

基于卷积神经网络的雷达人体动作与身份多任务识别

侯春萍, 蒋天丽, 郎玥*, 杨阳

天津大学电气自动化与信息工程学院, 天津 300072

摘要 为弥补单任务识别未充分利用相关任务监督信息的缺陷, 提出了一种基于卷积神经网络的多任务识别模型。该模型引入注意力机制, 对任务共享层的特征进行重校正, 并结合多尺度结构进行特征融合, 最后在任务特定层上进行多任务识别。针对共享特征空间内类分布不紧凑导致的模型泛化性能降低问题, 本文在模型中引入中心损失函数与均方误差损失函数, 与传统的交叉熵损失函数相结合, 共同优化模型。实验结果表明: 所提模型在人体 6 个动作类别和 15 个身份类别上的最高识别准确率分别可达 100% 和 99.93%, 两种任务上识别的总准确率可达 99.93%, 均优于任务独立识别时的各项准确率, 说明所提模型能更有效地同时完成人体动作及身份识别任务。

关键词 图像处理; 动作识别; 身份识别; 雷达时频图像; 卷积神经网络; 多任务

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP57.021009

Human Activity and Identity Multi-Task Recognition Based on Convolutional Neural Network Using Doppler Radar

Hou Chunping, Jiang Tianli, Lang Yue*, Yang Yang

School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China

Abstract A multitask recognition model based on convolutional neural network is proposed to avoid single task recognition ignoring supervision information of related tasks. The proposed model introduces an attention mechanism to perform feature recalibration of the task shared layer and combines the multiscale structure for feature fusion. Finally, multi-task recognition is performed on the task-specific layers. Center loss and mean square error loss functions are employed together with the traditional cross entropy loss function to solve the generalization degradation problem caused by uncompact class distribution in the shared feature space. Experimental results on 6 human activities and 15 identities show that the model can achieve the maximum recognition accuracies of 100% and 99.93% on each task, respectively, and the multitask accuracy is up to 99.93%. The results are better than those obtained by the single task models. This shows that the model can simultaneously perform human activity and identity recognition more effectively.

Key words image processing; activity recognition; identity recognition; radar time-frequency spectrograms; convolutional neural network; multi-task

OCIS codes 100.3008; 100.4996; 100.5010; 150.1135

1 引言

利用雷达进行人体动作识别与身份识别是近年来的研究热点, 在灾难营救、医疗监控等领域都有广泛的应用前景^[1]。雷达具有高距离分辨率和强穿透力, 能以非接触的方式对人体目标进行探测^[2]。与其他传统人体行为探测方式^[3-5]相比, 雷达探测不仅

不受环境条件限制, 还可抑制杂波, 更利于检测目标人体微小而复杂的行为变化。人体躯干及四肢的微弱动作会在雷达信号回波上引起微多普勒频移, 这种现象可反映在雷达信号的可视化谱图上^[6], 谱图中的微多普勒特征可被用于人体的动作识别与身份识别。

基于雷达谱图的人体动作识别与身份识别任务

收稿日期: 2019-06-10; 修回日期: 2019-06-28; 录用日期: 2019-07-01

基金项目: 国家自然科学基金(61520106002, 61731003)

* E-mail: langyue@tju.edu.cn

已受到广泛关注。2009年, Kim等^[7]分析了雷达时频图像特点, 提取时变特征实现了人体动作分类。2015年, Ricci等^[8]利用形状谱特征识别不同的人体身份。这些传统分类方法虽然取得了较好效果^[9-11], 但是需要人工提取判别特征, 因此在其他任务上应用性不强。卷积神经网络(CNN)是识别领域中应用最广泛的深度学习模型, 它可以自动学习图像特征完成识别。2016年, Kim等^[12]率先将CNN应用于雷达人体动作识别。2018年, 基于微多普勒雷达图像, Cao等^[13]利用CNN实现了4个人的身份识别。同年, Chen等^[14]在CNN浅层共享多基地雷达的各方向信息, 提升了人体动作识别与步态识别的效果, 但多基地雷达成本高, 不利于实际应用。这些深度学习方法证明了CNN模型在雷达人体目标识别相关任务上的优越性^[15-17]。但为了达到当前任务的最佳性能, 其都是优化单一模型, 易导致单任务上的过拟合。

多任务学习是一种归纳迁移机制, 它通过学习多个任务上的共享知识来获得更好的泛化性能, 可提升模型在多个任务上的有效性。文献^[18]将自底向上的动态加宽网络用于实现多任务学习, 但算法难以达到全局最优。文献^[19]提出的深度关系网络通过对全连接层(FC)添加矩阵先验知识来学习任务间的关系, 但在新任务上的应用性不强。多任务学习在自然语言处理^[20]、实例分割^[21]等领域都有广泛应用, 雷达时频图像上的多任务识别方法仍有

待研究。

本文提出了一种基于CNN的雷达人体动作与身份多任务识别方法。利用单基地雷达采集的人体回波信号, 生成雷达时频图像数据集。构建多任务识别模型, 通过对CNN特征的重校正和多尺度融合, 进行两个相关识别任务的共享特征表示, 最后分别在各自的特定任务层完成相关目标的识别。此外, 在传统交叉熵损失函数基础上引入中心损失函数和均方误差损失函数, 共同优化多任务模型参数。本文所提算法基于深度学习, 采用雷达实测数据进行模型训练来完成识别任务, 避免了对不同人体进行建模的复杂过程, 更具有实际应用意义, 其在雷达时频图像数据集上的识别结果验证了本方法的有效性。

