSI 数据切片流进行更改数据格式，然后通过卷积神经网络对给定格式的数据进行迭代学习训练，得到使损失值 loss 降到最低的卷积神经网络的最佳神经元参数取值。最后模型训练和验证的准确率变化可视化输出到界面中的数据显示区域，有了准确率变化曲线，WCDR 模型对网络学习过程具有一定的可解释性。

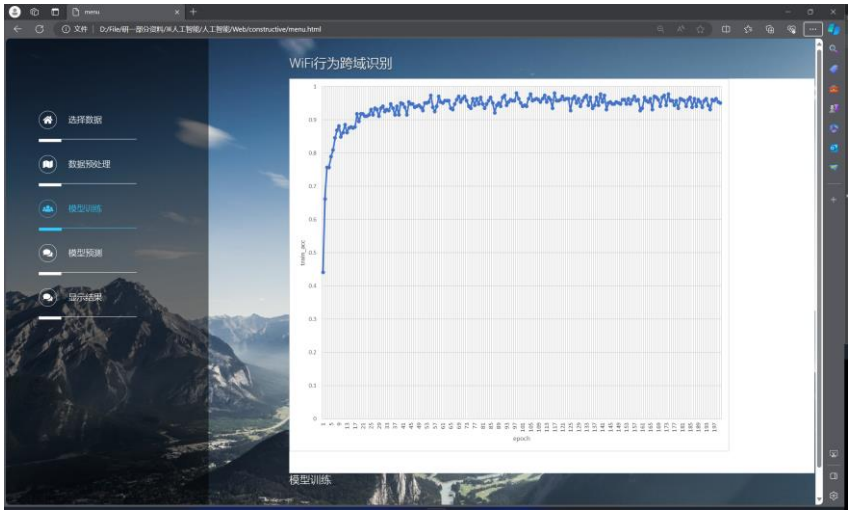


图 4.4　模型训练界面

4.模型预测：如图 4.5 所示，当卷积神经网络对 CSI 数据切片流训练完毕时，点击模型预测按钮，Python 后端就会调用 predict 函数，其过程就是用带有最优神经元参数取值的卷积神经网络对 CSI 数据切片流的测试集进行测试，卷积神经网络会输出对测试集样本的行为预测，然后 WCDR 模型会将预测的标签和样本真实标签进行比对，得出测试集上的识别准确率 accuracy 和识别损失值 loss。模型将准确率和损失值保存到文本文件中，点击显示结果按钮，可将文本中保存的跨域识别准确率的变化曲线显示到界面中的数据显示区域。

图 4.5　模型预测界面

5.显示结果：如图 4.6 所示，在每个模型模块完成后，都可以点击显示结果按钮将每一步的数据格式输出到界面的显示区域中，这使得 WCDR 模型具有一定的可解释性。当然，如果需要将过程进一步可视化，让 WCDR 模型具有可解释性，就可以在模型训练或模型预测完成后，点击显示结果按钮，Python 后端会将模型执行时中间生成的识别准确率变化图像和损失值变化图像等等相关过程结果显示到界面中的显示区域。通过 accuracy 和 loss 的变化可以看出卷积神经网络在学习样本训练的过程中的在迭代寻找最优权重使得准确率变大的同时降低损失值的一个过程，这样模型就具有了一定的可解释性。

