

文章编号: 2095-2163(2020)06-0180-03

中图分类号: TP311.5

文献标志码: A

# 基于深度学习的教师注意力识别与应用

刘荣华<sup>1</sup>, 范 诚<sup>2</sup>

(1 云南大学 职业与继续教育学院, 昆明 650091; 2 昆明理工大学 农业与食品学院, 昆明 650504)

**摘要:** 教师注意力的分配影响着课堂教学效果, 为了对课堂教学中教师注意的方向进行检测和分析, 本文基于深度学习YOLOv3算法, 使用课堂实录视频, 进行了教师注意力识别实验。结果表明, 该识别方法具有较高的准确率和可行性, 能精准的反馈课堂中教师注意力方向, 分析教师每节课中的注意力变化趋势。

**关键词:** 教师注意力分析; 深度学习; 图像识别

## Teacher Attention Recognition and Application based on Deep Learning

LIU Ronghua<sup>1</sup>, FAN Cheng<sup>2</sup>

(1 College of vocational and continuing education, Yunnan University, Kunming, Yunnan 650091, China;

2 College of agriculture and food, Kunming University of Science and Technology, Kunming, Yunnan 650504, China)

**【Abstract】** The allocation of teacher attention affects the effect of classroom teaching. In order to detect and analyze the direction of teachers' attention in classroom teaching, this paper, based on the deep learning YOLOv3 algorithm, uses classroom video recording to conduct an experiment of teachers' attention recognition. The results show that this recognition method has high accuracy and feasibility, can accurately feedback the attention direction of teachers in class, and analyze the change trend of teachers' attention in each class.

**【Key words】** analysis of teacher's attention; deep learning; image recognition

### 0 引言

人类的注意力体现着人们重要的动机和认知取向<sup>[1]</sup>, 师生互动是教师教育教学工作实现的重要载体<sup>[2]</sup>。师生互动中, 教师注意力的分配与转移影响着课堂教学效果, 而教师具有较强的注意分配能力是促使学生集中注意的一个必要条件<sup>[3]</sup>。因此, 对课堂教学环境下, 教师注意力的识别与分析具有重要意义。本研究基于课堂实录视频, 采用人工智能中的图像识别技术, 提出在教室自然场景下对教师注意力的识别方法, 并进行量化分析与评价, 进而提高教师对自身注意力的意识, 为培养教师较强的注意分配能力提供理论依据。

### 1 相关研究及进展

在教学过程中, 师生的课堂行为、学生的学习记录等环节, 都会产生大量的实时数据。利用深度学习和人工智能技术, 可以将这些数据充分利用起来, 优化教育的发展。陈靛影等人通过对头部姿态、面部表情、课堂互动多模态信息识别与融合, 提出了学生学习兴趣的智能化分析方法<sup>[4]</sup>; 徐振国等人提出了基于深度学习的学习者情感识别方法, 能够快速、准确的识别学习者情感<sup>[5]</sup>; 郭晓旭等人采

集了真实课堂环境下的学生数据, 提出了基于微表情学生课堂专注度识别<sup>[6]</sup>。以上研究可以看出: 关注学生课堂行为数据的研究中, 大多集中在学生学习兴趣、课堂专注度和学习者情感等方面, 而关注教师行为数据的研究较少; 其次, 深度学习在教育领域的研究已经取得一定的进展, 基于深度学习的教师注意力的识别具有可行性。然而, 在真实课堂环境下, 教育场景的复杂性带来了背景信息的冗余, 教师注意力的变化涉及头部、姿势、目光等多维度信息。为解决以上问题, 本研究采集了多种真实课堂环境下的教师图像数据集, 进一步提高了识别模型的泛化能力; 由于YOLOv3算法对小目标物体检测具有较高的精度和速度<sup>[7]</sup>, 本研究采用该算法来满足实时检测需求; 头部姿态在注意力方向中的贡献占40.3%<sup>[8]</sup>, 人的注意方向可以根据头部姿势预测<sup>[9]</sup>。本文在头部姿势的基础上增加了手势、动作等多维度识别特征, 提高了教师注意力识别的可靠性。

### 2 注意力识别方法与模型

#### 2.1 深度学习

深度学习是人工智能领域一个研究方向, 它可以高效的提取复杂数据的特征, 学习不同类型的知

作者简介: 刘荣华(1996-), 女, 硕士研究生, 研究方向为教育信息化; 范 诚(1994-), 男, 硕士研究生, 研究方向为计算机视觉、深度学习。

通讯作者: 范 诚 Email: 782476237@qq.com

收稿日期: 2020-03-04

识<sup>[10]</sup>,被广泛应用于图像识别、语音识别、自然语言处理等方面。卷积神经网络(Convolutional Neural Network, CNN)是深度学习技术中具有代表性的一种

