

融合区域特征和表格线识别的医学化验单布局识别方法研究

张泽毅^{1,2}, 傅湘玲^{1,2*}, 郭辰仪³

¹北京邮电大学计算机学院(国家示范性软件学院), 北京

²北京邮电大学可信分布式计算与服务教育部重点实验室, 北京

³清华大学电子工程系, 北京

收稿日期: 2021年12月14日; 录用日期: 2022年1月11日; 发布日期: 2022年1月18日

摘要

为解决医学化验单版式繁多, 通用文字识别系统输出正确率低, 识别结果黏连严重的问题, 在医学化验单文字识别全过程中, 借鉴文档版面分析的思路, 提出医学化验单进行布局识别的研究, 以提高文字识别的输出正确率。针对医学化验单布局上属于少线表和其呈明显区域分布的特点, 本文提出融合区域特征和表格线识别的医学化验单布局识别算法。首先, 利用基于Unet的表格线提取网络, 提前化验单图像的表格线。其次, 引入Mask R-CNN区域特征提取网络, 将表格线特征与原始化验单图像一同作为网络的输入。最后, 实现对上述区域特征与表格线特征的融合, 对待识别区域之间的关系进行建模, 并生成最终的精确坐标和语义标签。实验表明, 本算法能够较为明显提高医学化验单布局识别的准确度。

关键词

医学化验单, 文字识别, 布局识别, 特征融合

Research on Medical Laboratory Sheet Layout Recognition Method Combining Regional Features and Form Line Recognition

Zeyi Zhang^{1,2}, Xiangling Fu^{1,2*}, Chenyi Guo³

¹School of Computer Science (National Pilot Software Engineering School), Beijing University of Posts and Telecommunications, Beijing

²Key Laboratory of Trustworthy Distributed Computing and Service (BUPT), Ministry of Education, Beijing

*通讯作者。

文章引用: 张泽毅, 傅湘玲, 郭辰仪. 融合区域特征和表格线识别的医学化验单布局识别方法研究[J]. 计算机科学与应用, 2022, 12(1): 63-71. DOI: 10.12677/csa.2022.121008

Abstract

In order to solve the problem that the general character recognition system has a low output accuracy rate due to the large number of formats of medical test sheets, and the recognition results are severely stuck. In the whole process of text recognition of medical test sheets, drawing lessons from the idea of document layout analysis, research on the layout recognition of medical test sheets is proposed to improve the output accuracy of text recognition. Aiming at the characteristics of a few-line tables and obvious regional distribution in the layout of medical test sheets, this paper proposes a medical test sheet layout recognition algorithm that integrates regional features and table line recognition. First, use the Unet-based table line extraction network to test the table line of the single image in advance. Second, the Mask R-CNN regional feature extraction network is introduced, and the table line features and the original test sheet image are used as the input of the network. Finally, the fusion of the above-mentioned regional features and table line features is realized, the relationship between the regions to be recognized is modeled, and the final precise coordinates and semantic labels are generated. Experiments show that this algorithm can significantly improve the accuracy of medical test sheet layout recognition.

Keywords

Medical Laboratory Sheet, Text Recognition, Layout Recognition, Feature Fusion

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着社会的发展和科学技术的进步，人们愈发关注自己的身体健康情况，但随之面临的是困扰已久的“看病难”问题。在整个就诊流程中，化验通常是整个流程最重要的一环，由此产生的医学化验单最能反映患者的身体健康情况。当前，医疗体系中信息系统种类繁多，虽然电子病历已有一定发展，但各医院之间的电子病历数据互不相同，给医患均带来了一定的负担。高效、准确的解读化验单，并以合理的方式去管理化验单是整个医疗行业所面临的重大问题。

同时，针对患者转院转诊、报销等一系列问题，电子病历也不能很好的解决上述问题。因此，将纸质化验单上的数据经识别转换为文字信息，并利用电子档案等方式对信息进行存储是解决上述痛点问题的最佳方式。将图片信息数字化，既降低了患者就诊、报销时的负担，也能够提高医生问诊过程的效率，进而解决“看病难”的问题。