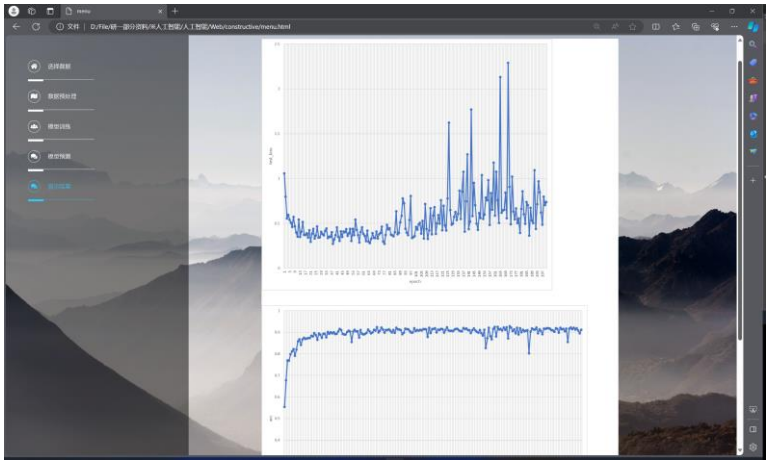


图 4.6　结果显示界面

4.2 WCDR 模型的实验与分析

4.2.1 模型参数设置

WCDR 模型通过对模型中的参数进行设定来提高模型的性能，不同的参数初值设置会对模型的性能具有提高或降低的作用。而对模型中的参数进行设置的原因就是一方面让模型发挥出最佳的性能，另一方面减小模型训练和测试所需要的时间成本。为了让模型具有最佳的性能，本节对模型中的一些重要的参数进行了预设置，若在实验中没有额外地说明，均使用本节所采用的参数设置。

模型在数据预处理阶段使用的是数据集筛选和特征提取算法。在数据集筛选中，函数 select_dataset()采用以原始 CSI 数据流的 dat 格式文件作为输入，经过多层循环来一次性选取同一个 WIFI 接收器接收到同一个动作发出者的发出动作的原始 CSI 数据流的 dat 格式文件。数据集筛选过程不需要通过大量的计算来进行优化，仅通过多层循环选出模型实验需要的训练数据集即可。在特征提取算法中，函数 extract()获取数据集筛选后的 CSI 数据流 dat 文件作为输入，函数最后返回形状为(60,180)的特征向量切片组。同样特征提取算法也不需要通过大量的迭代进行优化，仅通过一次提取已经可以提取出具有代表性的特征向量。

在跨域识别阶段，WCDR 模型使用卷积神经网络对数据预处理后的特征向量样本进行训练。经过训练确定好最优神经元参数取值，形成一个跨域稳健的卷积神经网络。接下来模型对测试集进行预测，并与真实标签相比对，确定模型预测的识别准确率和损失值。因此整个跨域识别模型中具有一定的模型参数，并在训练中卷积神经网络对其迭代优化。

WCDR 模型的训练参数如表 4.2 所示，表中设置了在训练模型中的一些参数的初始化值。

表 4.2　WCDR 模型训练参数表

参数名	参数值
WCDR	0.95
WIHF	0.9765
WiTransfer	0.99
Widar3.0	0.927
Chenetal.	-
DI	-
Nurmietal.	-
ColliGuillermoetal.	0.929

表 4.2 中有 8 个参数值。其中 epoch 代表的是模型对一个样本集的迭代训练

次数，取值为 100；batch_size 代表一个样本集中样本的数量，取值为 8；input_shape 是卷积神经网络中输入的张量形状，即 CSI 数据流切片形状，形状为(60,60,3)；output_shape 是卷积神经网络中输出的张量形状，即标签的种类数，形状为(10)；optimizer 代表的是优化器，取值为 Adam，表示模型使用 Adam 优化器，该优化器具有占内存小、高效的计算能力等特点；loss 代表的是损失函数，采用交叉熵损失函数；weight 和 bias 代表卷积神经网络中神经元的权重值和偏置值，其中 weight 采用 Xavier 初始化为正态分布的形式，bias 采用初始化为 0.0 的形式，并在训练中不断优化。此外，为了提高模型的计算速度，本实验采用 NVIDIAGeForceGTX1050Ti 高性能显卡进行计算。

4.2.2 WCDR 跨域识别模型训练

WCDR 模型中的两个重要参数是 weight 和 bias，卷积神经网络通过梯度下降和反向传播算法不断对 weight 和 bias 值进行更新，虽然这两个参数是重要的，但是在设置初始值的时候，并没有显得那么重要。因为在卷积神经网络进行数据训练的时候，网络会在不断减小训练损失的过程中不断更新 weight 和 bias 两个参数的取值，最后确定出可以使卷积神经网络损失值降到最低的同时识别准确率达到最高的一组 weight-bias 取值。因此，对网络起到重要作用的参数由 weight-bias 变成了 epoch 和 batch_size 的初始值。

