

基于高分辨率卷积神经网络的场景文本检测模型

陈淼妙 续晋华

(华东师范大学计算机科学与技术学院 上海 200062)

摘 要 卷积神经网络中常用的降采样在增大感受野的同时,减小了特征图的空间分辨率,导致大尺度文本在网络高层特征图中边界模糊,小尺度文本直接丢失。针对这种情况,设计一个主干网络,使其高层特征图具有较高的空间分辨率,用以提高模型对大尺度文本的定位能力和小尺度文本的分类能力,减少对背景误检的情况,从而提高检测的召回率和准确率。实验结果表明,采用该主干网络可以有效提升模型对多尺度文本的检测能力,F 值在 ICDAR2015 数据集上达到 81.89%。

关键词 场景文本检测 多尺度 卷积神经网络 特征融合 特征图分辨率

中图分类号 TP183 文献标志码 A DOI: 10.3969/j.issn.1000-386x.2020.10.022

SCENE TEXT DETECTION MODEL BASED ON HIGH RESOLUTION CONVOLUTIONAL NEURAL NETWORKS

Chen Miaomiao Xu Jinhua

(School of Computer Science and Technology, East China Normal University, Shanghai 200062, China)

Abstract The down-sampling method is commonly used in convolutional neural networks, which brings large valid receptive field, but reduces the spatial resolution of feature maps, resulting in the boundary blurring of large-scale texts and missing of small-scale texts in the high-level feature maps of the network. In view of this, we design a novel backbone network which has high spatial resolution in deeper layers. It improved the location ability of the model for large-scale texts and classification capability for small-scale texts, and reduced the background false detection, therefore improving both recall and precision of detection. Experimental results show that the backbone network can effectively improve the detection ability of the model for multi-scale texts, and achieve the F-measure of 81.89% on ICDAR2015 dataset.

Keywords Scene text detection Multi-scale Convolutional neural networks Feature fusion Feature map resolution

0 引 言

随着互联网和移动终端的飞速发展,越来越多的信息载体以图像的形式存在。自然场景中的文本具有丰富、明确的语义信息,利用计算机技术快速、准确地提取场景图像中的文本信息是计算机视觉和模式识别领域当下热门的研究课题之一。场景文本检测技术是文本识别的基础,在人们的日常生活和生产中有着广泛的应用。例如:手机设备上的拍照翻译软件,可以对异国街景中的文本进行即时的实景翻译,给出国旅游

带来极大的便利;汽车上的智能拍摄设备,可以对道路交通指示牌上的文本进行检测和识别,获取实时的地理位置信息和道路周边情况,实现辅助驾驶。除此之外,场景文本检测技术在盲人辅助、商业票据处理、工业自动化等领域也有着重要的应用价值。

与传统 OCR^[1]相比,自然场景图像中的文本检测面临背景复杂、文本尺度和字体多样、图像质量的不确定性等诸多困难和挑战。近年来,随着深度学习技术的迅速发展,深度学习的方法在文本检测任务上效果显著。现有的卷积神经网络已经具备很好的表征能力,但网络定位能力不强,对文本的边界定位不准确。特

收稿日期:2019-07-26。国家自然科学基金项目(61175116)。陈淼妙 硕士生 主研领域:计算机视觉 图形图像处理。续晋华,副教授。

征金字塔网络可以融合不同尺度的特征,但小尺度文本的高级语义信息在网络高层已经丢失,导致模型对多尺度文本的检测能力不强。

鉴于此,本文设计专门用于文本检测的主干网络,通过引入空洞卷积层,提高网络高层特征图的空间分辨率,以学习更精细的文本位置信息,提高模型的多尺度检测性能。

1 研究现状

场景文本检测是从自然场景图像中定位出文本的位置。近年来,随着基于深度学习的目标检测和图像分割算法的快速发展,场景文本检测的研究也进入了一个新阶段,涌现出大量基于深度学习的方法。这些方法主要分为三类:基于回归的文本检测方法、基于分割的文本检测方法和基于混合的文本检测方法。

1.1 基于回归的文本检测方法

受通用目标检测框架的启发,基于回归的文本检测方法通常采用回归文本边界框的方式获取文本区域,主要分为两类:

