

# 一种基于卷积神经网络深度学习的人体行为识别方法

王忠民 曹洪江 范琳  
(西安邮电大学计算机学院 西安 710121)

**摘要** 为提高基于智能终端的人体行为识别的准确率,提出一种基于卷积神经网络深度学习人体行为识别方法。该方法将原始数据进行简单处理,直接作为输入数据输入到卷积神经网络中,由卷积神经网络进行局部特征分析,得到特征输出项,直接输入到 Softmax 分类器中,可识别走路、跑步、上下楼梯、站立等 5 种动作。对比实验结果表明,其对不同的实验者的识别率达到 84.8%,证明了该方法的有效性。

**关键词** 行为识别,深度学习,卷积神经网络

中图分类号 TP391.4 文献标识码 A

## Method on Human Activity Recognition Based on Convolutional Neural Networks

WANG Zhong-min CAO Hong-jiang FAN Lin

(School of Computer Science and Technology, Xi'an University of Posts and Telecommunications, Xi'an 710121, China)

**Abstract** In order to improve the accuracy of human activity recognition based on intelligent terminal, we proposed a recognition method based on convolution neural network. We pre-process the raw acceleration data, and input the processed data directly into the convolution neural network to do local feature analysis. After processing, we got the characteristic output items, which can be directly inputted into the Softmax classifier, which can recognize five activity, such as walking, running, going downstairs, going upstairs and standing. By comparing the experimental results, the recognition rate of different experimenters is 84.8%, which proved that the method is effective.

**Keywords** Human activity recognition, Deep learning, Convolutional neural networks

## 1 引言

随着人们生活水平的不断提高、工作压力的不断增大,人们越来越需要健康的生活方式,对于健康颇有好处的体育运动成为人们日益关注的焦点。运动管理系统能够督促人们每天保持一定的运动量并提供心跳、血压等信息。在这种需求环境中信息科技和智能穿戴获得了迅速的发展。随着信息科学和智能穿戴的不断进步,基于传感器的人体行为识别开始走向社会舞台。其实时采集传感器数据,进行数据的预处理,通过数据的分析和运算从而判断出用户的当前活动状态,例如智能家居、老人跌倒识别系统等。目前,随身携带的智能手机内置加速度、陀螺仪、磁力计、方向传感器等多种传感器,它们可以应对不同的场景产生不同的数据,例如加速传感器可以就走路、跑步、上下楼等不同行为产生不同的较为准确的三维加速度数据,这样就为基于便携式智能手机应用于行为识别提供了可能性。

## 2 已有研究

### 2.1 深度学习

深度学习是机器学习理论中的一个新的研究领域,它的动机在于建立、模拟人脑进行分析学习的神经网络,它模仿人

脑的机制来解释数据,是无监督学习的一种。目前深度学习已经在图片识别、声音识别、自然语言处理等领域取得了巨大的成功。2013 年陈先昌等人通过改进卷积神经网络的模型进行手写体数值的识别,将 60000 个手写体数据用于训练 10000 个手写体数据进行测试,实验结果表明识别率达到 92%<sup>[12]</sup>。2008 年赵志宏等人将卷积神经网络用于车牌字符识别,取得了 98.68% 的识别率<sup>[15]</sup>。2012 年 O. Abdel-Hamid 等人将卷积神经网络用于声音信号的识别系统中,获得了较为理想的识别率<sup>[2]</sup>。

### 2.2 行为识别

国内外在人体行为识别方面已经有若干研究,早期的研究是可穿戴计算,即将传感器固定在特定的位置进行人的行为和手势识别。2004 年 MIT 的媒体实验室利用人体 5 个位置携带的加速度计来识别日常生活中的 20 种常见的行为,连续采集人体上不同部位携带的传感器数据,利用模式识别用户行为感知,分别对平均值、运动能量、频域熵等使用不同的分类器进行了测试,获得了较为准确的识别<sup>[11]</sup>。2010 年王喜昌等人将加速度计置于上肢,将获取的三轴加速度信息用于上肢动作的识别,获得了较为理想的识别准确率<sup>[11]</sup>。

随着智能手机风靡全球,基于智能手机的人体行为识别迅速开始兴起并取得了迅速发展。智能手机的人体行为识别

本文受国家自然科学基金项目(61373116)、陕西省教育厅产业化培育项目(2012JC22)、陕西省教育厅项目(15JK1653)、西安邮电大学青年教师科研基金资助项目(ZL2014-29)资助。

