

文章编号: 2095-2163(2019)06-0278-05

中图分类号: TP391.41

文献标志码: A

# 基于 ARM 智能视频监控人脸识别系统设计

郑嘉诚

(西安石油大学 计算机学院, 西安 710065)

**摘要:** 针对传统视频监控系统存在的不足,提出了一个基于 ARM 的人脸识别智能视频监控系统。本系统以 Raspberry Pi 为硬件平台,由视频捕捉设备获取视频流,采用 HOG 算法检测出人脸区域。通过 WiFi 模块上传给服务器,服务端程序使用 ERT 算法修正脸部姿势,采用 FaceNet 算法从对齐的人脸图像中提取 128 个 Embedding 特征测量值。通过与数据库中已存在的人脸特征值信息进行匹配,以实现人脸识别。检测到未知人员时通过微信报警。经测试该系统具有较高的稳定性和实用性。

**关键词:** 人脸识别; HOG; ERT; FaceNet

## Design of Intelligent Video Surveillance Face Recognition System Based on ARM

ZHENG Jiacheng

(School of Computer Science, Xi'an Shiyou University, Xi'an 710065, China)

**[Abstract]** Aiming at the shortcomings of traditional video surveillance systems, an ARM face recognition intelligent video surveillance system is proposed. The system uses Raspberry Pi as the hardware platform to obtain the video stream from the video capture device, and uses the HOG algorithm to detect the face region and upload it to the server through the WiFi module. The server program uses the ERT algorithm to correct the face pose, and then uses the FaceNet algorithm to align the face. The feature values of 128 Embedding are extracted from the face image, and matched with the existing face feature value information in the database to realize face recognition. WeChat alarm is triggered when an unknown person is detected. The system has been tested to have high stability and practicality.

**[Key words]** face recognition; HOG; ERT; FaceNet

## 0 引言

自美国“9.11”事件和西班牙马德里列车爆炸案、英国伦敦地铁大爆炸等恐怖袭击发生后,世界范围内对视频监控系统的需求也呈现了爆发式增长,各国部署的摄像头越来越密集,监控系统也日益庞大,全球视频监控市场获得了长足的发展<sup>[1]</sup>。据 Markets And Markets 预测,全球视频监控市场将从 2013 年的 139.8 亿美元增加至 2020 年的 420.6 亿美元。随着市场的快速增长、数据的积累、硬件能力的提升、算法的优化,视频监控技术需要从“高清化”向“智能化”发展。

视频监控的主要目标是人,所以人脸识别技术显得至关重要,它是后续各种更高级功能的基础。从 20 世纪 50 年代兴起了人脸识别的研究,作为一种重要的生物特征识别技术,一直被研究者所关注。人脸识别技术在安防和金融支持等方面有许多实际的应用,如视频监控、智能支持、访问控制等,是目前机器学习和模式识别中最热门的研究方向。

在嵌入式系统中应用计算机视觉(Computer

Vision, CV) 技术至今为止存在着多方面的难题。其中,如何使用嵌入式设备有限的计算资源实现例如人脸识别和姿态估计等较为复杂的功能。本系统针对以上问题,进行了初步的研究和探索。提出了以 Raspberry Pi 为硬件平台,以 Raspbian 为操作系统,通过 HOG 算法从摄像头的视频帧中提取检测到的人脸区域,通过 WiFi 模块将图像上传至服务器,在服务器上利用 ERT 算法将图像进行人脸对齐,通过采用 FaceNet 网络从修正过的人脸图像中提取 128 个嵌入人脸空间的特征测量值,并且与数据库存在的特征点进行对比,当两个特征向量之间的距离达到设定阈值 0.6 即可认为是同一人。本系统在 Pi-Motion 硬件平台上取得了较为理想的功能,是深度学习在实际应用中的一次有益尝试。

## 1 系统功能论述

本文的研究方案是设计一个基于 Raspberry Pi 平台下,经济实用、性能稳定和操作简单远程人脸识别系统。该系统可对监控区域内人员进行实时的身份识别,实现智能监控的目的。该系统主要具有

