

# 基于深度学习的学生课堂异常行为检测与分析系统

燕山大学信息科学与工程学院 廖 鹏 刘宸铭 苏 航 李启芳 韩延中

**【摘要】**视频分析技术已越来越多的用于视频监控等领域,为了对学生听课状况进行有效的监督和管理,项目构建了一个检测、统计学生课堂异常行为的辅助教学管理系统。通过基于VGG预训练网络模型迁移学习,来提取学生课堂异常行为特征,实现对玩手机、睡觉等异常行为的检测分析。系统在测试集上的平均识别正确率达到了85.2775%,其中识别睡觉95.1510%,识别正常90.5490%,识别玩手机70.1334%。基于背景差分提取目标区域,使系统可以识别视频中多个目标;通过连续的识别和统计,界定学生异常行为,最后自动生成课堂行为分析报告。结果表明,卷积神经网络能够对目标特征进行特征提取,适应复杂的环境。该方法可实现对学生课堂异常行为的准确监测,具有适用性强、抗干扰的特点。

**【关键词】**辅助教学管理系统;异常行为分析;卷积神经网络;模式分类

DOI:10.19353/j.cnki.dzsj.2018.08.054

## 1 引言

基于深度学习的学生课堂异常行为检测与分析系统将学生课堂行为模式分为听课、睡觉、玩手机三类,将图像素材进行预处理后制作成数据集,系统使用Matconvnet第三方工具,基于ImageNet预训练网络模型进行参数微调,通过修改vgg<sup>[3]</sup>网络结构,迁移到我们要解决的模式识别问题上来,并基于预训练的初始化网络参数,用自己的数据集来训练,从而使系统快速收敛。由于卷积神经网络只能识别包含单个目标的固定格式图片,因此在进行训练和测试前将图片需进行预处理。

## 2 材料与方法

### 2.1 制作数据集

试验视频采集自燕山大学第四教学楼,视频分辨率为1280(水平)×720(垂直)。拍摄对象共10个学生,每个学生单独拍摄5-6分钟视频。使学生分别坐在面对摄像机的4个不同位置上,在每个位置上分别做出听课、睡觉、玩手机的动作,每个动作持续60-120秒。在每个学生每个拍摄角度的每个行为模式下随机抽取25帧图像,得到每个学生300张图片,合计3000张图片,组成训练集。同理得到360验证集,360张测试集。

为了避免出现过拟合、泛化误差大的情况,通过两种方式进行数据增强:(1)实验采集了10名学生的各种各样的睡觉、玩手机、正常听课姿势,增加实验数据包含的信息量;(2)水平翻转来生成图像,使数据库增大了2倍。实验中为了减少训练样本之间的关联性,我们让每个学生三种行为模式的图片随机打乱排放,并生成标签,组成训练数据集。

### 2.2 卷积神经网络

#### 2.2.1 VGG网络

VGG由牛津大学视觉几何组提出,通过加深网络深度极大提升了系统准确率。实验中使用的VGG-F网络,网络结构如图1所示:

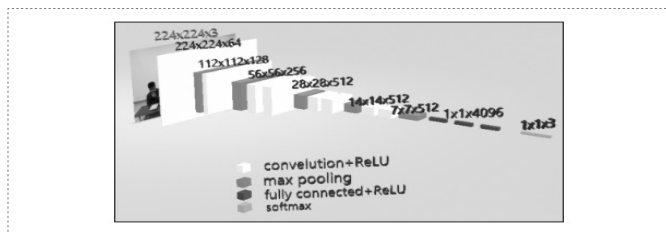


图1 VGG-F网络结构

VGGNet拥有5段卷积,其经常出现多个完全一样的3\*3的卷积层堆叠在一起的情况,通过卷积层串联拓宽感受野<sup>[3]</sup>。3个串联的3\*3的卷积层,拥有比1个7\*7的卷积层更少的参数量,拥有比1个7\*7的卷积层更多的非线性变换,使CNN对特征的学习能力更强。在预测时,VGG采用卷积层使用滑窗的方式进行分类预测<sup>[6]</sup>,将不同窗口的分类结果平均,再将不同尺寸Q的结果平均得到最后结果,来提高图片数据的利用率并提升预测准确率。

VGG的拓展性很强,容易迁移到其他图片数据上<sup>[4]</sup>。由于使用小卷积核和更深的网络进行的正则化,并且使用了预训练得到的数据进行参数的初始化,VGG网络相较于其他网络模型而言具有更高的准确性和更快的速度。