王忠民(1967—),男,博士,教授,主要研究方向为智能信息处理、人工智能, E-mail: zmwang@xupt.edu.cn; 曹洪江(1989—),男,硕士生,主要研究方向为移动情景感知; 范琳(1982—),女,硕士,副教授,主要研究方向为移动情景感知。

可以被看作是一种分类问题,其中输入数据是时域信号而输出结果是分类标签对应的类别。美国高通公司的 Gropop 等人将手机放到人体的不同部位,包括口袋、裤兜和手中等进行(行为和手机位置)二元组的识别,并获得了 92.6%和 66.8%的识别率<sup>[11]</sup>。国内 Chen 等人选择融合加速度计和 GPS 信息进行交通模式的识别,用于区分是否处于交通工具中的行为,结果难以区分静止时是否是处于交通工具上的静止。

### 3 深度学习理论研究

卷积神经网络是一种深度学习网络,包括卷积层和池化层,其先进行卷积过程再进行池化过程。其中卷积过程的输出项作为池化层的输入项,然后将池化层结果作为下次卷积层的输入项依次类推。卷积神经网络的关键部分如图 1 所示。

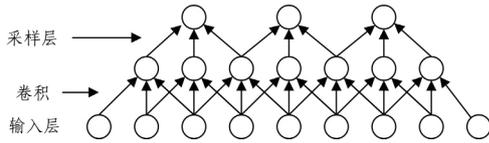


图 1 CNN 关键部分

#### 3.1 卷积层

源于对生物视觉细胞局部感受野的理解,卷积神经网络采用局部过滤器,进行卷积过程,即取输入项的局部子矩阵与局部过滤器进行内积运算,输出结果作为卷积输出矩阵的对应维数的值。为了形成更好的数据表现形式,卷积层提供多个这样的局部过滤器形成多个输出矩阵,每个输出矩阵的大小是  $(N-m+1)$ ,详细运算过程如式(1)所示。

$$X_i^{l,j} = f(b_j + \sum_{a=1}^m \omega_a^j x_{i+a-1}^{l-1,j-1}) \quad (1)$$

其中,  $X_i^{l,j}$  中的  $l$  表示第  $l$  个卷积层,  $i$  表示第  $i$  个卷积输出矩阵的某个值,  $j$  表示对应的输出矩阵的编号,从左往右依次表示为 0 到  $N$ ,  $N$  表示卷积输出矩阵的个数。  $f$  表示非线性函数,这里采用 sigmoid 函数。如图 1 所示,第一输入层作为输入项在第一个过滤层的作用下的输出矩阵的第 1 个位置上的结果如式(2)所示。

$$X_1^{1,1} = f(b_1 + \omega_1^{1,1} x_1^{0,1} + \omega_2^{1,1} x_2^{0,1} + \omega_3^{1,1} x_3^{0,1}) \quad (2)$$

#### 3.2 池化层

CNN 的池化层是一个进一步缩小矩阵的维数,同时不破坏数据内在联系的一个过程。池化层可以通过平均值构造也可通过最大值的方法来构造,本实验采用平均值构造,故也可以将该实验池化层称为均值池层。均值池层的输入来源于上一层的卷积层,输出结果作为下一层卷积的输入层。

均值池采用局部求均值的方式进行降维。详细运算过程如式(3)所示。

$$X_i^{l,j} = \frac{1}{N} * (\sum_{i=1, j=1}^n x_{i,j}) \quad (3)$$

其中,  $X_i^{l,j}$  表示进行池化过程后的局部的一个输出项,它是由上一层大小为  $n * n$  的局部小矩阵进行求均值所得到的。本实验中采用的均值池化过程由  $1/n$  的方阵与  $n * n$  的方阵调用卷积函数实现。

## 4 基于深度学习的行为识别模型

实验是在 Matlab 环境下进行的,涉及过程包括数据采集系统过程、基于卷积神经网络的数据处理过程和 softmax 分类器进行分类的过程。

### 4.1 数据采集系统

本文将 ARM 处理器的智能手机作为硬件平台,基于 Android 2.3 系统,搭建轻量级行为数据采集器,能实时进行加速度信息的采集。采集 15 名用户的行为数据,年龄分布在 20—50 岁,其中 20—30 岁 5 人,30—40 岁 8 人,40—50 岁 2 人,基本覆盖智能手机用户群;采集的行为信息分为 5 种,包括静止、走路、跑步、上楼、下楼。用户的手机穿戴位置明细为:包里、裤兜、手中。每个用户针对每种行为和位置采集 10 次运动过程中的加速度信息,每次采集时间为 10s。整个项目组采集的数据总和为 2250 组,其中一种行为有 450 组,每个位置有 750 组,如表 1、表 2 所列。

