

# 引入注意力机制的自然场景文本检测算法研究

牛作东 李捍东\*

(贵州大学电气工程学院 贵州 贵阳 550025)

**摘要** 随着深度学习、神经网络的兴起与发展,对于图像中的目标检测已经取得了巨大的进展。但是自然场景下的文本信息具有多样的形式和复杂的特点,通用的目标检测算法无法取得理想的效果,因此自然场景下的文本检测在计算机视觉以及机器学习领域仍然是一项具有挑战性的问题和未来的热点研究方向。根据当前学术界针对自然场景下的文本检测问题所提出的算法和思路,在 EAST 算法的主干网络 PVANet 的基础上通过引入注意力机制模块,使得提取文本目标特征时更加关注有用信息和抑制无用信息,从而有效改善原算法在预测长文本方向信息时视野不足的问题。实验结果显示,该方法在没有损失检测效率的同时提高了原算法的检测精度,并在一定程度上优于当前针对自然场景下的文本检测算法。

**关键词** 注意力机制 文本检测 深度学习 特征提取

中图分类号 TP301.6

文献标识码 A

DOI:10.3969/j.issn.1000-386x.2019.09.035

## NATURAL SCENE TEXT DETECTION ALGORITHM WITH ATTENTION MECHANISM

Niu Zuodong Li Handong\*

(College of Electrical Engineering, Guizhou University, Guiyang 550025, Guizhou, China)

**Abstract** With the rise and development of deep learning and neural network, object detection in images has made great progress. However, text information in natural scenes has various forms and complex characteristics, so universal object detection algorithms cannot achieve ideal results. Therefore, text detection in natural scenes is still a challenging problem and a hot research direction in the field of computer vision and machine learning. This paper studied the current academic algorithms and ideas for text detection in natural scenes. Based on the network-PVANet of EAST algorithm, we introduced the attention mechanism module to make the extraction of text target features pay more attention to useful information and suppress useless information, so as to effectively improve the problem of insufficient field of vision of EAST in predicting long text direction information. The experimental results show that compared with EAST, this method improves the detection accuracy without losing the detection efficiency, and to some extent it is superior to the current text detection algorithm for natural scenes.

**Keywords** Attention mechanism Text detection Deep learning Feature extraction

## 0 引言

人工智能的兴起促进着当前社会生产活动中各项产业的革新与发展,以机器学习、深度学习和计算机视觉为代表的人工智能技术已经得到了很大程度的推广和应用。比如:无人驾驶中的高级驾驶辅助系统通过计算机视觉来实现对于动态物体的识别、侦测与追踪;

以深度学习加强神经网络训练的阿尔法围棋已经被围棋界公认为超过了人类职业围棋顶尖水平。此外,深度学习在自然语言处理、语音识别、人脸识别、物体检测等诸多方面都有了广泛的应用,已经成为未来科技进步和发展的一个热点方向。

自然场景下的物体检测问题吸引世界各地的专家学者的关注和研究,传统的目标检测算法主要是基于手工特征的选取和滑动窗口的方式来检测目标物体,

其中最具代表性的算法主要包括 VJ、HOG 和 DPM 等。由于手工设计目标特征,因此带来了很多的缺点,比如在特定条件下存在着特征值不具有代表性和鲁棒性等问题,另外通过滑动窗口的方式来提取目标框进而进行判断的过程非复杂,算法量大,制约着系统的运行效率。DPM 算法提出之后,取得前所未有的目标检测效果,同时也代表着传统的目标检测方法已经遇到了无法突破的瓶颈。之后 Girshick 等<sup>[1]</sup>提出了 R-CNN 模型,开启了神经网络应用于目标检测的先河。随着 2012 年 Krizhevsky 等<sup>[2]</sup>使用扩展了深度的 CNN 在 ImageNet 大规模视觉识别挑战竞赛 (ImageNet Large Scale Visual Recognition Challenge, LSVRC) 中取得了当时最佳成绩的分类效果,使得 CNN 越来越受研究者的重视。随后,基于 CNN 模型的改进的算法 Fast R-CNN<sup>[3]</sup>、Faster R-CNN<sup>[4]</sup>、SSD<sup>[5]</sup>、YOLO<sup>[6]</sup>、Retina-Net 以及 Pyramid Networks 等又进一步推动了基于深度学习的神经网络模型在自然场景下目标检测领域的快速发展。相对于传统的目标检测方法,深度学习目标检测方法通过深度神经网络来学习到的目标特征更鲁棒,采用 Proposal 或者直接回归的方式来获得候选目标的过程更加高效。在完成目标分类的过程中同样使用深度网络实现了端到端的方式直接训练模型,从而使得深度学习目标检测方法准确度高实时性好。通用目标检测算法的基本流程如图 1 所示。