#### 2.2.2 实验结果

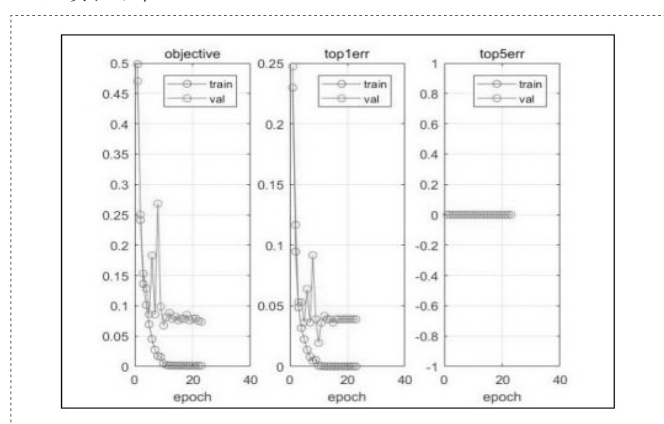


图2 训练结果

将训练集图像输入到网络,并用验证集验证,通过曲线观察系统的收敛情况。从图2中可以看出训练集很好的收敛,测试集开始时收敛很快,接着有一个震荡的过程,最终也收敛。当训练到第23代时,目标错误率收敛至0.0001;在验证集上,网络错误率收敛至0.07,分类到第一类错误的概率为0.04;用测试样本集进行验证,结果如表1所示。由表1可知,网络能够以较高的可靠性(准确性85.2778%)识别出未知图像中目标模式。

表1 测试集准确率

识别类型	睡觉	正常	玩手机
准确率	95.1510%	90.5490%	70.1334%
平均准确率	85.2778%		

测试集中较高的准确率说明我们方案是有效的。图3所示为两帧识别结果:



图3 识别结果

基金项目:2017年燕山大学“大学生创新创业训练计划”国家级立项(项目编号:201710216019)。

### 3 行为识别过程

#### 3.1 图像预处理

通过图像背景差分、滤波、计算连通域定位目标区域,将多个目标分割为单个目标,再送入网络识别。

将分割出的连通域分别截取其原图图像,还原到对应空的教室背景图中的目标位置,使其成为单个目标在教室中的完整图像,并转化为适合输入CNN分类的图片。

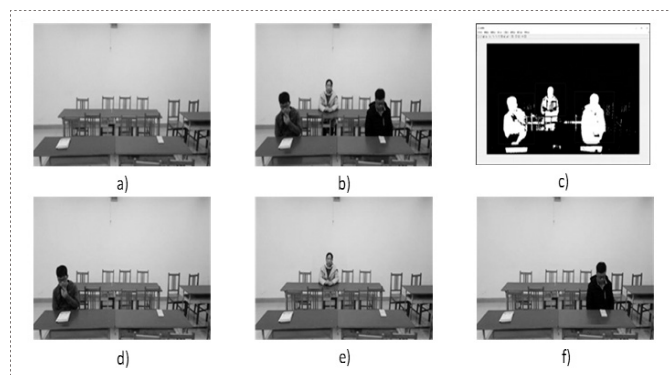


图4 图像预处理示意图

#### 3.2 行为判定

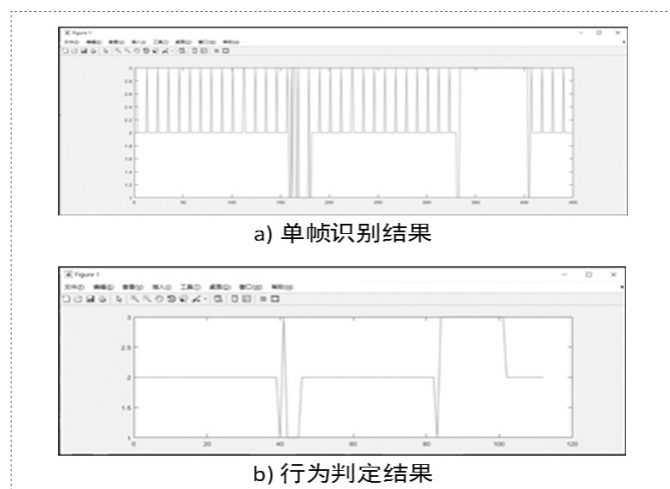


图5 行为判定结果图

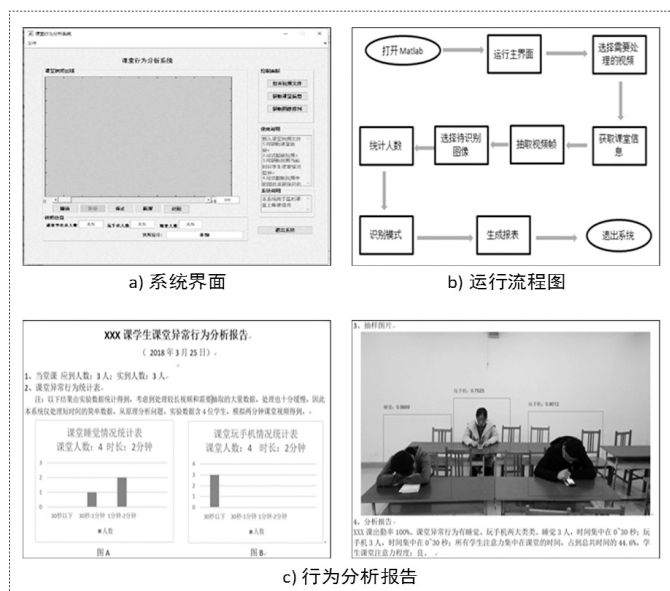


图6 系统界面及分析报告

由于网络存在识别错误,在曲线中形成了若干的尖峰(图5a)。单帧图片的识别结果具有强烈的波动,无法准确判断和描述学生在当前时刻的行为状态。我们通过将4个结果中识别最多的一类行为模式,判定为1秒内目标的行为(例:若1秒钟的4帧图像有1帧被判定为睡觉,1帧被判定为玩手机,2帧被判定为听课,则将此1秒内的行为模式全部判定为听课状态),此判定策略有效的解决了识别错误情况过多的问题(图5b)。

### 4 识别系统

课堂行为分析系统包含读入课堂视频及相关信息、统计出勤率、连续识别模式、自动生成行为分析报告等主要功能,如图6a)所示。

识别结束后系统会自动生成word格式的课堂行为报告,包含出勤率、异常行为的时间分布、典型异常截图、总结等内容,供老师和辅导员等人员查看,如图6c)所示。

### 5 结论

本文详细介绍了基于卷积神经网络对学生课堂异常行为的检测方法,并对传统的K最近邻法和卷积神经网络方法在课堂场景下进行行为模式识别的特点进行了分析比较和总结。设计了适用于课堂场景下的行为识别系统。得出了以下结论:卷积神经网络在特征提取方面具有优越的性能,抗干扰能力强,在实际生活的应用中具有十分广阔的前景和十分突出的效果。实践证明,采用预训练模型进行fine-tune(参数微调),或者运用迁移学习的方法,可以有效提高训练效率,减少过拟合,是工程实践的有效方法。在研究的过程中,也发现了亟待解决的问题。如何缩小目标区域,再送入神经网络训练,使训练结果的准确率更高。如何提高系统对细节特征,即玩手机模式的特征提取,以提高识别玩手机模式的准确率。将在后续研究中探讨这些问题。

深度学习网络与视频监控的结合使教学管理变得简单而高效。希望本研究可以为视频分析领域提供帮助。

#### 参考文献

- [1]管皓,薛向阳,安志勇.深度学习在视频目标跟踪中的应用进展与展望[J].自动化学报,2016,42(06):834-847.
- [2]朱煜,赵江坤,王逸宁,郑兵兵.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(06):848-857.
- [3]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(03):799-805.
- [4]赵凯旋,何东健.基于卷积神经网络的奶牛个体身份识别方法[J].农业工程学报,2015,31(05):181-187.
- [5]T Xiang, S Gong.Video Behavior Profiling for Anomaly Detection[J].IEEE Transactions on Pattern Analysis and Machine Intelligence,2008,30(5):893-908.
- [6]K.Simonyan and A.Zisserman, "Very deep convolutional networks for large-scale image recognition," in International Conference on Learning Representations(ICLR),2015.
- [7]A.Krizhevsky,I.Sutskever, and G.Hinton, "Imagenet classification with deep convolutional neural networks," in Neural Information Processing Systems(NIPS),2012.

#### 作者简介:

廖鹏(1996—),男,四川成都人,现就读于燕山大学信息科学与工程学院,主要研究方向为目标跟踪。

刘宸铭(1997—),男,河北唐山人,现就读于燕山大学信息科学与工程学院,主要研究方向为图像处理。

苏航(1997—),男,河北承德人,现就读于燕山大学信息科学与工程学院,主要研究方向为信号处理。

李启芳(1997—),男,海南儋州人,现就读于燕山大学信息科学与工程学院,主要研究方向为信号处理。

韩延巾(1997—),女,河北邯郸人,现就读于燕山大学信息科学与工程学院,主要研究方向为语音信号处理。

指导老师:胡正平(1970—),男,四川仪陇人,教授,博士生导师,目前为中国电子学会高级会员,中国图像图形学会高级会员,目前研究方向为视频图像。