

基于串联式一维神经网络的毫米波雷达动态手势识别方法

靳标* 彭宇 邝晓飞 张贞凯
(江苏科技大学电子信息学院 镇江 212003)

摘要: 现有的基于雷达传感器的手势识别方法,大多先利用雷达回波对手势的距离、多普勒和角度等信息进行参数估计,得到各种数据谱图,然后再利用卷积神经网络对这些谱图进行分类,实现过程较为复杂。该文提出一种基于串联式一维神经网络(1D-ScNN)的毫米波雷达动态手势识别方法。首先基于毫米波雷达获取动态手势的原始回波,然后利用一维卷积和池化操作对手势特征进行提取,并将这些特征信息输入一维Inception v3结构。最后在网络的末端接入长短期记忆(LSTM)网络来聚合一维特征,充分利用动态手势的帧间相关性,提高识别准确率和训练收敛速度。实验结果表明,该方法实现过程简单,收敛速度快,识别准确率可以达到96.0%以上,高于现有基于数据谱图的手势分类方法。

关键词: 雷达目标识别; 手势识别; 卷积神经网络; 毫米波雷达

中图分类号: TN911.73; TP391.4

文献标识码: A

文章编号: 1009-5896(2021)09-2743-08

DOI: 10.11999/JEIT200894

Dynamic Gesture Recognition Method Based on Millimeter-wave Radar by One-Dimensional Series Neural Network

JIN Biao PENG Yu KUANG Xiaofei ZHANG Zhenkai

(College of Electronic Information, Jiangsu University of Science and Technology, Zhenjiang 212003, China)

Abstract: For the most of the existing gesture recognition methods based on the radar sensor, the parameters such as the distance, Doppler, and angle are estimated using the radar echo at first. And then the obtained data spectra are inputted into the convolutional neural networks to classify the gestures. The implementation process is complicated. A dynamic gesture recognition method is proposed based on the millimeter-wave radar using the One-Dimensional Series connection Neural Networks (1D-ScNN) in this paper. Firstly, the original echo of dynamic gesture is obtained by the millimeter-wave radar. The gesture features are extracted by the one-dimensional convolution and pooling operations, and then are inputted into the one-dimensional inception v3 structure. In order to aggregate the one-dimensional features, the Long Short-Term Memory (LSTM) modular is connected to the end of the network. The inter-frame correlation of dynamic gestures echo is fully utilized to improve the recognition accuracy and the convergence speed of training. The experimental results show that the proposed method is simple in implementation and has a fast convergence speed. The classification accuracy can reach more than 96.0%, which is higher than the traditional gesture classification methods.

Key words: Radar target recognition; Gesture recognition; Convolutional Neural Networks(CNN); Millimeter-wave radar

1 引言

非接触式的手势识别在智能交互领域具有广泛的应用,如汽车的辅助驾驶,应用于游戏的增强现

实和可穿戴智能设备的人机交互等^[1]。毫米波雷达具有捕获细微动作的能力,而且可以在无光照环境下工作,通常被选作无接触式手势识别的传感器。

目前,基于雷达传感器的手势识别方法大多先利用雷达回波对手势的距离、多普勒和角度等信息进行参数估计,得到各种数据谱图,然后将特征谱图以图片数据集的形式输入支持向量机(Support Vector Machine, SVM)^[2]或卷积神经网络(Convolutional Neural Networks, CNN)等分类器进行识

收稿日期: 2020-10-19; 改回日期: 2021-01-30; 网络出版: 2021-02-24

*通信作者: 靳标 biaojin@just.edu.cn

基金项目: 国家自然科学基金(61701416, 61871203)

Foundation Items: The National Natural Science Foundation of China (61701416, 61871203)

别^[3-6]。北京航空航天大学的王俊等人^[3]提出了一种基于雷达距离-多普勒(Range-Doppler, RD)信息的手势识别方法。该方法通过快速傅里叶变换和相干累积获取手势的RD图像,以2维RD图像构建数据集,并输入CNN进行训练。重庆邮电大学的王勇等人^[4]采用时频分析和多重信号分类方法估计手势的距离、速度和角度信息,然后将这3维参数在时间上进行积累得到距离-速度谱图和角度-时间谱图,再建立双流融合神经网络(Two-Stream Fusion Neural Network, TS-FNN)进行手势特征提取。复旦大学的夏朝阳等人^[5]提出一种微动手势特征提取的最优参数设计准则,根据手势运动过程中的振幅和速度差异,利用联合的多特征谱图进行微动手势识别。然而,这类基于雷达数据图像的手势识别方法具有一定的局限性。第一,在将雷达数据图像输入神经网络之前,需要进行参数估计、特征融合和像素化等一系列预处理步骤,实现过程较为复杂。第二,雷达特征图谱在2维的CNN中被看作典型的静态光学图像,忽略了动态手势的帧间相关性。第三,这种先进行参数估计后进行特征融合的方法,所提取的特征维数有限。如果采用多维特征进行融合,则增加了算法实现的复杂程度。文献^[7]提出一种高维特征融合方法,将3类数据集输入距离-多普勒-角度(Range-Doppler-Angle, RDA)的多维参数卷积神经网络。这类方法增加了网络的空间复杂度,需要更多的训练数据来拟合网络^[8,9]。如果能将雷达原始回波输入卷积神经网络,让网络提取特征并训练,则有可能在不增加数据量的前提下提高识别速度和分类精度。