表 1 不同行为的样本数

行为样本	数量
Stay	450
Walk	450
Run	450
Upstairs	450
Downstairs	450
Total	2250

表 2 不同穿戴位置的样本数

位置样本	数量
Pocket	750
Hand	750
Bag	750
Total	2250

### 4.2 基于卷积神经网络的数据处理

本文整个实验流程包括对数据的预处理、数据的卷积神经网络处理过程和 softmax 分类过程,具体实验数据处理过程如图 2 所示。

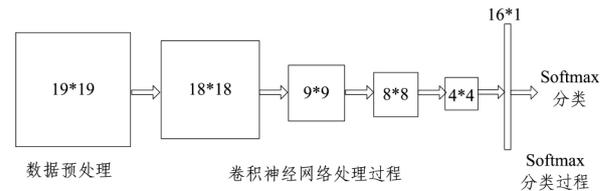


图 2 具体实验数据的处理过程

#### 4.2.1 数据的预处理

实验是采用上述数据采集系统所采集的数据进行的。在输入卷积神经网络进行训练和测试之前需要先对输入的数据进行预处理。采集的数据首先处理大小为  $120 \times 3$  的矩阵  $a = (a_x, a_y, a_z)$ ,其中  $a_x$  表示手机坐标系  $x$  方向上的加速度信息,  $a_y$  表示手机坐标系  $y$  方向上的加速度信息,  $a_z$  表示手机坐标系  $z$  方向上的加速度信息。实验中首先将矩阵  $a$  依次按手机坐标系  $x, y, z$  变化为  $1 \times 360$  的行向量  $b$ ,然后取出  $a_z$  中的最后一项添加到行向量  $b$  中形成  $1 \times 361$  的行向量  $c$ ,最后将行向量  $c$  处理为  $19 \times 19$  的矩阵  $d$ 。具体变化如图 3 所示,其中  $\triangle$  表示  $a_x$ ,  $\circ$  表示  $a_y$ ,  $\square$  表示  $a_z$ 。

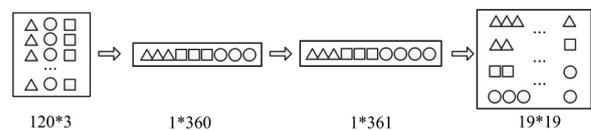


图 3 输入数据的变化过程

#### 4.2.2 数据的卷积神经网络处理过程

数据经过预处理过程后进入到卷积神经网络的第一个卷

卷积,卷积窗口大小为  $2 \times 2$ ,有 6 个这样的有着不同值的卷积窗口,于是得到 6 个不同的卷积映射矩阵,每个矩阵大小为  $18 \times 18$ 。中间处理过程包括局部卷积窗口的卷积过程和 sigmoid 函数的处理过程。卷积过程的处理过程如图 4 所示。

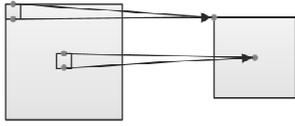


图 4 卷积处理示意图

数据经过卷积神经网络的第一层后进入池化层,池化层处理数据的方式有最大值方式和平均值方式等,其中最大值方式是将局部小矩阵的最大值作为映射层的输出,平均值方式是将局部小矩阵的平均值作为映射层的输出,本文采用平均值的方式进行池化层的采样过程。实验中的处理过程采用一个卷积窗口与要处理的矩阵进行卷积来完成,其中卷积函数的每项值都为  $1/n$ ,其中  $n$  为卷积窗口中元素的总个数。实验中经过池化层处理后的矩阵大小变为  $9 \times 9$ 。

数据经过卷积神经网络的第二层的池化层后进入第三层的卷积过程,第三层的卷积窗口共有 12 个,卷积窗口大小为  $2 \times 2$ ,第三层的输出为 12 个不同的  $8 \times 8$  的矩阵,其中每个矩阵是就卷积神经网络第二层的池化层中按一定的规律选取矩阵与对应的卷积窗口进行卷积获得的,如表 3 所列。第三层的第一个卷积层的输出是由前 3 个矩阵分别与第一个卷积窗口进行卷积然后求和所得,卷积的具体过程同卷积神经网络的第一层的卷积过程一样,其他过程同上。

