

基于目标均衡度量的核增量学习跌倒检测方法

忽丽莎¹, 王素贞¹, 陈益强^{2*}, 胡春雨², 蒋鑫龙², 陈振宇³, 高兴宇⁴

(1. 河北经贸大学 信息技术学院, 石家庄 050061; 2. 中国科学院计算技术研究所, 北京 100190;

3. 中国电力科学研究院, 北京 100192; 4. 中国科学院微电子研究所, 北京 100029)

(* 通信作者电子邮箱 yqchen@ict.ac.cn)

摘要: 针对增量学习模型在更新阶段的识别效果不稳定的问题, 提出一种基于目标均衡度量的核增量学习方法。通过设置经验风险均值最小化的优化目标项, 设计了均衡度量训练数据个数的优化目标函数, 以及在增量学习训练条件下的最优求解方案; 再结合基于重要性分析的新增数据有效选择策略, 最终构建出了一种轻量型的增量学习分类模型。在跌倒检测公开数据集上的实验结果显示: 当已有代表性方法的识别精度下滑至 60% 以下时, 所提方法仍能保持 95% 以上的精度, 同时模型更新的计算消耗仅为 3 ms。实验结果表明, 所提算法在显著提高增量学习模型更新阶段识别能力稳定性的同时, 大大降低了时间消耗, 可有效实现云服务平台中关于可穿戴设备终端的智能应用。

关键词: 增量学习; 神经网络; 核函数; 跌倒检测; 可穿戴设备

中图分类号: TP181 **文献标志码:** A

Objective equilibrium measurement based kernelized incremental learning method for fall detection

HU Lisha¹, WANG Suzhen¹, CHEN Yiqiang^{2*}, HU Chunyu², JIANG Xinlong², CHEN Zhenyu³, GAO Xingyu⁴

(1. Institute of Information Technology, Hebei University of Economics and Business, Shijiazhuang Hebei 050061, China;

2. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China;

3. China Electric Power Research Institute, Beijing 100192, China;

4. Institute of Microelectronics, Chinese Academy of Sciences, Beijing 100029, China)

Abstract: In view of the problem that conventional incremental learning models may go through a way of performance degradation during the update stage, a kernelized incremental learning method was proposed based on objective equilibrium measurement. By setting the optimization term of “empirical risk minimization”, an optimization objective function fulfilling the equilibrium measurement with respect to training data size was designed. The optimal solution was given under the condition of incremental learning training, and a lightweight incremental learning classification model was finally constructed based on the effective selection strategy of new data. Experimental results on a publicly available fall detection dataset show that, when the recognition accuracy of representative methods falls below 60%, the proposed method can still maintain the recognition accuracy more than 95%, while the computational consumption of the model update is only 3 milliseconds. In conclusion, the proposed method contributes to achieving a stable growth of recognition performance as well as efficiently decreasing the time consumptions, which can effectively realize wearable devices based intellectual applications in the cloud service platform.

Key words: incremental learning; neural network; kernel function; fall detection; wearable device

0 引言

随着医疗技术的发展与人民生活水平的提高, 人类社会日益步入老龄化阶段。如何实现健康老龄化, 提高老年人的生活质量是全社会亟待解决的难题。老年人在日常生活中由于平衡能力差、身体素质弱等因素, 极易发生跌倒^[1], 跌倒后如不能得到及时救治, 将会给老人带来严重的伤害, 甚至危及生命, 因此跌倒检测是面向老年人监护的一项重要研究内容^[2-3]。

已有的跌倒检测方法主要分为三大类^[4]: 1) 基于视频分析的方法。利用一个^[5]或多个^[6]相机和专用服务器实时跟踪目标, 视频资源上传至服务器并进行个体图像分割, 通过判定目标姿势的变化来检测是否发生了跌倒。该方法可在不影响用户的正常生活的同时实现对多位用户的跌倒检测监控; 但因相机运行范围受限、设备布置代价昂贵、隐私问题、相机盲点和对动态环境的闭塞等问题, 该方法仍存在许多不足。2) 基于声音与环境传感器的方法。利用红外线^[7]和麦克风^[8]等获取环境中的数据, 基于跌倒时产生的振动波频率等

收稿日期: 2017-09-25; 修回日期: 2017-11-20。 基金项目: 国家自然科学基金资助项目(61702491); 河北经贸大学校内科研基金资助项目(2016KYZ05); 国家电网公司总部科技项目(5442DZ170019); 中国电科院科技创新基金资助项目(5242001600H5)。

作者简介: 忽丽莎(1986—), 女, 河北石家庄人, 讲师, 博士, 主要研究方向: 机器学习、可穿戴计算; 王素贞(1964—), 女, 河北石家庄人, 教授, 博士, 主要研究方向: 移动云计算、大数据处理; 陈益强(1973—), 男, 湖南湘潭人, 研究员, 博士, 主要研究方向: 普适计算、人机交互; 胡春雨(1990—), 女, 山东威海人, 博士研究生, 主要研究方向: 机器学习、可穿戴计算; 蒋鑫龙(1988—), 男, 甘肃兰州人, 博士研究生, 主要研究方向: 机器学习、可穿戴计算; 陈振宇(1985—), 男, 湖南长沙人, 高级工程师, 博士, 主要研究方向: 机器学习、可穿戴计算; 高兴宇(1985—), 男, 湖南长沙人, 副研究员, 博士, 主要研究方向: 机器学习、可穿戴计算、多媒体。

检测跌倒。与基于视频分析的方法相比,该方法所需的硬件较为简单、低廉,也可实现多人检测,一定程度地缓和了隐私问题;但是由于使用场景的限制,以及真实环境中的噪声干扰等问题,严重影响了跌倒检测的精度。3) 基于可穿戴设备的方法。基于嵌入在用户衣物^[9]、手表^[10]、眼镜^[11]中的运动传感器采集数据,经过分析处理后检测是否发生了跌倒。与前面两种方法相比,基于可穿戴设备的跌倒检测方法具有价格低廉、配置简单、易携带等优点。本文重点研究基于可穿戴设备的跌倒检测方法。