(1) 间接回归。Faster R-CNN^[2]的思想进入文本检测领域,将目标检测的各个阶段整合到深度神经网络中。间接回归首先通过卷积神经网络得到整幅图像的特征图,再通过锚点(anchor)机制或区域建议网络(Region Proposal Network)获得文本候选框,最后使用回归方法或其他方式精细调整候选框得到最后的文本检测结果。CTPN^[3](Connectionist Text Proposal Network)提出了一种垂直锚点机制,将文本区域细分为多个固定宽度、预测高度的文本段,以提取多尺度的文本特征。RRPN^[4](Rotation Region Proposal Networks)利用任意角度的区域建议网络,可以检测任意方向的场景文本。

(2) 直接回归。YOLO^[5]属于直接回归的目标检测框架。受此启发,直接回归的文本检测方法直接用一个神经网络预测给定点到文本边界框的偏移量,不需要进行二次回归。FCRN^[6](Fully-Convolutional Regression Network)利用人工合成的场景文本训练基于YOLO的深度神经网络。He等^[7]基于直接回归,采用多任务学习的框架,缓解了锚点机制不能有效检测倾斜长文本的问题。

这类方法采用整体化思想,计算速度较快,但在一些情况下检测出的文本边界框不够精确且无法调整。

1.2 基于分割的文本检测方法

文本检测的核心是设计特征来区分出文本和背

景。基于分割的文本检测方法将文本检测问题转化为文本与背景的语义分割问题,首先通过分割网络提取文本块区域,再进行后处理获取文本边界框。大多数方法采用全卷积网络框架(FCN^[8])。Yao等^[9]将文本置信度、字符置信度和相邻字符连接方向多个通道的信息在一个网络中进行端到端训练,以精确分割文本与背景。He等^[10]先利用一个多尺度全卷积网络提取文本块区域,再级联一个实例感知分割网络得到精细的文本实例。PixelLink^[11]基于实例分割,先进行文本类别和连通预测,再利用实例分割得到最后的检测结果。这些方法可以检测倾斜或呈曲线的文本行,对文本边界定位较为准确,但对小尺度文本的检测能力不强,而且离得较近的文本行不容易区分开。

1.3 基于混合的文本检测方法

混合的思想是将回归方法与分割方法相结合,融合两种方法的优点。Zhou等^[12]提出了一个简单、高效的文本检测器(EAST),直接在整幅图像上预测任意方向的场景文本,整个流程没有冗余的中间步骤。Lyu等^[13]提出基于角点检测与区域分割的文本检测方法,以准确检测长宽比变化较大的文本实例,解决了相邻文本行的分割结果粘连的问题。

以上这些方法更多地受益于通用目标检测框架和图像分割方法,较少针对文本的尺度变化研究专门用于场景文本检测的主干网络。因此,本文基于混合思想,设计更关注文本的尺度变化和边界位置的网络结构,在不增加计算量的前提下提高模型的检测性能。

本文主要贡献是研究高分辨率特征图对文本检测结果的影响,因此本文采用了EAST^[12]中的流程(Pipeline),设计新的高分辨率主干网络。实验结果表明,新的主干网络与原始的EAST模型相比,提高了召回率,取得了更好的检测性能。

2 方法原理

本文提出的模型是一个端到端可训练的场景文本检测器,该模型由两部分组成:深度神经网络和后处理。受DetNet^[14]启发,通过维持较高的网络高层特征图的空间分辨率,让网络学习到文本的多尺度特征和精细的边界位置,从而提高模型的分类、定位能力。

2.1 网络结构

模型的主体是深度神经网络,网络结构如图1所示。网络由三部分构成:主干网络、特征融合层和输出层。图1中:Conv代表卷积操作,后面的参数分别为卷积核尺寸、通道数和下采样倍数;Max pool代表最大