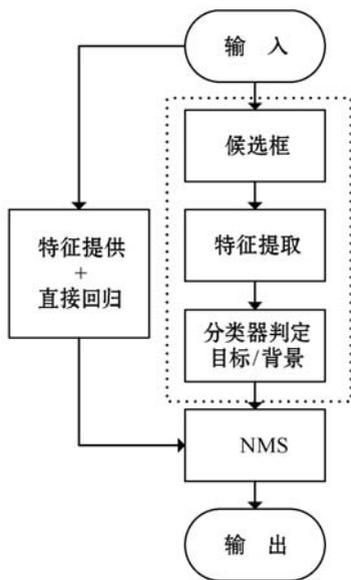


图 1 目标检测算法基本流程图

基于深度学习的目标检测算法在用于通用物体检测方面取得了很好的效果,但是面对自然场景下的文本信息的检测任务时却存在着许多问题,主要原因有:(1) 文本长宽比不定并且相关较小,与常规物体检测存在着差异,这个差异使得在选取候选框的长宽比时要做相应的处理。(2) 自然场景下的文本普遍存在着

水平、倾斜、弯曲等各种不定的形状和方向,这些差异性较丰富的文体状态导致了采用通用的目标检测算法很难沿着某个方向或者倾斜角度去完成文本检测的任务。(3) 场景下的文本很容易受到其中某些物体的局部图像影像影响其文本特征信息。(4) 自然场景下的文本内容存在许多艺术字、手写体等。(5) 自然场景下的文本容易受到自然环境的干扰,比如背景、光照影响等。针对上述问题,在设计文本检测算法的时候需要对常规物体检测方法进行改进的方向主要包括特征提取算法、区域建议网络(RPN)对候选区域进行推荐、多任务网络训练、损失函数 Loss 改进、非极大值抑制(NMS)、半监督学习等。因此,自然图像中的文本检测仍然是一项非常具有挑战性的研究工作,加上受到许多现实应用实例的驱动影响,比如地理位置和图像检索<sup>[7]</sup>、照片 OCR<sup>[8]</sup>等,此方面的研究工作也具有现实意义。

## 1 自然场景下的文本检测算法

### 1.1 研究现状

自然场景下的文本检测方法通常由两个主要部分组成,分别是文本检测和文本识别。文本检测主要以字边界框的形式定位图像中的文本。文本识别对裁剪后的单词或者文字图像进行编码,生成机器可解释的字符序列。本文的研究重点主要放在了文本的检测上。本文研究的目标检测算法中,可以将文本检测器分为基于原始检测目标的分类策略和基于目标边界框形状的分类策略,每个检测策略又有各自的特点和不同类别的算法。

基于原始检测目标的分类策略主要有:(1) 基于角色的检测算法。其过程为:首先检测单个字符或文本的一部分;然后将其分组成一个单词,以 Neumann<sup>[9]</sup>为代表的研究提出通过对极值区域进行分类来定位字符;最后通过穷举搜索方法对检测到的字符进行分组。(2) 基于字的检测方法。它是与一般物体检测类似的方式直接提取文本,Gomez<sup>[10]</sup>提出了一个基于 R-CNN 的框架,其中候选词首先由类别不可知的提议生成器生成,随后是随机森林分类器,再采用用于边界框回归的卷积神经网络来细化边界框。(3) 基于文本行的检测算法。该方法首先检测文本行,然后再分逐个分成单词,Long<sup>[11]</sup>提出利用文本的对称特征来检测文本行,然后通过使用全卷积神经网络来定位文本行。