常规的跌倒检测模型一般是基于大量用户的跌倒与非跌倒数据经过离线训练生成的^[12-13]。训练过程一结束,模型将被直接嵌入可穿戴设备中,从而为用户提供跌倒检测服务。一般情况下,这种静态式的通用模型不能很好地拟合用户的穿戴方式、行为等个性化特点。用户佩戴上设备以后,设备中的跌倒检测模型需要针对用户大量的个性化数据相应地进行自我调整,使之逐渐拟合该目标用户的个性化特点。为了达到这一目的,一些研究工作构建了基于可穿戴设备的增量学习跌倒检测模型^[14];此外,可穿戴设备的资源受限性使得模型的训练与测试过程均不能涉及大量的计算工作^[15]。

在众多的机器学习方法中,超限学习机(Extreme Learning Machine, ELM)模型因其训练速度快和泛化能力高等优点得到了广泛应用^[16-17]。以ELM为基础提出的正则化超限学习机(Regularized ELM, RELM)模型则成为解决两类分类、多类分类和回归问题的统一模型框架^[18]。当隐含层的映射函数未知或难以选择最优的映射函数时,可将RELM与核函数进行有效结合,从而构建一种核正则化超限学习机(Kernelized RELM)模型^[19]。以ELM、RELM或Kernelized RELM模型为基础的多种增量学习模型详见文献[20-33]。

本文重点研究以批量模型Kernelized RELM为基础构建的增量学习模型KB-ELM(Kernel Based Incremental ELM)^[22]和OKRELM(Online Kernelized RELM)^[34]。本文的主要工作如下:

1) 提出了一种核化的经验风险均值最小化超限学习机模型ALELM(Average-Lossed ELM),用于解决RELM模型推广至增量学习时出现的性能退化问题。

2) 以ALELM模型为基础,提出了一种性能鲁棒的增量学习模型Passive OALELM(Passive Online ALELM)和计算轻量的增量学习模型Active OALELM,后者可大幅减少模型更新和预测的时间消耗。

3) 在跌倒检测专用的公开数据集上的实验验证了本文的两种增量学习模型的识别精度和计算速度。

1 基础知识

1.1 Kernelized RELM

RELM^[16-17]是一个单隐含层前馈式神经网络,可用于解决两类分类、多类分类以及回归问题^[18]。结合使用核函数的Kernelized RELM模型则可以有效应用于那些隐含层的映射函数未知或难以选择的场景。本文重点关注利用Kernelized RELM模型解决分类问题。

Kernelized RELM只能用于离线训练的场景,是一种批量的学习模型。一个通用的行为识别模型可以基于大量用户的日常行为数据以离线的方式进行训练。通用模型到个性化模

型的转变,则需要根据持续采集的目标用户的个性化样本进行不断的更新操作,且更新过程是在线进行的。

1.2 KB-ELM

KB-ELM^[22]是一种将批量模型Kernelized RELM推广至增量式训练场景的增量学习模型,用于解决训练样本依次到达时的模型构建问题。KB-ELM模型的训练过程包含两个阶段:1) 初始构建阶段,基于初始数据集完成输出函数的初始化操作;2) 模型更新阶段,针对源源不断的新到数据完成输出函数的更新,最终实现分类模型的增量更新训练场景。

1.3 OKRELM

在KB-ELM模型的更新阶段,持续到达的新数据逐渐扩大了训练集的规模,同时也带来了模型更新阶段计算与存储需求的不断增长。此外,新数据中的一些冗余数据对模型的更新效果并不十分明显,然而这些数据在参与模型更新过程时带来了大量不必要的计算与存储消耗。针对KB-ELM模型的这一问题,增量学习模型OKRELM^[34]的核心思想是每次从新的数据块中挑选一些十分有价值的数据,仅基于挑选出的数据执行模型更新过程的计算。

2 基于目标均衡度量的核增量学习模型

2.1 ALELM

给定由 N 个训练样本组成的训练集合 $\{(x_i, t_i) | x_i \in \mathbf{R}^d, t_i \in \mathbf{R}^m, i = 1, 2, \dots, N\}$,其中:样本 x_i 表示 d 维的特征向量;向量 t_i 表示样本 x_i 的类标,且向量 $t_i = (t_{i1}, t_{i2}, \dots, t_{im})^T$ 中存在唯一的一个分量 t_{ik} 取值为1,表示样本属于第 k 类; t_i 其余的分量 $t_{ij}(j \neq k)$ 取值均为-1; N 和 m 分别表示训练样本个数和样本所属的类别个数。

为使更新后模型的性能不随训练样例个数的增加而出现不稳定的情况,本文首先提出一种新的批量学习模型ALELM。ALELM模型的优化问题如下:

$$\min_{\beta, \xi} \frac{1}{2} \|\beta\|^2 + \frac{C}{2N} \sum_{i=1}^N \|\xi_i\|^2 \quad (1)$$

$$\text{s. t. } \beta^T h(x_i) = t_i - \xi_i, i = 1, 2, \dots, N$$

其中: β 表示需要求解的 $L \times m$ 维输出变量; ξ 表示 $m \times N$ 维的松弛变量;矩阵 ξ 的第 i 列 ξ_i 表示样本 x_i 的训练误差; $h(x)$ 表示将 d 维样本 x 映射到某个未知的高维Hilbert空间的核函数 $k(\cdot, \cdot)$ 的映射函数,两者之间的关系为 $k(x, y) = h(x)^T h(y)$; C 是惩罚参数。

根据KKT(Karush-Kuhn-Tucker)定理,求解优化问题(1)等价于求解对应的对偶问题。其对偶问题的目标函数为:

$$D = \frac{1}{2} \|\beta\|^2 + \frac{C}{2N} \sum_{i=1}^N \|\xi_i\|^2 - \sum_{i=1}^N \sum_{j=1}^m \alpha_{ij} (\beta_{:,j}^T h(x_i) - t_{ij} + \xi_{ji}) \quad (2)$$