2 雷达时频图像数据集构建

雷达时频图像数据集的构建流程如图1所示。首先对受试者的各类别动作进行实测数据采集, 得到原始的人体回波信号。然后对回波信号进行短时傅里叶变换(STFT), 生成时频图像^[22]。为探究噪声对识别的影响, 在回波数据上加入不同量级的高斯白噪声, 将有噪声的回波数据进行特定窗长下的STFT, 获得具有5种信噪比的雷达时频图像; 为探究STFT窗长对识别的影响, 对无噪声回波数据在不同窗长下进行STFT, 获得对应5种窗长的雷达时频图像。最后对时频图像进行身份及动作的类别标定, 得到雷达时频图像数据集。

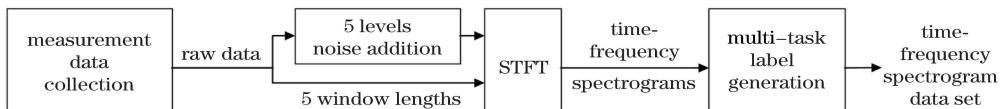


图1 雷达时频图像数据集构建流程

Fig. 1 Flow chart of building radar time-frequency image data set

2.1 数据采集

本文使用超带宽雷达模块 PulsON 440 进行人体行为的实测数据采集。雷达的工作频率为 3.1 GHz 到 4.8 GHz, 采样频率设置为 16 GHz。采集实验在室内环境下进行。将两个定向天线的高度调整为 1.2 m, 用其发射雷达信号并接收人体回波信号。实验共有十五位被试者参与。在雷达的径向方向上对每位受试者的 6 种典型人体动作进行采集, 所选取的动作分别为行走、拳击、地面爬行、潜行、前向立定跳、跑步。每种动作由每位被试者重复 10~15 次, 每次采集时间为 7 s 左右。

2.2 雷达时频图像生成

对采集的雷达原始回波信号进行 1024 点的

STFT 可得到雷达时频图像。STFT 将非平稳过程看作一系列短时平稳信号的叠加, 先在时间维度上对信号进行加窗, 再对窗内信号作傅立叶变换, 得到信号的时变频谱, 表达式为

$$X(\tau, \omega) = \int_{-\infty}^{\infty} x(t)w(t - \tau)\exp(-j\omega t) dt, \quad (1)$$

式中: $w(t)$ 为窗函数; $x(t)$ 为雷达人体回波信号; t 为时变信号的时间信息; τ 为信号被截取的某个特定时间段。

为探究噪声对识别算法的影响, 在雷达回波信号中加入均值为零的高斯白噪声得到 0 dB、10 dB、20 dB、-10 dB、-20 dB 这 5 种信噪比的有噪声信号, 对有噪声信号采用长度为 1.0 s 的滑动窗口生成

对应各噪声等级的 5 组时频图像;为探究 STFT 窗长对识别算法的影响,对无噪声信号分别采用长度为 0.2,0.4,0.6,0.8,1.0 s 的滑动窗口生成对应各窗长的 5 组时频图像。最后对生成的时频图像进行身份类别标定和动作类别标定。STFT 慢时间窗口长

度为 1.0 s 时,6 种人体动作对应的无噪雷达时频图像如图 2 所示。时频图像中,不同的颜色代表不同的雷达回波强度,其中最强的回波来自人体躯干,围绕躯干的微多普勒调制则来自于人体手臂和腿部的运动。