为了确定最优的 epoch 取值，在设置 epoch 的参数时，从 10 开始，一直不断增加到 200。对其他参数采用控制变量法，即只更改 epoch 的取值，观测最终模型可以在同躯体位置域和同面部方向域下达到的识别损失，并记录不同 epoch 情况下对应的模型损失值的变化，模型损失值随 epoch 取值的变化曲线如图 4.7 所示。

图4.7　模型在两种训练集上不同 epoch 取值的训练损失变化

从图 4.7 可以看出，在同躯体位置域的训练集上，epoch 从 10 到 200 的 20 组不同的取值中，卷积神经网络的损失值随着 epoch 的不断增大而渐渐减小，最终模型的训练损失会从 0.076 降低到 0.019，可以看出不断增大 epoch 的取值，会不断减小模型的训练损失。虽然 epoch 从 200 继续增大会继续减小训练损失，但是训练时间已经达到了 8 小时，可以看出，评价模型性能并不可以只从模型的损失值的高低来判定，还需要综合其他因素考虑，如模型消耗的时间成本。而在同面部方向域的训练集上，epoch 从 10 到 200 的 20 组不同的取值中，最终模型的训练损失会从 0.067 降低到 0.011。重要的是，在 epoch 从 170 增大到 200 的过程中，模型的训练损失从 0.012 增大到 0.016 再减小到 0.011，虽然没有后续实验，但是已经可以看出模型已经出现了过拟合现象，然而 epoch 取 170 时前一个训练集却处于欠拟合状态。综合两种训练集来考虑，将 epoch 取为 200 是一个很好的选择，既保证了较低的损失值，也节省了大量的模型时间成本。

下面将详细说明 WCDR 模型在不同 epoch 取值的情况下，模型所消耗的时间成本的变化情况，即在其他参数均控制不变的情况下，记录模型训练所需要的时间，时间变化如表 4.3 所示。

表4.3　不同 epoch 参数下模型训练时间变化表

epoch	time/h
100	4.0
110	4.4
120	4.8
130	5.2
140	5.6
150	6.0
160	6.4
170	6.8
180	7.2
190	7.6
200	8.0

表 4.3 中记录的时间是指一次 WCDR 模型对所有 CSI 数据流样本进行训练、验证所需要的时间，单位是小时。使用模型进行 WIFI 行为跨域识别时所消耗的时间成本主要花费在模型训练上，因此模型的训练时间对整个模型的运行时间成本具有重要的影响。需要特殊说明的是，在 epoch 值取小于 100 时，模型的时间也是呈正比增加，但是模型在该组 epoch 取值下的识别准确率较低并且损失值较高，所以不予列出。即表 4.3 只列出了在模型训练损失、验证损失较低的情况下(一般不高于 0.020)对应的 epoch 取值下的模型训练所花时间成本。

从表中可以看出 epoch 取 200 时，模型训练时间达到 8 小时，这虽然会让实验者们等待一段漫长的时间，但是会达到一个良好的识别效果。当 epoch 从 170 增大到 200 时可以不仅出现了过拟合现象，而且模型所花费的时间成本也在线性增大，因此 200 是可以设置为模型训练过程中 epoch 参数的最大接受值。另一方面，虽然 epoch 取小于 200 时模型花费更少的时间成本，但是模型训练损失、验证损失均在一个较高的水平，很明显也不适合作为模型训练过程中的参数。为了可以更直观地观察到模型花费时间成本随 epoch 取值的变化，将表 4.3 绘制到图 4.8 中。综合模型训练的时间成本和模型训练、验证损失考虑，将 epoch 设为 200 是最优解。