基于目标边界框形状的分类策略的检测方法可以分为两类。第一类为水平或接近水平的检测方法,这

类方法专注于检测图像中的水平或接近水平的文本。例如:Yi<sup>[12]</sup>提出一个三阶段框架,它由边界聚类、笔划分割和字符串片段分类组成;Tian 等<sup>[13]</sup>提出的方法检测几乎水平的文本部分,然后将它们链接在一起开成单候选区域;Cao 等<sup>[14]</sup>尝试使用去模糊技术来获得更强大的检测结果。第二类为多方向的检测方法。与水平或接近水平检测方法相比,多方向的文本检测更加稳健,因为自然场景下的文本可以在图像中处于任意方向。这类的主要研究方法有利用检测多方向文本的旋转不变特征,首先在特征计算之前估计检测目标的中心、比例、方向信息,然后根据尺寸变化、颜色自相似性和结构自相似性进行链级特征。此外还有基于纹理的纹理分类器用于区分文本和非文本候选区<sup>[17]</sup>。文献[18]提出了四边形滑动窗口、蒙特卡罗方法和平滑  $L_n$  损失来检测定向文本,这在检测十分复杂的场景时是有效的。

## 1.2 EAST 算法

本文重点研究了 Zhou 等<sup>[19]</sup>提出的 EAST(An Efficient and Accurate Scene Text Detector)算法,经验证,该算法在准确性和效率总体方面明显优于其他的方法。在 EAST 算法中提出了一种快速、准确的场景文本检测流水线,该流水线只有两个阶段。管道采用完全卷积网络(FCN)模型,直接生成字或文本行级别的预测,不包括冗余和缓慢的中间步骤。生成的文本预测,可以是旋转的矩形或四边形,发送到非最大抑制以产生最终结果,算法总体结构如图 2 所示。

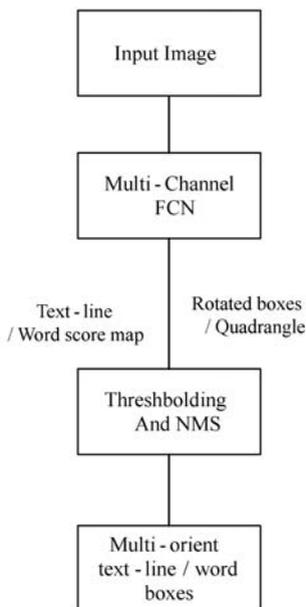


图 2 EAST 算法结构图

(1) EAST 算法管道和网络设计 从 EAST 算法结构图可以看出该算法遵循 Densebox<sup>[20]</sup>的总体设计,将图像输入到 FCN 中,生成多个像素级文本分数图和

几何图形通道。其中一个预测通道是一个分数图,其像素值在  $[0,1]$  范围内。其余通道表示从每个像素的视图中包含单词的几何图形。分数代表在同一位置预测的几何图形的可信度。

在进行网络设计时,为处理文本字体大小以及各个形态的特征差异,该算法采用了 U 形<sup>[21]</sup>的思想,在保持上采样分支较小的同时,逐渐合并特征图,最终形成了一个既能利用不同层次特性又能保持较小计算成本的网络。

(2) EAST 算法几何图形的生成 为不失一般性,该算法只考虑了四边形的情况,最终生成的几何图是 RBox 或者 Quad 之一。对于文本区域标注为四元样式的数据集,首先生成一个旋转矩形,该矩形覆盖区域最小。然后,对于每个具有正分数的像素,计算其到文本框的 4 个边界的距离,并将它们放到 RBOX 地面真值的 4 个通道中。对于四边形地真值,8 通道几何图中每个带正分数的像素的值是其从四边形的 4 个顶点的坐标偏移,进而确定文本区域。

## 2 注意力机制

视觉注意已经在各种结构的神经网络对于目标检测和预测任务中有所应用,比如图像或视频字幕<sup>[20-21]</sup>和视觉问题回答<sup>[24-26]</sup>。它的可行性主要归功于合理的假设,即人类视觉不会立即整体处理整个图像。相反,人们只在需要的时间和地点专注于整个视觉空间的选择性部分<sup>[27]</sup>。具体而言,注意力不是将图像编码成静态矢量,而是允许图像特征从手边的句子上下文演化,从而导致对杂乱图像的更丰富和更长的描述。通过这种方式,视觉注意力可以被视为一种动态特征提取机制,它随着时间的推移结合了上下文定位。