池化操作; Concat 代表将两个特征图按通道维度进行拼接; Unpool 代表上采样, 直接使用双线性插值。

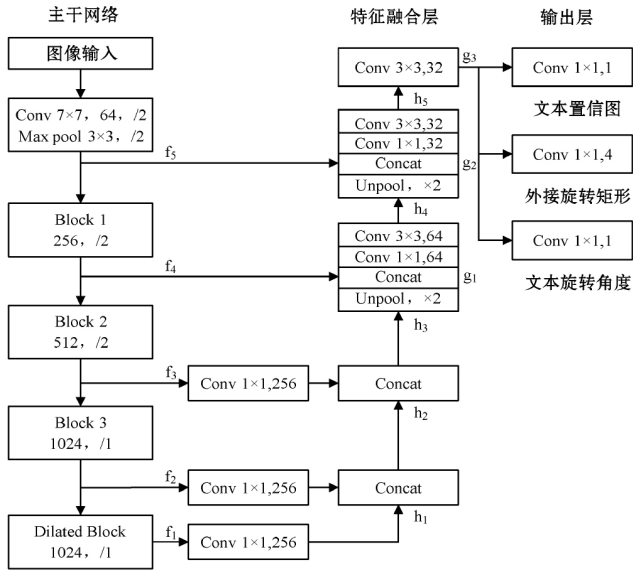


图1 网络结构

主干网络: 基于 ResNet50, 改进 Block 3 使得其特征图分辨率是输入的 1/16 (而不是原来的 1/32), 以提取到合适的多尺度特征和精细的文本位置。相应地改进 Block 4 为 Dilated Block, 该结构块由一个带有 1×1 卷积项的空洞卷积^[15]结构(图 2(a))开始, 再堆叠两个无 1×1 卷积项的空洞卷积结构(图 2(b)), 以增大网络的有效感受野, 捕获高级语义信息, 增强网络的分类能力。同时, Dilated Block 的网络宽度与 Block 3 保持一致, 以减少网络的计算代价。

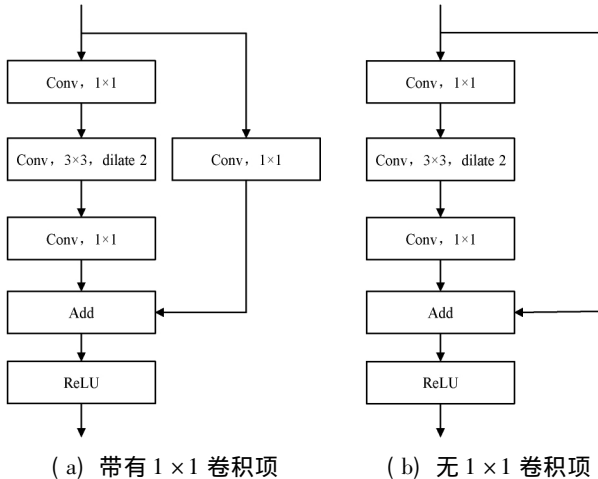


图2 空洞卷积结构

特征融合层: 从主干网络的最高层开始, 将其中的 5 层特征融合, 以增强网络低层的表征能力, 充分利用网络各层提取到的特征来应对文本的尺度变化。融合方式如下:

$$h_i = \begin{cases} conv_{1 \times 1}(f_i) & i = 1 \\ [h_{i-1}, conv_{1 \times 1}(f_i)] & 1 < i \leq 3 \\ conv_{3 \times 3}(conv_{1 \times 1}([g_{i-3}, f_i])) & 3 < i \leq 5 \end{cases} \quad (1)$$

$$g_i = \begin{cases} unpool(h_{i+2}) & i \leq 2 \\ conv_{3 \times 3}(h_{i+2}) & i = 3 \end{cases} \quad (2)$$

式中: $[\cdot, \cdot]$ 代表拼接操作。

输出层: 特征融合层得到原图四分之一大小的特征图, 利用 1×1 卷积操作得到 6 个通道的输出, 其中一个通道预测每个像素属于文本区域的置信度, 四个通道分别预测每个像素距离文本外接矩形 4 条边的距离, 另一个通道预测该文本外接矩形的旋转角度。

2.2 标签生成

标签生成过程如图 3 所示, 图中虚线四边形为 ground truth 对文本区域的标注。将虚线四边形的 4 条边均向内缩减 0.3 倍的短边长得到一个内部四边形, 目的是缓解紧挨文本区域存在的分割结果粘连的问题; 将内部四边形内的像素视作正样本, 真值为 1, 其余区域值设为 0; 生成虚线四边形的最小外接矩形; 仅对内部四边形内的像素计算损失, 以图 3 中一个像素点为例, 4 个箭头代表该像素点距离外接矩形 4 条边的距离; 用 θ 表示外接矩形的旋转角度, 最终表示出文本区域的外接旋转矩形。