**作者简介:** 郑嘉诚(1994-),男,硕士研究生,主要研究方向:管理信息系统。

**收稿日期:** 2019-09-25

以下几个功能<sup>[2]</sup>。

(1) 人脸检测功能。摄像头在监控范围内发现人脸信息会上传至服务器, 并实现初步预警, 同时也为后续身份识别做准备。

(2) 视频流无线传输功能。解决传统视频监控通过布线方式的高安装成本, 可以在移动端或 PC 端方便查看。

(3) 人脸识别功能。能识别闯入的未知身份人员, 自动提示用户处理。

(4) 微信报警功能。如果监控区域被非法闯入自动记录, 并启动微信报警。

## 2 系统算法设计

在实际应用场景中人脸识别需要考虑受到人面部表情、装扮、姿态、角度和遮挡等原因引起的类内变化, 及来自外界光照、背景等不同引起的类间变化, 这两种变化的分布是高度复杂且非线性的。基于浅层学习的传统人脸识别方法, 在面对类内和类间这两种变化的复杂分布和非线性的情况无法达到理想效果<sup>[3]</sup>。本系统的人脸识别算法基于深度学习框架, 可以学习到对于光照、表情、角度等不变的特性。可以用来解决人脸识别中的类内和类间的变化分布这一难题。本系统算法主要包含人脸检测、人脸对齐、人脸识别三个主要流程。其算法如下。

### 2.1 人脸检测

人脸识别算法的第一步首先就是要从图像序列中检测出人脸部分, 通常人脸检测的算法目前主要有 PCA 算法、卷积神经网络算法、Adaboost 等算法。本系统主要采用性能较好的方向梯度直方图 (Histogram of Oriented Gradients, HOG) 特征检测算法<sup>[4]</sup>。通过 HOG 特征训练 SVM 分类器, 以实现从视频中人脸检测。

#### 2.1.1 HOG 特征计算过程

##### 2.1.1.1 颜色空间归一化

(1) 图像灰度化。将 RGB 图像转化为灰度图, 其式如下:

$$H(i, j) = 0.299 \times R(i, j) + 0.578 \times G(i, j) + 0.114 \times B(i, j). \quad (1)$$

(2) Gamma 校正。为了减少光照因素的影响, 首先需要将整个图像进行正则化。实验证明  $\gamma$  参数为 0.5 时有较高的性能提升。Gamma 正则化公式如下:

$$H(x, y) = H(x, y)^\gamma. \quad (2)$$

##### 2.1.1.2 单个 cell 单元计算梯度

(1) 将人脸的灰度图像以  $8 \times 8$  的像素邻域网格

(block) 作为采样窗口, 以没有重叠的方式遍布整个人脸<sup>[5]</sup>。

(2) 将每个 block 上以 cell (将 block 分为 4 个  $4 \times 4$  的块) 为单位计算梯度方向和幅值。方法用中心对称梯度算法  $[-1, 0, 1]$ , 梯度方向和幅值的计算公式为:

$$G_y(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2}, \quad (3)$$

$$\theta(x, y) = \tan^{-1}\left(\frac{G_y(x, y)}{G_x(x, y)}\right), \quad (4)$$

公式中的  $G_x(x, y)$  表示像素点  $(x, y)$  的水平方向梯度,  $G_y(x, y)$  表示像素点  $(x, y)$  的垂直方向梯度。其计算方法如下, 其中  $H(x, y)$  为点  $(x, y)$  的像素值:

$$G_x(x, y) = H(x + 1, y) - H(x - 1, y), \quad (5)$$

$$G_y(x, y) = H(x, y + 1) - H(x, y - 1). \quad (6)$$

##### 2.1.1.3 块内归一化梯度直方图

(1) 每一个 cell 内的相同梯度方向的幅值按照权值为 1 相加, 然后将 block 内各 cell 的直方图连成一个向量得到 HOG 特征。

(2) 将 block 内的直方图向量进行一次  $L2 - norm$  标准化, 设  $v$  是标准化前的特征向量,  $\|v\|_k$  表示  $k$  范数, 其中,  $k = 1, 2$ , 令  $\varepsilon$  为一个很小的标准化常量以防止除数为 0:

$$v \rightarrow \frac{v}{\sqrt{\|v\|_2^2 + \varepsilon^2}}, \quad (7)$$

(3) 经过  $L2 - norm$  后将  $v$  的阈值设置为 0.2, 将  $v \geq 0.2$  的分量赋值为 0.2 再做一次  $L2 - norm$  标准化:

$L1 - norm$ :

$$v \rightarrow \frac{v}{\|v\|_1 + \varepsilon}, \quad (8)$$

$L1 - sqrt$ :

$$v \rightarrow \sqrt{\frac{v}{\|v\|_1 + \varepsilon}}, \quad (9)$$

(4) 将所有标准化后的直方图向量组成一个  $n \times m$  的矩阵, 用于表示人脸 HOG 特征, 其中,  $n$  是 block 内直方图向量维数;  $m$  是整个人脸要计算的 block 个数。

#### 2.1.2 基于 HOG 的 SVM 分类器训练过程

(1) 提取正负样本的 HOG 特征。

(2) 投入 SVM 分类器训练, 得到 model。

(3) 由 model 生成检测子。

(4) 利用检测子检测负样本, 得到 hard example。

(5) 提起 hard example 的 HOG 特征并合并第一

步中的特征一起投入训练,得到最终检测子。

## 2.2 人脸对齐

本系统采用基于回归树的 ERT 算法完成人脸对齐。这种方法通过建立一个级联回归树来逐步修正人脸关键点的位置。在已知这些关键点的前提下就可以通过仿射变换使得人脸在图片中居中,使得下一步人脸识别的结果更加精确。其算法核心是使用两层的回归来建立数学模型<sup>[6]</sup>。

首先,有训练数据集  $(I_1, S_1), \dots, (I_n, S_n)$ , 其中  $I_i$  为图片,  $S_i$  为人脸关键点的位置。在第一层的回归训练中,数据组织形式可以写为  $(I_{\pi_i}, \hat{S}_i^{(t)}, \Delta S_i^{(t)})$ ,  $\hat{S}_i^{(t)}$  是第  $t$  层的预测关键点位置,  $\Delta S_i^{(t)}$  是这一层回归结果与真实值的差值。其迭代公式为:

$$\hat{S}_i^{(t+1)} = \hat{S}_i^{(t)} + r_t(I_{\pi_i}, \hat{S}_i^{(t)}), \quad (10)$$

$$\Delta S_i^{(t+1)} = S_{\pi_i} - \hat{S}_i^{(t+1)}. \quad (11)$$

按照这样的方式不断进行迭代,当第一层回归级联层数设置为  $K$  层时,就会产生  $r_1, r_2, \dots, r_k$  这些回归器,这  $K$  个回归器就是通过训练所得到的回归模型。

在第二层回归,也就是  $r_t$  内部也是一次回归过程,采用的是梯度提升树(Gradient Boosting Tree)<sup>[6]</sup>的方法来得到一系列的回归树最终完成第二级的回归。

每一个回归树的每一个叶子节点上都存储着一个残差回归量,当输入落到了一个节点上时,就将残差加到改输入上,起到回归的目的,最终将所有残差叠加在一起,最终得到人脸的关键点。

## 2.3 人脸识别

本文中关于人脸识别的部分采用 FaceNet<sup>[7]</sup> 的卷积神经网络的方法,其主要思想是把人脸图像映射到一个多维空间,比较两张图像所对应的特征欧式距离直接对应两个图像的相似度。同一个人脸图像的空间距离比较小,不同人脸图像的空间距离比较大。这样通过人脸图像的空间映射就可以实现人脸识别。

FaceNet 中采用基于深度神经网络的图像映射方法和基于 Triplet Loss 的损失函数训练神经网络,网络直接输出为 128 维度的向量空间,其结构如图 1 所示。

其主要思想如下:

(1) 对人脸样本图像经过深度学习框架进行特

征提取,主体模型采用一个深度网络 Inception-ResNet。

(2) 对特征进行 L2 归一化,使得特征的二范式满足式(12)条件,使得所有图像特征都映射到一个超球面上,规避样本的成像环境带来的差异:

$$\|f(x)\|_2 = 1. \quad (12)$$

(3) 经过 L2 归一化后得到图像的 128 维嵌入空间特征向量(embedding),采用 Triplet Loss 作为损失函数进行模型优化,更新 embedding。

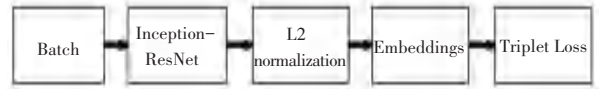


图 1 FaceNet 模型结构

Fig. 1 FaceNet model structure

## 3 系统设计与实现

系统硬件以 YAHBOOM 公司的 Pi-Motion 作为开发平台,子模块主要包括电源、USB 摄像头、WiFi 通信模块、BST-AI 拓展版、Raspberry Pi 开发板、扬声器、顶部舵机、底部舵机,如图 2-图 5 所示。