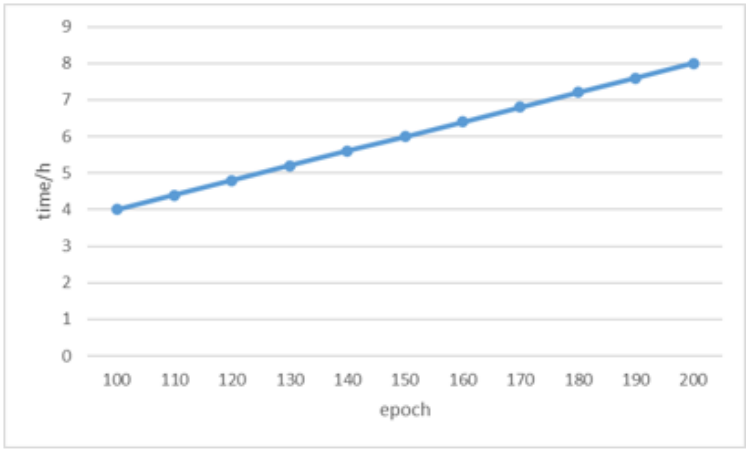
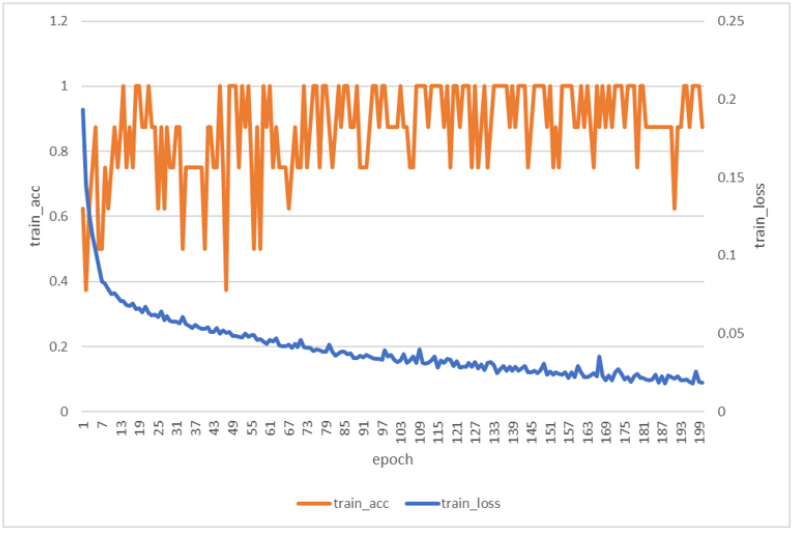
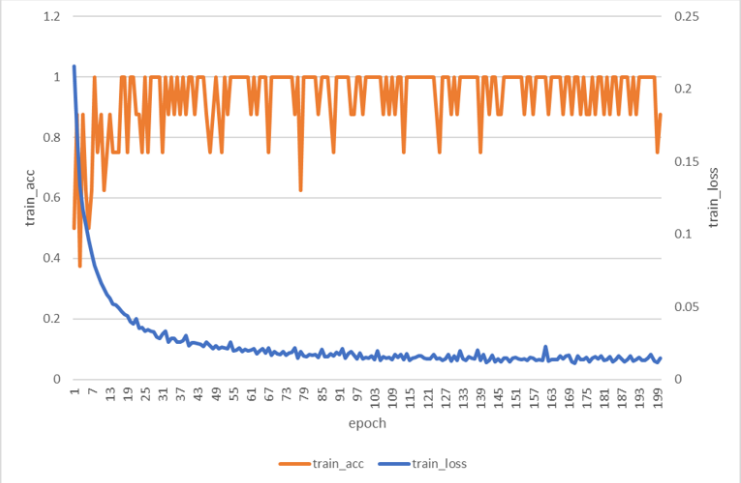