当在描述图像中检测目标的特征和信息的图像处理任务中加入注意力机制,注意力模块需要处理的特征信息包含明确的序列项  $a = \{a_1, a_2, a_3, \dots, a_L\}$ ,  $a_i \in \mathbf{R}^D$ , 其中  $L$  代表特征向量的个数,  $D$  代表的是空间维度。因此所采用的注意力机制需要计算出当前时刻  $t$  每个特征向量  $a_i$  的权重  $\alpha_{t,i}$ , 公式如下:

$$e_{ti} = f_{\text{att}}(a_i, h_{t-1}) \quad (1)$$

$$\alpha_{t,i} = \frac{\exp(e_{ti})}{\sum_{k=1}^L \exp(e_{tk})} \quad (2)$$

式中:  $f_{\text{att}}()$  代表多层感知机;  $e_{ti}$  代表中间变量;  $h_{t-1}$  代表的是上个时刻的隐含状态;  $k$  代表特征向量的下标。计算出权重后,模型就可以对输入的序列  $a$  进行筛选,得到筛选后的序列项为:

$$\hat{Z}_i = \mu(\{\alpha_{i,j}\}, \{a_i\}) \quad (3)$$

最终由函数  $\mu$  来决定该注意机制是硬注意力还是软注意力<sup>[28]</sup>。

在图像检测和目标识别的深度学习领域,文献[30]在目标检测框架中引入与子区域特征和宽高比特性相关的注意力特征库,并生成注意力特征图对原始的 ROI 池化特征进行优化,模型的检测精度和检测速度有明显提升。文献[28]设计了一种引入注意力机制的 AttenGAN 模型,该模型包括一个生成器和一个判别器,生成器根据行人过去的轨迹概率性地对未来进行多种可能性预测。文献[31]介绍了一种新的卷积神经网络,称 SCA-CNN,它在 CNN 中结合了空间注意力和通道方向注意力机制,并取得了良好的效果。

### 3 引入注意力机制的 EAST 算法

#### 3.1 Attention-EAST 算法结构

针对 EAST 算法文本检测器可处理的最大文本实例大小与网络的接收字段成比例,其限制了网络预测更长文本区域的能力。为了使文本检测器 PVANet 网络保证在提取精确的文本特征和位置特征的基础上,扩大目标检测视野,本文在 EAST 算法的基础上引入了注意力机制,构成 Attention-EAST 算法,该算法的核心框架如图 3 所示。

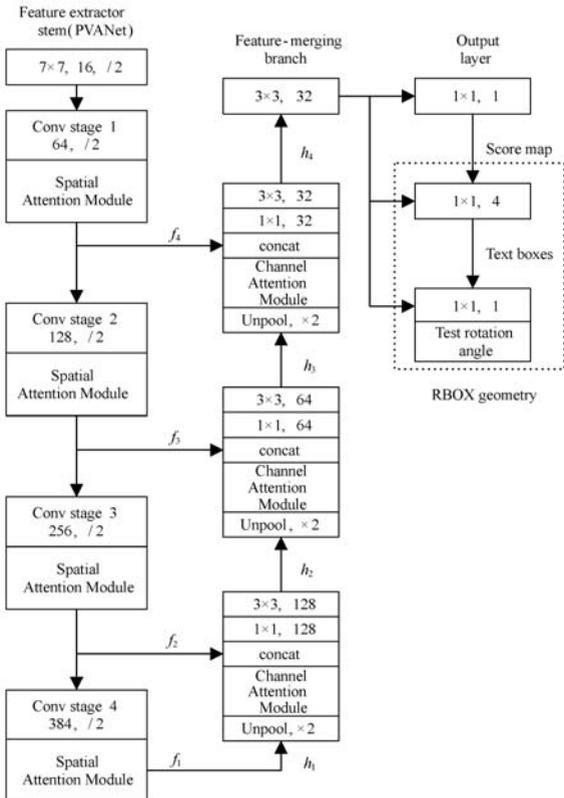


图3 Attention-EAST 算法结构图

在利用 PVANet 网络进行下采样的过程中,通过中间的文本特征信息的空间关系生成空间注意力模块,其主要功能是用来捕捉二维空间中对于目标区判定的重要性信息,每次卷积生成的特征信息为  $I \in \mathbf{R}^{1 \times H \times W}$ ,并经过 sigmoid 函数激活,其表达式为:

$$W_s(I) = \sigma f^{7 \times 7} Pool(I) \quad (4)$$

式中:  $f^{7 \times 7}$  为卷积操作,卷积核为  $7 \times 7$  的卷积层。在上采样的过程中通过 unpool 池化的方式提取特征用于对目标位特征的逼近生成通道注意力模块,然后经过共享网络 MLP 进行调整,其表达式为:

$$W_c(I') = \sigma MLP(unpool(I)) = \sigma W_1 W_0 I' \quad (5)$$

式中:  $\sigma$  为 sigmoid 激活函数;  $W_0 \in \mathbf{R}^{C/r \times C}$  和  $W_1 \in \mathbf{R}^{C \times C/r}$  分别为 MLP 的权重。最后在特征融合的过程中,将通道注意力权重和空间注意力权重构成整个软件分支注意力模型,其过程可表示为:

$$I' = (W_s(I) + 1) \odot I \quad (6)$$

$$I'' = (W_c(I') + 1) \odot I' \quad (7)$$

式中:  $\odot$  为对应矩阵元素相乘,由于每个模块最后都需要使用 sigmoid 函数来激活,使注意力通道每个元素值在  $[0, 1]$  之间,达到注意力模块强化有用图像信息和抑制无用信息的效果。

#### 3.2 损失函数设计

损失函数可表示为:

$$L = L_s + \lambda_g L_g \quad (8)$$

式中:  $L_s$  和  $L_g$  分别表示分数图和几何图形的损失,而  $\lambda_g$  表示两个损失之间的重要性。本文将  $\lambda_g$  设为 1。为了简化训练过程,本文算法借鉴了文献[32]中引入的类平衡交叉熵:

$$L_s = \text{balanced-xent}(\hat{Y}, Y^*) = -\beta Y^* \log(\hat{Y}) - (1 - \beta)(1 - Y^*) \log(1 - \hat{Y}) \quad (9)$$

式中:  $\hat{Y} = F_s$  是分数图的预测值;  $Y^*$  是基本的真实值。参数  $\beta$  是正负样本之间的平衡因子:

$$\beta = 1 - \frac{\sum_{y^* \in Y^*} y^*}{|Y^*|} \quad (10)$$

为了使大文本区域和小文本区域生成精确的文本几何预测,保持回归损失尺度不变,旋转矩形框 RBBox 回归部分采用 IoU 损失函数,因为它对不同尺度的对象是固定,其表达式为:

$$L_R = -\log(\text{IoU}(\hat{R}, R^*)) = -\log \frac{|\hat{R} \cap R^*|}{|\hat{R} \cup R^*|} \quad (11)$$

式中:  $\hat{R}$  表示为预测的几何形状;  $R^*$  是其对应的真实形状;相交矩形  $|\hat{R} \cap R^*|$  的宽度和高度分别为:

$$\omega_i = \min(\hat{d}_2, d_2^*) + \min(\hat{d}_4, d_4^*) \quad (12)$$

$$h_i = \min(\hat{d}_1, d_1^*) + \min(\hat{d}_3, d_3^*)$$

式中:  $d_1$ 、 $d_2$ 、 $d_3$  和  $d_4$  分别表示像素到其对应矩形的上、右、下和左边界的距离。联合区由以下公式给出:

$$|\hat{R} \cup R^*| = |\hat{R}| + |R^*| - |\hat{R} \cap R^*| \quad (13)$$

由此可以轻松计算交叉或者联合区域。接下来, 旋转角损失计算如下:

$$L_\theta(\hat{\theta}, \theta^*) = 1 - \cos(\hat{\theta} - \theta^*) \quad (14)$$