针对以上问题,本文提出一种基于串联式1维神经网络(One-Dimensional Series connection Neural Networks, 1D-ScNN)的毫米波雷达动态手势识别方法。本方法的创新点为:

(1)针对雷达原始回波为复数序列的特性,构建1维雷达回波的实部和虚部双通道输入序列,无

需复杂的数据预处理过程,可以提高动态手势数据集的处理速度和模型的训练速度。

(2)基于递归神经网络具有记忆功能的特性,在1维卷积神经网络中串联长短期记忆(Long Short-Term Memory, LSTM)网络层,用于处理动态手势的帧间相关性,以激活手势的时序特征,可以提高网络的收敛速度和分类精度。

2 雷达回波的预处理

2.1 雷达回波解析

本文采用TI公司的77~81 GHz毫米波雷达捕获手势信息。该雷达系统有3个发射天线和4个接收天线,发射信号为线性调频连续波(Linear Frequency Modulated Continuous Wave, LFM-CW),发射功率为12 dBm。通过数据采集卡与终端设备进行通信,终端读取数据并进行解析,具体的解析流程如图1所示。首先将采集的雷达原始回波(以ADC_data.bin文件存储)以16位整型(int16)读入,然后按照接收天线的个数(本文接收天线的个数 $N_{Rx}=4$)重组为数据矩阵。其中,每个接收通道对应的数据长度为采样点数(n_samples)、chirp数(n_chirps)和采集帧数(frames)的乘积。最后,通过标准化和Reshape操作按照实部(I路)和虚部(Q路)输出为1维雷达原始回波序列。

2.2 回波数据标准化

由于设备噪声和背景杂波的干扰,采集到的雷达原始回波在某些时刻可能存在突变。如果某个特征的方差远大于其他特征的方差,那么它将会在学习过程中占据主导,导致学习机制难以学习其他特征,使得模型难以收敛。因此需要对这样的特征数据进行标准化处理,更新函数为

$$\hat{x}_{ij} = \frac{x_{ij} - \mu}{\sigma} \quad (1)$$

其中, \hat{x}_{ij} 表示更新后数据矩阵的元素, x_{ij} 表示原接收数据矩阵的元素。根据原矩阵中所有元素求得

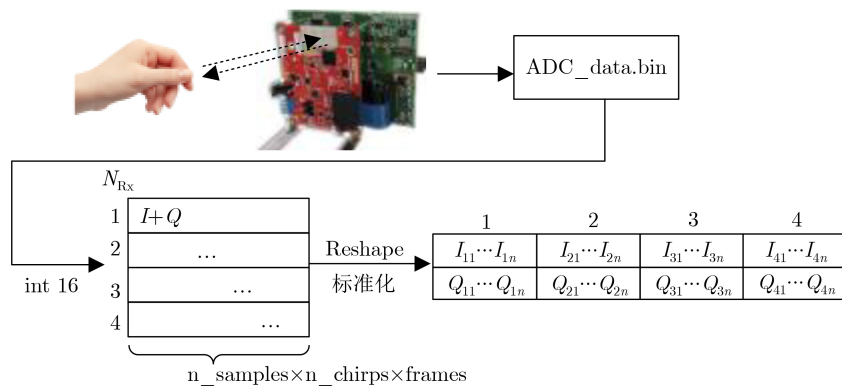


图1 雷达原始回波解析流程图

均值 μ 和方差 σ ，使得输入数据服从标准正态分布 $N(0,1)$ ，更新后的1维数据将直接作为1D-ScNN的输入。

3 串联式1维神经网络(1D-ScNN)

1D-ScNN的总体结构如图2所示。首先输入1维雷达原始回波序列(Raw data)，前端是3个1维Inception v3结构^[10,11]，其最小单元均采用1维卷积算子。然后串联LSTM层以聚合前端的时序特征。最后通过稠密层(Dense)将前面提取的特征经过非线性变换映射到标签集(Label)，并输出识别结果。

