

基于卷积神经网络的手机行为识别

张项博¹ 刘宜林¹ 王浩然¹ 周禾玺¹ 王宇轩²

1. 中国农业大学 山东烟台 100083

2. 高丽大学 韩国 02841

摘要: 随着智能手机运算能力的不断加强和越来越多的传感器装入智能手机,手机也不局限于通信功能,在人机交互、跌倒检测、信息安全等领域也产生了巨大的应用价值。基于手机传感器数据进行人体行为识别也成为研究热点。本文基于来自百度飞浆的手机行为识别数据通过构造卷积神经网络,并使用批归一化方法加快收敛速度,最终准确率达到96.7。提出的几种算法均能对六种行为有效识别。最后对几种模型进行分析,讨论了它们在行为识别中的应用优势和不足之处。

关键词: 行为识别; MLP; CNN; 批归一化

Mobile Phone Behavior Recognition Based on Convolutional neural network

Xiangbo Zhang¹, YiLin Liu¹, Haoran Wang¹, Hexi Zhou¹, Yuxuan Wang²

1. China Agricultural University, Yantai 100083, Shandong Province.

2. Korea University South Korea 02841

Abstract: With the continuous enhancement of computational capabilities in smartphones and the increasing integration of various sensors into them, smartphones have transcended their conventional role in communication and gained significant application value in areas like human-computer interaction, fall detection, and information security. Utilizing smartphone sensor data for human activity recognition has also become a prominent research focus. Based on mobile phone behavior recognition data from Baidu Feizi, this paper constructs convolutional neural network and uses batch normalization method to accelerate the convergence speed, and the final accuracy rate reaches 96.7. The algorithms proposed in this paper are effective in recognizing six distinct behaviors. The study concludes with an analysis of various models, discussing their strengths and weaknesses in the context of behavior recognition applications.

Keywords: Behavior Recognition; MLP; CNN; Batch Normalization

一、绪论

1. 背景和意义

近年来,智能手机的普及,使传感器获取数据更方便,相较于计算机视觉方法,基于传感器的行为识别可以更好地跟踪人类行为,反应更快,应用范围更广,实用价值更高。同时,智能手机的便携性也避免给受试者带来不必要的麻烦,可以在获取手机数据权限时,从大众日常生活中获取数据。因此,研究者们对智能手机集成的多种传感器数据及其应用价值愈发关注。

基于手机传感器对人体行为识别,其目的是通过传感器的观察数据,对人动作类型,行为模式等进行分析

和识别并使用自然语言将其描述或使智能设备完成相应行为,以更好的服务人们的日常生活。目前,基于传感器的行为识别已在医疗健康、智能环境等方面有重要应用。本文基于百度飞浆提供的手机行为识别数据集,使用深度学习方法进行行为识别,达到较好准确率。

2. 文献综述

(1) 手机行为识别研究概况

随着网络通讯技术和微电子技术的发展,智能手机集成了如加速度传感器、陀螺仪等多种传感器,对周围环境具有强大感知能力。目前,利用传感器进行行为识别较为常用的方法是从加速度传感器和陀螺仪进行数据

采集,从中特建立行为识别模型。杨晨晨等结合加速度和陀螺仪传感器,实时读取人体运动的加速度和角速度信息,从而识别跌倒行为。

(2) 深度学习在手机行为识别中的研究现状

深度学习方法与传统机器学习方法的不同在于,改变了传统方法的提取和特征选择步骤,在分类模型训练时可以自动生成特征,而非采用手工设计的特征,针对复杂场景和数据类型能更准确地表征真实人体行为特征。

石代伟等利用CNN自动提取特征,结合Softmax分类器进行分类识别,在人体6种行为模式上取得了较高的识别率。蒋翔宇设计了三种方案,通过改变卷积层卷积过程获得了更高的准确率。于涛基于RNN(Recurrent Neural Network,循环神经网络)及LSTM(Long Short Time Memory,长短期记忆)网络能够提取样本中的时间依赖特征,提出了基于强化学习的自适应采样率算法,在保证模型识别准确率的同时,尽量降低能耗。

金博等在时间卷积神经网络(TCN)基础上引入两类注意力机制,在公共人体行为数据集PAMAP2上准确率高达98.25%,相较原始TCN有较大提升。李东东在RNN和LSTM模型基础上加入了注意力机制,准确率得到显著提升。黄俊强使用transformer,基于多传感器信息融合方法有效提高了准确率。