式中:  $\hat{\theta}$  是对旋转角度的预测;  $\theta^*$  表示实际值。最后可计算出总的几何损失为:

$$L_g = L_R + \lambda_\theta L_\theta \quad (15)$$

在实验过程中本文将  $\lambda_\theta$  设置成 10。由此整个算法的损失函数搭建完毕。

### 3.3 模型训练

对于本文提出的模型按照 EAST 算法的训练方式采用 Adam<sup>[21]</sup> 优化器对网络进行端到端的训练。为了加快学习速度, 将原始图像  $512 \times 512$  的训练样本每次统一打包成 24 个进行批处理。Adam 的学习率从  $1e^{-3}$  开始, 每 27 300 个小批量下降到十分之一, 停在  $1e^{-5}$ , 对网络进行训练, 直到性能改善趋于平稳。

## 4 实验与分析

### 4.1 实验环境

本次实验是在 Ubuntu18.04 LTS 操作系统上进行, 开发语言为 Python 3.6, 集成开发环境为 Pycharm, 深度学习框架是 GPU 版本的 TensorFlow。硬件配置 CPU 为四核八线程的 i7-6700k, 其主频 4 GHz, 内存为 32 GB, GPU 为 NVIDIA GTX 1080T, 显存 11 GB。

### 4.2 实验结果

本次实验采用的数据集为 ICDAR 挑战赛所用的数据集, 该数据集也是当文本目标检测算法中比较流行的数据集, 共有 1 500 张图片, 其中 1 000 张图片用于模型训练, 其余图片用于测试集。其文本区域由四边形的四个顶点进行注释, 对应于目标文本中的四边几何图形。这些图片均由手机或相机随机拍摄, 因此, 场景中的文本信息是任意方向的, 而且可能受到自然环境的影响, 这些特征有利于对文本检测算法的估计检验。

本文引入注意力机制的 Attention-EAST 算法与 EAST 算法在处理自然场景下的长文本的检测结果对比如图 4 所示。



(a) EAST 算法长文本检测效果图



(b) 本文算法长文本检测效果图

图 4 长文本检测对比图

可以看出, 通过加入注意力机制对于提取文本和方位的特征信息的增强, 提高了文本检测视野, 有效改善了对于长文本的检测效果。同时, 本文使用召回率 (Recall)、准确率 (Precision) 和加权调和平均值 F-measured 三个指标来评价本文算法在 ICDAR 数据集上的训练效果。并选择了当前文本检测领域比较有代表性的算法进行对比, 包括以 R-CNN 为框架的文献 [10], 以 SSD 为框架的文献 [15] 和以 Faster R-CNN 为框架的文献 [16], 实验结果如表 1 所示。实验结果表明, 本文提出的引入注意力机制的方法相比原 EAST 算法在文本检测性能指标均有所提升, 并且相对优于其他主流的文本检测算法。

表 1 各文本检测算法实验结果对比数据

算法	Recall	Precision	F-measure
Attention-EAST	0.790 2	0.840 1	0.814 4
EAST	0.783 1	0.822 4	0.802 2
文献 [10]	0.677 5	0.798 5	0.733 0
文献 [15]	0.586 6	0.721 4	0.647 1
文献 [16]	0.523 9	0.732 1	0.610 7

为分析引入注意力模块后在检测效率上对于原 EAST 算法的影响, 在本文的实验环境下采用每秒帧率

(Frame Per Second, FPS) 这一指标来评价本文算法和原 EAST 算法的检测效率,表示每秒处理的图片数量,将测试集 500 张检测图片随机分成 5 份分别进行测试。实验结果如表 2 所示,可以看出加注注意力模块后,并没有损失原算法的检测效率。

表 2 两种算法文本检测效率对比数据(FPS)

算法	1st	2nd	3rd	4th	5th	Average
EAST	4.83	4.86	4.90	4.82	4.86	4.854
Attention-EAST	4.86	4.84	4.82	4.87	4.90	4.858

## 5 结 语

本文研究了当前计算机视觉领域中中文本检测的算法,其中针对 EAST 算法在对于文本方向特征提取时视野受限的问题。通过在主干网络 PVANet 中引入注意力机制,提出一种 Attention-EAST 算法,使得训练模型在提取文本目标特征时更加关注有用信息和抑制无用信息。实验证明,该方法有效提升了 EAST 算法检测长文本的能力,在没有损失检测效率的同时提升了检测精度。本文研究通过引入注意力机制的方法可为自然环境下的文本检测研究领域提供一定的参考。