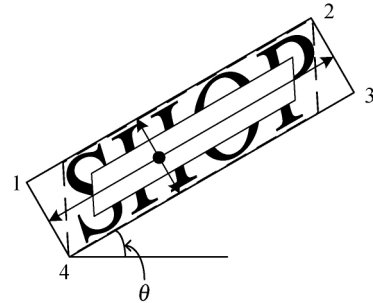


图3 标签生成示意图

2.3 损失函数

网络的损失函数是分类损失与回归损失的加权和:

$$L = L_{cls} + \lambda_{reg} L_{reg} \quad (3)$$

式中: L_{cls} 代表分类损失; L_{reg} 代表回归损失; 平衡系数 λ_{reg} 设为 1。与 EAST 中的交叉熵损失不同, 本文采用二分类的 Dice 损失作为分类损失 L_{cls} :

$$L_{cls} = 1 - \frac{2 \sum Y^* \hat{Y}}{\sum Y^* + \sum \hat{Y}} \quad (4)$$

式中: Y^* 代表真值; \hat{Y} 代表预测的每个像素属于文本区域的置信度。每个像素对应的外接旋转矩形的损失采用 P_{IoU} (预测位置与真实位置面积交叠率) 的相反数来进行计算:

$$L_{geo} = -\log P_{IoU}(R^*, \hat{R}) = -\log \frac{|R^* \cap \hat{R}|}{|R^* \cup \hat{R}|} \quad (5)$$

式中: R^* 代表真实的文本区域的外接旋转矩形; \hat{R} 代表预测的文本区域的外接旋转矩形。外接旋转矩形的旋转角度损失为:

$$L_{\theta} = 1 - \cos(\hat{\theta} - \theta^*) \quad (6)$$

式中: $\hat{\theta}$ 代表文本旋转角度的预测值; θ^* 代表文本旋转角度的真值。最终, 回归损失为:

$$L_{\text{reg}} = L_{\text{geo}} + \lambda_{\theta} L_{\theta} \quad (7)$$

式中: λ_{θ} 是几何损失和角度损失的平衡系数, 实验中设置为 20。

2.4 非极大值抑制算法

非极大值抑制算法 (Non-Maximum Suppression, NMS) 的本质是局部极大值搜索, 抑制非极大值元素。该算法在文本检测任务的后处理中被广泛应用, 目的是去除冗余的检测结果, 得到最佳的文本边界框。

模型的后处理采用融合 NMS^[12] 算法步骤如下:

(1) 对模型生成的文本边界框按行从左向右依次进行融合操作, 计算两个文本边界框的面积交叠率 (P_{IoU}), 计算公式如下:

$$P_{\text{IoU}} = \frac{\text{area}(P) \cap \text{area}(Q)}{\text{area}(P) \cup \text{area}(Q)} \quad (8)$$

式中: $\text{area}(\cdot)$ 代表文本边界框的面积。如果交叠率大于设定的阈值, 对文本边界框的 4 个顶点坐标分别进行相应的权重平均, 得到融合后的文本边界框, 其置信度设为原始文本边界框的置信度之和; 否则, 保留原始文本边界框。

(2) 将经过融合操作后的所有文本边界框按照置信度从高到低排序, 并将第一个文本边界框作为当前抑制文本边界框。

(3) 非极大值抑制: 将其他文本边界框作为被抑制的元素, 计算当前抑制文本边界框与被抑制文本边界框的面积交叠率。如果交叠率大于设定的阈值, 去除该文本边界框。

(4) 按照步骤 (2) 中的排序结果, 取下一个未被抑制的文本边界框作为当前抑制文本边界框, 执行步骤 (3), 重复步骤 (4) 直至剩下最后一个文本边界框, 算法结束。