令 D 关于各自变量的偏导数为0,可以获得式(3)~(5):

$$\frac{\partial D}{\partial \beta_{:,j}} = \beta_{:,j} - \sum_{i=1}^N \alpha_{ij} h(x_i) = 0; \forall j \Leftrightarrow \beta = H^T \alpha \quad (3)$$

$$\frac{\partial D}{\partial \xi_{ji}} = \frac{C}{N} \xi_{ji} - \alpha_{ij} = 0; \forall i, j \Leftrightarrow \xi^T = \frac{N}{C} \alpha \quad (4)$$

$$\frac{\partial D}{\partial \alpha_{ij}} = \beta_{:,j}^T h(x_i) - t_{ij} + \xi_{ji} = 0; \forall i, j \Leftrightarrow H\beta - T + \xi^T = 0 \quad (5)$$

其中: T 表示 $N \times m$ 维的类标矩阵, $T_{i,j} = t_i^T$ 。

将式(3)和(4)代入式(5), α 可按式(6)进行求解:

$$\alpha = A^{-1}T \quad (6)$$

其中: $A = \Omega + \frac{N}{C}I$; Ω 表示数据块对应的 $N \times N$ 维 Gram 矩阵, Ω 中第 i 行且第 j 列的元素 $\Omega_{ij} = k(x_i, x_j)$; I 表示单位矩阵。

将式(6)代入式(3), β 的求解方式如下:

$$\beta = H^T \alpha \quad (7)$$

最后, ALELM 模型输出函数的计算方式为:

$$f(x) = \beta^T h(x) = T^T A^{-1} [k(x_1, x) \ k(x_2, x) \ \cdots \ k(x_N, x)]^T \quad (8)$$

ALELM 模型的训练过程见算法 1, 其中步骤 3 需要计算 $N \times N$ 维矩阵 A 的逆矩阵 A^{-1} , 复杂度为 $O(N^3)$, 是算法 1 中最为耗时的一个步骤。

在测试阶段, 针对任意测试样例 x , $f(x)$ 表示 $m \times 1$ 维的向量。ALELM 模型关于测试样例 x 的预测结果见式(9):

$$Pre(x) = \arg \max_{k \in \{1, 2, \dots, m\}} f^k(x) \quad (9)$$

其中: $f^k(x)$ 表示向量 $f(x)$ 的第 k 个分量。

算法 1 ALELM 模型的训练过程。

输入 包含 N 个训练样本的数据 (X, T) , 核函数 $k(\cdot, \cdot)$, 惩罚参数 C 。

输出 输出函数 $f(x)$ 。

- 1) 计算 Gram 矩阵 Ω 。
- 2) 计算矩阵 $A = \Omega + \frac{N}{C}I$ 。
- 3) 计算矩阵 A 的逆矩阵 A^{-1} 。
- 4) 计算输出 $f(x) = T^T A^{-1} [k(x_1, x) \ k(x_2, x) \ \cdots \ k(x_N, x)]^T$ 。

2.2 Passive OALELM

以批量模型 ALELM 为基础, 本文提出了经验风险均值最小化增量学习模型 (Passive OALELM), 用于解决当训练数据先后到达时模型的训练问题。与其他的增量学习模型类似, Passive OALELM 模型的训练过程也包含初始构建和模型更新两个阶段。下面对 Passive OALELM 的这两个阶段分别进行介绍:

1) 初始构建阶段。给定包含 N_0 个训练样本的初始数据块 (X_0, T_0) , 则 Lagrange 矩阵 α^0 和输出函数 $f^0(x)$ 的求解方式如下:

$$\alpha^0 = A_0^{-1} T_0 \quad (10)$$

$$f^0(x) = (\alpha^0)^T [k(x_1, x) \ k(x_2, x) \ \cdots \ k(x_{N_0}, x)]^T \quad (11)$$

其中: $A_0 = \Omega_0 + \frac{N_0}{C}I$ 。

2) 模型更新阶段。假定新到达的数据块 (X_1, T_1) 包含 N_1 个样本, 此时 Lagrange 矩阵 α^1 表示为:

$$\alpha^1 = \begin{bmatrix} A_0^{-1} & U_1 \\ U_1^T & \Omega_1 + \frac{N_0 + N_1}{C}I \end{bmatrix}^{-1} \begin{bmatrix} T_0 \\ T_1 \end{bmatrix} \quad (12)$$

其中: U_1 表示 $N_0 \times N_1$ 维的混合 Gram 矩阵, 且 $(U_1)_{ij} = k(x_i, x_{N_0+j})$, $i = 1, 2, \dots, N_0$, $j = 1, 2, \dots, N_1$ 。矩阵 A_0 和 D_1 的解释如下:

$$A_0 = A_0 + \frac{N_1}{C}I \quad (13)$$

$$D_1 = \Omega_1 + \frac{N_0 + N_1}{C}I - U_1^T (A_0')^{-1} U_1 \quad (14)$$

根据分块矩阵的求逆公式^[35], Lagrange 矩阵 α^1 可按如下方式进行求解:

$$\alpha^1 = \begin{bmatrix} (A_0')^{-1} + (A_0')^{-1} U_1 D_1^{-1} U_1^T (A_0')^{-1} & -(A_0')^{-1} U_1 D_1^{-1} \\ -D_1^{-1} U_1^T (A_0')^{-1} & D_1^{-1} \end{bmatrix} \quad (15)$$

此时, 输出函数 $f^1(x)$ 表示为:

$$f^1(x) = (\alpha^1)^T [k(x_1, x) \ k(x_2, x) \ \cdots \ k(x_{N_0+N_1}, x)]^T \quad (16)$$

根据式(15)~(16)可以发现, Passive OALELM 的模型更新过程涉及 $(A_0')^{-1}$ 和 D_1^{-1} 两个矩阵的求逆计算; 而矩阵 A_0 和 A_0^{-1} 已在初始构建阶段求出, 则 $(A_0')^{-1}$ 可以直接求逆, 也可进行迭代求解^[36]。

Passive OALELM 模型的训练过程见算法 2。初始构建阶段的计算复杂度与 ALELM 模型相同, 为 $O(N_0^3)$; 在模型更新阶段, 当包含 $N_{s+1} - N_s$ 个训练样本的新数据块达到时, 需要分别计算矩阵 A_s' 和 D_{s+1} 的逆矩阵, 计算复杂度为 $O(N_s^3) + O(N_{s+1} - N_s)^3$ 。

