

文章编号: 2095-2163(2020)02-0158-03

中图分类号: TP311.52

文献标志码: A

# 基于 CNN 深度学习模型的大学生课堂行为检测研究

左国才<sup>1</sup>, 苏秀芝<sup>1</sup>, 王海东<sup>2</sup>, 吴小平<sup>1</sup>

(1 湖南软件职业学院, 湖南 湘潭 411100; 2 湖南大学, 长沙 410082)

**摘要:** 为了切实提高职业院校教师课堂教学质量, 对学生在专业课程课堂中表现出的各类上课行为进行检测分析, 从而量化判断学生是否专注于课堂和教师的实际教学效果, 达到提高教师课堂教学质量和督促学生专注课堂学习的目的。本文提出将卷积神经网络 CNN 深度学习模型应用于学生课堂行为检测识别, 实现对是否专注课堂学习的行为进行分类。实验证明, 该卷积神经网络能够对检测目标特征进行深度特征提取, 并且对学生上课中的课堂行为检测取得良好的识别效果。

**关键词:** 卷积神经网络; 深度学习; 课堂行为

## Research on college students' classroom behavior detection based on CNN deep learning model

ZUO Guocai<sup>1</sup>, SU Xiuzhi<sup>1</sup>, WANG Haidong<sup>2</sup>, WU Xiaoping<sup>1</sup>

(1 Hunan Vocational Institute of Software, Xiangtan Hunan 411100, China; 2 Hunan University, Changsha 410082, China)

**[Abstract]** In order to improve the classroom teaching quality of teachers in vocational colleges, the various classroom behaviors of students in professional courses are tested and analyzed, which could quantify whether students focus on the classroom and the actual teaching effect of teachers, therefore achieve the purpose of improving the classroom teaching quality of teachers and supervising students to concentrate on classroom learning. The paper proposes applying the Convolutional Neural Network in-depth learning model to the detection and recognition of students' classroom behavior, so as to classify whether the students are focused on classroom learning. Experiments show that the Convolution Neural Network can extract the depth features of the detected target features, and achieve good recognition results for students' classroom behavior detection in class.

**[Key words]** Convolutional Neural Network; deep learning; classroom behavior

## 0 引言

为了客观量化地评价高职院校课堂教学质量与学生学习情况, 对高职院校学生在上课时的课堂行为进行检测分析, 客观评估课堂教与学的效果。由于卷积神经网络 (Convolutional Neural Network, CNN)<sup>[1]</sup> 在深度特征提取方面有很强的表现能力, 越来越多地应用于各种目标任务检测, 刘智等人<sup>[2]</sup> 设计出构建多尺度深度卷积神经网络行为识别模型; 王军等人<sup>[3]</sup> 提出基于深度学习模型提取深度特征的异常行为检测算法; Xiang 等人<sup>[4]</sup> 提出基于异常检测的视频行为分析; 廖鹏等人<sup>[5]</sup> 设计实现基于深度学习的学生课堂异常行为检测与分析系统。本文总结以上研究成果及研究经验, 提出将深度学习卷积神经网络 CNN 经典模型 VGG16<sup>[6]</sup> 应用于大学生课堂行为检测。基于 VGG16 模型的大学生课堂行为检测将学生在上课中的课堂行为表现分成 2 类, 即: 专注学习与不专注学习。

## 1 VGG16 模型原理

2014 年牛津大学计算机视觉组和 Google DeepMind 公司的研究员一起研发出了新的深度卷积神经网络: VGGNet。该模型参加 2014 年的 ImageNet 图像分类与定位挑战赛, 取得了优异成绩: 在分类任务上排名第二, 在定位任务上排名第一。VGGNet 可以看成是加深版本的 AlexNet, 都是由卷积层、全连接层两大部分构成。VGG 由 5 层卷积层、3 层全连接层、softmax 输出层构成, 层与层之间使用 max-pooling(最大化池) 分开, 所有隐层的激活单元都采用 ReLU 函数。VGG16 的结构图如图 1 所示。VGG16 执行过程如图 2 所示。

高职院校学生上课中的课堂行为检测分析, 由于摄像头拍摄时存在一定的角度差, 因此教室里的学生课堂监控视频图像不能获取检测目标(学生)的正面信息, 大多为侧面信息, 同时由于学生课堂视频数据集的限制及拍摄效果不佳, 增加了学生在教室上课时课堂教学视频中行为检测的难度, 因此本