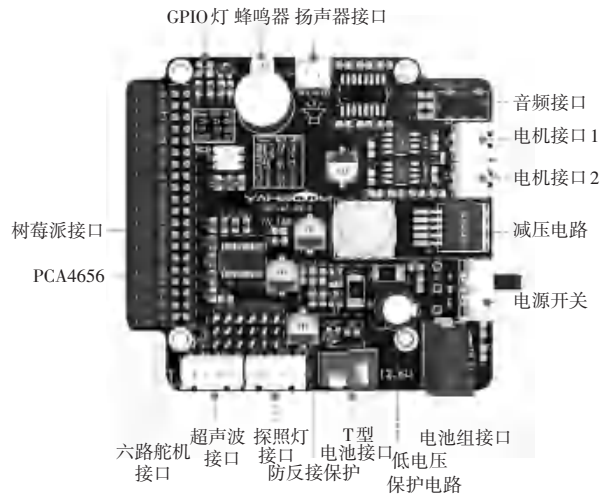


图 2 BST-AI 拓展板

Fig. 2 BST-AI expansion board

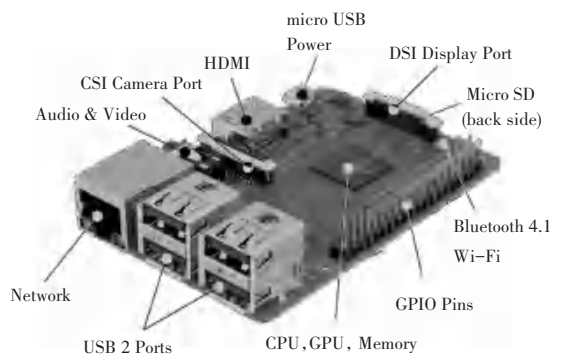


图 3 Raspberry Pi 开发板

Fig. 3 Raspberry Pi development board



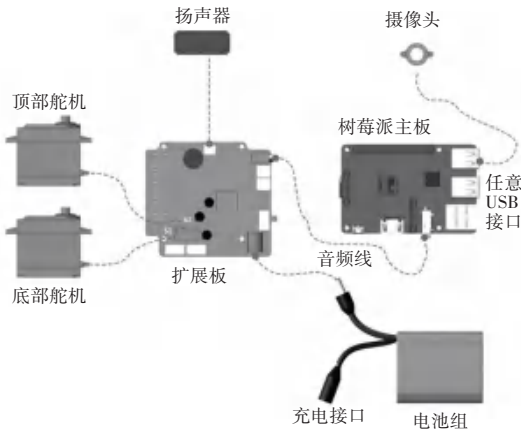


图 4 硬件接线图

Fig. 4 Hardware wiring diagram

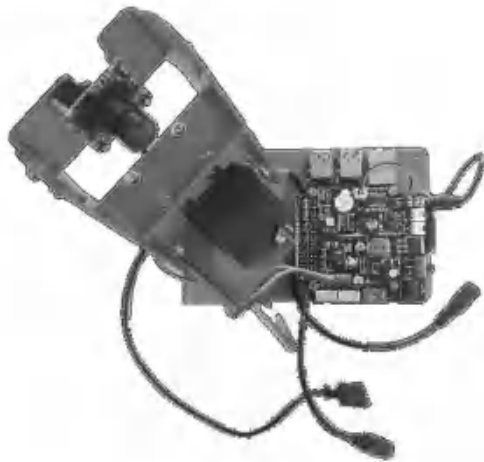


图 5 硬件平台

Fig. 5 Hardware platform

硬件平台以 Raspberry Pi 为核心。图像采集通过 USB 摄像头。在实际监控环境下,侧脸人像的识别要远远难于正脸识别。而且可调角度的监控摄像头可以获得更多可视区域。所以硬件平台采用了两个舵机,为摄像头提供了两个自由度。Raspberry Pi 可通过 Adafruit-PCA9685 驱动操作 GPIO 接口,控制 BST-AI 拓展板以实现控制舵机。扬声器可提供报警功能,WIFI 模块完成视频流、控制指令、识别图像的传输。

