

基于 9 轴姿态传感器的 CNN 旗语动作识别方法



钟岳 方虎生 张国玉 王钊 朱经纬

陆军工程大学野战工程学院 南京 210042

(1286486130@qq.com)

摘要 区别于传统光纤传感器、图像识别和 Kinect 深度图像的旗语动作识别方法,提出了一种基于 9 轴姿态传感器的旗语动作识别方法。该方法通过佩戴在手腕处的 9 轴姿态传感器来采集旗语动作的 3 轴加速度、3 轴角速度以及 3 轴磁偏角数据;在运用卷积神经网络(Convolutional Neural Network, CNN)分类模型的基础上对其中的数据进行预处理,并通过分类识别算法对其进行改进;在数据预处理阶段,利用小波分解与重构函数对采集到的 9 轴数据进行高频去噪和低频信息提取,通过时间序列加窗进行分割处理,对各动作样本进行维度和长度统一;在特征提取阶段,采用构建的双卷积层、单池化层、单全连接层网络模型对重构数据进行特征提取;在分类识别阶段,提出一种 CrossEntropy-Logistic 联合损失函数来对 5 种动作进行迭代训练。实验结果表明,所提方法利用 detcoef 小波分解与重构函数对信号进行低频细节系数提取并采用一维 CNN 对降噪后的数据进行特征提取,通过 CL 联合损失函数对预测损失值和预测概率进行融合,分析所得到的训练准确率与测试准确率,在与各类方法的对比中取得了最高值,其平均训练识别率可达 99% 以上,测试准确率可达 94%。

关键词: 卷积神经网络; 9 轴姿态传感器; 旗语动作识别; CrossEntropy-Logistic; 小波分解与重构

中图分类号 TP212.9

Method of CNN Flag Movement Recognition Based on 9-axis Attitude Sensor

ZHONG Yue, FANG Hu-sheng, ZHANG Guo-yu, WANG Zhao and ZHU Jing-wei

College of Field Engineering, Army Engineering University of PLA, Nanjing 210042, China

Abstract Different from the traditional method of flag movement recognition of optical fiber sensor, image recognition and kinect depth image, this paper proposes a method of flag movement recognition based on 9-axis attitude sensor. The data of 3-axis acceleration, 3-axis angular velocity and 3-axis magnetic decrement are collected by wearing a 9-axis attitude sensor at the wrist; based on the CNN classification model, the algorithm of data preprocessing and classification recognition is improved; in the data preprocessing stage, the wavelet decomposition and reconstruction functions are used to carry out the high-frequency denoising and low-frequency information extraction of the collected 9-axis data, and the dimension and the length of each action sample are unified through time series windowing and segmentation; in the feature extraction stage, the constructed network models of double convolution layer, single pooling layer and single full connection layer are used to extract the features of the reconstructed data; in the stage of classification and recognition, a CrossEntropy-Logistic joint loss function is proposed to carry out iterative training for 5 actions. The experimental results show that the use of wavelet decomposition and reconstruction detcoef function coefficient of low frequency detail of signals is extracted by using one-dimensional CNN data feature extraction. The training and testing accuracy obtained by the fusion analysis of the predicted loss value and the predicted probability through CL joint loss function is the highest in comparison with various methods. The average training recognition rate can reach more than 99% and the testing accuracy can reach 94%.

Keywords CNN, 9-axis attitude sensor, Flag action recognition, CrossEntropy-Logistic, Wavelet decomposition and reconstruction

1 引言

在计算机虚拟仿真训练领域,人体动作的实时捕捉与智

能识别是人机自然交互的研究热点。旗语动作识别技术被广泛应用于交通、建筑、航海、军事等协同仿真训练中,其部署的灵活性、识别的准确率对于系统人机交互的性能和训练仿真

到稿日期:2020-05-06 返修日期:2020-08-06 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家重点研发计划(2016YFC0802904);国家自然科学基金(61671470)

This work was supported by the National Key R&D Program of China(2016YFC0802904) and National Natural Science Foundation of China(61671470).