3.1 1维卷积模块

为了与1维双通道输入数据匹配，本文在整个网络中主要采用1维卷积层(Conv1D)进行卷积计算。本文采用不同尺寸的Conv1D，并将Conv1D和1维池化层以串联和构建1D-Inception子网的方式搭建深度网络。Conv1D主要从数据中获得浅层特征，尽可能保留序列信息^[12]。1维池化层完成下采样，减少计算量的同时，保证模型的泛化能力。以卷积核大小为3的1维卷积为例，假设卷积的权值 $w=(w_1, w_2, w_3)$ ，输入 x 是一个有限长的1维向量。卷积算子 $y=Conv(w, x)$ 可以写成

$$y_i = w_1x_{i-1} + w_2x_i + w_3x_{i+1} \quad (2)$$

这里可以将卷积运算理解为移位与乘累加，即将输入 x 分别移位 $-1, 0, +1$ ，再与 w 相乘，最终求和得到 y 。移位操作的具体形式为

$$x_i^{-1} = x_{i-1}, x_i^0 = x_i, x_i^{+1} = x_{i+1} \quad (3)$$

乘累加的具体形式为

$$y = w_1x^{-1} + w_2x^0 + w_3x^{+1} \quad (4)$$

3.2 串联的1D-Inception子网组

1维卷积层在处理长序列数据时，对于数据中的高层特征学习能力不足，需要更复杂的网络结构进一步提取。因此，本文以Inception v3作为1维卷积神经网络模块的基础结构，搭建串联的1D-Inception子网组^[13]。1D-Inception结构如图3所示，将1维输入分为5个分支，其1维卷积形式为卷积核尺寸*数量。其中，平均池化下采样层的池化尺寸为3， k 为4个分支的滤波器总数。网络使用了3个相似的1D-Inception结构， i 是控制聚合算子输出时序维度的尺寸。 i 的初始化不同，输出的时序维度和张量维度也不同。如本文输入时序长度为1024，1D-Inception的 i 分别取4, 6和7，则输出的时序长度分别为256, 64和16。

本文改进之后的1D-Inception结构在不改变网络宽度的同时，可以进一步降低模型的时间复杂度。1D-Inception的时间复杂度分析可以类比2维卷积，其计算公式为^[14]

$$Time \sim O\left(\sum_{l=1}^D M_l \cdot K_l \cdot C_{l-1} \cdot C_l\right) \quad (5)$$

其中， D 为1D-Inception中卷积层的总数， M_l 为第

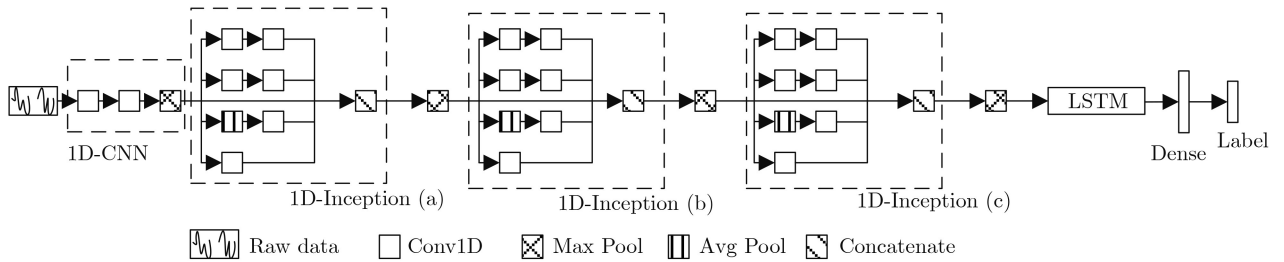


图2 1D-ScNN结构图

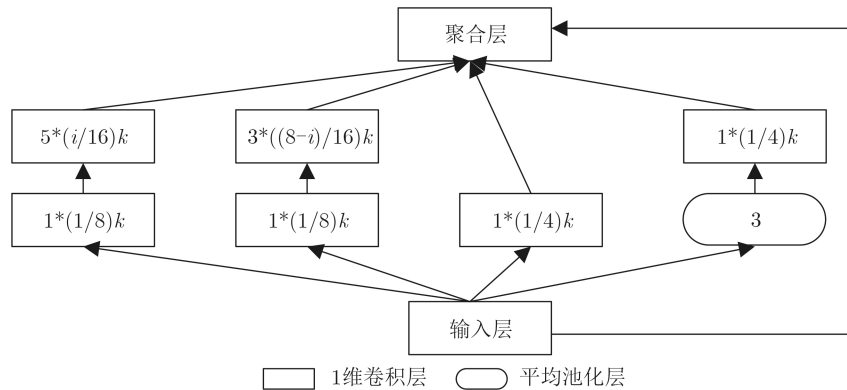


图3 1D-Inception结构

l 个卷积核输出特征序列的长度, K_l 为第 l 个卷积核的长度, C 为卷积层的输出通道数。假设本文1D-Inception上层输入尺寸为 (M_{in}, C_{in}) 。根据式(5)易得, 该结构的时间复杂度约为

$$T_{1D-Inception} = M_{in} \left[\frac{1}{4}k \cdot C_{in} + k^2 \cdot \left(\frac{12+i}{8} \right) \right] \quad (6)$$