二、相关技术综述

1. 基于传感器的人体行为识别框架

基于传感器的人体行为识别技术在各种研究中所使用传感器、算法、针对目标行为等略有不同,但其整体的研究框架基本类似。技术框架如下图2-1所示。

首先是数据采集。然后对收集到的序列数据进行校正、滤波、重采样等步骤去除噪声和误差。

基于机器学习技术需要先对原始序列数据进行特征提取,如时域、频域、时频域等,若提取到特征维度庞大,还需要进行特征降维,如使用LDA和PCA来提取和选择对目标影响大的变量,最后使用机器学习方法建模、优化参数等方法完成分类任务。

而基于深度学习技术的行为识别则通常可以通过浅层网络自动提取特征,直接进行模型构建即可。

最后在模型训练阶段,将预处理后的训练样本送入模型进行学习,并在测试样本上进行测试,进行模型比较和优化,选择出满足需求、性能优秀的模型。

2. 相关技术概述

(1) 卷积神经网络(2D-CCN)

卷积神经网络(CNN)是模仿人类视觉产生,可以

对图片等多维数据进行学习。其局部感知,权重共享和池化三个特性使CNN可以关注到空间局部性信息。局部感知使用卷积核实现,卷积核按照一定步长移动,将局部信息提取到下一层;权重共享使用一个卷积核完成一个二维张量的特征提取,减少卷积核的数量和需要学习的参数;池化提高网络的特征提取能力降低网络计算规模的同时,保留局部特征,扩大了下层网络的感受野。CNN主要由输入层、卷积层、池化层、全连接层与输出层组成。如图2-8是一个CNN结构图。对于本文分类任务而言,只需在最后一层添加soft Max函数即可。

(2) 本文模型评估标准

①准确率:指分类正确的样本占总样本数的比例,具体评价指标是测试集上的预测正确个数占总测试集样本数的百分比,取值在0-1之间。

②损失函数:损失函数是评估模型预测结果与实际结果之间差异的指标,由于是多分类问题,本文采用的损失函数为交叉熵。即对每个样本,计算真实标签的概率分布与模型预测标签的概率分布之间的交叉熵,然后将所有样本的交叉熵求平均。交叉熵损失函数越小,表示模型预测结果与实际结果之间的差异越小,模型的性能越好。

③F1值(F1 Score):F1值是一种用于评估分类模型性能的指标,是精确率和召回率的加权调和平均值。它综合了精确率和召回率的优点,可以更全面地评估模型的性能。其计算公式为:

$$F1 = 2 * (\text{精确率} * \text{召回率}) / (\text{精确率} + \text{召回率})$$

其中,精确率是指预测为正例且实际为正例的样本占预测为正例的样本比例,召回率是指实际为正例且被预测为正例的样本占实际为正例的样本比例。F1值的取值范围为0-1,值越大表示模型性能越好。

三、数据集的特征提取与预处理

本文实验数据原始数据集是UCI-HAR数据集,该数据集是由30名19-48岁年龄段的志愿者在腰部佩戴智能手机(三星Galaxy S II)进行六项活动(步行、步行楼上、步行楼下、坐、站、躺)采集得到。实验以50Hz的恒定速率捕获3轴线性加速度和3轴陀螺仪角速度的数据。共六个维度的序列数据。

基于该原始数据集,将原始数据三轴加速度和角速度六维变量各看作一个序列,基于收集到一段时间内数据对每个维度变量进行时域和频域的特征提取。在这些特征提取之后,最终每个样本生成并保留了561个特征和1个分类标签。从中按照均等概率保留8000个训练集样本

和2000个测试集样本。形成了本文实验所用数据集。

通过箱线图去除训练集中每个标签中的异常值，将结果保存为 tensor，float32 格式数据，并对特征数据进行标准化处理，绘制箱线图的观察：

由上图可以清晰看出各种运动中存在的离群值，但是离群值频繁出现的两个行为是躺和坐两个静止行为，其异常可能由于加速度计静止而角速度计有差异导致。例如人在坐时，可能会左右进行侧身等。所以这些异常值的存在可能是一种模型需要学习到的知识。为了验证上述想法是否正确，下文将在去除异常值和不去除异常值两种数据集上进行学习，最终比较其在测试集上的准确率。最后按照4：1的比例将8000个训练集样本分为训练集和验证集。