数据经过第三层池化层后进入卷积神经网络的第四层池化过程,第四层的处理方式与卷积神经网络的第二层的池化过程是一样的。

表 3 卷积神经网络的第三层矩阵选取规律

	0	1	2	3	4	5	6	7	8	9	10	11	
0	X				X	X	X				X	X	X
1	X	X				X	X	X				X	X
2	X	X	X				X	X	X				X
3		X	X	X			X	X	X	X			
4			X	X	X			X	X	X	X		
5				X	X	X			X	X	X	X	

### 4.3 基于 softmax 的分类模型

经过卷积神经网络的数据处理过程后,数据被处理为一个  $16 \times 1 \times 12$  的列矩阵,输入到 softmax 分类器进行分离。softmax 分类器是将要分离的信号映射到相应的标签上。训练时,信号经过卷积神经网络的数据处理过程得到一个分类结果,将其与相应的标签数据进行比较算出相应的相对误差,通过训练一定的次数不断修正卷积神经网络中的卷积窗口上的权值使得相对误差不断减小,最后趋于收敛,然后将测试集输入到该网络中进行测试分类。本文的训练次数是 2000,得到的分类误差曲线如图 5 所示。

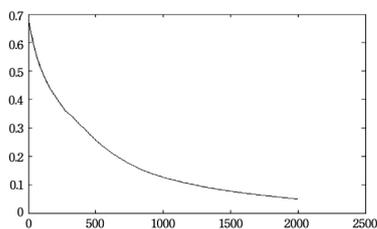


图 5 卷积神经网络训练误差曲线

## 5 实验结果与分析

课题组采用多分类算法进行行为识别,分别为 BP 神经网络、SVM 算法和卷积神经网络。分别采用两种方法进行测试。

(1)样本集测试。基于份样本进行建模,再取出其中的 450 份进行测试,实验结果显示卷积神经网络识别准确率达 92%。

(2)交叉验证。将整个样本随机分成 3 份取其中的 2 份进行训练,1 份进行测试,即随机选择 300 份样本进行训练,剩余的 150 份样本进行测试,其中 5 种行为各自随机取 30 个,交叉验证具有比样本测试集低的准确率,但有着更接近未标注样本的分类准确度。

交叉验证实验结果及准确率如表 4、表 5 所列。

表 4 具体实验结果

		待预测类别				
		Run	Walk	Stay	Upstairs	Downstairs
实验结果	Run	28	2	0	0	0
	Walk	2	21	0	4	3
	Stay	0	0	30	0	0
	Upstairs	0	3	0	27	0
	Downstairs	0	3	0	6	21

表 5 各种行为的识别率

行为样本	识别准确率(%)
Stay	100
Walk	70
Run	93.3
Upstairs	90
Downstairs	70

各种模型准确率如表 6 所列。通过对实验准确率的对比可知,卷积神经网络的识别率最高。因此,基于卷积神经网络建模更适合进行基于多特征的行为识别。

表 6 3 种常见行为识别模型的准确率对比

机器学习算法	测试集准确率(%)	交叉验证准确率(%)
传统神经网络	52.11	65.83
MVP	58.09	75.76
卷积神经网络	92	84.7

结束语 本文针对 5 种日常行为,将手机分别放在 3 种不同位置的情况进行了用户行为加速度信息采集,将采集到的数据按 3 种不同方法进行识别分类。与传统的 BP 神经网络、SVM 相比,卷积神经网络无需进行人为的特征优选过程且有着更高的识别率,其准确率能够达到 84.7%,是一种理想的分类模型,但走路、上下楼 3 种相似的行为信号的识别率较易区分的跑步和静止两种行为的识别率偏低。

## 参考文献

- [1] Zeng M, Nguyen L T, et al. Convolutional Neural Networks for Human Activity Recognition using Mobile Sensors[C]// Mobo SCASE. 2014:197-205
- [2] Abdel-Hamid O, Mohamed A R, Jiang H, et al. Applying convolutional neural networks concepts to hybrid nn-hmm model for speech recognition [C] // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2012:4277-4280