通信作者:方虎生(fhssxxll@sohu.com)

水平具有重要影响。国内外对旗语动作识别开展了多方面的技术研究,早期技术采用光纤传感器,通过检测上肢的光纤弯曲角度对动作进行识别;Zhang 等^[1]提出了一种单目视觉识别算法,即利用单个摄像机实现多个旗语动作的识别与跟踪;西南交通大学 Li 等^[2]提出了一种基于 Kinect 体感技术的旗语识别方法,该方法基于 Kinect 感应器进行动作识别。传统的旗语动作识别方法对数据采集和设备要求较高,且在实际应用中旗语指挥者无法在场地走动或改变身体方位,导致设备的部署不够灵活,识别效果有待提升。本文将小波分析技术、深度学习技术和穿戴式技术相结合,利用 9 轴数字姿态传感器实时采集旗语指挥者动作的多元信息,研究小波分解重构和改进 CNN 方法,建立实验系统,通过信号滤波处理、特征提取和分类识别,实现对旗语动作的准确识别。

2 基于可穿戴式 9 轴姿态传感器的旗语动作捕捉

可穿戴技术(wearable technology)是一种将多媒体、传感器和无线通信等技术嵌入衣物以获得实时信号的信息采集技术。Rezaie 等^[3]利用 4 个加速度传感器组成人体行为识别网络用于健康监测;Ma 等^[4]利用多个传感器组合表示群体环境中的不变特征,提出了具有时间标签的异构数据表示学习方法来提取典型特征,以提高对人体常规动作的识别效率;Kim 等^[5]利用佩戴在人体全身的 17 个传感器采集建筑工人的实时动作,以检测其身体工作状况和预测危险系数,但多个传感器存在设备功耗、硬件成本较高以及动作相互干扰等问题^[6]。基于此,本文采用单个佩戴在人体左手腕处的 9 轴姿态传感器来采集 5 种上肢旗语动作数据,传感器如图 1 所示。实验所用 WT901C 数字姿态传感器集成高精度的 3 轴陀螺仪、3 轴加速度计和 3 轴地磁场传感器,采用高性能微处理器和先进动力学解算能够快速求解出模块当前的实时运动姿态。静态时传感器的姿态测量精度为 0.05° ,动态时为 0.1° ;加速度量程为 $\pm 2/4/8/16\text{ g}$;角速度为 $\pm 250/500/1\ 000/2\ 000^\circ/\text{s}$;角度 $\pm 180^\circ$;数据输出频率为 $0.1\sim 200\text{ Hz}$ 。



图 1 WT901C 数字姿态传感器示意图

Fig. 1 Schematic diagram of WT901C digital attitude sensor

实验阶段,将传感器佩戴于实验者的左手手腕处并通过 USB 接口连接至采集设备(计算机),将实时采集到的各类信号以文本形式传输至计算机,传感器动作采集接收平台以及佩戴方式如图 2 所示;数据监控端如图 3 所示。实验采集了 3 轴的加速度、角速度和磁偏角度(横滚角(x 轴)、俯仰角(y 轴)和航向角(z 轴))共计 9 个通道的数据,共计 5 种类别的旗语动作,将其分别编号为 A,B,C,D,E,其动作示范以及动作要领分别如图 4 所示和表 1 所列。



图 2 实验平台以及传感器佩戴方式

Fig. 2 Experimental platform and sensor wearing mode

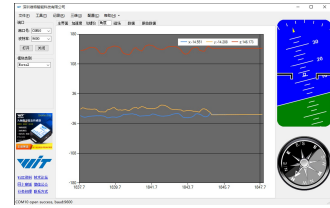
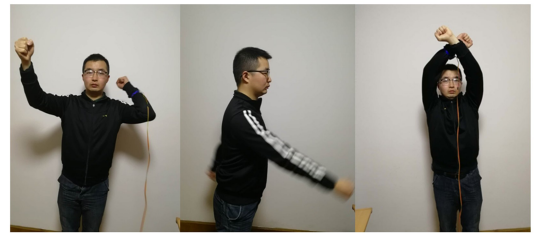


图 3 数据传输监控端

Fig. 3 Data transmission monitoring terminal



动作A

动作B

动作C



动作D

动作E

图 4 待测试旗语动作示意图

Fig. 4 Diagram of flag action to be tested

表 1 待测试旗语动作要领

Table 1 Movement essentials of flag to be tested

上肢动作名称	动作要领
A	大臂与身体垂直,肘部呈弯曲状,小臂置于头部两侧,交叉前后摆动
B	双臂自然落下,贴于身体两侧,前后摆动
C	双臂垂直举过头顶,手臂反复交叉与分开
D	双臂伸直举过头顶,双臂平行左右摇摆
E	左臂伸直,在身体正前方逆时针画圈