四、模型构建

1.模型设计

进一步对特征进行分析可知，特征中对于x，y，z三轴特征分别包含66个加速度计特征，40个角速度计特征，和1个综合特征。于是可以将特征基于x，y，z三轴的加速度、角速度特征划分为3个通道，每个通道（14，14）的数据形式。其中对每个通道数据使用0填充将187个特征填充为196个特征，使每个通道都是“正方形图片”数据格式。基于二维卷积神经网络，按照3个卷积层，两个全连接层，是否使用批归一化层为差别构建了三个模型，本模型搭建中，基于一维卷积神经网络模型经验，全部使用平均池化层，具体原因在将在结论中分析。如表1所示：

表1 二维卷积神经网络模型结构图

模型编号	模型结构	组合方式	模型训练
模型6	卷积层1 卷积层2 卷积层3 卷积层4 全连接层1 全连接层2	Conv2D1+ReLU+AvgPooling Conv2D2+ReLU+AvgPooling Conv2D3+ReLU+AvgPooling Conv2D4+ReLU+AvgPooling Linear1+Relu Linear2	步长：0.001 训练轮次：100 损失函数：交叉熵 优化算法：Adam
模型7	卷积层1 卷积层2 卷积层3 卷积层4 全连接层1 全连接层2	Conv2D1+ReLU+BN+AvgPooling Conv2D2+ReLU+BN+AvgPooling Conv2D3+ReLU+BN+AvgPooling Conv2D4+ReLU+BN+AvgPooling Linear1+Relu Linear2	步长：0.01 训练轮次：100 损失函数：交叉熵 优化算法：Adam (添加批归一化层，可以选择更大步长)
模型8	卷积层1 卷积层2 全连接层1 全连接层2	Conv2D1+ReLU+MaxPooling Conv2D2+ReLU+MaxPooling Linear1+Relu Linear2	步长：0.001 训练轮次：100 损失函数：交叉熵 优化算法：Adam

2.实验结果总览

将上述相应的数据送入10个模型分别训练后，性能

表2 实验结果一览

模型类别	模型序号	模型特点	验证集准确率 (单位：%)	F1值 (单位：%)	测试集准确率 (单位：%)
二维卷积神经网络	模型6	网络更深+平均池化层	95.15	95.1	95
	模型7	网络更深+平均池化层+批归一化层	96.9	96.9	96.7
	模型8	网络较浅+最大池化层	93.2	93.2	92.55

对比标准是在验证集上的准确率和F1值和和在测试集上的准确率。模型序号和其特点即三个指标得分如表2所示，其单位为%：

对于卷积神经网络而言，在较深的网络结构中加入批归一化层可以保留靠近输入端的特征，避免网络变深导致的梯度消失，也增强了模型计算性能，其准确率也是所有模型中最好的。同样地，平均池化优于最大池化且增加模型深度可以获得更高的准确率。

此外，对于第三章中提出的坐、躺动作中存在的离群值是否需要处理问题，结论是其中确实包含了静止动作的特有信息，不做删除处理在测试集上得分略高。对模型4和模型7的实验结果如下表3所示：

表3 是否处理异常值的对比

模型类别	模型序号	模型特点	去除异常值准确率 (单位: %)	不去除异常值准确率 (单位: %)
二维卷积神经网络	模型7	网络更深+平均池化层+批归一化层	95.85	96.7

五、结论和展望

如第三、四、五章所述,在进行较好的特征提取之后,卷积神经网络由于可以提取多个传感器之间的依赖信息而表现最优。多层感知机模型由于模型很难关注到不同传感器数据间的相互依赖,且计算基于整体而表现欠佳。循环神经网络在特征提取后的数据中可能因为多

个传感器变量输入中模型只能学习序列数据,忽视不同传感器之间差异,表现欠佳,循环神经网络和LSTM模型可能更适合对原始序列进行学习。

参考文献:

[1]孙晓杰.基于时空特征融合网络的人体行为识别技术研究[D].山东大学,2022.

[2]杨博.基于智能移动终端的人体运动识别技术研究与应用[D].西南交通大学,2017.

[3]蒋翔宇.利用CNN实现基于智能手机传感器的人体行为识别[D].兰州大学,2019.

备注:项目源码已发布在:https://blog.csdn.net/Hongrui_Chang;