(下转第 87 页)

$$\begin{aligned}
& |a_i \wedge b_i| \\
= & 2 - \frac{1}{n^m} \sum_{i=1}^{n^m} |a_i - b_i| - \left| \frac{1}{n^m} \sum_{i=1}^{n^m} (a_i - b_i) \right| \\
= & 1 - \frac{1}{n^m} \sum_{i=1}^{n^m} |a_i - b_i| + 1 - \left| \frac{1}{n^m} \sum_{i=1}^{n^m} a_i - \frac{1}{n^m} \sum_{i=1}^{n^m} b_i \right| \\
= & \xi_1(A, B) + \xi_3(A, B) \\
= & \frac{1}{2} (2 - \xi_1(A, B) - \xi_3(A, B)) \\
= & \frac{1}{2} ((1 - \xi_1(A, B)) + (1 - \xi_3(A, B))) \\
= & \frac{1}{2} (\rho_1(A, B) + \rho_3(A, B))
\end{aligned}$$

定义 6 设  $A, B \in F(S)$ ,  $\rho_i(A, B) = 1 - \xi_i(A, B)$ ,  $i = 1, 2, 3$ , 称  $\rho_i(A, B)$  为  $F(S)$  上的伪距离,  $(F(S), \rho_i)$  为伪度量空间。

注 1:  $\rho_1(A, B) = \frac{1}{n^m} \sum_{i=1}^{n^m} |a_i - b_i|$ ,  $\rho_3(A, B) = \left| \frac{1}{n^m} \sum_{i=1}^{n^m} (a_i) - \frac{1}{n^m} \sum_{i=1}^{n^m} (b_i) \right|$ , 由定理 10 和定义 6 直接可得。

定理 15 设  $A, B \in F(S)$ ,  $k = 1, 2, 3$ , 则

- (i)  $\rho_k(A, A) = 0$ ;
- (ii)  $\rho_k(A, B) = \rho_k(B, A)$ ;
- (iii)  $\rho_k(A, C) \leq \rho_k(A, B) + \rho_k(B, C)$ 。

证明: 由定理 11 和定义 6 直接可得。

定理 16 设  $A, B \in F(S)$ ,  $k = 1, 2$ , 则

- (i)  $\rho_k(A \vee B, A) = 1 - \tau(B \rightarrow A)$ ;
- (ii)  $\rho_k(A \wedge B, A) = 1 - \tau(A \rightarrow B)$ ;
- (iii)  $\rho_3(A \vee B, A) \leq 1 - \tau(B \rightarrow A)$ ;
- (iv)  $\rho_3(A \wedge B, A) \leq 1 - \tau(A \rightarrow B)$ 。

证明: 由定理 12 和定义 6 得

- (i)  $\rho_k(A \vee B, A) = 1 - \xi_k(A \vee B, A) = 1 - \tau(B \rightarrow A)$ ;
- (ii)  $\rho_k(A \wedge B, A) = 1 - \xi_k(A \wedge B, A) = 1 - \tau(A \rightarrow B)$ ;
- (iii)  $\rho_3(A \vee B, A) = 1 - \xi_3(A \vee B, A) \leq 1 - \tau(B \rightarrow A)$ ;
- (iv)  $\rho_3(A \wedge B, A) = 1 - \xi_3(A \wedge B, A) \leq 1 - \tau(A \rightarrow B)$ 。

定理 17 设  $A, B \in F(S)$ , 则

- (i)  $\rho_1(A, B) \geq \rho_2(A, B) \geq \rho_3(A, B)$ ;
- (ii)  $\rho_2(A, B) = \frac{1}{2} (\rho_1(A, B) + \rho_3(A, B))$ 。

证明: 由定理 13、定理 14 和定义 6 得

$$\begin{aligned}
& \text{(i) } 1 - \xi_1(A, B) \geq 1 - \xi_2(A, B) \geq 1 - \xi_3(A, B), \text{ 则 } \rho_1(A, B) \geq \rho_2(A, B) \geq \rho_3(A, B) \\
& \text{(ii) } \rho_2(A, B) = 1 - \xi_2(A, B) \\
& \quad = 1 - \frac{1}{2} (\xi_1(A, B) + \xi_3(A, B)) \\
& \quad = \frac{1}{2} \times 2 - \frac{1}{2} \xi_1(A, B) - \frac{1}{2} \xi_3(A, B)
\end{aligned}$$

结束语 本文利用公式的赋值及赋值顺序给出了  $n$  值 Łukasiewicz 命题逻辑系统中公式真度的定义, 该定义与文献 [1] 中的定义等价; 同时介绍了 Łukasiewicz 命题逻辑系统中公式间的 3 种相似度的性质和关系, 并由此得到了 3 种伪距离, 建立了较广泛意义上的逻辑度量空间。本文的研究思想在其它多值命题逻辑系统中的性质将在另文中讨论。