3.3 长短期记忆(LSTM)网络层

为了有效聚合雷达回波相邻帧之间的信息, 文献[4]在3维卷积层(Conv3D)中接入LSTM层, 用于提取并联融合特征的时序信息。但是将3维卷积层与1维的LSTM对接, 需要通过Flatten层降维过渡, 增加了算法的复杂度^[15,16]。本文在1D-Inception结构后直接串联LSTM层, 无需降维过渡, 避免了维度变换过程中可能丢失时序信息现象。LSTM算子的运算过程如图4所示, 主要包括遗忘、选择记忆和输出阶段。

遗忘阶段: 这部分操作通过一个称为遗忘门的网络层来处理, 以决定丢弃哪些信息。它通过查看 h_{t-1} 和 x_t 来输出一个 $0 \sim 1$ 之间的向量, 以决定序列 C_{t-1} 中需要丢弃哪些信息。其函数形式为

$$f_t = \sigma(w_f \cdot [h_{t-1}, x_t] + b_f) \quad (7)$$

其中, w_f 为遗忘门权值, b_f 为偏置, $\sigma(x)$ 采用hard_sigmoid激活函数。

选择记忆阶段: 通过一个称为输入门的操作来决定更新哪些信息。然后通过tanh函数得到新的候选序列 \tilde{C}_t , 并将旧的序列 C_{t-1} 更新为序列 C_t 。其函数形式为

$$\left. \begin{aligned} i_t &= \sigma(w_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t &= \tanh(w_c \cdot [h_{t-1}, x_t] + b_c) \\ C_t &= f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \end{aligned} \right\} \quad (8)$$

其中, $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$, w_i 和 w_c 分别为输入门和更新算子的权值, b_i 和 b_c 为对应的偏置。

输出阶段: 更新完序列状态后, 根据输入的 h_{t-1} 和 x_t 判断输出序列的状态。这里需要经过一个称为输出门的sigmoid层得到判断条件 o_t , 然后通过tanh函数得到元素值在 $-1 \sim 1$ 之间的向量。将该向量与 o_t 相乘, 得到最终该LSTM单元的输出。具体函数形式为

$$\left. \begin{aligned} o_t &= \sigma(w_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t &= o_t \cdot \tanh(C_t) \end{aligned} \right\} \quad (9)$$

其中, w_o 和 b_o 分别为输出门的权值和偏置。

由式(9)可知, 当前时间步的输出 o_t 受隐藏层输出 h_t 和输入 x_t 的影响, 隐藏层 h_t 又受到过去所有输入时间步的影响, 影响程度由遗忘门 f_t 控制。因此, 在LSTM中利用时间相关性的过程主要由遗忘门控制, 根据过去时间的隐藏层输出动态调整当前时间步的权值, 使得有效手势片段所在时间步的激活程度最高。

4 实验分析与讨论

4.1 实验环境

本文实验采用TI公司的AWR1443毫米波雷达传感器和DCA1000 EVM采集卡。雷达回波通过采集卡传入PC端, 在CPU上进行数据解析。1D-ScNN基于TensorFlow2.0+keras深度学习框架, 在配置为Intel E5-2640处理器和NVIDIA Quadro K620显卡的服务器上训练。表1给出了雷达系统的配置参数, 其中每个手势的采集时间约为1.3 s (采集帧数 \times 帧时间)。为了控制单个样本的规模并提高回波解析的效率, 需要将采样点数(n_{samples})设置得尽可能小, 同时满足最小采样时间大于 f_s/n_{samples} 。

4.2 动态手势的定义

本文采用AWR1443雷达传感器配套的软件采集原始回波, 自建动态手势的原始回波数据集。本文共设计5种手势, 包括上下按压、前后推拉、手

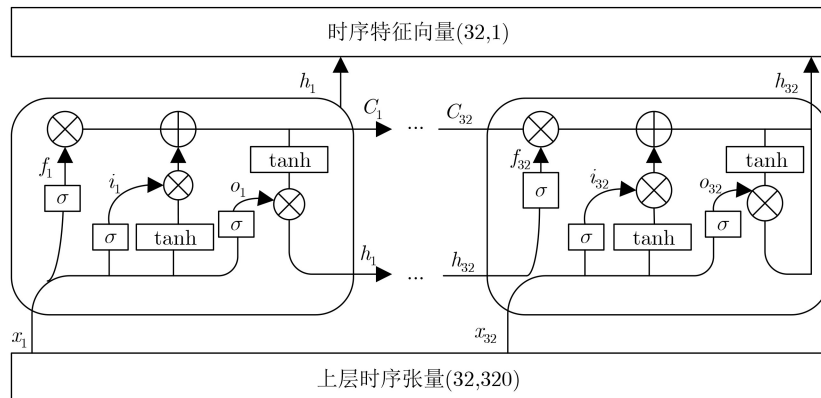


图4 LSTM算子运算过程

表 1 雷达传感器参数

参数	数量
发射天线数量(个)	3
接收天线数量(个)	4
采集帧数(帧)	32
帧时间(ms)	40
Chirp数(个)	32
带宽(MHz)	1798.92
采样点数	64
采样率(MHz)	10