当第 n 个数据块 (X_n, T_n) 到达时, 依照算法 2 中模型更新阶段的计算步骤将输出函数更新为 $f^n(x)$ 。

算法 2 Passive OALELM 模型的训练过程。

初始构建阶段: $s = 0$, 包含 N_s 个训练样本的初始数据块 (X_s, T_s) , 核函数 $k(\cdot, \cdot)$, 惩罚参数 C 。

1) 计算 Gram 矩阵 Ω_s 。

2) 计算矩阵 $A_s = \Omega_s + \frac{N_s}{C}I$ 。

3) 计算矩阵 A_s 的逆矩阵 A_s^{-1} 。

4) 计算 $f^s(x) = T_s^T A_s^{-1} [k(x_1, x) \ k(x_2, x) \ \cdots \ k(x_{N_s}, x)]^T$ 。

模型更新阶段: 新到达包含 $N_{s+1} - N_s$ 个训练样本的数据块 (X_{s+1}, T_{s+1}) 。

1) 计算 Gram 矩阵 Ω_{s+1} 与混合 Gram 矩阵 U_{s+1} 。

2) 根据 A_s^{-1} 计算 $(A_s')^{-1}$ 。

3) 计算矩阵

$$D_{s+1} = \Omega_{s+1} + \frac{N_0 + N_1 + \cdots + N_{s+1}}{C}I - U_{s+1}^T (A_s')^{-1} U_{s+1}$$

4) 计算矩阵 D_{s+1} 的逆矩阵 D_{s+1}^{-1} 。

5) 计算矩阵 A_{s+1} 的逆矩阵

$$A_{s+1}^{-1} = \begin{bmatrix} (A_s')^{-1} + (A_s')^{-1} U_{s+1} D_{s+1}^{-1} U_{s+1}^T (A_s')^{-1} & -(A_s')^{-1} U_{s+1} D_{s+1}^{-1} \\ -D_{s+1}^{-1} U_{s+1}^T (A_s')^{-1} & D_{s+1}^{-1} \end{bmatrix}$$

6) 更新输出函数

$$f^{s+1}(x) = [T_0^T \ T_1^T \ \cdots \ T_{s+1}^T] A_{s+1}^{-1} [k(x_1, x) \ k(x_2, x) \ \cdots \ k(x_{N_{s+1}}, x)]^T$$

7) $s = s + 1$

2.3 Active OALELM

为了尽可能地降低模型更新阶段的计算复杂度, 借鉴 OKRELM 模型的核心思想, 本文提出在 Passive OALELM 模型的更新阶段仅利用错分的数据参与更新, 并将此模型称之为 Active OALELM, 训练过程见算法 3。对应地, 将 2.2 节中所有数据参与更新的模型称之为 Passive OALELM。

Active OALELM 模型初始构建阶段的计算复杂度与 Passive OALELM 模型相同; 在模型更新阶段, 当包含 $N_{s+1} - N_s$ 个训练样本的新数据块达到时, 仍需要分别计算矩阵 A_0' 和

D_{s+1} 的逆矩阵; 但此时由于矩阵 D_{s+1} 的维数仅与选择出的有效样例个数 (用 $|\Delta|$ 表示) 有关, 与 Passive OALELM 模型的模型更新阶段相比, 有效降低了计算量。Active OALELM 的计算复杂度为 $O(N_s^3) + O(|\Delta|)^3$ 。

算法3 Active OALELM 模型的训练过程。

初始构建阶段: $s = 0$, 包含 N_s 个训练样本的初始数据块 (X_s, T_s) 核函数 $k(\cdot, \cdot)$, 惩罚参数 C 。

执行算法2中初始构建阶段的步骤1) ~ 4), 获得矩阵 A_s^{-1} 和输出函数 $f^s(x)$ 。

模型更新阶段: 新到达包含 $N_{s+1} - N_s$ 个训练样本的数据块 (X_{s+1}, T_{s+1})。

1) 计算 Gram 矩阵 Ω_{s+1} 与混合 Gram 矩阵 U_{s+1} 。

2) 计算当前输出函数 $f^s(x)$ 关于 X_{s+1} 的输出 $f^s(X_{s+1}) = [T_0^T, T_1^T, \dots, T_s^T] A_s^{-1} U_{s+1} = F$ 。

3) 计算类别预测结果:

$$Pre(x_j) = \arg \max_{k \in \{1, 2, \dots, m\}} F_{kj}; j = 1, 2, \dots, N_{s+1} - N_s$$

4) 生成由所有被错分的数据组成的指标集

$$\Delta = \{j | Pre(x_j) \neq \arg \max_{i \in \{1, 2, \dots, m\}} t_{ij}; j = N_s + 1, N_s + 2, \dots, N_{s+1}\}$$

5) 更新矩阵 Ω_{s+1} 与 U_{s+1} 只与指标集中的数据有关, 即 $(\Omega_{s+1})_{ij} = k(x_i, x_j); i, j \in \Delta; (U_{s+1})_{ij} = k(x_i, x_j); i \in \{1, 2, \dots, N_s\}, j \in \Delta$ 。

6) 执行算法2中模型更新阶段的步骤2) ~ 6)。

3 实验及结果分析

3.1 数据集

为验证本文方法的性能, 在跌倒检测专用的公开数据集^[37]上开展了实验。该数据集包含32位用户(28男4女)4种朝向(前、后、左、右)的跌倒行为、1种疑似跌倒行为(失衡)和7种非跌倒日常行为(站立、平躺、走路、蹲伏、爬楼梯、坐床、坐椅子)。数据采集设备穿戴于用户的胸部和腿部, 每个部位分别包含一个3轴加速度计和一个3轴陀螺仪, 采样率均为100 Hz。数据集的规模介绍见表1中“特征提取前”这一阶段的内容。