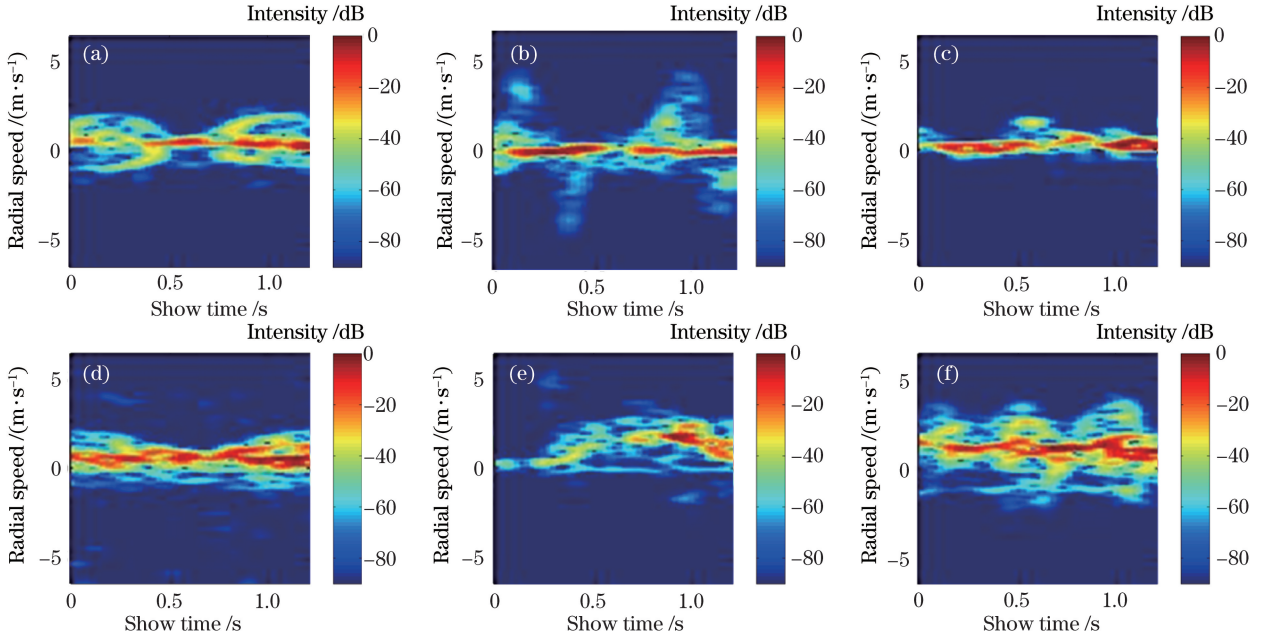


图 2 6 种人体动作下的雷达时频谱图。(a)行走;(b)拳击;(c)地面爬行;(d)潜行;(e)前向立定跳;(f)跑步
Fig. 2 Radar time-frequency spectrograms of radar for six human activities. (a) Walking; (b) boxing; (c) crawling on the ground; (d) diving; (e) standing forward and jumping; (f) running

在生成的雷达时频谱图中,具有相同信噪比和相同 STFT 窗口长度的每种图像数量为 180000 张。在 180000 张的单种图像数据中,每个受试者的单个动作种类下含雷达时频图像 2000 张。每张图像对应一个身份类别标签和一个动作类别标签,尺寸为 120 pixel×120 pixel。

3 基于 CNN 的雷达人体动作与身份多任务识别

3.1 基于注意机制的特征重校正结构

注意力机制的本质是将处理资源分配给更具价值的输入信息,并抑制无用信息以提高任务处理的效率与准确性。压缩激励(SE)模块^[23]是一种基于特征通道的注意力加权操作,可以通过筛选特征信息来获取更有效的共享特征表示,从而提升多任务识别效果,其结构如图 3 所示,其中 X 表示输入的特征图, Y 表示输出后的特征图。SE 模块在卷积层(Conv)输出的特征图上提取全局信息,并据此建立各通道之间的依赖关系,在学习过程中自适应地生成注意力权重,从而加强和抑制各通道信息,实现特

征的重校正。

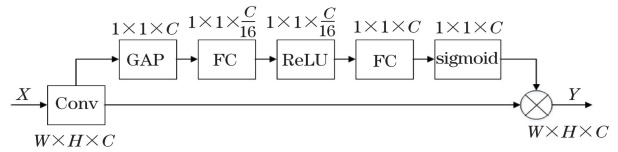


图 3 SE 模块结构示意图

Fig. 3 Structural diagram of SE module

SE 模块对维度为 $W \times H \times C$ 的卷积层输出特征图进行全局池化(GAP),即通过全局平均将特征图每个通道上的二维特征表征为一个实数值,从而压缩各通道的特征图。所得输出特征的维度与特征图的通道数 C 相对应,每个通道上的输出值都具有全局的感受野,表示该通道的全局特征信息。

本结构还引入了源于循环神经网络的门激励机制^[24],通过全连接层学习各通道之间的非互斥关系,并将激活函数作为门控,用于获取每个通道上的激励。SE 模块利用一个全连接层将输入特征的维度降低为原来的 1/16,使用一个修正线性单元(ReLU)激活后,利用第二个全连接层将特征升维至原始输入维度。两个全连接层构成的瓶颈结

构使特征权重生成过程中的模型参数量减少,降低了计算复杂度,且增加了模型的非线性。采用 sigmoid 激活函数在网络的末端生成各通道权重,此权重值与可学习的模型参数相关,通过对卷积后的特征图进行乘法加权来自动完成特征重校正。