掌翻转、手指摩擦和抓握5个手势动作，具体动作如图5所示。每一种手势重复采集数据800次，共4000个雷达原始回波数据文件。

4.3 1D-ScNN初始化

本文构造的1D-ScNN参数初始化主要针对1维卷积进行配置。1维卷积参数配置情况如表2所示。其中1D-Inception结构的卷积核形式是 $k*i$ ， k 控制该结构的1维滤波器总数， i 控制不同尺寸滤波器的分配比例。Conv1D的卷积核形式为卷积核数量*卷积核尺寸。

Conv1D输入的序列维度是一个多参数的乘积，包括采样点数、采集帧数(32帧)和Chirp数。考虑到Conv1D-2的输出尺寸会影响1D-Inception模块的计算量，本文讨论了不同Conv1D-1滤波器尺

寸对测试精度的影响，滤波器的尺寸从32开始测试，以3和16为增量单位对网络进行调参。实验结果如图6所示。从图6可以看出，在保证Conv1D-2输出尺寸尽可能小的前提下，滤波器的最优尺寸为48。确定了滤波器尺寸后，可以通过式(7)、式(8)计算出1D-CNN模块和1D-Inception子网组的时间复杂度分别为 5.03×10^8 FLOPs和 5.84×10^4 FLOPs。具体每一层卷积的时间复杂度在表2中已经给出，这里不再赘述。

本文将数据集划分为训练集和测试集，以数据集的20%定义训练测试集。采用早停法(early stopping)监听测试集的准确率，利用Compile和Fit库函数定义初始学习率和损失，在泛化效果变差时结束模型训练。为了探索如何使模型训练的收敛速度更快，设置迭代次数为100，观测不同学习率下训练损失的变化。图7给出了不同学习率下的训练精度和训练损失。由图7可知，当学习率为 3×10^{-4} 时，训练损失和精度达到最佳状态。

4.4 与传统方法的比较结果

将本文所提方法与文献[4]、文献[7]和文献[17]中的手势识别方法进行对比。文献[4]采用3D-CNN和2D-CNN对距离-速度谱和角度-时间谱进行特征提取，并使用LSTM层提取融合特征图中的时序信息。文献[7]采用单参数网络并联融合，得到多维参数卷积神经网络。文献[17]将时间分布(time dis-

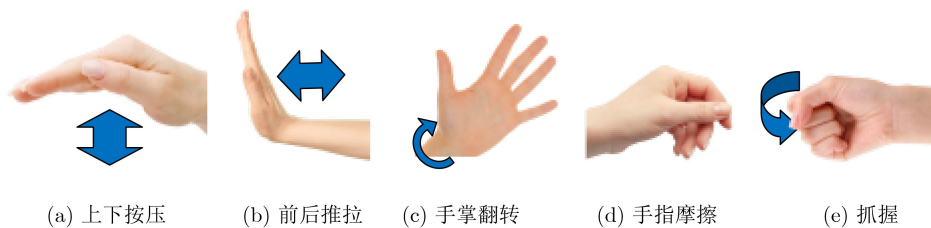


图 5 动态手势定义

表 2 1维卷积参数配置

类型	卷积核+步长	参数量	输出尺寸	时间复杂度(FLOPs)
Input	-	0	(8, 262144, 2)	-
Conv1D-1	64*48+8	6208	(8, 32768, 64)	2.01×10^8
Conv1D-2	128*9+8	73856	(8, 4095, 128)	3.02×10^8
MaxPool1D	1*4+4	0	(8, 1024, 128)	-
1D-Inception(a)	64*4+1	7248	(8, 1024, 192)	1.43×10^4
MaxPool1D	1*4+4	0	(8, 256, 192)	-
1D-Inception(b)	64*6+1	10448	(8, 256, 256)	2.05×10^4
MaxPool1D	1*4+4	0	(8, 64, 256)	-
1D-Inception(c)	64*7+1	13584	(8, 64, 320)	2.36×10^4
MaxPool1D	1*4+2	0	(8, 32, 320)	-