3 实验阶段

3.1 数据预处理

3.1.1 基于小波分解重构的信号去噪

数据预处理阶段,采用小波消噪方式对传感器采集的原始数据进行过滤,小波消噪的特点在于时频分辨率随分解尺度的变化而变化,对于信号的低频成分,其时间分辨率高但频率分辨率低;对于信号的高频成分,其时间分辨率低但频率分辨率高,适用于一维动作信号^[7]。目前小波去噪的方法主要包括:小波分解与重构、模极大值法、相关法、小波阈值法、平移不变量法等^[8]。由于实验中采集的信号大多为低频信号,选择小波分解重构方法更为合适,其具体分解过程为:设信号

S 的频率范围为 $[0, f]$, S 经多分辨率分析的第一层分解后得到两部分,即高频部分 D_1 和低频部分 A_1 ;其中,高频部分对应频率 $[f/2, f]$,低频部分对应频率 $[0, f/2]$;同时再将 D_1 和 A_1 做二次分解,分别得到 AA_2, DA_2 和 AD_2, DD_2 , 对应的频率分别为 $[0, f/4], [f/4, f/2], [f/2, 3f/4], [3f/4, f]$ 。以此类推,将原信号进行逐层分解,分解关系如式(1)所示^[9],小波包分解结构示意图如图 5 所示。