## 参 考 文 献

- [1] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2014: 580–587.
- [2] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks[C]//Proceedings of the Advances in Neural Information Processing Systems. Lake Tahoe, USA, 2012: 1097–1105.
- [3] Girshick R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2015: 1440–1448.
- [4] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2015, 1: 91–99.
- [5] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector[C]//Proceedings of European Conference on Computer Vision. Aire-la-Ville: Eurographics Association Press, 2016: 21–37.
- [6] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2016: 779–788.
- [7] Jaderberg M, Simonyan K, Vedaldi A, et al. Reading text in the wild with convolutional neural networks[J]. International Journal of Computer Vision, 2016, 116(1):1–20.
- [8] Bissacco A, Cummins M, Netzer Y, et al. PhotoOCR: Reading text in uncontrolled conditions[C]//2013 IEEE International Conference on Computer Vision (ICCV). IEEE, 2013.
- [9] Neumann L, Matas J. Real-time scene text localization and recognition[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2012.
- [10] Gomez L, Karatzas D. TextProposals: A text-specific selective search algorithm for word spotting in the wild[J]. Pattern Recognition, 2017, 70: 60–74.
- [11] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 39(4):640–651.
- [12] Chucai Y, Yingli T. Localizing text in scene images by boundary clustering, stroke segmentation, and string fragment classification[J]. IEEE Transactions on Image Processing, 2012, 21(9):4256–4268.
- [13] Tian Z, Huang W, He T, et al. Detecting text in natural image with connectionist text proposal network[C]//Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part VIII. 2016:56–72.
- [14] Cao X, Ren W, Zuo W, et al. Scene text deblurring using text-specific multiscale dictionaries[J]. IEEE Transactions on Image Processing, 2015, 24(4):1302–1314.
- [15] Liao M, Shi B, Bai X, et al. TextBoxes: A fast text detector with a single deep neural network[C]//Thirty-First AAAI Conference on Artificial Intelligence, 2016.
- [16] Ma J, Shao W, Ye H, et al. Arbitrary-oriented scene text detection via rotation proposals[J]. IEEE Transactions on Multimedia, 2017, 20(11):3111–3122.
- [17] Yao C, Bai X, Sang N, et al. Scene text detection via holistic, multi-channel prediction[EB]. arXiv: 1606. 09002, 2016.
- [18] Liu Y, Jin L. Deep matching prior network: Toward tighter multi-oriented text detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
- [19] Zhou X, Yao C, Wen H, et al. EAST: An efficient and accurate scene text detector[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
- [20] Huang L, Yang Y, Deng Y, et al. DenseBox: Unifying landmark localization with end to end object detection[EB]. eprint arXiv:1509.04874, 2015.

法在挖掘时间和内存占用上均有较高的效率。将该算法应用于对某高校专任教师的基本信息和人才类型进行关联规则挖掘上,从中提取潜在的有价值的信息,发现影响人才发展的因素,并且为高校人才引进和选拔提供科学依据,从而促进高校的全面发展,提高高校的竞争力。