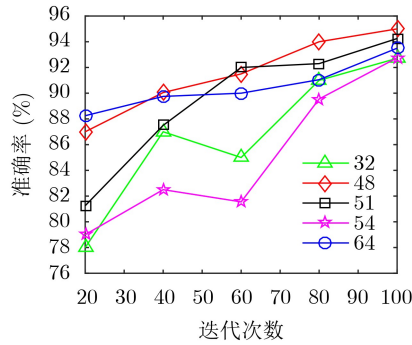


图6 滤波器尺寸对测试精度的影响

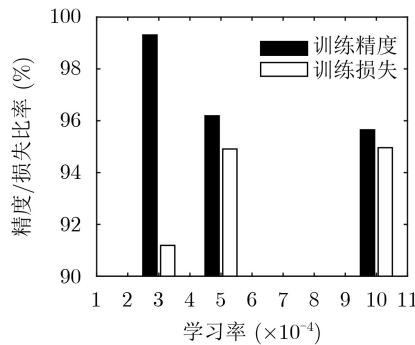


图7 不同学习率的训练精度和训练损失

tributed)层应用于2维CNN实现时序建模,对距离-多普勒图像序列进行分类。为了验证LSTM层对1D-ScNN时序关系的影响,本文根据文献[13]的方法构造了端到端的1维全卷积神经网络(1D-ID-CNN),并将1D-ScNN的初始化参数赋给1D-ID-CNN,同时把LSTM层替换成了Dense层(输出维度与激活函数不变),最后使用softmax分类器输出手势类别标签。同时,为了评估1D-ScNN模型的稳定性,本文还通过增加数据量的方式试探模型的平均准确率上限。这里分别测试了2000,3600和4000个样本的数据集,设置迭代上限为250次,并重复训练和测试过程。在2000个样本的实验中,1D-ScNN的训练和测试准确率如图8所示。从图8可以看出,

本文所提1D-ScNN方法只需要140次左右的迭代即可收敛,而文献[4]和文献[7]方法由于数据量较多、网络结构复杂等原因,则需要超过500次迭代后才能收敛。

表3对比了输入分别为传统数据谱图和原始回波两类网络模型的各项参数指标。通过对比可知:一方面,RDA-T模型比TS-FNN模型和LRACN模型的平均准确率提高了3.24%和0.96%,比TS-FNN模型的卷积层高出一个维度,说明多维参数谱图构建的数据集能够提供更丰富的信息,但是却需要以牺牲空间复杂度来提高识别的准确率;另一方面,LRACN模型以数据谱图序列为输入形式,本质上仍是对每一帧的谱图进行学习。相比于数据谱图的输入形式,LRACN模型为了保证模型的准确率,增加了手势的采集时间和卷积的维度,导致模型的参数量和计算复杂度进一步加大。通过对原始回波输入模型的分析可知,本文提出的1D-ScNN与不加入LSTM层的1D-ID-CNN相比,准确率提高了2.25%,参数量减少了56.9%。这是由于1D-ID-CNN捕获的特征没有顺序,而加入LSTM能充分利用动态手势的帧间相关性,提高手势识别的准确率。同时,对比传统谱图输入的网络模型,本文所提1D-ScNN模型在时间复杂度上比RDA-T模

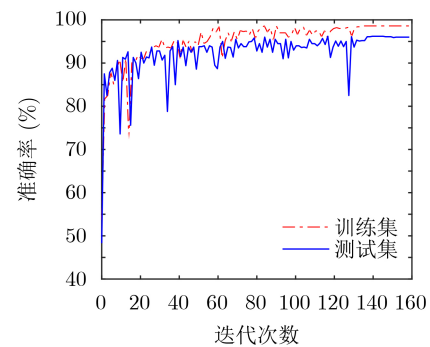
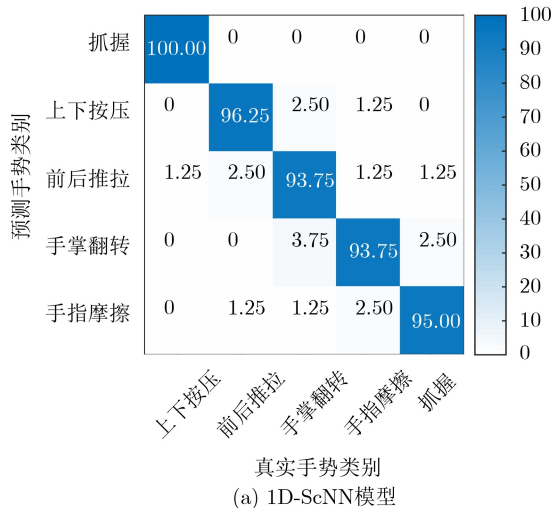


图8 1D-ScNN识别准确率

表3 不同模型比较结果

输入类型	模型	数据量	参数量	单样本采集帧数	时间复杂度(FLOPs)	平均准确率(%)
数据谱图	TS-FNN ^[4]	4000	-	32	3.20×10^{10}	92.06
	RDA-T ^[7]	3600	-	32	2.11×10^9	95.30
	LRACN ^[17]	750	1116277	100	8.74×10^8	94.34
原始回波	1D-ID-CNN	2000				93.75
		3600	275381	32	5.03×10^8	94.00
	1D-ScNN(本文方法)	4000				93.02
		3600	156693	32	5.03×10^8	95.75
		4000			96.01	