表1 特征提取前后跌倒检测数据集的规模

Tab. 1 Scale of fall detection dataset before and after feature extraction

阶段	行为类别	数据个数	行为执行次数	用户个数
特征提取前	跌倒	240 717	448	32
	疑似跌倒	101 637	183	32
	非跌倒	2 203 310	1 162	42
	合计	2 545 664	1 793	42
特征提取后	跌倒	2 346	444	32
	疑似跌倒	689	178	32
	非跌倒	3 459	121	41
	合计	6 494	743	42

3.2 数据预处理与特征提取

图1展示了跌倒行为通常包含的三个阶段, 其中纵轴表示三轴加速度计的合成加速度(用均方根(Root Mean Square, RMS)表示), g 为重力加速度($1g = 9.8 \text{ m/s}^2$)。针对上述跌倒检测原始数据集的预处理操作包含三个步骤: 1) 将加速度计和陀螺仪的三轴数据通过计算均方根(RMS)的方式分别合并为一个轴, 构成合成加速度和合成角速度, 消除设备穿戴朝向对传感器读数的影响; 2) 针对跌倒这种特殊的异常行为本身的特点(见图1), 根据合成加速度数据识别跌倒过程中

的撞击前、撞击和撞击后三个阶段^[38]; 3) 基于每个阶段的加速度计和陀螺仪数据分别提取最小值、最大值、均值、方差和平方根五种简单的时域特征, 共计60维(5种特征 \times 3个阶段 \times 2种传感器 \times 2个穿戴位置)特征, 特征提取后生成的样例集合见表1中“特征提取后”这一阶段的内容。

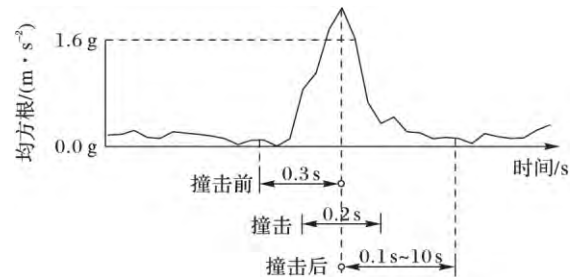


图1 一次跌倒行为的三个阶段

Fig. 1 Three phases during a fall activity

3.3 参数选择

为了验证本文方法的有效性, 核函数统一选取 Gaussian 核 $k(x_i, x_j) = \exp(-\sigma \|x_i - x_j\|^2)$ 。这样, 四种增量学习模型 KB-HELM、OKRELM、Passive OALELM 和 Active OALELM 均包含惩罚参数 C 、核参数 σ 、初始数据块大小 N_0 和新增块大小 $Block$ 四个参数。参数 C 和 σ 的取值范围均设置为 $\{2^{-24}, 2^{-23}, \dots, 2^{24}, 2^{25}\}$; N_0 的取值范围为 $\{100, 200, \dots, 900, 1000\}$; $Block$ 的取值范围为 $\{50, 100, \dots, N_0\}$ 。利用 RRSS(Repeated Random Sub-Sampling)^[39] 方式对数据集进行划分, 选出的最优取值情况见表2。

表2 四种增量学习模型的最优参数取值

Tab. 2 Optimal parameter values for four incremental learning models

模型	惩罚参数 C	核参数 σ	初始块大小 N_0	新增块大小 $Block$
KB-HELM	2^7	2^3	600	200
OKRELM	2^7	2^3	600	200
Passive OALELM	2^{19}	2^3	600	200
Active OALELM	2^{19}	2^3	600	200

3.4 跌倒检测精度对比

为了对比 KB-HELM、OKRELM、Passive OALELM 和 Active OALELM 四种增量学习模型跌倒检测的精度, 本文在每次模型更新之后记录上述四种模型关于跌倒行为的查准率(Precision)和查全率(Recall)。训练集和测试集的划分方式见3.3节, 每种模型的参数取值见表2。图2显示了四种增量学习模型10次实验结果后 Precision 和 Recall 的平均值。需要强调的是, 这里的 Precision 和 Recall 指标仅用于衡量跌倒这一类行为的识别精度。

从图2可以看出: 随着后续新增数据块的陆续到达, KB-HELM 模型 Precision 和 Recall 的取值均下滑至60%以下, 并出现上下波动的情况; OKRELM 模型的 Precision 和 Recall 分别下滑至30.8%和50.23%, 也出现了性能退化和不稳定的现象。

与 KB-HELM 和 OKRELM 模型不同, Passive OALELM 和 Active OALELM 模型在整个模型更新阶段的 Precision 和 Recall 始终保持在95%以上, 均展示了精准鲁棒的识别结果, 从而显示了 ALELM 模型优化问题中经验风险均值最小化策略的优越性。

进一步, Passive OALELM 模型在 Precision 和 Recall 上均

优于 Active OALELM 模型,但在时间消耗方面 Passive OALELM 模型远比 Active OALELM 模型的代价高(见图4),因此可以根据需求的不同分别选择合适的模型。

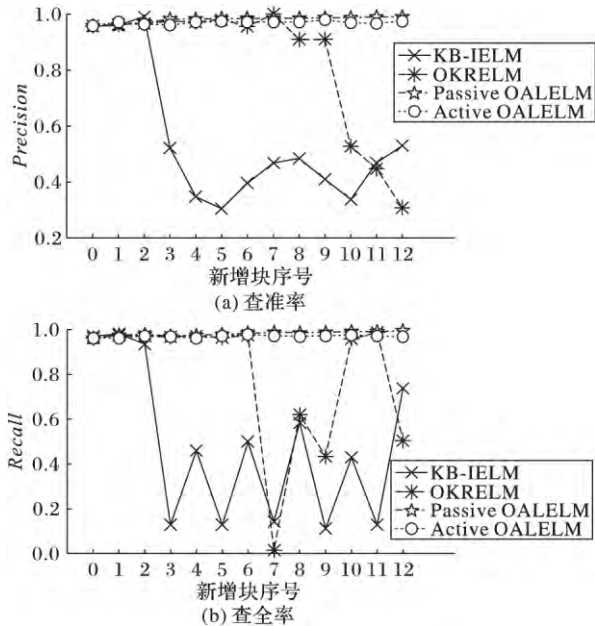


图2 四种增量学习模型关于跌倒行为的 Precision 与 Recall 结果对比
Fig. 2 Precision and Recall comparisons of four incremental learning models on the fall activity datasets