$$S = AAA_3 + DAA_3 + ADA_3 + DDA_3 + AAD_3 + DAD_3 + ADD_3 + DDD_3 \quad (1)$$

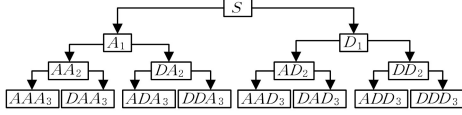


图 5 小波包分解结构示意图

Fig. 5 Schematic diagram of wavelet packet decomposition structure

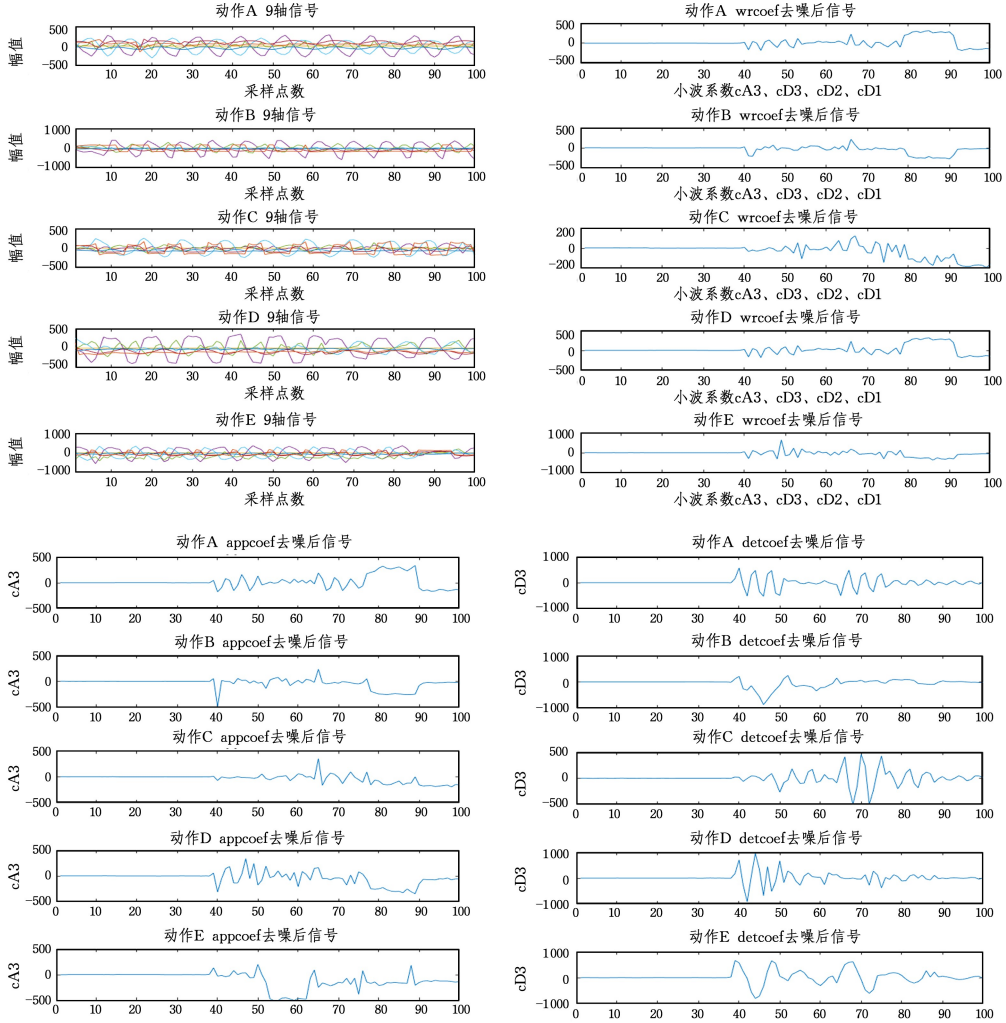


图 6 5 类动作原始信号及 3 种小波重构函数去噪结果

Fig. 6 5 kinds of original signal and 3 kinds of wavelet reconstruction function denoising results

3.1.2 样本数据分割以及数据标签

实验采集的是连续的动作时间序列,在特征提取前需对去噪后的数据进行加窗分割处理。在已采集样本数据中,5 种动作的样本数各为 100,样本总数为 500;同时,为使输入样本的时间长度一致,将每个动作的采样点数设为 100,因此输

实验选择了 3 种较为常见的小波重构函数,分别为 wrcoef 函数、appcoef 函数和 detcoef 函数。其中,wrcoef 是一种小波(单支)重构函数,通过分解 $[c, l]$ 组构中 A_1, A_2, A_3 系数得到各自的重构信号,适用于多尺度的一维小波变换;appcoef 函数和 detcoef 函数重构分解算法与 wrcoef 函数类似,区别在于 appcoef 主要通过离散小波变换(Discrete Wavelet Transform, DWT)提取小波的逼近系数以获得信号的高频信息,而 detcoef 通过 DWT 提取小波的细节系数以获得信号的低频信息^[10]。实验对数据均进行了在“ cA_1, cA_2, cA_3 ”这 3 种尺度下的小波分解,小波基函数采用‘db2’,提取了 3 种尺度系数下的小波重构结果。图 6 给出了原始信号及 3 种重构函数对信号提取后的结果。通过比较原始信号与 3 类小波重构信号可以发现,经小波包重构后的信号不仅更加完整平滑,同时也将带有一定周期性的原信号转化为更具差异性的融合信号。

入的单个样本可被视为大小为 $1 \times 100 \times 9$ 的矩阵。将 500 个动作均分为 20 组,每一组包含 25 个动作数据,每种动作各 5 个样本,将 20 组独立数据根据十折交叉验证方式划分为训练集和测试集,分别记为 $U01, U02, \dots, U020$,并将不同的动作数据按照 One-Hot 数据类型进行标记。样本标记如表 2 所列。

表2 上肢动作的数据结构

Table 2 Upper limb movement data structure

动作名称	类别标记	One-Hot 数据
A	0	(1,0,0,0,0)
B	1	(0,1,0,0,0)
C	2	(0,0,1,0,0)
D	3	(0,0,0,1,0)
E	4	(0,0,0,0,1)

3.2 CNN 工程平台搭建

3.2.1 总体网络结构

输入样本维度设为 $1 \times 100 \times 9$, 样本数量为 500, 将输入通道数设为 9。通过设置的两层卷积层和一层池化层对数据进行特征提取, 提取特征后, 采用 Relu 激活函数, 对每个动作样本在各个类别中的概率进行 (0,1) 预测。全连接层中, 将神经元个数设为 1000, 因此得到的数据大小为 1000×1 。将所得到的 1000×1 维数据作为 softmax 层的输入, 通过设定的损失函数 Loss 采用梯度下降函数来进行优化, 得到最终的识别结果。其总体结构如图 7 所示。

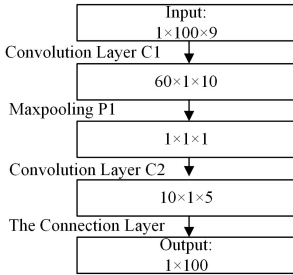


图7 CNN 总体网络结构图

Fig. 7 Overall network structure of CNN

3.2.2 卷积神经网络特征提取

卷积神经网络中, 卷积层的作用是对输入的数据进行特征提取, 通过设置的卷积核对输入的样本数据进行卷积操作, 偏置项对卷积后的数据进行误差修正, 最后使用激活函数得到输出特征^[11]; 每次的输入数据可以通过多个卷积核进行卷积, 公式如下:

$$x_j^l = f(u_j^l) = f\left(\sum_{i \in M_l} x_i^{l-1} k_{ij}^l + b_j^l\right) \quad (2)$$