## 参考文献

- [1] 王国俊. 数理逻辑引论与归结原理(2版)[M]. 北京: 科学出版社, 2006
- [2] 王国俊. 计量逻辑学(1)[J]. 工程数学学报, 2006, 23(2): 191-215
- [3] 惠小静. 模糊逻辑系统中广义有效推理的真度递减定理[J]. 模糊系统与数学, 2013, 27(4): 36-41
- [4] 周建仁, 吴洪博. Łukasiewicz 命题逻辑系统中真度的等价定义及相关性质[J]. 工程数学学报, 2013, 30(4): 580-590
- [5] 王国俊, 刘保翠. 4 种命题逻辑中公式的相对  $\Gamma$ -重言度理论[J]. 工程数学学报, 2007, 24(4): 598-610
- [6] 周红军. Łukasiewicz 命题逻辑中命题的 Borel 概率真度理论和极限定理[J]. 软件学报, 2012, 23(9): 2235-2247
- [7] 折延宏, 贺晓丽. 粗糙逻辑中公式的 Borel 型概率粗糙真度[J]. 软件学报, 2014, 25(5): 970-983
- [8] 惠小静, 王国俊. 经典推理模式的随机化研究及其应用[J]. 中国科学: E 辑, 2007, 37(6): 801-812
- [9] 惠小静. 三值  $R_0$  命题逻辑系统的随机化[J]. 应用数学学报, 2009, 32(1): 19-27
- [10] 王国俊, 惠小静. 概率逻辑学基本定理的推广[J]. 电子学报, 2007, 35(7): 1333-1340
- [11] 左卫兵. 多值逻辑系统中公式的  $\mu$ -真度理论[J]. 系统科学与数学, 2011, 31(7): 879-892
- [12] 马巧云, 吴洪博. 经典逻辑系统中公式的真度及公式间伪距离的一种等价定义[J]. 模糊系统与数学, 2013, 27(1): 28-33
- [13] 贺锦瑞, 惠小静, 双靖宁. 三值 Łukasiewicz 逻辑系统中公式的向量表示及其研究[J]. 计算机科学与探索, 2015, 9(9): 1147-1152

(上接第 58 页)

- [3] Bagci U, Bai L. A comparison of daubechies and gabor wavelets for classification of mr images[C]// IEEE International Conference on Signal Processing and Communications, 2007 (ICSPC 2007). IEEE, 2007: 676-679
- [4] Bengio Y. Learning deep architectures for ai [J]. Foundations and trends R in Machine Learning, 2009, 2(1): 1-127
- [5] Bhattacharya S, Nurmi P, Hammerla N, et al. Using unlabeled data in a sparse-coding framework for human activity recognition[J]. arXiv preprint arXiv:1312.6995, 2013
- [6] 王喜昌, 杨先军, 徐强, 等. 基于三维加速度传感器的上肢动作识别系统[J]. 传感器学报, 2010, 6(23): 816-819
- [7] 王忠民, 曹栋. 基于蚁群算法的行为识别特征优选方法[J]. 西安邮电大学学报, 2014, 9(1): 73-77
- [8] 衡霞, 王忠民. 基于手机加速度传感器的人体行为识别[J]. 西安

邮电大学学报, 2014, 19(6): 76-79

- [9] 徐姗姗. 卷积神经网络的研究与应用[D]. 南京: 南京林业大学, 2013
- [10] 陈先昌. 基于卷积神经网络的深度学习算法和应用研究[D]. 杭州: 浙江工商大学, 2013
- [11] 范琳, 王忠民. 穿戴位置无关的手机用户行为识别模型[J]. 计算机应用研究, 2015, 32(1): 63-66
- [12] 陈先昌. 基于卷积神经网络的深度学习算法与应用研究[D]. 杭州: 浙江工商大学, 2014
- [13] 杨心. 基于卷积神经网络的交通标识识别研究与应用[D]. 大连: 大连理工大学, 2014
- [14] 陆璐. 卷积神经网络的研究及其在车牌识别系统中的应用[D]. 合肥: 合肥工业大学, 2006
- [15] 赵志宏, 杨绍普, 马增强. 基于卷积神经网络 LeNet-5 的车牌字符识别研究[J]. 系统仿真学报, 2010(3): 638-641