型降低了约75.83%，比谱图序列输入的LRACN模型降低了约42.45%。在对模型稳健性的分析中，本文主要比较了相同数据量下模型平均准确率的变化。从对比结果可知，本文所提1D-ScNN模型在相同样本采集帧数和数据量下，平均准确率均优于传统模型。



4.5 神经网络的在线验证

为了验证本文所构建的神经网络的在线识别准确率，选取了未参与训练的400组数据作为测试集，具体分析模型对每种手势的识别表现。图9给出了1D-ScNN模型和1D-ID-CNN模型在测试集上得出的混淆矩阵。

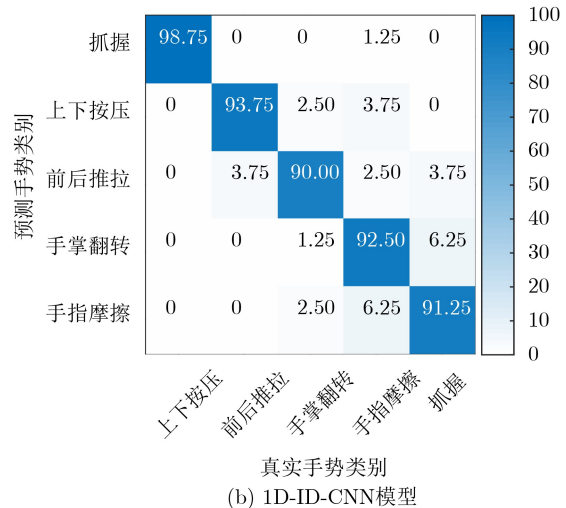


图9 1D-ScNN与1D-ID-CNN模型的混淆矩阵

从图9可以得出，1D-ID-CNN模型在准确率和召回率的表现均不如1D-ScNN模型。加入LSTM后对手势3(手掌翻转)和手势5(抓握)的识别准确率提高最为明显。这是因为两种手势运动的前后关联性更强，LSTM能够充分利用帧间相关性提高识别准确率。1D-ScNN模型在预测手势时，手势1(上下按压)和手势2(前后推拉)的表现最好，其准确率和召回率均在96%以上。手势3(手掌翻转)和手势4(手指摩擦)的准确率和召回率较差，主要是因为这两种手势的运动幅度小、特征不明显，不易识别。

5 结论

本文提出一种基于1D-ScNN的毫米波雷达动态手势识别方法。本方法不需要复杂的预处理过程，直接将雷达原始回波输入卷积神经网络，可以尽可能多地提取手势动作的特征。同时，所提方法将1D-CNN与LSTM层结合，能够有效提取向量的时序特征。实验结果表明，与其他方法相比，所设计的网络结构能够以更少的训练数据和更低的计算复杂度达到更高的识别精度，各方面性能均优于传统的手势识别方法。

参考文献

[1] LIEN J, GILLIAN N, KARAGOZLER M E, et al. Soli: Ubiquitous gesture sensing with millimeter wave radar[J].

ACM Transactions on Graphics, 2016, 35(4): 142: 1–142: 19. doi: [10.1145/2897824.2925953](https://doi.org/10.1145/2897824.2925953).

[2] KIM Y and 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. doi: [10.1109/TGRS.2009.2012849](https://doi.org/10.1109/TGRS.2009.2012849).

[3] 王俊, 郑彤, 雷鹏, 等. 基于卷积神经网络的手势动作雷达识别方法[J]. *北京航空航天大学学报*, 2018, 44(6): 1117–1123. doi: [10.13700/j.bh.1001-5965.2017.0397](https://doi.org/10.13700/j.bh.1001-5965.2017.0397).

WANG Jun, ZHENG Tong, LEI Peng, et al. Hand gesture recognition method by radar based on convolutional neural network[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2018, 44(6): 1117–1123. doi: [10.13700/j.bh.1001-5965.2017.0397](https://doi.org/10.13700/j.bh.1001-5965.2017.0397).

[4] 王勇, 王沙沙, 田增山, 等. 基于FMCW雷达的双流融合神经网络手势识别方法[J]. *电子学报*, 2019, 47(7): 1408–1415. doi: [10.3969/j.issn.0372-2112.2019.07.003](https://doi.org/10.3969/j.issn.0372-2112.2019.07.003).

WANG Yong, WANG Shasha, TIAN Zengshan, et al. Two-stream fusion neural network approach for hand gesture recognition based on FMCW radar[J]. *Acta Electronica Sinica*, 2019, 47(7): 1408–1415. doi: [10.3969/j.issn.0372-2112.2019.07.003](https://doi.org/10.3969/j.issn.0372-2112.2019.07.003).

[5] 夏朝阳, 周成龙, 介钧誉, 等. 基于多通道调频连续波毫米波雷达的微小手势识别[J]. *电子与信息学报*, 2020, 42(1): 164–172. doi: [10.11999/JEIT190797](https://doi.org/10.11999/JEIT190797).