图4.8　模型训练时间成本随 epoch 取值变化曲线

为了从整体上更加直观地看到实验中模型的训练效果，图 4.9(a)和(b)分别展示了实验中在同躯体位置域和同面部方向域的训练集上的训练损失和训练准确率的变化。



(a) 同躯体位置域中训练损失值变化曲线和识别准确率变化曲线



(b) 同面部方向域中训练损失值变化曲线和识别准确率变化曲线

图4.9　模型训练中损失值和识别准确率变化曲线

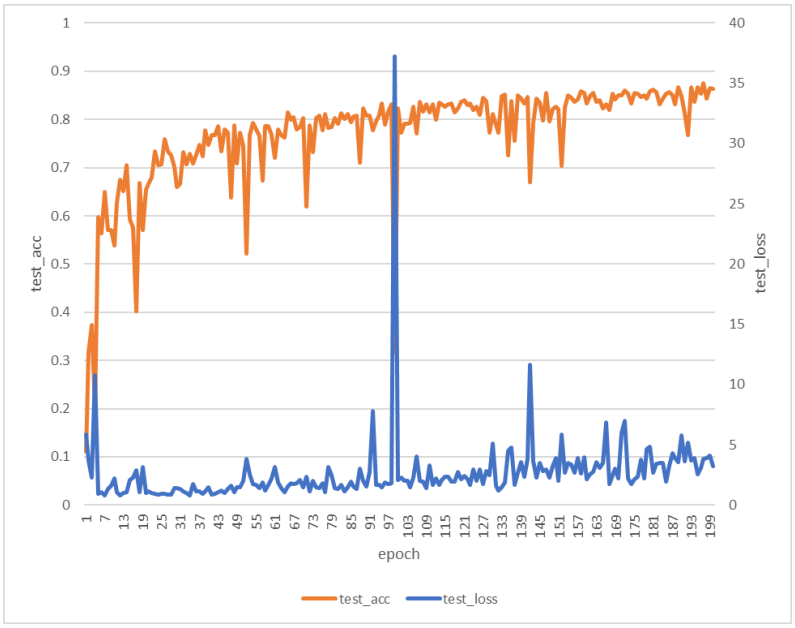
图 4.9 中的(a)图是模型在同躯体位置域下训练和验证损失和识别准确率的变化曲线，可以看出，在模型训练并进行 200 次训练验证的过程中，模型的损失值随训练迭代进行而降低并稳定到 0.01817 左右，模型的识别准确率随训练迭代进行而升高并稳定在 95%左右。图 4.9 中的(b)图是模型在同面部方向域下训练和验证损失和识别准确率的变化曲线，同样可以看出，在模型训练并进行 200 次训练验证的过程中，模型的损失值随训练迭代进行而降低并稳定到 0.01123 左右，模型的识别准确率随训练迭代进行而升高并稳定在 95%左右。

4.2.3 使用 WCDR 进行跨域识别预测

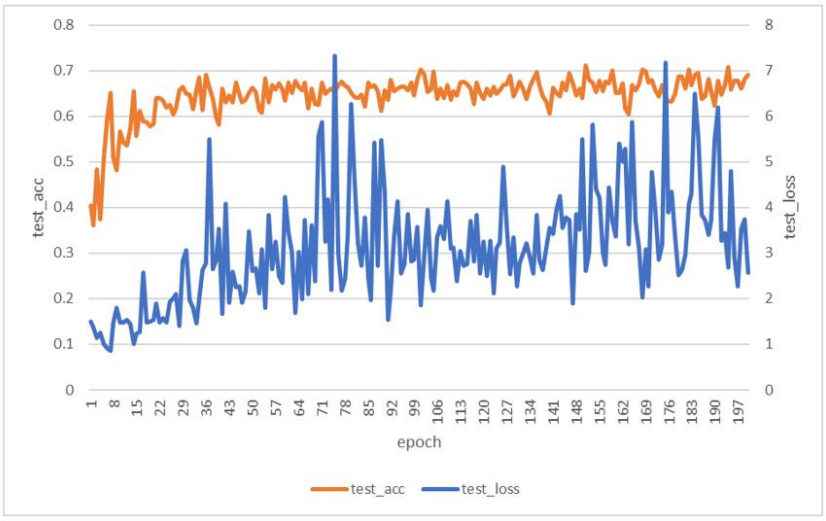
在面向 WIFI 行为跨域识别的领域中，有很多现有方法可以实现跨域识别，如基于渐进式注意力和分块遮挡的跨域行人重识别[30]。在这些方法中，每个方法都有自己的优势和局限性。总的来说，一个公平良好的跨域识别算法应该具有两条性质：效果和稳健性。算法要有效果是说算法要可以实现在更少的时间成本下将损失降到更低，同时保持更高的准确率。算法要稳健是说算法不仅仅在一个目标域下具有良好的效果，还要在不同的目标域下也能实现良好的效果。本节将对模型的效果和稳健性两方面对 WCDR 模型进行测试。