在医学化验单 OCR 识别的全过程中，布局识别是其中的关键步骤之一。通常来说，医学化验单由表格辅助线可以分为三大区域：医院信息区、患者信息区和化验项区。上述三大区域，每部分所包含的信息不同，且每部分可以独立看作少线表。因此，在 OCR 识别过程中，首先需要对化验单进行布局识别，将不同的区域切分，并利用相应的方法进行后续识别工作。布局识别的目标是识别非结构化医学化验单

中不同信息的区域,并识别每个区域的类别。由于医学化验单布局的多样性和复杂性,此任务难度较大。

目前,针对医学化验单布局识别问题,尚未提出相关研究方法。与本任务相似任务主要是文档的布局识别方法研究,针对文档布局识别任务,有部分学者提出了一些解决此任务的深度学习模型。已提出的研究一般从任务的视觉特征出发,采用自顶向下的策略将布局识别任务定义为目标检测或图像分割任务,但此类模型的实验结果较差。

针对医学化验单是少线表的特点,可以引入表格线关系作为分析特征。传统的表格线检测方法主要使用手工设计的特征,如利用隐式马尔科夫模型进行检测。这些特征对于布局变化不具有鲁棒性,尤其是对于版式繁多的医学化验单更加不适用。基于深度学习的表格线检测方法主要是基于 VGG 和 GNN 的方法,相较于传统方法,他们对于样式变换的鲁棒性更强,识别准确度更高。但单一的表格线特征不足以完成医学化验单布局识别任务。

为解决上述问题,在本文中,我们采用融合区域特征和辅助表格线特征提取的方法,提出了用于医学化验单布局识别的模型。首先,医学化验单以图像的形式作为模型的输入;其次,采用基于 VGG 的 Unet 模型提取医学化验单图像的表格线特征,并将其输出与原始医学化验单图像一同送入基于 Mask R-CNN 的深度学习网络提取图像的区域特征,实现上述特征的融合,对待识别区域之间的关系进行建模,并生成最终的精确坐标和语义标签。

本文的工作主要有三点贡献:

- 1) 提出了融合区域特征和辅助表格线的布局识别框架,是一种全新的医学化验单布局识别办法;
- 2) 提出了一个全新的多布局的医学化验单数据集;
- 3) 与先前的模型对比,本文提出的方法在识别准确率上有较大提升。

2. 相关工作

针对医学化验单的布局识别的相关研究较少,与本任务相关的工作主要有文档布局识别任务和图像分类任务。

2.1. 文档布局识别

已提出的文档布局识别研究根据利用特征的数目,可以分为利用一维特征模型和多维特征模型两大类。有研究提出将视觉特征与语义特征相结合,首先,通过卷积神经网络提取视觉特征,然后通过文本嵌入图引入语义特征;然后,使用检测或分段模型(例如 Mask RCNN)生成布局目标候选区域。虽然基于视觉特征和语义特征的方法能够较好地捕获空间信息,但仍然有两大方面的局限性:第一,语义信息有限。不同粒度的语义所表达的信息不同,以单一粒度嵌入的文本对于识别各部分元素的贡献是有限的。第二,缺乏对文档各区域之间关系的建模。文本图像中不同区域之间有较为密切的关联关系,正确对其关系进行建模可以显著提高布局分析的准确度。

一维特征模型主要从视觉特征或者语义特征进行文档布局识别方法的研究[1] [2] [3]。利用图像视觉建模的研究主要使用了基于 CNN 的模型[4],主要完成了对图像中的文本段落[5] [6],标题行[5]和文档中表格[7]的识别任务。同时,利用语义特征建模的研究主要利用文档的语义信息,结合 NLP 相关任务,来解决文档布局识别的任务[8] [9]。但是,上述两种研究都仅利用了单一模态的特征,任务效果也与特征工程有较强的关联性,因此模型的鲁棒性较差,对于版式变化繁多的医学化验单并不适用。