其中, u_j^l 为卷积层 l 的第 j 个节点在特征映射被激活前的值, 它由对前一层的输出特征 x_i^{l-1} 进行卷积操作后与偏置项相加所得; x_j^l 为卷积层 l 的第 j 个特征映射的输出; M_l 为第一层输入特征子集; k 为卷积核矩阵; b 为对卷积后的特征映射的偏置项; f 为激活函数, 本实验中采用 Relu 函数, 其公式为:

$$f(x) = \max(0, x) \quad (3)$$

设置池化层进一步对矩阵进行降维, 池化层的输入来自上一层的卷积层, 输出结果为下一层卷积层的输入; 采用最大池化层, 计算过程如式(4)所示:

$$X_i^{l,j} = \frac{1}{N} * \left(\sum_{i=1, j=1}^n x_{i,j} \right) \quad (4)$$

其中, $X_i^{l,j}$ 表示进行池化后局部的第一个输出项, 其为上一层卷积后的输出求最大值, 经过多次卷积和池化操作后, 加入全连接层, 最后经 Softmax 层进行分类输出。第一个卷积层的

layer1-conv1 中, 设置 60 个 1×10 的 9 通道卷积核, 偏置项根据上一层的卷积核数量设置为 60, 池化层采用 1×1 的池化核, 步长设为 20, 采用全 0 填充方式; 在第二个卷积层 layer2-conv2 中, 参数设为 10 个 1×5 的 20 通道卷积核, 偏置项个数设置为 10, 不采用池化层; 全连接层中, 设置了 1000 个隐藏节点。该方法的模型结构及参数设置如表 3 所列。

表3 CNN 模型的结构及参数情况

Table 3 Structure and parameters of CNN model

结构层(类型)	输出大小	参数数量	下一连接层
Layer1-conv1 (Convolution1D)	(None, 20, 60)	6 653	Layer1-conv1_ input_1[0][0]
Layer-pool1 Maxpooling1D)	(None, 9, 60)	0	Conv1d_1[0][0]
Layer3-conv2 (Convolution2D)	(None, 5, 10)	56 164	Layer3-conv2_ [0][0]
Layer4-fc1 (Full Connection)	(None, 1000)	183 825	Layer3-conv2_ [0][0]
Dropout(Dropout)	(None, 500)	0	Layer4-fc1_2[0][0]
Output(Dense)	(None, 5)	7 546	Dropout[0][0]
参数总数: 254 188			

3.2.3 基于 CrossEntropy-Logistic 联合损失函数的分类识别任务

CNN 模型通常使用损失函数对卷积操作后的各个类别的实际值与预测值差距建立模型, 基于传统的交叉熵(Cross-Entropy)损失函数对小规模数据集进行分类时易出现过拟合现象^[12]。针对这一问题, Zhang 等^[13]采用一种 CenterLoss-Softmax(CS)和 A-Softmax(AS)联合优化损失函数, 对学校新生的人脸小数据集进行识别分类, 其平均识别率能够达到 98% 以上。这验证了基于联合损失函数在小规模数据集上的应用效果。

本文实验中对传统的交叉熵损失函数进行改进, 建立了一种基于 CrossEntropy 与 Logistic 加权联合的损失函数。CrossEntropy 损失函数被广泛应用于多分类任务中, 表征了预测输出在多个维度上与真实值的差距, Bosman 等^[14]提出了一种基于梯度的随机抽样方法对二次损失和熵损失两种不同误差度量产生的损失曲面进行实证研究, 证实了与二次损失相比, 熵损失具有更强的梯度和更少的固定点, 这验证了熵损失函数具有更优的全局搜索性; 同时在神经网络的应用中, 成功捕获到损失曲面中的全局最小值; 逻辑回归(Logistic)损失函数被广泛应用于二分类任务中, 表征了对于单个样本属于某种类别的 (0,1) 属性, 应用 Logistic 损失函数在经典的模式识别数据集 Iris 上基于部分特征和全部特征进行建模和分类预测, 证实了 Logistic 损失函数在较高权重特征组合的建模中可以获得较高的分类准确率^[15]。其中, Logistic 与 Cross-Entropy 损失函数分别如式(5)和式(6)所示:

$$\text{loss}(h_\theta(x), y) = \sum_{i=1}^m [-y_i \lg(h_\theta(x)) - (1-y_i) \lg(1-h_\theta(x))] \quad (5)$$

$$\text{loss}(x_i, y) = \sum_{i=1}^m (-x_i + \log \sum_j e^{(x_j - y_j)}) \quad (6)$$