系统包含了人脸检测、人脸识别、人脸追踪、微信报警、语音报警等功能。其主要的人脸识别功能的流程如图 6 所示。

Raspberry Pi 移植了 Paspbian 操作系统,以及为控制舵机及采集视频图像所需的 PCA9685 驱动、V4L2 驱动等。

因为 Raspberry Pi 有限的算力和存储空间,所以本系统只有将人脸检测部分的程序移植到

Raspberry Pi 上,以及其所依赖的 Dlib 等模块和运行环境。

Raspberry Pi 启动时会运行/etc/init.d/目录下的服务脚本文件从而运行监控和控制舵机的服务端等两个进程。当摄像头采集到图像后会采用 mjpeg-stream 程序实现实时的网络监控。

本系统中的 HOG 算法、ERT 算法和 FaceNet 的训练过程放在上位机完成。随后将训练好的 HOG+SVM 分类器模型移植到 Raspberry Pi 硬件平台上。

下位机负责完成视频采集和通过 HOG 算法检测人脸,如视频帧中含有人脸信息则将人脸部分通过 WiFi 模块上传给服务器,服务器会根据人脸部分通过 ERT 算法和 FaceNet 网络提取人脸的 128 个特征点信息与 MySQL 数据库中的已存在的人脸信息进行匹配,当两个特征值距离小于设定阈值就可以认为是同一个人,如果出现未知人脸信息,上位机就通过 itchat 模块发送微信报警。

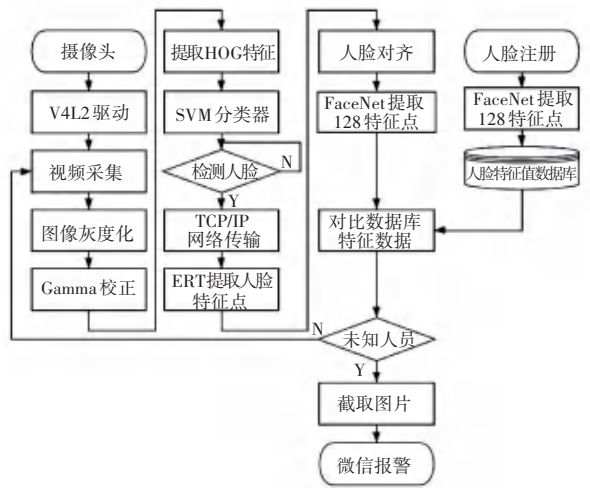


图 6 系统总体流程

Fig. 6 Overall flow chart of system

### 4 系统实验结果

上位机硬件环境为: CPU i7-4720HQ, 内存 16GB 1600MHz 双通道,GPU NVIDIA GTX 970M,主程序采用 Python3 编写,图像界面采用 PyQt 编写。

系统测试采用 LFW 人脸库和测试人员自建的人脸库混合,建立了一个 100 张人脸照片的小型人脸库。传统的嵌入式平台将人脸特征信息存储在本地,存储的记录数量受限,且查询速度较慢,无法实现信息共享,本系统将每个人脸的 128 个特征点信息存储在数据库服务器中,能够快速的匹配人脸信息,在测试中准确率达到 97%。

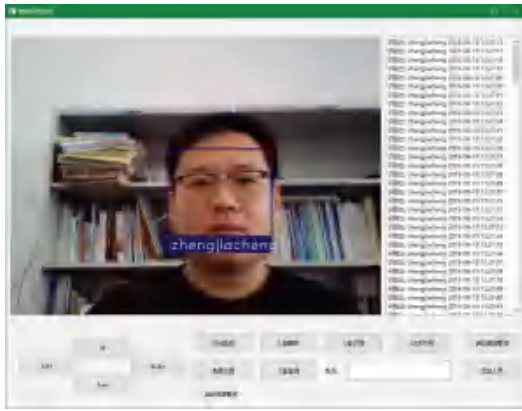


图7 实时监控画面

Fig. 7 Real-time monitoring screen

## 5 结束语

本文主要设计并实现了一种基于 Raspberry Pi 平台下人脸识别智能监控系统。阐述了系统的总体设计方案。本设计首先在视频监控中检测人脸,并将视频帧中含有的人脸信息通过 WiFi 模块上传给服务器,服务器会根据人脸部分通过 ERT 算法和 FaceNet 网络提取人脸的 128 个特征点信息与数据库中的已存在的人脸信息进行匹配,解决了嵌入式平台算法不足和存储空间有限而无法实现实时的、

大量数据下的人脸识别的问题。实验验证了系统的可行性,为智能视频监控系统的的设计提供了一种解决方案,随着监控行业的不断发展,本文设计的人脸识别智能监控系统将有广泛的应用前景。