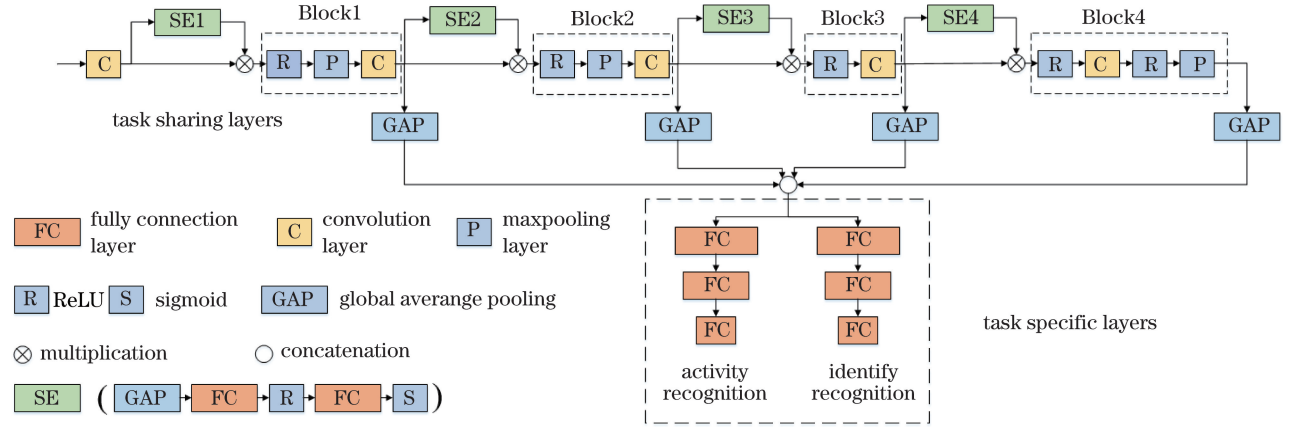


图4 基于CNN的多任务识别模型结构

Fig. 4 Multi-task recognition model structure based on CNN

任务共享部分以传统的CNN结构为基础进行特征提取,含有五个卷积层,三个池化层。第一个卷积层使用96个 11×11 大小的卷积核;第二个卷积层使用256个 5×5 大小的卷积核;第三、四、五个卷积层分别使用384、384、256个 3×3 大小的卷积核。池化层都以 3×3 的窗口尺寸为基准进行最大池化操作。后四个卷积层分别与池化层和ReLU构成4个块(Block),由浅到深地提取微多普勒图像上不同层次的特征。保留最后一个卷积层的输出,在网络的前4个卷积层之后都加入SE模块,通过学习权重参数,对各层级的特征图进行重校正,提升相关特征的重要性,同时抑制不重要的特征。包含后四个卷积层的Block1~Block4分别放置于每个重校正结构之后,对校正后的特征图进行进一步抽象处理。

网络浅层所输出的底层特征具有雷达时频图像上的细节信息,如几何特征等,但缺少了上下文信息。网络深层输出的高层特征包含更抽象的语义信息,但丢失了较多空间位置信息。本文方法使用多尺度特征融合结构,保留多尺度特征图上的各通道全局信息,对CNN模型生成的各尺度特征进行串行连接(concat),丰富了共享特征的层次。对Block1~Block4输出的四个尺度特征图分别进行全局池化操作,可以得到四个特征向量,特征向量表征了特征图各个通道上的全局信息,其

3.2 基于CNN的多任务识别模型

基于CNN的多任务识别模型由任务共享层与任务特定层构成,具体结构如图4所示。任务共享层利用特征重校正结构和多尺度特征融合结构得到动作识别任务与身份识别任务的共享特征;任务特定层以多任务的共享特征为输入,在各自的任务层上完成识别。

维数与该尺度特征图的通道数相同。将四个尺度下的特征向量沿各自的通道维度进行串行连接。多尺度特征融合后的特征向量即为任务共享层输出的共享特征。

任务特定部分由每个任务上的三层全连接层构成。第一层全连接层输入1280维的特征向量,经过多次实验将输出维度设为96,第二层全连接层输入输出维度相同,第三层全连接层输入的特征向量维度为96,输出为相应任务类别上的识别结果。Dropout是一种随机失活操作。在深度学习网络的训练过程中,Dropout按照一定的概率将神经网络单元暂时从网络中丢弃。每两个全连接层中间添加ReLU和Dropout以增加非线性和防止过拟合。人体动作识别与身份识别都以多任务的共享特征为输入,在对应任务的特定层上获取识别结果。

3.3 多任务识别的联合损失函数

多任务识别的联合损失函数由交叉熵损失函数、中心损失函数、均方差损失函数构成。三种损失函数的联合监督扩大了样本在特征空间中的类间距离,减小了类内距离,均化了样本在识别上的错分概率,从而使多任务模型具有更强的识别能力。

交叉熵损失 L_E 是分类问题中广泛使用的损失函数,表达式为

$$L_E = - \sum_{i=1}^m \ln \frac{\exp(\mathbf{W}_j^T \mathbf{x}_i + b_j)}{\sum_{j=1}^n \exp(\mathbf{W}_j^T \mathbf{x}_i + b_j)}, \quad (2)$$

式中: m 和 n 代表每个训练批次的样本数和样本类别数; \mathbf{x}_i 为类别 j 下的样本 i 生成的特征矩阵; \mathbf{W} 、 \mathbf{b} 为模型的参数矩阵和偏置项。交叉熵损失通过简单的计算结构衡量类分布之间的差异,可避免梯度弥散。在训练过程中,交叉熵损失函数仅注重于扩大样本的类间距离,但对相同类别下的样本分布没有进行约束,这种优化方式得到的样本分布可能存在类内间距大于类间间距的情况。为优化样本分布,在交叉熵损失函数的基础上加入中心损失函数与均方差损失函数,以进行联合优化。

中心损失函数 $^{[25]}L_C$ 在单个识别任务的每个类别 j 上动态更新一个类中心 \mathbf{c}_j ,根据样本 i 与其类中心的距离计算中心损失函数,其表达式为

$$L_C = \frac{1}{2} \sum_{i=1}^m \|\mathbf{x}_i - \mathbf{c}_j\|_2^2, \quad (3)$$

类分布在损失优化下的变化过程可表示为

$$\frac{\partial L_C}{\partial x_{ij}} = x_{ij} - c_{ij}, \quad (4)$$

式中: x_{ij} 、 c_{ij} 表示矩阵 \mathbf{x}_i 和类中心 \mathbf{c}_j 的矩阵元。

最小化中心损失有利于类内样本的聚合,可以形成更紧凑的类分布。训练数据量较大时,无法直接获取类中心,因此先对类中心进行一个随机初始化。在计算每个类中心的更新量时,只考虑和中心类别相同的样本。以 α 作为类中心的学习效率,通过多次迭代对类中心进行更新。