具体实验过程是将在 Widar3.0 数据集中 CSI 数据集上训练好的卷积神经网络迁移

到不同躯体位置和面部方向领域下获取的 CSI 数据上进行跨域行为识别任务，由于人体位置、人体朝向这两种领域因素不同，实现跨域识别。为了从整体上更加直观地看到实验中模型的测试效果，图 4.10(a)和(b)分别展示了实验中模型在躯体位置和面部方向上进行跨域的识别损失和跨域识别准确率的变化。



(a) 躯体位置上跨域的损失值变化曲线和识别准确率变化曲线



(b) 面部方向上跨域的损失值变化曲线和识别准确率变化曲线

图4.10　模型在跨域识别中损失值和识别准确率变化曲线

实验对 Widar3.0 数据集中 CSI 数据中在 user1 中不同躯体位置和面部方向上的 10 种行为进行了跨域测试，并记录 200 次迭代测试中对用户 user1 的人体行为进行跨域识别的预测损失和预测准确率。从图 4.10(a)可以看出模型在躯体位置上进行跨域的测试损失和测试准确率的变化趋势。模型经过 200 次迭代训练后，进行跨域识别的准确率以大概率稳定在 88%左右。从图 4.10(b)可以看出模型在面部方向上进行跨域的测试损失和测试准确率的变化趋势。不难得出，WCDR 模型在第 170 次迭代训练时出现了过拟合现象，具体表现为模型在测试集上的损失值出现了剧烈的抖动。但是从跨域识别测试的准确率变化曲线中又可以看出模型可以在过拟合阶段的识别准确率以大概率稳定在 71%左右。综上所述，实验证明 WCDR 模型在躯体位置和面部方向上进行跨域识别内具有良好

的效果和较强的鲁棒性。

4.2.4 模型性能分析

设计 WCDR 模型的出发点就是实现跨域识别，因此模型具有较强的泛化性。实验在 Widar3.0 数据集上不同领域因素之间测试准确率保持较高的水平，模型具有较好的识别效果。此外，模型实现了迁移学习，这通常比从零开始学习的方式更好[21]，具体表现为初始模型的性能要比随机初始化的模型要好，学习速度更快、收敛更快，最终性能更好、泛化性更强。总体来说，模型具有较强的综合性能。

4.2.5 实验结果分析

表 4.4 分别从训练准确率、测试准确率和总体准确率三个方面展示了 WCDR 模型和一些现有模型的性能比较，其中总体准确率代表了介绍该模型的论文未指出该准确率是训练还是测试的准确率。表 4.4 将每一中准确率中最高的准确率加粗显示。从表中不难看出，WiTransfer 模型实现了最高训练准确率，Wi-SL 模型实现了最高总体准确率，而在给出跨域识别准确率的模型中，WCDR 模型实现了较高的跨域准确率，也即 WCDR 模型具有较强的跨域鲁棒性。

表4.4　WCDR 与现有模型跨域识别效果对比

模型名称	训练准确率	跨域准确率	平均准确率
WCDR	0.95	0.8762、0.7117	-
WIHF	0.9765	0.8967	-
WiTransfer	0.99	0.9156	-
Widar3.0	0.927	0.826-0.924	-
Chenetal.	-	-	<=0.95
DI	-	-	0.9445
Nurmietal.	-	-	0.8
Siamaketal.	0.928	0.75	-
ZhangMenghuanetal.	0.96	-	-
FaSee	0.9475	-	-
YanYulingetal.	0.8126	-	-
WiDG	0.957	-	-
WiGAN	0.956	-	-
Wi-SL	-	-	0.958
ColliGuillermoetal.	0.929	-	-

实验证明，模型参数的取值对模型的性能具有非常重要的影响，所以要将模型的参数设置为能使模型在综合条件下最优化的取值组合。通过实验数据可

以得到以下结论，使用 WCDR 模型进行 WIFI 行为跨域识别时具有较高的识别准确率和较强的泛化能力，模型对于躯体位置和面部方向这两种领域因素上不具有严重的依赖性，可以作为跨域识别的算法进行扩展。