3 实验

3.1 数据集

为了验证模型的有效性, 实验在公开的场景文本检测数据集 ICDAR2015 上进行性能评估。ICDAR2015

数据集包含 1 000 幅训练图像和 500 幅测试图像, 图像相对模糊且文本为任意方向、包含畸变, 标注真值为文本区域外接四边形的 4 个顶点坐标。

3.2 网络参数设置

实验中使用 Adam 优化器来进行网络训练, 训练数据为 ICDAR2013 训练集中的 229 幅图像和 ICDAR2015 训练集中的 1 000 幅图像。为了加速网络收敛, 训练前将图像随机裁剪缩放 (缩放比例为 0.5、1.0、2.0、3.0), 并用黑色填充以重新将图像大小调整为 512×512 , Batch Size 设为 16。初始学习率为 0.000 1, 之后每 1 万步以 0.96 的比例衰减, 总迭代次数为 9 万次。

3.3 性能指标

研究采用标准的自然场景文本检测评估指标: 准确率 (P)、召回率 (R) 和 F 值 (F)。它们的计算公式分别为:

$$P = \frac{|TP|}{|D|} \quad (9)$$

$$R = \frac{|TP|}{|G|} \quad (10)$$

$$F = \frac{2 \times P \times R}{P + R} \quad (11)$$

式中: TP 表示检测正确的文本框数量; D 表示模型检测出的文本框数量; G 表示 ground truth 中真实的文本框数量。若检测结果与 ground truth 之间的交并比大于阈值 (一般设为 0.5) 并且得分也大于相应阈值, 则认为该检测结果是正确的。F 值是准确率和召回率的调和平均数, 以综合评估模型的性能。

3.4 实验结果

为了验证专为文本检测任务设计的主干网络 (TextNet) 的性能, 首先将提出的模型 (HRTexDet) 在 ICDAR2015 数据集上进行测试, 在输入图像为单尺度的条件下, 与原始模型 (EAST^[12]) 进行实验对比。如表 1 所示, 本文模型最大的优势在于召回率得到显著的提升, 达到 79.35%, 比原始模型提高 5.88%; 准确率达到 84.60%, 比原始模型提高 1.03%; F 值达到 81.89%, 提升 3.69%。

表 1 HRTexDet 与原始模型的性能对比 %

模型	准确率	召回率	F 值
PVNet2x (EAST)	83.57	73.47	78.20
PVNet2x_MS (EAST)	83.27	78.33	80.72
HRTexDet50 (本文模型)	84.60	79.35	81.89

注: MS 代表多尺度测试, 50 是网络深度, PVNet2x 的通道数是 PVNet 的两倍

此外,原始模型 PVANet2x(EAST) 使用多尺度测试 PVANet2x_MS(EAST) 后,召回率得到极大的提升,说明原始模型尽管融合了高低层特征,仍无法应对文本剧烈的尺度变化。然而,本文模型使用单尺度测试与使用多尺度测试的 PVANet2x_MS(EAST) 相比,在准确率、召回率和 F 值上分别仍有 1.33%、1.02% 和 1.17% 的提升。这是因为专为文本检测设计的主干网络,能够提取到文本的多尺度特征和精细的边界位置。图 4 为原始 EAST 模型和提出模型的结果对比图。图 4(a) 中,原始 EAST 模型在多尺度文本检测上能力不足,而提出的模型具有较高的网络高层特征图的空间分辨率,在不均匀照明、低分辨率、背景复杂等各种有挑战的场景中,仍能有效地检测出多尺度文本;图 4(b) 中,原始 EAST 模型对文本的边界定位不准确,而本文模型能准确地定位出文本的位置;图 4(c) 中,原始 EAST 模型存在背景误检的情况,而本文模型能有效地保留高级语义信息,具有较强的分类能力,能减少对背景误检的情况。



(a) 多尺度检测



(b) 边界定位



(c) 背景误检

图 4 原始 EAST 模型(左)和本文模型(右)的实验结果对比

为了探讨主干网络的网络深度对模型性能的影响,改变主干网络的网络深度,在 TextNet50 的最高层堆叠一个 Dilated Block,得到 TextNet59。在 ICDAR2015 数据集上进行单尺度测试,并与 HRTextDet50 模型进行实验对比。如表 2 所示,主干网络为 TextNet59 的模型在准确率、召回率和 F 值上与主干网络为 TextNet50 的模型相比均有所下降。这表明主干网络的深度不是越