多维特征模型将视觉特征和语义特征相融合,进行文档布局识别任务。根据主要依赖的方法,可以分为两大类:基于 NLP 的方法和基于 CV 的方法。在基于 NLP 的方法中, PubLayNet [10]是文档图像布局识别的大型数据集,包含超过 36 万个文档图像,用多边形边框分割标注文档各区域布局情况。MMPAN

[11]是主要依赖与 NLP 模型进行文档布局识别任务的框架之一,但上述方法表现出利用视觉信息不足的缺点,导致其模型效果欠佳。基于 CV 的模型主要利用了基于目标检测或图像分割的方法, MFCN [12]在上述方法的基础上,引入了文本的语义信息,并在网络的最后一层引入了字符级的语义信息。虽然上述方法效果较好,但也有较为明显的缺陷:使用的语义有限,模型融合的策略较为简单,各模块之间相互依赖较少。

2.2. 图像分类

图像分类任务是指利用某种深度学习模型。在输入图像后,图像分类器模型根据图像输出对应的分类标签。Alex Krizhevsky 在 2012 年 ILSVRC 提出的 CNN 模型取得了历史性的突破,效果大幅度超越传统方法,获得了 ILSVRC2012 冠军,该模型被称作 AlexNet [13]。这也是首次将深度学习用于大规模图像分类中。ResNet (Residual Network)是 2015 年 ImageNet 图像分类、图像物体定位和图像物体检测比赛的冠军。针对随着网络训练加深导致准确度下降的问题,ResNet 提出了残差学习方法来减轻训练深层网络的困难[14]。

3. 研究方法

3.1. 模型结构简介

我们提出融合区域特征和表格线识别的模型(Fusion of Regional Features and Table Line Recognition Model, FRFTLR),由三部分组成:区域特征卷积网络、表格线提取网络和特征融合网络(如图 1 所示)。首先,基于医学化验单是少线表的特征,使用表格线提取网络(第 3.2 节),提取化验单图像的表格线特征,其中以医学化验单图像作为输入;其次,将表格线提取网络的输出与原始医学化验单图像一起送入到区域特征卷积网络(第 3.3 节),实现特征融合的目的,并作为最后的输出,生成最终结果。