实验也体现出了 WCDR 模型的一些不足。首先模型所耗费的训练时间成本仍然较高，从模型读取数据、预处理到模型训练这一过程相比其他模型仍然具有较高的时间成本。其次模型跨域准确率当前仅为 87.62%和 71.17%，仍有较大的提升空间。最后，由于模型在训练域中出现了过拟合现象，仍需继续优化来消除模型训练中出现的过拟合现象。

结 论

本文充分地调研了面向 WIFI 信号的人体行为跨域识别研究领域的背景、研究意义和相关工作，并对相关解决方法的优缺点进行了详细地总结，进而提出了基于卷积神经网络的人体行为跨域识别模型 WCDR。该模型将人体行为跨域识别分为两个阶段：数据预处理阶段和跨域识别阶段。其中数据预处理阶段包括数据集筛选和特征提取算法，在跨域识别阶段本文设计了一个带有 7 层卷积层的卷积神经网络。通过上述两个阶段，WCDR 模型在模型性能上取得了良好的效果。

本文将 Widar3.0 上的 CSI 数据集对 WCDR 模型进行训练，通过实验探究将模型性能发挥最好的参数组合，在模型训练时间成本和模型训练效果上对 WCDR 进行了评估。此外，本文将不同躯体位置和面部方向领域下的 Widar3.0 中 CSI 数据集对 WCDR 模型进行测试，通过迁移学习进行实验探究模型的跨域能力和鲁棒性。实验结果表明，WCDR 模型具有较高的人体行为识别准确率并对跨域识别具有一定的鲁棒性。

总体来说，在跨域识别领域的现有模型有很多，在使用 WIFI 信号识别人体行为的方法也有很多，本文正是借鉴两个领域中性能较好的方法，结合机器学习方法设计出了一个跨域识别模型，并且达到了一定的效果，但是也有待提高的地方：模型训练时间成本仍然较高，模型跨域识别的准确率也仍有提升的空间。因此，在未来中，应该注重探索能缩减模型时间成本的同时降低拟合程度来提高跨域识别准确率的方法。

参 考 文 献

[1] Chenning L, Manni L, Zhichao C. WiHF: Gesture and User Recognition With WiFi[J]. 2022,21(2):757-768. DOI:10.1109/TMC.2020.3009561.

[2] Yi Z, Yue Z, Kun Q, et al. Widar3.0: Zero-Effort Cross-Domain in Gesture Recognition with Wi-Fi[J]. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 2021.

[3] Kaixuan C, Lina Y, Dalin Z, et al. Distributionally Robust Semi-Supervised Learning for People-Centric Sensing[C]. The Thirty-Third AAAI Conference on Artificial Intelligence, Honolulu, 2019:3321-3328.

[4] Jianwei L, Jinsong H, Feng L, et al. Adversary Helps: Gradient-based Device-Free Domain-Independent Gesture Recognition[C/OL]// Computer Vision and Pattern Recognition, Washington State Convention Center, Seattle, Wshington, June, 2020[2022,01,07]. https://arxiv.org/abs/2004.03961.

[5] Yuanrun F, Biyun S, Haiyan W, et al. WiTransfer: A Cross-scene Transfer Activity Recognition System Using WiFi[C/OL]//ACM Turing Celebration Conference - China, Hefei, China, May 22-24, 2020[2022,01,07]. https://doi.org/10.1145/3393527.3393538.

[6] Yue Z, Yi Z, Kun Q, et al. Zero-Effort Cross-Domain Gesture Recognition with Wi-Fi[C/OL]//The 17th Annual International Conference on Mobile Systems, Applications and Services, Seoul, Republic of Korea, June 17-21, 2019[2022,01,07]. https://doi.org/10.1145/3307334.3326081.

[7] Jie Z, Petteri N, Zheng W, et al. CrossSense: Towards Cross-Site and Large-Scale WiFi Sensing[C/OL]//The 24th Annual International Conference on Mobile Computing and Networking, New Delhi, India, October 29-November 2, 2018: Session: Take Me Back to School: Learning and Sensing[2022,01,07]. https://doi.org/10.1145/3241539.3241570.

[8] Siamak Y, Hirokazu N, Sankalp D, et al. A Survey of Human Activity Recognition Using WiFi CSI[EB/OL]. (2017,08,23)[2022,04,23].