XIA Zhaoyang, ZHOU Chenglong, JIE Junyu, et al. Micro-

- motion gesture recognition based on multi-channel frequency modulated continuous wave millimeter wave radar[J]. *Journal of Electronics & Information Technology*, 2020, 42(1): 164–172. doi: [10.11999/JEIT190797](https://doi.org/10.11999/JEIT190797).
- [6] ZHANG Zhenyuan, TIAN Zengshan, and ZHOU Mu. Latern: Dynamic continuous hand gesture recognition using FMCW radar sensor[J]. *IEEE Sensors Journal*, 2018, 18(8): 3278–3289. doi: [10.1109/JSEN.2018.2808688](https://doi.org/10.1109/JSEN.2018.2808688).
- [7] 王勇, 吴金君, 田增山, 等. 基于FMCW雷达的多维参数手势识别算法[J]. 电子与信息学报, 2019, 41(4): 822–829. doi: [10.11999/JEIT180485](https://doi.org/10.11999/JEIT180485).
WANG Yong, WU Jinjun, TIAN Zengshan, *et al.* Gesture recognition with multi-dimensional parameter using FMCW radar[J]. *Journal of Electronics & Information Technology*, 2019, 41(4): 822–829. doi: [10.11999/JEIT180485](https://doi.org/10.11999/JEIT180485).
- [8] LI Gang, ZHANG Rui, RITCHIE M, *et al.* Sparsity-driven micro-doppler feature extraction for dynamic hand gesture recognition[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2018, 54(2): 655–665. doi: [10.1109/TAES.2017.2761229](https://doi.org/10.1109/TAES.2017.2761229).
- [9] RYU S J, SUH J S, BAEK S H, *et al.* Feature-based hand gesture recognition using an FMCW radar and its temporal feature analysis[J]. *IEEE Sensors Journal*, 2018, 18(18): 7593–7602. doi: [10.1109/JSEN.2018.2859815](https://doi.org/10.1109/JSEN.2018.2859815).
- [10] SZEGEDY C, VANHOUCKE V, IOFFE S, *et al.* Rethinking the inception architecture for computer vision[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 2818–2826. doi: [10.1109/CVPR.2016.308](https://doi.org/10.1109/CVPR.2016.308).
- [11] SZEGEDY C, LIU Wei, JIA Yangqing, *et al.* Going deeper with convolutions[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 2015: 1–9. doi: [10.1109/CVPR.2015.7298594](https://doi.org/10.1109/CVPR.2015.7298594).
- [12] DUMOULIN V and VISIN F. A guide to convolution arithmetic for deep learning[J]. arXiv: 1603.07285, 2018.
- [13] CHEN Haiquan and YE Wenbin. Classification of human activity based on radar signal using 1-D convolutional neural network[J]. *IEEE Geoscience and Remote Sensing Letters*, 2020, 17(7): 1178–1182. doi: [10.1109/LGRS.2019.2942097](https://doi.org/10.1109/LGRS.2019.2942097).
- [14] KONSTANTINIDIS E and COTRONIS Y. A quantitative roofline model for GPU kernel performance estimation using micro-benchmarks and hardware metric profiling[J]. *Journal of Parallel and Distributed Computing*, 2017, 107: 37–56. doi: [10.1016/j.jpdc.2017.04.002](https://doi.org/10.1016/j.jpdc.2017.04.002).
- [15] TRAN D, BOURDEV L, FERGUS R, *et al.* Learning spatiotemporal features with 3D convolutional networks[C]. 2015 IEEE International Conference on Computer Vision, Santiago, Chile, 2015: 4489–4497. doi: [10.1109/ICCV.2015.510](https://doi.org/10.1109/ICCV.2015.510).
- [16] LIN J, GAN Chuang, and HAN Song. TSM: Temporal shift module for efficient video understanding[C]. 2019 IEEE/CVF International Conference on Computer Vision, Seoul, Korea (South), 2019: 7082–7092. doi: [10.1109/ICCV.2019.00718](https://doi.org/10.1109/ICCV.2019.00718).
- [17] HAZRA S and SANTRA A. Robust gesture recognition using millimetric-wave radar system[J]. *IEEE Sensors Letters*, 2018, 2(4): 7001804. doi: [10.1109/LSSENS.2018.2882642](https://doi.org/10.1109/LSSENS.2018.2882642).
- 靳 标: 男, 1986年生, 博士, 副教授, 研究方向为雷达目标跟踪与识别、MIMO雷达发射波形设计等。
- 彭 宇: 男, 1996年生, 硕士生, 研究方向为雷达目标识别、深度学习。
- 邝晓飞: 男, 1996年生, 硕士生, 研究方向为MIMO雷达资源分配。
- 张贞凯: 男, 1982年生, 博士, 副教授, 研究方向为雷达通信一体化、认知雷达。

责任编辑: 余 蓉