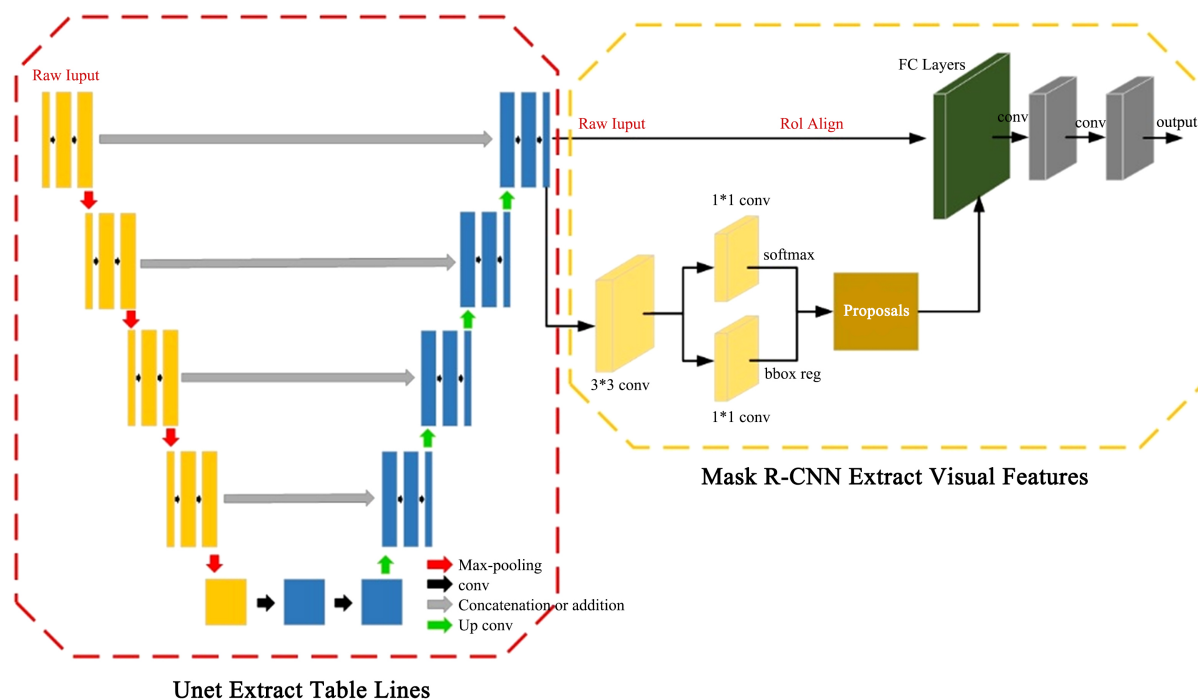


Figure 1. Model architecture overview
图 1. 模型整体结构图

3.2. 表格线提取 Unet 网络

Unet 属于全卷积网络(Fully Convolutional Networks, FCN)的一种变体。作者在设计 Unet 模型时, 主要目的是解决医学生物图像相关问题, 由于 Unet 实验效果确实很好, 研究者将其广泛应用与语义分割的任务上, 语义分割即将一张图像作为输入, 模型的输出为图像中分割物体的准确轮廓。常见的语义分割任务主要有表格线检测任务, 工业瑕疵检测分割, 医学图像检查分割任务等。Unet 是典型的编码-解码(Encoder-Decoder)结构, 主要结构简洁, 即上采样、下采样和跳跃连接三部分具体来说, 使用 Encoder 部分来实现特征提取, 每经过一次下采样, 通道数翻倍; Decoder 部分主要包含由一个 2×2 的上采样卷积层(ReLU)、Concatenation 以及 2 个 3×3 的卷积层(ReLU)反复构成; 在最后一层通过一个 1×1 卷积将通道数变成期望的类别数。Unet 网络结构如图 2 所示, 由于 Unet 的网络结构是对称的, 且形似英文字母 U 所以被称为 Unet。