其中, $h_\theta(x)$ 为二分类任务中每个样本的观测值; y 为每个样本的实际值; y_i 为每类样本的观测值; x_j 为一个样本观测值

中,每一类别的预测概率; m 为样本总数; i 为类别数, x_i 为多分类任务中每个样本的观测值,取值在(0,1)之间; y_j 为每个样本在不同类别下的实际分类情况,为 one-hot 值。当数据直接线性可分时,Logistic 回归函数通过数据的原始线性特征可以找到一条线性决策边界,但对于加速度、角速度和磁偏角融合数据无法实现直接线性可分,因此需要将特征映射到高维数据空间进行线性切分;交叉熵损失函数虽然适用于多类数据融合的分类模式,但易出现特征映射维度过高所导致的“维度灾难”问题,从而导致分类结果的过拟合现象^[16]。因此,在实验分类识别阶段,本文采用了一种 Softmax 与 Logistic 函数线性相加的联合损失函数,如式(7)所示,其中, a, b 分别为二分类和多分类任务的权重系数,即 $a+b=1$ 。

$$\text{loss}(h_{\theta}(x), x_i, y) = a\text{loss}(h_{\theta}(x), y) + b\text{loss}(x_i, y) \quad (7)$$

基于 CL 联合损失函数算法的分类步骤及思想如下所示。

(1) 利用 Logistic 二分类函数,对整个数据集进行二维线性分类,利用两条相交线性函数将动作数据集划分为 4 个初始类别,分别为 C_1, C_2, C_3, C_4 。

(2) 将类间距离问题转变为角度问题,通过对卷积神经网络全连接层权重与偏置值初始化的不断更新,来调整角度间隔参数,以此来不断对类间夹角进行调整,具体公式如下:

$$f_k^i = \begin{cases} \|F(x_i)\| \cos p\theta_{(k,i)}, & k=i \\ \|F(x_i)\| \cos \theta_{(k,i)}, & k \neq i \end{cases} \quad (8)$$

其中, $F(x_i)$ 为第 i 个样本的所有特征,对其取 2 范数; f_k^i 为 i 个样本属于第 k 类的特征; $\theta_{(k,i)}$ 为全连接层权重 w 的第 k 列与第 i 个样本的夹角; p 为角度间隔参数。每一次参数更新都对权重 w 的第 k 列进行归一化操作,同时设置偏置值 b 为 0。

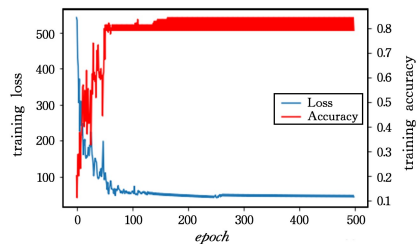
(3) 对 Logistic 初始分类数据进行高维特征映射,将平面数据转化为多维数据。

(4) 利用 CrossEntropy 多分类函数,对映射后的 4 类初始分类数据在高维数据空间中再进行类别 C_5 与 C_1, C_2, C_3, C_4 的线性分类任务。

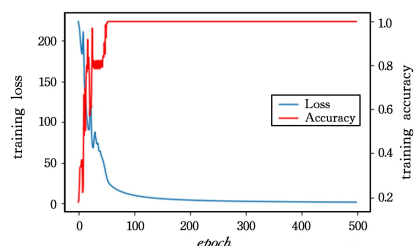
为避免损失函数在迭代中产生振荡以及局部搜索最优值等问题^[17],在确定 loss 损失函数形式后,采用 Adagrad 梯度下降法对 loss 函数的系数进行训练。表 4 给出了几种不同精度下,各类数据动作的平均损失值,经过 1000 次迭代后发现,当 a 取 0.63, b 取 0.37 时,训练效果达到最佳。为比较 CrossEntropy-Logistic 联合损失函数与单一的交叉熵损失函数在训练集上的效果,实验同时采用两种损失函数对预处理后的各类数据进行分类,其中,分别选取了两类损失函数在不同数据中的最高识别率进行对比。图 8 给出了使用 CrossEntropy-Logistic 损失函数前后的效果变化,使用 CrossEntropy-Logistic 函数前,loss 值分布于 0~500 的范围内,且震荡幅度剧烈,训练识别率(training accuracy)仅在 80%上下浮动;通过 CrossEntropy-Logistic 联合损失函数对 loss 值进行二分类与多分类任务加权融合,可以大幅度减少训练中带来的预测损失,将 loss 值稳定在 50 以内,并将训练识别率大幅提升至 99.4%。