网络结构,主要用于图像识别。本研究提出使用基于 TensorFlow 开源框架的 YOLOv3 神经网络,来实现教师注意力的识别。YOLOv3 网络结构如图 1 所示。

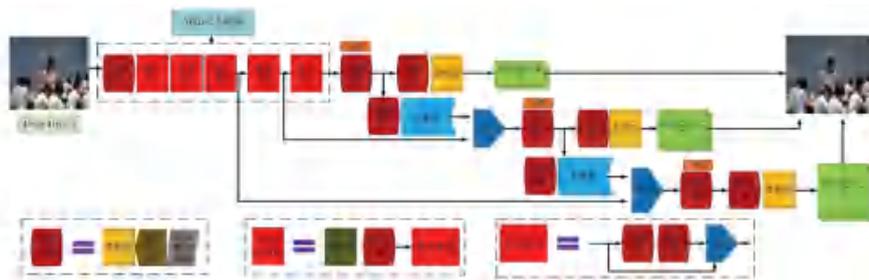


图 1 YOLOv3 网络结构

Fig. 1 YOLOv3 network structure

YOLOv3 主干网络采用 53 层卷积层,包含了多组残差模块。在训练过程中接收  $416 \times 413 \times 3$  大小的图像作为输入。具体流程如下:



图 2 教师注意力检测流程图

Fig. 2 Flow chart of teacher's attention detection

(1)通过卷积块提取特征,经过残差模块进一步提取深层特征,分别生成  $13 \times 13, 26 \times 26, 52 \times 52$  三种特征层。其中,上层特征层的网络会进行上采样后与下层网络进行融合,能有效保留物体的浅层信息,有利于提升物体的检测精度。

### 3 数据采集与处理

#### 3.1 特征定义

在日常教学中,经常会发现有些教师的注意力只集中于授课内容,而忽视了讲台下的学生学习活动,因此很容易造成学生缺乏学习兴趣的现象。因此,本研究提出 4 种对教学效果影响较大的典型课堂注意对象类别:首先将教师的注意力对象分为教学内容和学生表现两个方面;其次将教学内容细分为板书和课件两种;将学生表现分为学生听课状态和学生练习情况。在视频观察的基础上,总结出每类教师注意对象的行为特征,包括头部姿势、手势、动作等特征,并且对每类注意行为进行编码,如表 1 所示。

(2)通过对特征层的融合进行图像的识别定位。  
(3)通过 softmax 激活函数对物体进行分类。

由于本研究教师注意行为有 4 种,故网络模型的输出有 4 种类别。

#### 2.2 教师注意力识别

通过采集真实课堂中教师 4 类注意行为的图像作为数据集,进行 YOLOv3 网络模型的训练,生成权重文件;对识别结果进行统计,分析教师每节课中的注意力变化趋势,进而传输到相应的教师移动端,供教师参考和反思。教师注意力检测分析流程如图 2 所示。

表 1 教师注意力分布编码表

Tab. 1 Teacher attention distribution code table

教师注意对象	编码	教师行为表述	教师行为特征
教学内容	1	板书,通过在黑板或者在白板上书写展示教学内容	面对黑板或者白板,站在讲台上,右手抬起
	2	一边看黑板或者多媒体课件,一边讲授	侧对学生,站在讲台上,有时用手指向课件
学生表现	3	只通过语言讲授或者提问	面对学生,讲授或者提问,可以时刻观察学生听课状
	4	巡视或者个别指导	在学生中巡视,弯腰或者低头

#### 3.2 数据采集

由于课堂教学环境差异较大,为获取真实环境下的数据,本研究在国家教育资源服务平台中选取了 40 节中小学课堂教学实录视频。其中包括不同年龄、不同性别、不同科目的教师课堂实录视频,每间隔 3s 左右采集课堂教学实录视频中的一帧图像作为数据样本,将其中模糊或者不包含教师的图像

删除,获得包含 4 种编码行为的教师注意力图像。

#### 3.3 数据预处理

通过对原始图像进行裁剪、旋转及镜像等方式来扩大数据集,最终获得的教师注意力数据集共 3 600 张图片,每种编码行为约 900 张图片,将该数据集按比例分为训练集与测试集。按照教师注意力分布编码表,使用 Labellmg 标注工具对图像中的关

键特征用矩形框进行标注,标注后的图像数据保存为 PASCAL VOC 格式。

## 4 实验

### 4.1 实验配置与步骤

(1)实验环境。本实验的主要设备配置包括: GTX1060(GPU)、6G 显存、I7-8750(CPU);软件环境采用 TensorFlow 开源深度学习框架、OpenCV、Cuda、Cudnn、Python 等完成。