## 参考文献

(上接第 277 页)

未来 LMS 的首要目标是将重点从支撑管理任务转移到深化学习行为。第二目标是开发更灵活的空间,支持个性化,并参与到过程性学习评价中,而传统的 LMS 功能仍然是生态系统的一部分。例如,目前如火如荼的大规模开放在线课程 MOOC。未来的 LMS 应当以类似“乐高”的方式进行搭建,让教师和学生能够灵活地创建或适应独特需求的学习环境,无缝的定制虚拟环境,整合资源,利用大量开放的教育资源和学习分析来提供个性化学习体验。

## 5 结束语

随着科学技术的不断发展,人工智能也将为高等教育的进一步发展提供新机遇,从而在各个层面实现教育的信息化、全面化、个性化。然而,在拥有巨大前景的同时也面临潜在的危险。如何适应人工智能环境下教师角色的转变,教育手段,学习方式的变化,个性化学习和深度学习是现代学习型社会的必然需要,也需要尝试更多的新方法,新手段去适应。

## 参考文献

[1] National Science and Technology Council. The national artificial

- [1] 蒋馨. 浅析国外智能视频监控技术的发展及应用[J]. 中国安防, 2011(10):105-108.
- [2] 江烂达, 储曙, 缪君. ARM 平台下人脸识别智能监控系统[J]. 计算机工程与设计, 2018.
- [3] 杨巨成, 刘娜, 房珊珊, 等. 基于深度学习的人脸识别方法研究综述[J]. 天津科技大学学报, 2016, 31(6):1-10.
- [4] DALAL N, TRIGGS B. Histograms of Oriented Gradients for Human Detection[C]// 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). IEEE, 2005.
- [5] 向征, 谭恒良, 马争鸣, et al. HOG 在人脸识别中的性能研究[J]. 计算机工程, 2012, 38(15):194-196.
- [6] KAZEMI V, SULLIVAN J. One Millisecond Face Alignment with an Ensemble of Regression Trees[C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2014.
- [7] SCHROFF F, KALENICHENKO D, PHILBIN J. [IEEE 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) - Boston, MA, USA (2015.6.7-2015.6.12)] 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) - FaceNet: A unified embedding for face recognition and clustering[J]. 2015:815-823.

- [1] intelligence research and development strategic plan[EB/OL]. [2016-11-24]. [http://www.360doc.com/content/16/10/15/20/37334461\\_598685262.shtml](http://www.360doc.com/content/16/10/15/20/37334461_598685262.shtml).
- [2] 新一代人工智能发展规划[EB/OL]. [2017-11-19]. [http://www.gov.cn/zhengce/content/2017-07/20/content\\_5211996.htm](http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm).
- [3] 中华人民共和国中央人民政府. 国家中长期教育改革和发展规划纲要(2010-2020年)[EB/OL]. 2010. [http://www.gov.cn/jrzq/2010-07/29/content\\_1667143.htm](http://www.gov.cn/jrzq/2010-07/29/content_1667143.htm), 2014-02-12.
- [4] 余凯, 贾磊, 陈雨强, 徐伟. 深度学习的昨天、今天和明天[J]. 计算机研究与进展, 2013, 50(9):1799-1804.
- [5] 孙志远, 鲁成祥, 史忠植, 马刚. 深度学习研究与进展[J]. 计算机科学, 2016, 43(2):1-8.
- [6] R.基思·索耶. 剑桥学习科学手册[M]. 北京:教育科学出版社, 2010.
- [7] 余明华, 冯翔, 祝智庭. 人工智能视域下机器学习的教育应用与创新探索[J]. 远程教育杂志, 2017(3):11-21.
- [8] S·亚当斯贝克尔, M·卡明斯, A·戴维斯, A·弗里曼, C·霍尔给辛格, V·安娜塔娜娅娅; 殷丙山, 高茜, 任直, 刘鑫驰等(译). 新媒体联盟地平线报告:2017 高等教育版[J]. 开放学习研究, 2017(2):1-20, 62.
- [9] 牟智佳. “人工智能+”时代的个性化学习理论重思与开解[J]. 远程教育杂志, 2017(3):22-30.
- [10] 闫志明, 唐夏夏, 秦旋等. 教育人工智能(EAI)的内涵、关键技术与应用趋势——美国《为人工智能的未来做好准备》和《国家人工智能研发战略规划》报告解析[J]. 远程教育杂志, 2017(1):26-35.