**基金项目:** 湖南省教育科学规划课题研究成果(XJK19CZY018)。

**作者简介:** 左国才(1978-), 女, 硕士, 副教授、高级工程师, 主要研究方向: 计算机视觉、深度学习。

**收稿日期:** 2019-07-27

文采用卷积神经网络 CNN 经典的 VGG16 深度学习模型对学生课堂行为进行检测分析, 以提高课堂行为检测识别率。

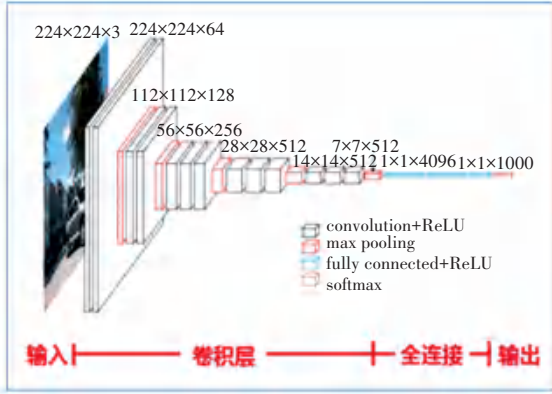


图 1 VGG16 结构图

Fig. 1 Structure of VGG16

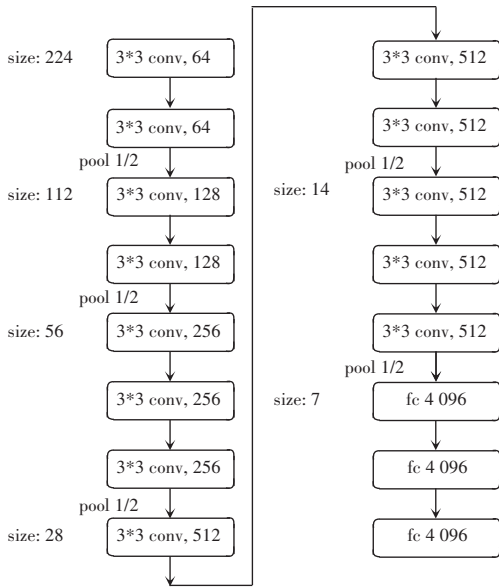


图 2 VGG16 执行过程

Fig. 2 VGG16 execution process

## 2 基于 VGG16 模型的课堂行为检测分析

### 2.1 获取训练数据集

本文所采用的测试数据集为高职院校学生上课时的真实课堂监控视频, 由于拍摄设备性能不高, 拍摄的效果不佳。拍摄设备安装在教室投影仪左右两边以及教室最后两边, 采集到的课堂教学视频数据集均有一定倾斜角度, 由于班级人数较多, 教室较大, 学生上课时坐的位置不同, 有的集中、有的分散, 每个课堂教学视频中拍摄的学生人数约为 10~30 个左右。从整个拍摄的课堂教学视频来看, 学生在上课时的课堂行为变化不多, 测试时, 从拍摄的课堂教学视频中以 4 s 的时间间隔抓取一帧图像, 随机

抽取图像帧选取不同角度、不同教室的实验数据集总共 4 000 张, 进行训练。

### 2.2 课堂行为检测分析

基于 CNN 深度学习模型的大学生上课中课堂行为检测与分析, 将大学生在上课中的课堂行为分为专注学习与不专注学习两类, 将学生课堂视频图像进行相关预处理操作后, 作为实验数据集。

由于卷积神经网络 CNN 模型只能识别包含单个检测目标的固定格式图片, 因此在进行网络模型训练和测试前, 需要将图像进行预处理。

本文中使用的深度学习框架工具 Matcovnet, 采用大型可视化数据库 ImageNet<sup>[7]</sup> 预训练网络模型, 在线微调各种参数, 修改 CNN 深度学习模型 VGG16 网络结构, 并且基于已经进行过预训练的初始化网络参数, 用学生课堂教学视频提取的课堂行为图像数据集来进行训练, 使系统达到快速收敛效果。学生课堂行为检测分析流程如图 3 所示。