深越好,与文本检测可能不需要太深的高级语义信息这一猜想吻合。

表 2 不同深度 TextNet 的性能对比 %

模型	准确率	召回率	F 值
TextNet59	83.99	78.57	81.19
TextNet50	84.60	79.35	81.89

为了验证 Dilated Block 设计中 1×1 卷积项具有开启新的语义阶段的作用,实验使用无 1×1 卷积项的空洞卷积结构开启 TextNet59 主干网络的第二个 Dilated Block,把该结构叫作 TextNet59_v0。在 ICDAR2015 数据集上用 TextNet59_v0 进行单尺度测试,并与主干网络为 TextNet59 的 HRTTextDet 模型进行实验对比。如表 3 所示,主干网络为 TextNet59_v0 的模型在 F 值上与主干网络为 TextNet59 的模型相比略有下降,这表明 1×1 卷积项可以开启新的语义阶段。

表 3 Dilated Block 中有/无 1×1 卷积项的性能对比 %

模型	准确率	召回率	F 值
TextNet59_v0	84.82	77.47	80.98
TextNet59	83.99	78.57	81.19

图 5 给出了更多检测结果,其中:(a)是检测成功的例子;(b)是检测失败的例子。图像中的文本太过模糊时,对文本检测器是个很大的挑战。



(a) 检测成功示意图



(b) 检测失败示意图

图 5 HRTTextDet 模型检测结果示意图

最后,把 HRTTextDet 模型与常用的自然场景文本检测方法进行实验对比,在 ICDAR2015 数据集上的实验结果如表 4 所示。可以看出,本文模型 F 值达到

81.89%,比之前最好的方法提高 1.19%。HRTTextDet 模型的最大的优势在于召回率得到显著的提升,比之前最好的方法 SegLink^[16]的 76.80% 提高 2.55%。

表 4 在 ICDAR2015 数据集上的实验结果 %

模型	准确率	召回率	F 值
CTPN ^[3]	51.56	74.22	60.85
He 等 ^[10]	76.00	54.00	63.00
RRPN ^[4]	79.41	70.00	74.41
SegLink ^[16]	73.10	76.80	75.00
EAST ^[12]	83.57	73.47	78.20
Lyu 等 ^[13]	94.10	70.70	80.70
HRTTextDet50(本文模型)	84.60	79.35	81.89

4 结 语

本文提出一个具有高空间分辨率的文本检测模型,通过专为文本检测任务设计的主干网络使得模型可以提取到文本的多尺度特征和精确的边界位置。实验结果表明,在不均匀照明、低分辨率、背景复杂等各种有挑战的场景中,本文模型能有效地应对文本剧烈的尺度变化,准确地检测出场景文本。本文模型能检测任意方向的场景文本,由于采用旋转矩形来表示文本区域,在一定程度上限制了模型对弯曲文本的检测性能。因此,下一步的工作将尝试采用多边形来表示文本区域,使得模型能检测任意方向和任意形状的场景文本。

参 考 文 献

- [1] 朱利娟,云中,边巴旺堆. 基于极坐标变换的脱机手写藏文字符特征提取方法[J]. 计算机应用与软件, 2018, 35(3): 162-166.
- [2] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 39(6): 1137-1149.
- [3] Tian Z, Huang W, He T, et al. Detecting text in natural image with connectionist text proposal network[C]//European Conference on Computer Vision, 2016.
- [4] Ma J, Shao W, Ye H, et al. Arbitrary-oriented scene text detection via rotation proposals[J]. IEEE Transactions on Multimedia, 2018, 20(11): 3111-3122.
- [5] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016.
- [6] Gupta A, Vedaldi A, Zisserman A. Synthetic data for text

localisation in natural images [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition 2016.