均方差损失函数 L_M 通常在回归问题中用于计算预测值与真实值间的距离。在识别任务中,可利用均方差损失函数来衡量模型的预测分布与真实分布之间的差异。在优化过程中,均方差损失函数与非样本所属类别的预测结果相关,可均化错分类别。将样本 i 的真实标签进行独热编码,得到的编码向量 \mathbf{h}_i 视为真实分布,维度等于识别任务中的类别数。向量仅在样本真实类别所对应的位上取1,其余位为0。模型在特定识别任务上对样本 i 的输出 \mathbf{z}_i 是各类别上的预测概率,视为预测分布。故均方差损失函数的表达式为

$$L_M = \sum_{i=1}^m \|\mathbf{z}_i - \mathbf{h}_i\|_2^2. \quad (5)$$

多任务识别下的联合损失函数生成过程如图5所示。FC_A1~FC_A3表示动作识别的全连接层,FC_D1~FC_D3表示身份识别的全连接层, $L_{C,A}$ 、 $L_{E,A}$ 、 $L_{M,A}$ 分别表示动作识别任务下的中心损失、交叉熵损失、均方差损失。 $L_{C,D}$ 、 $L_{E,D}$ 、 $L_{M,D}$ 分别表示身份识别任务下的中心损失、交叉熵损失、均方差损

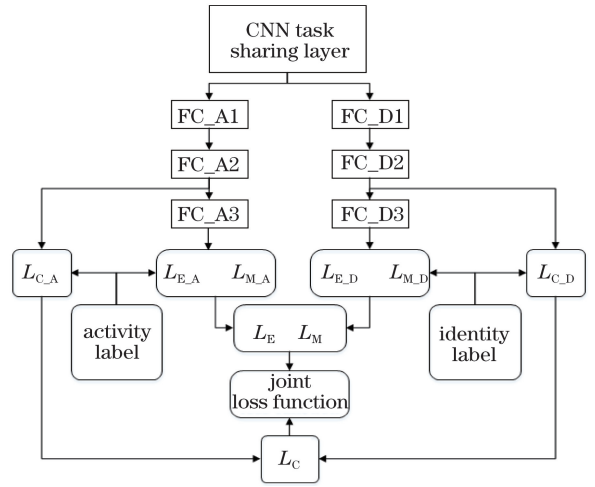


图5 联合损失函数生成示意图

Fig. 5 Diagram of joint loss function generation

失。利用任务特定层中最后一级全连接层上的输出计算交叉熵损失函数和均方差损失函数;利用倒数第二级全连接层上的输出计算中心损失函数。将两个任务上的损失函数按类别相加,分别利用 λ_E 、 λ_C 、 λ_M 加权后得到最终的联合损失函数,将联合损失回传至各个网络层以更新模型参数。联合损失函数表达式为

$$L = \lambda_E L_E + \lambda_C L_C + \lambda_M L_M. \quad (6)$$

4 分析与讨论

实验使用64位的Ubuntu14.04系统,处理器型号为Intel(R)Core(TM)i7-7700K CPU@4.20GHz,内存8G,显卡型号为GeForce GTX 1080Ti,显存为11G。算法的实现基于python语言和Pytorch深度学习框架,优化器使用Adam梯度下降算法。

4.1 多任务识别与单任务识别分析

实验采用慢时间窗口长度为1.0s的无噪雷达时频图像(180000张),其中每个受试者的单个动作种类包含2000张雷达时频图像。将180000张时频图像数据按照受试者类别和动作类别进行分层随机抽样,可得到4个数量相等的子集(45000张)。每个子集中,单个受试者的单个动作类别包含时频图像500张。每次取3个子集作为训练集(135000张),一个子集作为测试集(45000张),得到交叉验证集(cross validation)1~4。采用本文提出的模型结构和联合损失优化方法,进行人体动作(HA)与人体身份(HD)的多任务识别(MR)以及单任务识别(SR)。在单任务识别中,利用对应任务上的联合损失分别训练各任务的对应模型。

训练中学习率为 0.0001, 迭代次数为 20。经多次实验, 设定 $\lambda_E, \lambda_M, \lambda_C$ 为 0.01, α 为 0.005。表 1 为识别结果, 其中总识别准确率 (HT) 表示在两个任务上都识别正确的样本比例。图 6 显示了训练次数增加时, 交叉验证集 4 上各任务损失值和识别准确率的变化情况。

实验结果表明, 四折交叉验证上的多任务识别准确率均高于单任务独立识别的准确率。多任务下的总识别准确率平均值比单任务识别高出 1.7%, 证明了雷达人体动作识别任务与身份识别任务之间具有相互促进的作用, 同时验证了本文方法的有效性。与单任务识别相比, 多任务识别中的总任务识别和

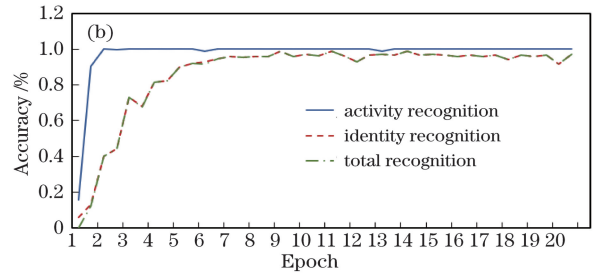
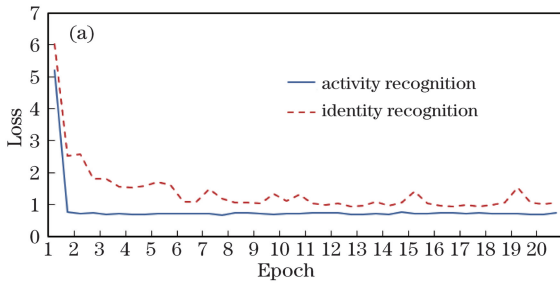