3.5 行为识别精度对比

为了评估模型关于跌倒与非跌倒二类分类问题的分类精度,本文统计了 KB-IELM、OKRELM、Passive OALELM 和 Active OALELM 四种增量学习模型在每次更新计算后的 $F1$ -Score 和 Accuracy。实验设置与 3.3 节一致,实验结果见图3。

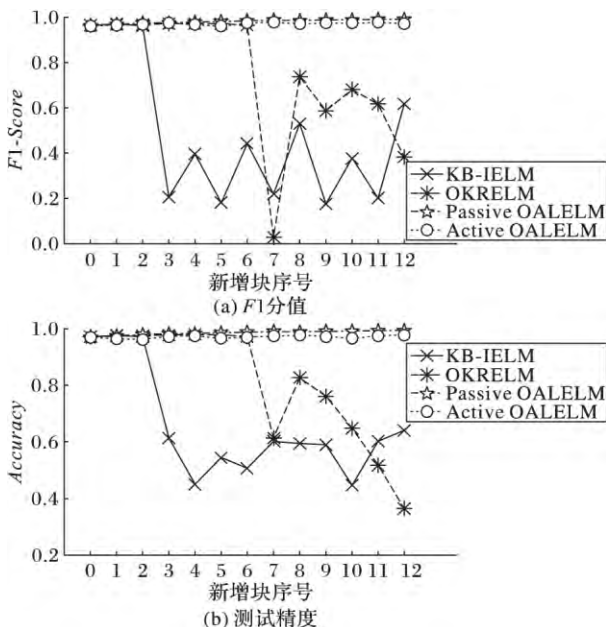


图3 四种增量学习模型关于跌倒行为的 $F1$ -Score 与 Accuracy 结果对比
Fig. 3 $F1$ -Score and Accuracy comparisons of four incremental learning models on the fall activity datasets

根据图3的实验结果可以看出,在第7个新增数据块更新结束后 OKRELM 模型的 Accuracy 为 61.47%,但 $F1$ -Score 仅为 2.98%。结合图2中的结果,这就意味着 OKRELM 模型在第7次更新结束后几乎将所有的跌倒样例分错,从另外一

个角度验证了 OKRELM 模型会出现性能不稳定的情况。

另一方面,Passive OALELM 模型的性能在整个模型更新阶段均超越了 KB-IELM 模型,从而显示了 Passive OALELM 模型在解决二类分类问题时的鲁棒性和有效性,而这正是得益于以之为基础的批量学习模型 ALELM 的优化问题中引入了经验风险均值最小化这一原则。总体而言,Passive OALELM 模型在 $F1$ -Score 和 Accuracy 两方面的性能均优于 Active OALELM 模型,因此 Passive OALELM 模型更适合于那些对鲁棒性和识别能力要求非常高的应用场景。

Passive OALELM 和 Active OALELM 两个模型的识别能力对比情况在图3中并不十分明显。3.6节将详细介绍 Passive OALELM 和 Active OALELM 模型在时间有效性方面的对比。

3.6 时间有效性对比

在训练时间消耗方面,KB-IELM 和 OKRELM 模型的初始构建阶段都是以 RELM 模型为基础,两者的初始训练时间相同,均为 0.04 s(见图4(a))。随着模型更新阶段新增数据块的不到达,KB-IELM 模型的更新时间消耗不断增加,并最终上升至 2.47 s;而 OKRELM 模型的更新时间消耗最终只增加至 0.11 s。与 KB-IELM 模型相比,Passive OALELM 模型由于在模型更新阶段需要额外计算一个逆矩阵(见 2.2 节),因此在时间有效性方面不如 KB-IELM 模型。而 Active OALELM 模型由于只利用少量被错分的样例参与模型的更新过程,因此继承了 OKRELM 模型更新速度快这一优势。事实上,Active OALELM 关于模型更新阶段的最长时间消耗仅为 0.09 s,更新速度甚至超过了 OKRELM 模型。

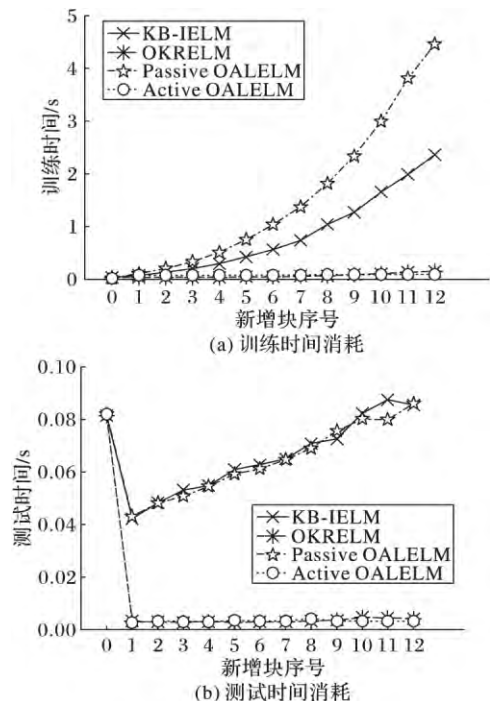


图4 四种增量学习模型的时间消耗情况对比
Fig. 4 Time consumption comparison of four incremental learning models

在测试时间消耗方面,经过最后一次更新后的 OKRELM 与 Active OALELM 模型的测试时长约为 0.003 s,而 KB-IELM 和 Passive OALELM 模型约为 0.08 s(见图4(b)),因此 Active OALELM 模型在测试阶段的时间优势也非常明显。