## 参 考 文 献

- [1] Wu X, Zhu X, Wu G Q, et al. Data mining with big data [J]. IEEE Transactions on Knowledge and Data Engineering, 2014, 26(1):97-107.
- [2] 崔妍,包志强. 关联规则挖掘综述[J]. 计算机应用研究, 2016, 33(2):330-334.
- [3] Jamsheela O, Raju G. Frequent itemset mining algorithms: A literature survey [C]//2015 IEEE International Advance Computing Conference (IACC). IEEE, 2015.
- [4] Agrawal R, Srikant R. Fast algorithms for mining association rules [C]//Proceedings of the 20th VLDB International Conference on Very Large Data Bases. CA, 1994:487-499.
- [5] Krupali R, Garg D. Survey on the techniques of FP-growth tree for efficient frequent item-set mining [J]. International Journal of Computer, 2017, 160(1):40-43.
- [6] Deng Z H, Wang Z H, Jiang J J. A new algorithm for fast mining frequent itemsets using N-lists [J]. Science China Information Sciences, 2012, 55(9):2008-2030.
- [7] Deng Z H, Lv S L. Fast mining frequent itemsets using Nodesees [J]. Expert Systems with Applications, 2014, 41(10):4505-4512.
- [8] Aryabarzan N, Minaei-Bidgoli B, Teshnehlab M. negFIN: An efficient algorithm for fast mining frequent itemsets [J]. Expert Systems with Applications, 2018, 105(12):129-143.
- [9] Vo B, Le T, Coenen F, et al. Mining frequent itemsets using the N-list and subsume concepts [J]. International Journal of Machine Learning & Cybernetics, 2016, 7(2):253-265.
- [10] 沈戈晖,刘沛东,邓志鸿. NB-MAFIA: 基于 N-List 的最长频繁项集挖掘算法 [J]. 北京大学学报(自然科学版), 2016, 52(2):199-209.
- [11] Dam T L, Li K, Fournier-Viger P, et al. An efficient algorithm for mining top-rank- k frequent patterns [J]. Applied Intelligence, 2016, 45(1):96-111.
- [12] 林晨,顾君忠. 基于 Nodesees 的最大频繁项集挖掘算法 [J]. 计算机工程, 2016, 42(12):204-207,216.
- [13] Wang Q, Zhang K K, Chen H T. Association rule mining for the talents introduction strategy: a case study of Zhejiang University of Finance & Economics [J]. American Journal of Applied Mathematics, 2018, 6(2):55-61.

## (上接第 203 页)

- [21] Kingma D P, Ba J. Adam: A method for stochastic optimization [EB]. eprint arXiv:1412.6980, 2014.
- [22] Xu K, Ba J, Kiros R, et al. Show, attend and tell: Neural image caption generation with visual attention [EB]. eprint arXiv:1502.03044, 2015.
- [23] Yao L, Torabi A, Cho K, et al. Describing videos by exploiting temporal structure [C]//2015 IEEE International Conference on Computer Vision (ICCV), 2015.
- [24] Chen K, Wang J, Chen L C, et al. ABC-CNN: An attention based convolutional neural network for visual question answering [J]. eprint arXiv:1511.05960, 2015.
- [25] Yang Z, He X, Gao J, et al. Stacked attention networks for image question answering [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.
- [26] Xu H, Saenko K. Ask, attend and answer: Exploring question-guided spatial attention for visual question answering [M]//Computer Vision—ECCV 2016. Springer International Publishing, 2016.
- [27] Corbetta M, Shulman G L. Control of goal-directed and stimulus-driven attention in the brain [J]. Nature reviews Neuroscience, 2002, 3(3):215-229.
- [28] Li C, Hou Y, Wang P, et al. Joint distance maps based action recognition with convolutional neural networks [J]. IEEE Signal Processing Letters, 2017, 24(5):624-628.
- [29] 孙萍,胡旭东,张永军. 结合注意力机制的深度学习图像目标检测 [J/OL]. 计算机工程与应用:1-7 [2019-04-29].
- [30] 孙亚圣,姜奇,胡洁,等. 基于注意力机制的行人轨迹预测生成模型 [J]. 计算机应用, 2019, 39(3):668-674.
- [31] Chen L, Zhang H, Xiao J, et al. SCA-CNN: spatial and channel-wise attention in convolutional networks for image captioning [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017:6298-6306.
- [32] Xie S, Tu Z. Holistically-Nested Edge Detection [J]. International Journal of Computer Vision, 2015, 125(1/3):3-18.

## (上接第 248 页)

- [10] 周丽丽,王金玉,孙晶. 大数据中位置服务的隐私保护方法的研究 [J]. 计算机应用, 2016, 35(8):37-42.
- [11] 张琳,刘彦,王汝传. 位置大数据服务中基于差分隐私的数据发布技 [J]. 中通信学报, 2016, 37(9):46-54.
- [12] 周志华,杨强. 机器学习及其应用 [M]. 清华大学出版社, 2011.
- [13] MARY S S, Selvi R T. A study of K-means and cure clustering algorithms [J]. International Journal of Engineering Research and Technology, 2014, 3(2):1985-1987.
- [14] Dwork C. The promise of differential privacy: a tutorial on algorithmic techniques [C]//Proceedings of Foundations of Computer Science. Piscataway: IEEE Press, 2011:1-2.