图 6 训练变化曲线。(a) 损失变化曲线; (b) 准确率变化曲线

Fig. 6 Loss and accuracy curves. (a) Loss curves; (b) accuracy curves

不同人体动作产生的雷达时频图像因具有一定的可视化差异易被识别, 而不同受试者在相同动作下产生的时频图像差异是细粒度的, 且待识别的身份种类多达 15 种, 因此身份识别在难度上高于动作识别, 动作识别任务在训练过程中具有更快的收敛速度和更高的准确率。

4.2 网络结构与联合损失对识别的影响分析

为证明网络结构各部分的有效性, 首先在本文模型的基础上去除 4 个放置于卷积层后的 SE 模块, 将得到的结构记为 D-SE。然后, 保留本文模型中的 CNN 特征提取部分, 去掉多尺度特征融合结构, 将任务特定部分的全连接层连接到 Block4 之后, 此时任务特定层的输入仅为 CNN 最深层的特征, 将此结构记为 D-MS。最后, 将去除 4 个 SE 模块后的 D-MS 结构记为 D-MSSE。三个模型都采用联合损失优化, 在交叉验证集 4 上的识别结果如表 2 所示。

Table 2 Recognition results of different network structures

| Model | Recognition rate / % | | |
|--------|----------------------|-------|-------|
| | HA | HD | HT |
| D-MSSE | 99.98 | 99.62 | 99.61 |
| D-MS | 100 | 99.88 | 99.88 |
| D-SE | 99.99 | 90.47 | 90.47 |

表 1 多任务识别与单任务识别结果

Table 1 Results of multi-task and single-task recognitions

| Cross validation set No. | Recognition rate of MR / % | | | Recognition rate of SR / % | | |
|--------------------------|----------------------------|--------------|--------------|----------------------------|-------|-------|
| | HA | HD | HT | HA | HD | HT |
| 1 | 99.94 | 98.78 | 98.74 | 99.70 | 95.99 | 95.72 |
| 2 | 99.94 | 98.80 | 98.77 | 99.88 | 96.70 | 96.93 |
| 3 | 99.97 | 98.90 | 98.89 | 99.90 | 97.91 | 97.84 |
| 4 | 100.00 | 99.93 | 99.93 | 99.96 | 99.37 | 99.34 |
| Mean | 99.96 | 99.10 | 99.08 | 99.86 | 97.49 | 97.38 |

身份识别的准确率差异更小, 即两个任务中被正确识别的样本重合度更高, 这表明多任务信息的共享使模型学习到了更具泛化性的特征表示。

果如表 2 所示。

利用 D-MS 网络对前几个卷积层的输出进行特征校正, 并将最高层的 CNN 特征用于特定的任务识别, 识别效果与 D-MSSE 相比有一定改善。D-SE 综合了多尺度特征图的各通道全局信息用于特定任务识别, 但因为失去了 SE 模块对各通道特征的校正作用, 各尺度上的通道全局信息缺乏准确性和一致性, 所以特征融合后的识别结果有所下降。本文模型的识别效果高于此三种模型结构, 证明了特征重校正结构和多尺度结构联合使用的必要性。两个结构共同利用了各通道特征的全局信息, 在特征提取、校正和融合的联合作用下, 能够提升识别效果。

本文模型以传统的交叉熵损失函数为基础, 加入不同损失函数进行优化后的识别结果如表 3 所示。结果表明, 加入均方差损失的联合优化结果与加入中心损失的联合优化结果均优于交叉熵损失单独优化的结果。其中, 中心损失在两个分任务上的优化效果更加明显。交叉熵损失函数扩大了样本在特征空间中的类间距离, 均方差损失函数均化了错分类别, 中心损失函数在优化过程中减小了样本的类内间距。联合损失函数可以生成更为紧致的类分布, 更利于共享特征空间的优化。

表3 不同损失优化下的识别结果

Table 3 Recognition results of different loss optimization

| Loss | Recognition rate /% | | |
|-------------|---------------------|-------|-------|
| | HA | HD | HT |
| L_E | 99.99 | 99.74 | 99.73 |
| $L_E + L_M$ | 99.99 | 99.86 | 99.86 |
| $L_E + L_C$ | 100 | 99.87 | 99.86 |

4.3 STFT 窗长、噪声、训练数据量对识别的影响分析

在 5 种 STFT 窗长(L)的无噪声雷达时频图像数据集上分别进行动作、身份类别的分层随机抽样,按照 3:1 的数量比例将样本划分为训练集和测试集,采用本文方法进行识别。表 4 为不同 STFT 窗长对识别准确率产生的影响。STFT 窗长较大时可获得较好的频域解析度,适当增加窗长有利于获取信号中丰富的微多普勒信息,从而提升识别效果。识别结果表明,STFT 窗长为 1.0 s 时有最佳识别效果;窗长为 0.2 s 时,雷达时频图像上的总识别准确率仍可保持 77.68%;窗长为 0.4 s 时,总准确率可接近 90%。本文模型在不同 STFT 窗长下的识别具有一定的稳定性。

表4 不同 STFT 窗长下的识别结果

Table 4 Recognition results under different STFT window lengths

| L / s | Recognition rate /% | | |
|---------|---------------------|-------|-------|
| | HA | HD | HT |
| 0.2 | 97.99 | 78.59 | 77.68 |
| 0.4 | 99.75 | 89.25 | 89.13 |
| 0.6 | 99.91 | 92.47 | 92.44 |
| 0.8 | 99.98 | 99.73 | 99.71 |