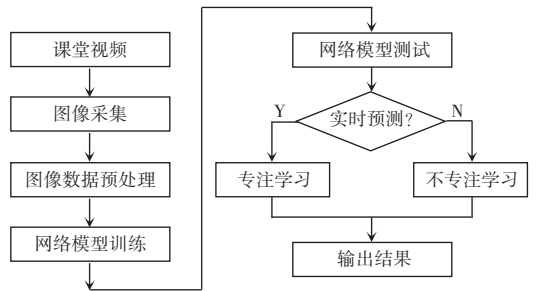


图 3 课堂行为检测分析流程图

Fig. 3 Flow chart of classroom behavior detection and analysis 部分代码如下:

```

model = VGG16(weights = 'imagenet',
include_top = True)
i_path = 'test.jpg'
img = image.load_img(i_path, target_size =
(224, 224))
featur = model.preict(x)
pre = decode_preictions(featur, top = 5)[0]
for element in pre:
    values.append(element[2])
    bar_label.append(element[1])
z = fig.add_subplot(111)
z.bar(range(len(values)), values, tick_label =
bar_label, width = 0.5, fc = 'g')
for a,b in zip(range(len(values)), values):
    z.text(a, b + 0.000 5, percent(b), ha =
'center', va = 'bottom', fontsize = 7)

```

### 3 实验结果分析

本文的图像数据集是从学生上课中的课堂监控视频中获取到的实验数据集。针对拍摄的学生上课中的课堂监控视频,完成数据集图像数据预处理、修改训练参数、训练网络模型,随机选择训练集和测试集,并设置训练集与测试集的比例为8:2。“不专注学习”课堂行为实验结果如图4所示,“专注学习”课堂行为实验结果如图5所示。



图4 “不专注学习”课堂行为

Fig. 4 "Unfocused learning" classroom behavior



图5 “专注学习”课堂行为

Fig. 5 "Concentrated learning" classroom behavior

实验结果表明,训练的网络模型对学生不专注学习行为的检测识别率较低,对抬头专注学习行为检测识别率较高。不专注学习行为的检测识别率低的原因是在初始的学生课堂教学视频图像训练集中,学生不专注学习行为的人数较少,同时所采集的课堂教学视频图像序列中不专注学习行为的学生的图片相应减少,不能从课堂教学视频图像训练集中

提取到足够的非专注学习行为的特征。用学生课堂教学视频图像测试集进行验证,测试结果见表1。

表1 测试结果

Tab. 1 Test results

识别类型	识别结果	平均准确率	%
专注学习	92.22	89.24	
不专注学习	83.67	80.16	

### 4 结束语

在学生上课时课堂行为检测分析中,原始课堂教学图像被用作整个网络输入。与传统的行为检测算法相比,避免了特征提取的过程。这些特征是由网络自动从训练数据集中提取出来的。在每帧课堂教学视频图像中,所设计的模型方法能够用矩形框自动定位到学生课堂行为检测区域,以及自动识别该课堂行为方式。由于课堂教学视频图像数据集数量不够多、学生课堂行为标记的准确度也有待改进,给实验的检测率造成了影响,实验结果表明,卷积神经网络CNN具有强大的深度特征提取能力,能够有效地检测到学生课堂教学视频中的课堂行为,取得了良好的检测识别效果。

### 参考文献

- [1] SIMONYAN K ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556,2014.
- [2] 刘智,黄江涛,冯欣. 构建多尺度深度卷积神经网络行为识别模型[J]. 光学精密工程,2017,25(3):799.
- [3] 王军,夏利民. 基于深度学习特征的异常行为检测[J]. 湖南大学学报(自然科学版),2017,44(10):130.
- [4] XIANG Tao, GONG Shaogang. Video behavior profiling for anomaly detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2008,30(5):893.
- [5] 廖鹏,刘宸铭,苏航,等. 基于深度学习的学生课堂异常行为检测与分析系统[J]. 电子世界,2018(8):97.
- [6] 朱煜,赵江坤,王逸宁,等. 基于深度学习的人体行为识别算法综述[J]. 自动化学报,2016,42(6):848.
- [7] KRIZHEVSKY A I. SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks [C]// International Conference on Neural Information Processing Systems. Lake Tahoe: NIPS Foundation, Inc., 2012:1097.

(上接第157页)

[9] 侯志强,赵梦琦,余旺盛,等. 基于SLIC与分水岭算法的彩色图像分割[J]. 光电工程,2019,46(6):73.

[10] ACHANTA R, HEMAMI S, ESTRADA F, et al. Frequency-

tuned salient region detection[C]//IEEE International Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE,2009:1597.