4 结语

跌倒作为一种特殊的异常行为, 严重伤害老年人的身心健康, 并给日益严重的老龄化社会带来异常沉重的经济负担, 因此, 跌倒检测技术的研究工作意义重大且十分关键。本文提出了一种基于目标均衡度量的核增量学习模型 OALELM, 用于面向可穿戴设备的跌倒检测; 并在跌倒检测公开数据集上验证了本文方法的有效性。实验结果显示了 OALELM 模型关于模型更新的鲁棒性。更进一步, Passive OALELM 模型在更新训练阶段显示了鲁棒精准的识别能力; Active OALELM 模型在一定程度地保证识别能力的同时有效提高了模型更新训练阶段和测试阶段的计算速度。

参考文献(References)

- [1] ROBIE K. Falls in older people: risk factors and strategies for prevention[J]. The Journal of the American Medical Association, 2010, 304(17): 1958–1959.
- [2] 沈秉乾, 武志勇, 贺国华, 等. 人体姿势状态判决的跌倒检测方法[J]. 计算机应用, 2014, 34(增刊1): 223–227, 264. (SHEN B Q, WU Z Y, HE Q H, et al. Falling detection method based on human body posture judgment[J]. Journal of Computer Applications, 2014, 34(S1): 223–227, 264.)
- [3] 罗丹, 罗海勇. 基于随机森林的跌倒检测算法[J]. 计算机应用, 2015, 35(11): 3157–3160, 3165. (LUO D, LUO H Y. Fall detection algorithm based on random forest[J]. Journal of Computer Applications, 2015, 35(11): 3157–3160, 3165.)
- [4] YU X. Approaches and principles of fall detection for elderly and patient[C]// HealthCom 2008: Proceedings of the 10th International Conference on E-health Networking, Applications and Services. Piscataway, NJ: IEEE, 2008: 42–47.
- [5] HU F, HAO Q, SUN Q, et al. Cyber physical system with virtual reality for intelligent motion recognition and training[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2017, 47(2): 347–363.
- [6] AUVINET E, MULTON F, SAINT-ARNAUD A, et al. Fall detection with multiple cameras: an occlusion-resistant method based on 3-D silhouette vertical distribution[J]. IEEE Transactions on Information Technology in Biomedicine, 2011, 15(2): 290–300.
- [7] TOREYIN B U, SOYER E B, ONARAN I, et al. Falling person detection using multi-sensor signal processing[J]. EURASIP Journal on Advances in Signal Processing, 2008, 2008: Article No. 29.
- [8] ZHUANG X, HUANG J, POTAMIANOS G, et al. Acoustic fall detection using Gaussian mixture models and GMM supervectors[C]// ICASSP 2009: Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway, NJ: IEEE, 2009: 69–72.
- [9] GJOESKI H, KOZINA S, GAMS M, et al. RAReFall — real-time activity recognition and fall detection system[C]// Proceedings of the 2014 IEEE International Conference on Pervasive Computing and Communications Workshops. Piscataway, NJ: IEEE, 2014: 145–147.
- [10] RAKHECHA S, HSU K. Reliable and secure body fall detection algorithm in a wireless mesh network[C]// BodyNets 2013: Proceedings of the 8th International Conference on Body Area Networks. Brussels, Belgium: ICST, 2013: 420–426.
- [11] CHEN Z, CHEN Y, HU L, et al. Leveraging two-stage weighted ELM for multimodal wearables based fall detection[C]// Proceedings of ELM-2014. Berlin: Springer, 2015, 2: 161–168.
- [12] QU W, LIN F, WANG A, et al. Evaluation of a low-complexity fall detection algorithm on wearable sensor towards falls and fall-like activities[C]// Proceedings of the 2015 IEEE Signal Processing in Medicine and Biology Symposium. Piscataway, NJ: IEEE, 2015: 1–6.
- [13] WANNENBURG J, MALEKIAN R. Physical activity recognition from smartphone accelerometer data for user context awareness sensing[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2016, 47(12): 3142–3149.
- [14] HU L, CHEN Y, WANG S, et al. Less annotation on personalized activity recognition using context data[C]// Proceedings of the 2016 International IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress. Piscataway, NJ: IEEE, 2016: 327–332.
- [15] HU L, CHEN Y, WANG S, et al. b-COELM: a fast, lightweight and accurate activity recognition model for mini-wearable devices[J]. Pervasive and Mobile Computing, 2014, 15: 200–214.
- [16] HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: theory and applications[J]. Neurocomputing, 2006, 70(1): 489–501.
- [17] HUANG G B, ZHU Q Y, SIEW C K. Universal approximation using incremental constructive feedforward networks with random hidden nodes[J]. IEEE Transactions on Neural Networks, 2006, 17(4): 879–892.
- [18] HUANG G B, ZHOU H, DING X, et al. Extreme learning machine for regression and multiclass classification[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2012, 42(2): 513–529.
- [19] SCARDAPANE S, COMMINELO D, SCARPINITI M, et al. Online sequential extreme learning machine with kernels[J]. IEEE Transactions on Neural Networks and Learning Systems, 2015, 26(9): 2214–2220.
- [20] LIANG N Y, HUANG G B, SARATCHANDRAN P, et al. A fast and accurate online sequential learning algorithm for feedforward networks[J]. IEEE Transactions on Neural networks, 2006, 17(6): 1411–1423.
- [21] GU Y, LIU J, CHEN Y, et al. Constraint online sequential extreme learning machine for lifelong indoor localization system[C]// Proceedings of the 2014 International Joint Conference on Neural Networks. Piscataway, NJ: IEEE, 2014: 732–738.
- [22] GUO L, HAO J H, LIU M. An incremental extreme learning machine for online sequential learning problems[J]. Neurocomputing, 2014, 128: 50–58.
- [23] GAO X, HOI S, ZHANG Y, et al. Sparse online learning of image similarity[J]. ACM Transactions on Intelligent Systems and Technology, 2017, 8(5): Article No. 64.
- [24] GAO X, CHEN Z, TANG S, et al. Adaptive weighted imbalance learning with application to abnormal activity recognition[J]. Neurocomputing, 2016, 173: 1927–1935.
- [25] JIANG X, LIU J, CHEN Y, et al. Feature adaptive online sequential extreme learning machine for lifelong indoor localization[J]. Neural Computing and Applications, 2016, 27(1): 215–225.
- [26] CHEN Z, CHEN Y, GAO X, et al. Unobtrusive sensing incremental social contexts using fuzzy class incremental learning[C]// Proceedings of the 2015 IEEE International Conference on Data Min-