在 5 种信噪比(N_{SR})的有噪声雷达时频图像数据集上分别进行动作、身份类别的分层随机抽样,按照 3:1 的数量比例将样本划分为训练集和测试集,采用本文方法进行识别。表 5 为不同级别噪声下的识别准确率。噪声量级越大,微多普勒图像的细节信息损失越严重,识别将受到更多干扰。其中 -10 dB、-20 dB 级别的噪声对识别干扰较大。在 -10 dB 的情况下,本文方法仍能维持 76.49% 的总识别准确率,信噪比在 0 dB 以上时,总识别准确率都高于 90%。

除此之外,本文还验证了数据量不足情况下的模型识别效果,从交叉验证集 4 的训练集中随机挑选数据,组成 6 种包含不同数据量的训练集。训练后的模型都在交叉验证集 4 的测试集上进行识别,实验结果如表 6 所示。表中记录的单类数据量

表5 不同信噪比下的识别结果

Table 5 Recognition results under different N_{SR}

| N_{SR} / dB | Recognition rate /% | | |
|---------------|---------------------|-------|-------|
| | HA | HD | HT |
| -20 | 67.53 | 24.4 | 19.06 |
| -10 | 96.91 | 77.63 | 76.49 |
| 0 | 99.91 | 90.8 | 90.75 |
| 10 | 99.99 | 99.5 | 99.49 |
| $+\infty$ | 100 | 99.93 | 99.93 |

表6 不同训练数据量下的识别结果

Table 6 Recognition results under different training data sizes

| Data size | Recognition rate /% | | |
|-----------|---------------------|-------|-------|
| | HA | HD | HT |
| 400 | 99.94 | 97.52 | 97.46 |
| 600 | 99.97 | 99.71 | 99.57 |
| 800 | 99.97 | 99.62 | 99.59 |
| 1000 | 99.99 | 99.76 | 99.75 |
| 1300 | 99.99 | 99.85 | 99.84 |
| 1500 | 100 | 99.93 | 99.93 |

(data size)表示训练集中一个受试者单个动作类别下的数据量。增加训练数据可以使模型学习更准确的共享特征,获得更加紧密的类分布,故可以明显改善识别效果。本文模型在多个训练数据量下的识别结果比较稳定,在单类数据量为 400 张时,动作识别准确率仍可达到 99.94%,总识别率可达 97.46%。

4.4 对比算法结果分析

本文采用 PRGC 算法^[13]和 MRN 算法^[18]在交叉验证集 4 上进行对比。PRGC 算法使用多个雷达基站的时频图像共同训练浅层 CNN 以提升识别准确率;MRN 算法主要通过在全连接层上进行信息迁移实现多任务学习。将本文提出的模型记为 MCNN,表 7 的实验结果显示,在 1.0 s STFT 窗长的无噪声数据集上,MCNN 的动作识别、身份识别及总识别的准确率均高于 PRGC 算法与 MRN 算法,且 MCNN 算法的总任务识别准确率与单独的身份识别任务准确率差距更小,验证了本文算法的有效性。

表7 对比算法结果

Table 7 Comparison of results of algorithms

| Algorithm | Recognition rate /% | | |
|-----------|---------------------|-------|-------|
| | HA | HD | HT |
| PRGC | 99.94 | 98.9 | 98.84 |
| MRN | 99.97 | 99.9 | 99.85 |
| MCNN | 100 | 99.93 | 99.93 |

PRGC 算法未对模型的共享特征提取结构进行

针对性改进,而 MRN 算法仅在全连接层上进行多任务的信息交互,在特征提取时仅使用了较为基础的 CNN 结构。本文方法关注于两个任务的潜在相关性,不仅改善了模型结构、加强了所提特征对各识别任务的有效性,还在优化方法上进一步约束了类分布以提高模型的泛化性能,从而综合地提升识别效果。本文提出的模型结构和联合损失的优化方法使网络能学习到雷达人体动作识别与身份识别任务上更具泛化性的共享特征表示,有效提升了识别的准确率。

5 结 论

本文对雷达在人体动作识别与身份识别上的应用进行研究,针对相关任务间的信息共享问题和样本类分布的约束问题,提出一种基于 CNN 的多任务识别模型。在模型结构方面,利用特征重校正结构和多尺度特征融合结构提取共享特征,在特定任务层上进行多任务识别;在模型优化方面,提出基于多任务联合损失函数的优化方法,引入中心损失函数聚合类内样本,引入均方差损失函数促进类间样本分离,以获得更加紧致的类分布。实验结果表明,本文方法可有效提升模型的泛化性能,从而提高识别准确率,总识别准确率最高可达 99.93%,优于任务独立识别结果以及其他算法的识别结果,且本文方法在不同的噪声量级、STFT 窗长和训练数据量上都具有一定的稳定性。今后将针对小数据集上的识别方法,尝试多种模型结构与优化方法,进一步提高雷达人体动作与身份识别的效率与准确度。

参 考 文 献

- [1] Fairchild D P, Narayanan R M. Classification of human motions using empirical mode decomposition of human micro-Doppler signatures[J]. IET Radar, Sonar & Navigation, 2014, 8(5): 425-434.
- [2] Zhang Y, Zhu Y P, Cheng Y Q, et al. Human target radar echo signal analysis based on micro-Doppler characteristic[J]. Signal Processing, 2009, 25(10): 1616-1623.
张翼, 朱玉鹏, 程永强, 等. 基于微多普勒特征的人体目标雷达回波信号分析[J]. 信号处理, 2009, 25(10): 1616-1623.
- [3] Huang Y W, Wan C L, Feng H. Multi-feature fusion human behavior recognition algorithm based on convolutional neural network and long short term memory neural network[J]. Laser & Optoelectronics

Progress, 2019, 56(7): 071505.