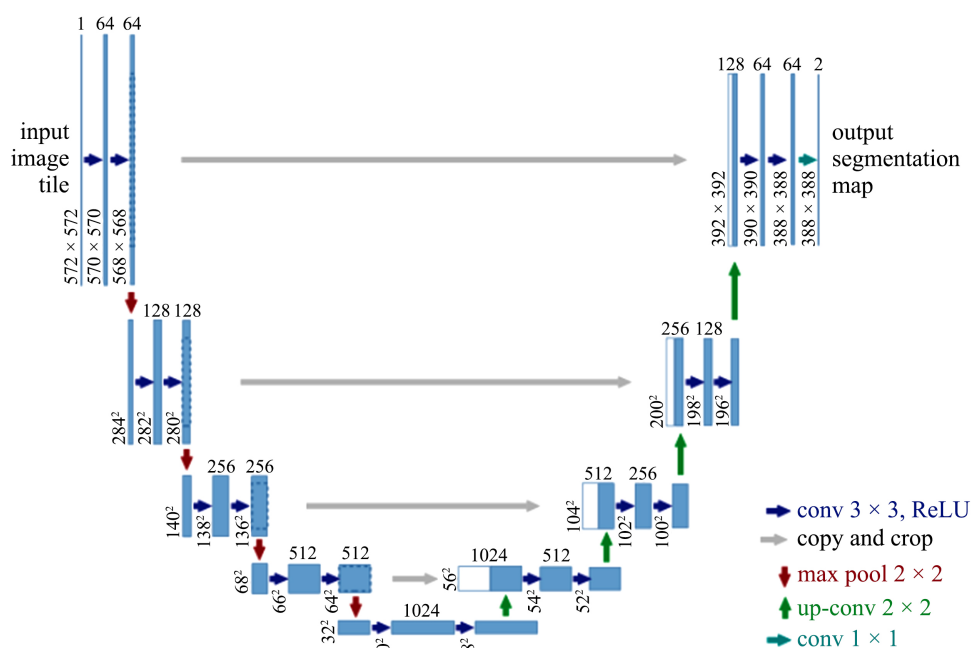


Figure 2. Table line extraction network Unet model structure

图 2. 表格线提取网络 Unet 模型结构

3.3. Mask R-CNN 区域特征提取网络

Mask R-CNN 目标检测算法是由何凯明等人对 Faster R-CNN 进行改进而提出的目标检测框架, 该算法可以同时完成目标检测与分割任务。Mask R-CNN 算法相较于 Faster R-CNN 算法, 选择基于 ResNet 网络替换后者的 VGG 网络。Mask R-CNN 实现目标检测任务由两大步骤实现: 第一个步骤是使用基于 FPN 网络, 对输入图片利用神经网络进行处理, 进而生成对应的候选框集和; 第二个步骤是对候选框集的分类, 并实现边框和 mask 回归任务。

Mask R-CNN 在 COCO 等自然场景图像数据集上表现出了良好的分割效果, 但对于医学化验单或文本图像等相关数据尚无相关实验结论。本文研究的医学化验单布局识别是要将输入的医学化验单实现对医院信息区、患者信息区、化验项及结果区这三个类别实现实例分割。

在进行提取医学化验单区域特征研究时, 首先将医学化验单图像和前序网格线提取网络输出的特征

图一起输入到基于 ResNet101 的卷积神经网络中提取特征；其次将由 ResNet 网络生成的特征图(Feature Map)送入到区域生成网络(Region Proposal Network, RPN)中, 生成提取候选框；随后将上述网络输出的 Feature Map 和提取候选框一起输入到 RoI Align 层, 由 RoI Align 层得到的特征图进行两大分支的处理, 第一将基于全卷积网络(Fully Convolutional Networks, FCN)对输出进行 mask 预测, 并利用全连接层(Fully Connected Layers, FC)完成 Softmax 分类和 Bbox 边框回归任务, 最终完成医学化验单布局识别任务。基于 Mask R-CNN 构建的医学化验单布局识别算法的网络结构如图 3 所示。