表 4 不同 a, b 权值下的 loss 值Table 4 Loss values under different weights of a and b

系数取值	Loss value
$a=0.9, b=0.1$	0.543
$a=0.8, b=0.2$	0.433
$a=0.7, b=0.3$	0.256
$a=0.6, b=0.4$	0.145
$a=0.5, b=0.5$	0.273



(a) 使用联合损失函数前的分类结果



(b) 使用联合损失函数后的分类结果

图 8 使用 CrossEntropy-Logistic 联合损失函数前后的效果对比

Fig. 8 Before and after the use of CrossEntropy-Logistic combined loss function

3.3 实验过程及结果分析

信号重构分解过程在 matlab 环境中进行,训练及测试部分采用深度学习框架 Tensorflow,集成开发环境为 Spyder 编辑器。设置卷积神经网络学习率为 0.0001,测试迭代次数设为 500,分类器为 CrossEntropy-Logistic 联合损失函数,对模型进行 1000 次测试得到了每种动作的平均测试识别率。为比较各小波重构函数去噪后数据的识别效果,对 3 类重构函数去噪后的数据进行特征提取与识别,得到了 3 类重构信号与原始信号的测试准确率(test accuracy)的对比,如表 5 所列。

表 5 基于 CrossEntropy-Logistic 损失函数的各类重构信号的测试准确率

Table 5 Test accuracy of various reconstructed signals based on Crossentropy-Logistic loss function

	A	B	C	D	E
原始信号	73.12	79.48	70.14	65.33	83.32
wrcoef	86.54	84.25	78.56	80.36	87.22
appcoef	95.11	94.23	88.42	87.24	96.54
detcoef	93.53	95.26	90.22	91.55	98.77

由表 5 可知,经小波重构后的信号在分类识别过程中的效果明显要优于原始信号;而在 3 类不同重构信号识别结果的对比中发现,基于 detcoef 小波重构函数的动作数据在后续的特征提取与识别过程中展现出较为优异的结果,经 wrcoef 重构后的信号在识别过程中的效果较差;相比 appcoef 与 det-

coef, wrcoef 重构过程缺少对重构信号特征系数的提取^[18], 因此分类识别过程中存在由特征“维度缺乏”导致的参数初始化范围过大的问题; 相比 5 类动作信号, 采集过程中因晃动而产生的噪声信号明显较强。因此, 利用 appcoef 对信号进行分解重构时, 由于其高频提取特性^[19], 在滤波过程中部分噪声信号仍未被滤除, 导致分类结果中预测误差高于 detcoef, 这验证了在旗语动作识别过程中, 使用 detcoef 小波分解重构函数较为合适。通过对 5 类动作的频率比较发现, 动作 $A, B > C, D > E$, 因此在动作 A 的识别结果中可发现基于 appcoef 重构信号的识别结果略优于 detcoef; 同时, 由表 5 可知, 动作 E 的平均识别率最高, 动作 C, D 的平均识别率较低, 这是因为在 5 类动作类别中, 动作 E 的差异性明显高于其余 4 类动作, 动作 C, D 存在角速度与磁偏角信号差异不明显的部分, 从而导致在分类识别过程中由 C, D 信号的互相混淆造成的误差。因此在下一阶段的实验中, 应重点对相似信号的特征差异性展开研究, 以提升类间相似信号的识别准确率。基于本文的研究方法, 针对 5 类旗语信号识别的准确率最高达到 98.77%, 最低为 90.22%, 平均为 99.4%。实验结果说明, 利用 detcoef 函数对采集样本进行小波分解重构, 并采用 CrossEntropy-Logistic 联合损失函数进行分类, 可以达到较为理想的测试准确率, 本文研究的基于 9 轴姿态传感器的 CNN 分类模型在旗语动作信号识别中具有较强的泛化能力。