(2)实验步骤。①数据采集。采集课堂实录中教师 4 中注意行为的图像。②数据预处理。对采集的图像进行预处理,主要包括数据标注和数据增强。③模型的训练与调整。搭建 YOLOv3 网络结构及设置相关参数后,在 TensorFlow 框架中迭代生成权值模型。④模型测试与评价。在实验过程中可通过记录每幅图像的检测时间来衡量模型的实时性。从图像测试集中随机抽取图片作为模型评价数据集,重复进行 3 次检测后对各变量取值计算平均值,计算每种注意行为的识别正确率与速率来评估算法性能。

### 4.2 实验结果与分析

图像检测效果如图 3 所示。图 3(a)用蓝色方框标记了教师板书的行为,此时注意对象为多媒体课件,该行为编码为 1;图 3(b)用紫色方框标记了教师侧对的行为,此时注意对象为讲授内容;图 3(c)用红色方框标记了教师面向学生讲授的行为,此时注意对象为学生的听课状态,编码为 3;图 3(d)用绿色方框标记了教师巡视和个别指导的行为,此时注意对象为学生练习情况,编码为 4。



(a) 教师板书行为

(b) 教师侧对行为

(a) Writing on the blackboard (b) Look at the courseware and talk



(c) 教师面向学生讲授行为

(d) 教师巡视/个别指导的行为

(c) Lecturing to students (d) Make an inspection

图 3 教师注意力检测结果示例

Fig. 3 Example of teacher attention detection results

为了达到最好的识别效果,比较多个权重模型,找到准确率、召回率、调和平均值以及精度均值最好的模型,选出最好的训练模型对真实的课堂教师注意力检测中进行识别与分析。本研究统计了三次随机实验的平均精度,采用  $F$  值作为综合评价指标,其算式如式(1)所示。式(1)中,  $RS$  表示正确识别图像中的目标,  $FS$  表示非正确识别的目标,  $FN$  表示非识别的目标。  $P$  表示正确率,  $R$  表示召回率,  $F$  为调和平均值。不同类别的检测正确率和速率如表 2 所示。

$$P = \frac{RS}{RS + FS}, R = \frac{RS}{RS + FN}, F = \frac{2PR}{P + R} \quad (1)$$

表 2 不同类别的检测效果比较

Tab. 2 Comparison of detection effects of different categories

注意对象类别	P	R	F	平均检测速率(ms · f <sup>-1</sup> )
板书内容	94.45	88.31	91.27	38.92
授课内容	91.46	84.73	87.98	38.46
学生听课状态	90.14	82.97	86.41	36.64
学生练习情况	93.22	86.86	89.93	36.26

## 5 结束语

综上所述,将基于深度学习的图像识别技术应用到课堂教学教师的注意力识别中,实现对教师注意力分配的智能分析与实时检测,有利于辅助教师教学的评价与培养教师注意分配能力,提高教师专业素养。本研究训练出的教师注意力识别模型,具有较高的准确率和识别速度。

## 参考文献

- [1] POSNER M I. Attention in cognitiveneuroscience: an overview[J]. The cognitive neurosciences, 1995, 615-624.
- [2] 李卉,王蔚. 基于 Kinect 体感交互的儿童注意力测评系统[J]. 现代教育技术, 2014, 24(7): 120-126.
- [3] 陈光全. 浅谈小学教师在课堂上的注意分配能力[J]. 小学教学研究, 1982(3): 3-4.
- [4] 陈颀影,罗珍珍,徐如意. 课堂教学环境下学生学习兴趣智能化分析[J]. 电化教育研究, 2018, 39(8): 91-96.
- [5] 徐振国,张冠文,孟祥增,等. 基于深度学习的学习者情感识别与应用[J]. 电化教育研究, 2019, 40(2): 87-94.
- [6] 郭晓旭. 基于微表情识别的学生课堂专注度分析系统研究[D]. 云南师范大学, 2019.
- [7] 刘学平,李珂乾,刘励,等. 自适应边缘优化的改进 YOLOv3 目标识别算法[J]. 微电子学与计算机, 2019, 36(7): 59-64.
- [8] PROUDLOCK F A, SHEKHAR H, GOTTLÖB I. Coordination of eye and head movements during reading [J]. Investigative ophthalmology & visual science, 2003, 44(7): 91-92.
- [9] FANELLI G, GALL J, VAN GOOL L. Real time head pose estimation with random regression forests[C]//Computer Vision and Pattern Recognition. Washington DC: IEEE, 2011: 617-624.
- [10] 刘小刚,范诚,李加念,等. 基于卷积神经网络的草莓识别方法[J]. 农业机械学报, 2020, 51(2): 237-244.