黄友文, 万超伦, 冯恒. 基于卷积神经网络与长短期记忆神经网络的多特征融合人体行为识别算法[J]. 激光与光电子学进展, 2019, 56(7): 071505.

- [4] Ikeda T, Ishiguro H, Miyashita T, et al. Pedestrian identification by associating wearable and environmental sensors based on phase dependent correlation of human walking[J]. Journal of Ambient Intelligence and Humanized Computing, 2014, 5(5): 645-654.
- [5] Xu H Y, Kong J, Jiang M, et al. Action recognition based on histogram of spatio-temporal oriented principal components[J]. Laser & Optoelectronics Progress, 2018, 55(6): 061009.
徐海洋, 孔军, 蒋敏, 等. 基于时空方向主成分直方图的人体行为识别[J]. 激光与光电子学进展, 2018, 55(6): 061009.
- [6] van Dorp P, Groen F C A. Feature-based human motion parameter estimation with radar[J]. IET Radar, Sonar & Navigation, 2008, 2(2): 135-145.
- [7] Kim Y, Ling H. Human activity classification based on micro-Doppler signatures using a support vector machine[J]. IEEE Transactions on Geoscience and Remote Sensing, 2009, 47(5): 1328-1337.
- [8] Ricci R, Balleri A. Recognition of humans based on radar micro-Doppler shape spectrum features[J]. IET Radar, Sonar & Navigation, 2015, 9(9): 1216-1223.
- [9] Li J, Phung S L, Tivive F H C, et al. Automatic classification of human motions using Doppler radar[C]// The 2012 International Joint Conference on Neural Networks (IJCNN), June 10-15, 2012, Brisbane, QLD, Australia. New York: IEEE, 2012: 12906230.
- [10] Javier R J, Kim Y. Application of linear predictive coding for human activity classification based on micro-Doppler signatures[J]. IEEE Geoscience and Remote Sensing Letters, 2014, 11(10): 1831-1834.
- [11] Tahmoush D. Review of micro-Doppler signatures[J]. IET Radar, Sonar & Navigation, 2015, 9(9): 1140-1146.
- [12] Kim Y, Moon T. Human detection and activity classification based on micro-Doppler signatures using deep convolutional neural networks[J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(1): 8-12.
- [13] Cao P B, Xia W J, Ye M, et al. Radar-ID: human identification based on radar micro-Doppler signatures

- using deep convolutional neural networks [J]. *IET Radar, Sonar & Navigation*, 2018, 12(7): 729-734.
- [14] Chen Z X, Li G, Fioranelli F, et al. Personnel recognition and gait classification based on multistatic micro-Doppler signatures using deep convolutional neural networks [J]. *IEEE Geoscience and Remote Sensing Letters*, 2018, 15(5): 669-673.
- [15] Seyfioglu M S, Ozbayoglu A M, Gurbuz S Z. Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2018, 54(4): 1709-1723.
- [16] Kim Y, Li Y. Human activity classification with transmission and reflection coefficients of on-body antennas through deep convolutional neural networks [J]. *IEEE Transactions on Antennas and Propagation*, 2017, 65(5): 2764-2768.
- [17] Klarenbeek G, Harmanny R I A, Cifola L. Multi-target human gait classification using LSTM recurrent neural networks applied to micro-Doppler [C] // 2017 European Radar Conference (EURAD), October 11-13, 2017, Nuremberg, Germany. New York: IEEE, 2017: 167-170.
- [18] Lu Y X, Kumar A, Zhai S F, et al. Fully-adaptive feature sharing in multi-task networks with applications in person attribute classification [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 1131-1140.
- [19] Long M S, Cao Z J, Wang J M, et al. Learning multiple tasks with multilinear relationship networks [C] // *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, December 4-9, 2017, Long Beach, CA, USA. Canada: NIPS, 2017: 1594-1603.
- [20] Hashimoto K, Xiong C M, Tsuruoka Y, et al. A joint many-task model: growing a neural network for multiple NLP tasks [J/OL]. (2017-07-24) [2019-06-09]. <https://arxiv.org/abs/1611.01587>.
- [21] Kendall A, Gal Y, Cipolla R. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 18-22, 2018, Salt Lake City, Utah. New York: IEEE, 2018: 7482-7491.
- [22] Liu Y P, Wang C, Xia H Y. Application progress of time-frequency analysis for lidar [J]. *Laser & Optoelectronics Progress*, 2018, 55(12): 120005. 刘燕平, 王冲, 夏海云. 时频分析在激光雷达中的应用进展 [J]. *激光与光电子学进展*, 2018, 55(12): 120005.
- [23] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019: 1.
- [24] Zhou G B, Wu J X, Zhang C L, et al. Minimal gated unit for recurrent neural networks [J]. *International Journal of Automation and Computing*, 2016, 13(3): 226-234.
- [25] Wen Y D, Zhang K P, Li Z F, et al. A discriminative feature learning approach for deep face recognition [M] // Leibe B, Matas J, Sebe N, et al. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9911: 499-515.