- ing. Piscataway, NJ: IEEE, 2015: 71–80.
- [27] XIANG Z, CHEN Z, GAO X, et al. Solving large-scale TSP using a fast wedging insertion partitioning approach[J]. *Mathematical Problems in Engineering*, 2015, 2015: Article ID 854218.
- [28] JIANG X, CHEN Y, LIU J, et al. Real-time and accurate indoor localization with fusion model of Wi-Fi fingerprint and motion particle filter[J]. *Mathematical Problems in Engineering*, 2015, 2015: Article ID 545792.
- [29] CHEN Z, CHEN Y, HU L, et al. ContextSense: unobtrusive discovery of incremental social context using dynamic bluetooth data[C]// *UbiComp 2014: Proceedings of the 2014 Adjunct Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. New York: ACM, 2014: 23–26.
- [30] ZHANG H, YUAN J, GAO X, et al. Boosting cross-media retrieval via visual-auditory feature analysis and relevance feedback[C]// *MM 2014: Proceedings of the 22nd ACM International Conference on Multimedia*. New York: ACM, 2014: 953–956.
- [31] GAO X, HOI S, ZHANG Y, et al. SOML: sparse online metric learning with application to image retrieval[C]// *AAAI 2014: Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*. Menlo Park: AAAI Press, 2014: 1206–1212.
- [32] CHEN Z, CHEN Y, WANG S, et al. Inferring social contextual behavior from bluetooth traces[C]// *UbiComp 2013: Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*. New York: ACM, 2013: 267–270.
- [33] CHEN Z, WANG S, SHEN Z, et al. Online sequential ELM based transfer learning for transportation mode recognition[C]// *Proceedings of the 2013 IEEE Conference on Cybernetics and Intelligent Systems*. Piscataway, NJ: IEEE, 2013: 78–83.
- [34] HU L, CHEN Y, WANG S, et al. OKRELM: online kernelized and regularized extreme learning machine for wearable-based activity recognition[J]. *International Journal of Machine Learning and Cybernetics*, 2016: 1–14.
- [35] HAGER W W. Updating the inverse of a matrix[J]. *SIAM Review*, 1989, 31(2): 221–239.
- [36] WANG W, HOU Z G, CHENG L, et al. Toward patients' motion intention recognition: dynamics modeling and identification of iLeg — an LLRR under motion constraints[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2016, 46(7): 980–992.
- [37] OJETOLA O, GAURA E, BRUSEY J. Data set for fall events and daily activities from inertial sensors[C]// *MMSys 2015: Proceedings of the 6th ACM Multimedia Systems Conference*. New York: ACM, 2015: 243–248.
- [38] OJETOLA O. Detection of human falls using wearable sensors[D]. Coventry, UK: Coventry University, 2014: 49–66.
- [39] ALTUN K, BARSHAN B. Human activity recognition using inertial/magnetic sensor units[C]// *HBU 2010: Proceedings of the First International Workshop on Human Behavior Understanding, LNCS 6219*. Berlin: Springer, 2010: 38–51.

This work is partially supported by the National Natural Science Foundation of China (61702491), the Research Foundation of Hebei University of Economics and Business (2016KYZ05), the Science and Technology Program of State Grid Corporation of China (5442DZ170019), the Science and Technology Innovation Program of China Electric Power Research Institute (5242001600H5).

HU Lisha, born in 1986, Ph. D., lecturer. Her research interests include machine learning, wearable computing.

WANG Suzhen, born in 1964, Ph. D., professor. Her research interests include mobile cloud computing, big data analysis.

CHEN Yiqiang, born in 1973, Ph. D., professor. His research interests include pervasive computing, human-computer interaction.

HU Chunyu, born in 1990, Ph. D. candidate. Her research interests include machine learning, wearable computing.

JIANG Xinlong, born in 1988, Ph. D. candidate. His research interests include machine learning, wearable computing.

CHEN Zhenyu, born in 1985, Ph. D., senior engineer. His research interests include machine learning, wearable computing.

GAO Xingyu, born in 1985, Ph. D., associate professor. His research interests include machine learning, wearable computing, multimedia.

(上接第 927 页)

- [24] ZHANG D, SUN L, LI B, et al. Understanding taxi service strategies from taxi GPS traces[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(1): 123–135.
- [25] RAO F, ZHANG X, DONG Y, et al. Understanding the mobility pattern of passenger-searching taxis[C]// *Proceedings of the 2014 IEEE 17th International Conference on Intelligent Transportation Systems*. Piscataway, NJ: IEEE, 2014: 290–295.
- [26] QIAO S, HAN N, ZHU W, et al. TraPlan: an effective three-in-one trajectory-prediction model in transportation networks[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(3): 1188–1198.
- [27] QIAO S, SHEN D, WANG X, et al. A self-adaptive parameter selection trajectory prediction approach via hidden Markov models[J]. *IEEE Transactions on Intelligent Transportation*, 2015, 16(1): 284–296.
- [28] 乔少杰, 韩楠, 李天瑞, 等. 一种基于空间编码技术的轨迹特征提取方法[J]. *中国科学: 信息科学*, 2017, 47(11): 1523–1537.
- (QIAO S J, HAN N, LI T R, et al. A trajectory feature extraction approach based on spatial coding technique[J]. *Chinese Science: Informationis*, 2017, 47(11): 1523–1537.)
- [29] QIAO S, HAN N, WANG J, et al. Predicting long-term trajectories of connected vehicles via the prefix-projection technique[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2017, PP(99): 1–11.
- SHANG Jiandong**, born in 1968, Ph. D., professor. His research interests include data mining, high performance computing.
- LI Panle**, born in 1992, M. S. candidate. His research interests include trajectory data mining.
- LIU Runjie**, born in 1972, Ph. D., associate professor. His research interests include characteristics of communication network, chaotic fractal method.
- LI Runchuan**, born in 1991, Ph. D. candidate. His research interests include intelligent medical treatment, machine learning.