[9] 张梦欢, 王亚刚. 基于 CNN 和超声传感的手势识别及辅助身份认证[J]. 传感器与微系统,2022,41(05):110-113+117.

[10] Yu G, Jiang'an L. A Novel WiFi Gesture Recognition Method Based on CNN-LSTM and Channel Attention[C]. Proceedings of 2021 3rd International Conference on Advanced Information Science and System(AISS 2021), Sanya, China, 2021:497-500.

[11] Xianjia M, Lin F, Hao C, et al. Just-in-Time Human Gesture Recognition Using WiFi Signals[J]. Chinese Journal of Electronics, 2021,30(6):1111-1119.

[12] Yuling Y, Lijun Z, Minye C. AGRMTS: A virtual aircraft maintenance training system using

gesture recognition based on PSO-BPNN model[J]. Computer Animation and Virtual Worlds, 2021,33(1):e2031.

[13]Zhengjie W, Xue S, Jingwen F, et al. WiDG: An Air Hand Gesture Recognition System Based on CSI and Deep Learning[C]. 第 33 届中国控制与决策会议论文集（8）, Kunming, China, 2021:450-455.

[14]Jianchun G, Jiannuan G, Lili W. Gesture Recognition Method Based on Attention Mechanism for Complex Background[J]. Journal of Physics: Conference Series, 2021,1873(1):012009.

[15]Dehao J, Mingqi L, Chunling X. WiGAN: A WiFi Based Gesture Recognition System with GANs[J]. Sensors, 2020,20(17):4757.

[16]Zhanjun H, Yu D, Xiaochao D, et al. Wi-SL: Contactless Fine-Grained Gesture Recognition Uses Channel State Information[J]. Sensors, 2020,20(14):4025.

[17]Colli G, Trejos L. User-Independent Hand Gesture Recognition Classification Models Using Sensor Fusion[J]. Sensors, 2022, 22(4):1321.

[18]Siamak Y, Hirokazu N, Sankalp D, et al. A Survey on Behaviour Recognition Using WiFi Channel State Information[J]. IEEE COMMUNICATION MAGAZINE, 2017,7:155986-156024.

[19]佚名. 蛋形矮胖子_eecb.阅读笔记：《Understanding and Modeling of WiFi Signal Based Human Activity Recognition》[EB/OL]. (2019,09,26)[2022,04,12]. https://www.jiansu.com/p/0b275d75d81b.

[20]Microstrong. 主成分分析（PCA）原理详解[EB/OL]. (2018,06,08)[2022,04,12]. https://zhuanlan.zhihu.com/p/37777074.

[21]邱锡鹏.神经网络与深度学习[M].北京:机械工业出版社,2020.

[22]Ethem A..Introduction to Machine Learning[M].3rd ed.北京:机械工业出版社,2015:27-29.

[23]Prateek J..Python 机器学习经典案例[M].北京:人民邮电出版社,2017:31-34.

[24]于营, 杨婷婷, 杨博雄. 混淆矩阵分类性能评价及 Python 实现[J]. 现代计算机,2021(20):70-73+79.

[25]罗冬日.TensorFlow 入门与实战[M].北京:人民邮电出版社,2018:64-72.

[26]Malik S., Muhammad K., Muhammad T., et al. LSTM Enabled Artificial Intelligent Smart Gardening System[J]. RACS '20: Proceedings of the International Conference on Research in Adaptive and Convergent Systems, 2020,136-141.

[27]Vishnu S..Pytorch 深度学习[M].北京:人民邮电出版社,2019:123-125.

[28]Qirong B., Xingxia M., Jingzhao H., et al. TransferSense: towards environment independent and one-shot wifi sensing.[J/OL]. Pers Ubiquit Comput, 2021[2022-04-22]. https://doi.org/10.1007/s00779-020-01480-6.

[29]明日科技.HTML5 从入门到精通[M].北京:清华大学出版社,2019:3-5.

[30]李云龙, 程德强, 等. 基于渐进式注意力和分块遮挡的跨域行人重识别[J/OL]. 北京航空航天大学学报:1-13, 2022[2022-05-18]. DOI:10.13700/j.bh.1001-5965.2022.0025.