结束语 本文基于 CNN 框架完成了对 5 种典型的旗语动作的识别准确率的研究分析, 通过 wrcoef, appcoef, detcoef 这 3 类小波分解重构对数据进行去噪、关键信息提取及获取相应重构信号, 提出了一种基于 CrossEntropy-Logistic 联合损失函数的分类模型。经过实验发现, 基于 CL 联合损失函数的 detcoef 小波分解重构在测试集上的实验获得了较高的测试准确率, 这验证了 CNN 用于旗语动作识别的可行性, 为后续工作的展开提供了方法和实验借鉴。

参 考 文 献

- [1] ZHANG J L, WANG L L, RUI T, et al. Flag Language Recognition System Based on ANN [J]. Journal of PLA University of Science and Technology(Natural Science), 2003(4): 65-68.
- [2] LI P F. The Research of Action Recognition based on Kinect and Application of Semaphore Training [D]. Chengdu: Southwest Jiaotong University, 2015.
- [3] REZAIE H, GHASSEMIAN M. An Adaptive algorithm to improve energy efficiency in Wearable activity recognition systems [J]. IEEE Sensors Journal, 2017, 17(16): 5315-5323.
- [4] MA C M, ZHU Q, WANG S, et al. Representation learning from time labelled heterogeneous data for mobile crowdsensing [J]. Mobile Information Systems, 2016, 306: 22-30.
- [5] KIM K, CHO Y K. Effective inertial sensor quantity and locations on a body for deep learning-based worker's motion recognition [J]. Automation in Construction, 2020, 113: 27-40.
- [6] NGUYEN N D, TRUONG P H, JEONG G M. Daily wrist activity classification using a smart band [J]. Physiological Measurement, 2017, 38(9): 10-16.
- [7] SIDDHARTH, SHEKAR, CHEN C. Wavelet denoising of high-bandwidth nanopore and ion channel signals [J]. Nano Letters, 2019, 2: 1090-1097.
- [8] ZHE S, WEI N J. Based on Wavelet Decomposition and Reconstruction Method in Extraction of In-seam Reflection Signal [J]. Coal Technology, 2019, 38(12): 55-57.
- [9] LI J J, YU J P, TAO J. Review of One-Dimensional Signal Wavelet De-Noiseing [J]. Journal of Chongqing University of Technology(Natural Science), 2016(12): 93-89.
- [10] YANG J B, NGUYEN M N, SAN P P, et al. Deep Convolutional neural networks on multichannel time series for human activity recognition [C] // International Conference on Artificial Intelligence, 2015: 3995-4001.
- [11] ZHANG Z, SABUNCU M R. Generalized Cross Entropy Loss for Training Deep Neural Networks with Noisy Labels [C] // 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada, 2018, 11: 155-168.
- [12] LI Z, KAMNITSAS K, GLOCKER B. Overfitting of neural nets under class imbalance: Analysis and improvements for segmentation [J]. arXiv:1907.10982, 2019.
- [13] ZHANG X Y, YOU M Y, ZHU J, et al. Face Recognition of Small-Scale Dataset Based on Joint Loss Functions [J]. Transactions of Beijing Institute of Technology, 2020, 40(2): 163-168.
- [14] BOSMAN A S, ENGELBRECHT A, HELBIG M. Visualising basins of attraction for the cross-entropy and the squared error neural network loss functions [J]. Neurocomputing, 2020, 400: 113-136.
- [15] NGUYEN-VUONG Q T, GHAMRI-DOUDANE Y, AGOUL-MINE N. On utility models for access network selection in wireless heterogeneous networks [C] // Processings of Network Operations and Management Symposium 2008. Salvador: IEEE, 2001: 144-151.
- [16] ZOU X H. Data classification based on Logistic regression [J]. Intelligent Computer and Applications, 2016, 6(6): 139-140, 143.
- [17] MEIGNAN D, KNUST S. A neutrality-based iterated local search for shift scheduling optimization and interactive reoptimization [J]. European Journal of Operational Research, 2019, 279(2): 320-334.
- [18] LIU H J, ZHANG X, LI M. Determination of the Number of Wavelet Denoising Decomposition Layers in Vehicle Driving dE-valuation [J]. Noise and Vibration Control, 2019, 39(4): 7-11.
- [19] QIANG L, ZHE W. Image Compression Based on Wavelet Analysis [J]. Communications Technology, 2010, 43(8): 236-238, 241.



ZHONG Yue, born in 1996, postgraduate. His main research interests include artificial intelligence and pattern recognition.



FANG Hu-sheng, born in 1979, Ph.D., associate professor, is a member of China Computer Federation. His main research interests include computer technology and so on.