- [7] He W , Zhang X Y , Yin F , et al. Deep direct regression for multi-oriented scene text detection [C]//2017 IEEE International Conference on Computer Vision , 2017.
- [8] Long J , Shelhamer E , Darrell T. Fully convolutional networks for semantic segmentation [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition , 2015.
- [9] Yao C , Bai X , Sang N , et al. Scene text detection via holistic multi-channel prediction [EB]. arXiv: 1606.09002 2016.
- [10] He D , Yang X , Liang C , et al. Multi-scale fcn with cascaded instance aware segmentation for arbitrary oriented word spotting in the wild [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition , 2017.
- [11] Deng D , Liu H , Li X , et al. Pixellink: detecting scene text via instance segmentation [C]//32nd AAAI Conference on Artificial Intelligence , 2018.
- [12] 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 , 2017.
- [13] Lyu P , Yao C , Wu W , et al. Multi-oriented scene text detection via corner localization and region segmentation [C]//2018 IEEE Conference on Computer Vision and Pattern Recognition , 2018.
- [14] Li Z , Peng C , Yu G , et al. Detnet: design backbone for object detection [C]//European Conference on Computer Vision (ECCV) , 2018.
- [15] Yu F , Koltun V. Multi-Scale context aggregation by dilated convolutions [EB]. arXiv: 1511.07122 , 2015.
- [16] Shi B , Bai X , Belongie S. Detecting oriented text in natural images by linking segments [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition , 2017.

(上接第 26 页)

着至关重要的作用。本文通过查询逆向工程的方法,根据用户提供的符合目标模式的实例数据以及源数据库,计算出源数据库生成给定实例的查询,在实例数据列和源数据库列之间建立映射关系,帮助用户自动完成数据库模式匹配,此方法不要求用户对源数据库有任何了解。依照定义提出的基于行的模式匹配方法在分别处理实例表中各数据行后取各数据行模式匹配结果的交集。基于重用的优化方法将上一行数据的模式匹配结果作为处理下一行数据的输入,减小数据搜索范围和生成的查询计划数量,提升执行效率。基于列的优化方法将整个实例表作为一个整体,以列为单位进行搜索和查询计划树的生成,实现了效率的大幅提升,能够很好地处理实例表中有多行数据的情况。实

验验证了本文提出的优化方法的高效性。未来研究仍有如下问题值得思考:在很多情况下,用户提供的样例表不一定完全准确无误,如何处理这些错误数据是一个巨大的挑战。此外,在模式匹配中处理经过加工或者格式转换的数据也是一个非常值得研究的部分。

参 考 文 献

- [1] Kolaitis P. Schema mappings , data exchange , and metadata management [C]//Proceedings of ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems 2005.
- [2] Popa L , Velegrakis Y , Fagin R. Translating web data [C]//International Conference on Very Large Data Bases. ACM , 2002.
- [3] Drumm C , Schmitt M , Do H , et al. Quickmig: Automatic schema matching for data migration projects [C]//Proceedings of the 16th Conference on Information and Knowledge Management 2007.
- [4] Melnik S , Garcia-Molina H , Rahm E. Similarity flooding: Aversatile graph matching algorithm and its application to schema matching [C]//Proceedings of the 18th International Conference on Data Engineering 2002.
- [5] Berlin J , Motro A. Database schema matching using machine learning with feature selection [C]//Proceedings of the 14th International Conference on Advanced Information Systems Engineering. Springer 2006.
- [6] Dhamankar R , Lee Y , Doan A , et al. iMAP: Discovering complex mappings between database schemas [C]//Proceedings of ACM SIGMOD International Conference on Management of Data 2004.
- [7] Doan A , Domingos P , Halevy A. Reconciling schemas of disparate data sources: A machine-learning approach [J]. ACM SIGMOD Record 2001 30(2) : 509 - 520.
- [8] Kalashnikov D , Lakshmanan L , Srivastava D. FastQRE: Fast query reverse engineering [C]//Proceedings of International Conference on Management of Data 2018.
- [9] Shen Y Y , Chakrabarti K , Chaudhuri S , et al. Discovering queries based on example tuples [C]//Proceedings of ACM SIGMOD International Conference on Management of Data , 2014.
- [10] Zhang M H , Elmeleegy H , Procopiuc C M , et al. Reverse engineering complex join queries [C]//Proceedings of ACM SIGMOD International Conference on Management of Data , 2013.
- [11] Buneman P , Tan W C. Data provenance: What next? [J] ACM SIGMOD Record 2019 47(3) : 5 - 16.
- [12] Hristidis V , Papakonstantinou Y. Discover: Keyword search in relational databases [C]//Proceedings of the 28th International Conference on Very Large Databases 2002.