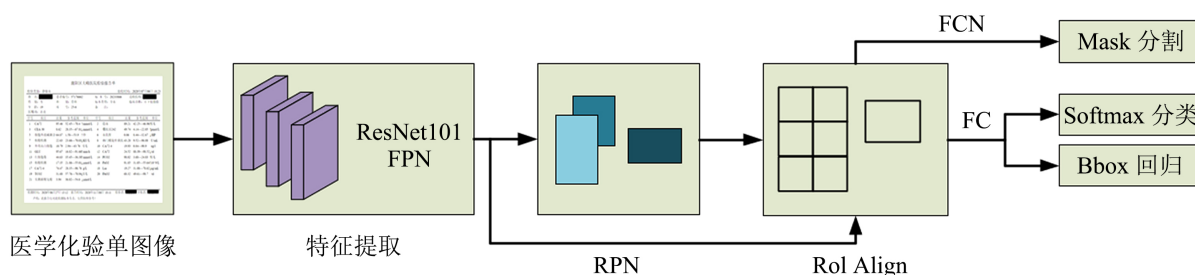


Figure 3. Regional feature extraction network Mask R-CNN model structure

图 3. 区域特征提取网络 Mask R-CNN 模型结构

4. 实验

4.1. 数据集

目前没有公开的医学化验单图像数据集来完成医学化验单布局识别研究任务。根据实际调研, 若采用人工标注海量化验单数据作为本文的训练集, 会面临费时费力且成本高的问题, 因此考虑采用人工合成数据的方法。基于以上思路构建了基于 SVG 的自标注医学化验单图像数据集。SVG 是一种图像文件格式, 它的英文全称为 Scalable Vector Graphics, 意思为可缩放的矢量图形。它是基于 XML (Extensible Markup Language), 由 World Wide Web Consortium (W3C) 联盟进行开发的, 严格来说应该是一种开放标准的矢量图形语言, 能够实现带有文字位置的自标注信息。在本模块中, 事先设定模板所需的行和列数目, 即可生成对应的数据; 此外, 考虑真实的医学化验单图片伴随倾斜、扭曲、畸变、变色等常见问题, 进而基于 OpenCV 实现对生成数据的加噪功能, 避免数据差异性的问题, 这部分数据将作为本文训练集的来源。

测试集采用从某三甲医院获取到的真实医学化验单图片, 使用 Labelme 图像标注工具对化验单进行区域标注, 并采用数据增广的方式, 进一步扩充测试集的样本数。

4.2. 实验参数

本文提出的模型是在 PyTorch 框架下实现的。实验参数设置如下, 在 RPN 中, 设定 7 个 anchor ratios, 分别为 0.02、0.05、0.1、0.2、0.5、1.0、2.0, 用来处理大小和比例不同的文档元素。由 SGD 优化器进行训练, 设置 $batchsize = 2$, $momentum = 0.9$, $weight-decay = 10^{-4}$ 。将初始学习率设置为 10^{-3} 。设置模型训练的 epoch 为 30。所有实验均在 Nvidia GeForce GTX 2080Ti GPU 上进行。

4.3. 实验结果

在医学化验单布局识别实验中, 为更有效的评估模型效果, 采用不同 IoU (Intersection over Union) 的平均准确率 mAP (mean Average Precision) 对模型的识别效果进行评价, 将医院信息区的 mAP 定义为 Hospital mAP、患者信息区的 mAP 定义为 Personal mAP、化验项及结果区的 mAP 定义为 Test item mAP。

定义基于 $\text{IoU} = 0.5$ 的 mAP 为 mAP_{50} ，基于 $\text{IoU} = 0.75$ 的 mAP 为 mAP_{75} ，定义 mAP 对医学化验单布局识别整体性能进行评价。 IoU 的计算如公式(1)所示， AP 值计算如公式(2)， mAP 值计算如公式(3)。

$$\text{IoU} = \frac{\text{area}(\text{P}) \cap \text{area}(\text{G})}{\text{area}(\text{P}) \cup \text{area}(\text{G})} \quad (1)$$

$$\text{AP} = \int_0^1 \text{P}(\text{R}) \text{dR} \quad (2)$$

$$\text{mAP} = \frac{\sum_{n=1}^N \text{AP}(n)}{N} \quad (3)$$

上式中， $\text{area}(\text{P})$ 指的是预测的 Bbox 集合或 Mask 集合； $\text{area}(\text{G})$ 指的是真实的 Bbox 集合或 Mask 集合； P 为精确率； R 为召回率； N 为训练集大小。其中精准率(P)和召回率(R)的计算如公式(4) (5)。

$$\text{P} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (4)$$

$$\text{R} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (5)$$

在上述公式中， TP 定义为模型能够识别正确的正样本数， TN 定义为模型能够正确识别的负样本数，定义 FP 为模型识别错误的正样本数，定义 FN 为模型错误识别的负样本数。本实验采用基于不同 IoU 阈值的 mAP 来评价医学化验单布局识别结果，对比四种不同的深度学习网络，得到的模型效果评价指标结果见表 1。

Table 1. Model comparison experiment results

表 1. 模型对比实验结果

模型	数据集	Hospital mAP	Personal mAP	Test item mAP	mAP_{50}	mAP_{75}	mAP
Faster RCNN	训练集	73.0	75.5	54.3	81.2	42.1	65.8
Mask RCNN	训练集	69.5	70.5	65.7	83.9	44.9	68.1
Unet	训练集	71.4	70.1	64.6	81.2	42.9	66.2
FRFTLR (Ours)	训练集	75.1	70.4	70.8	87.4	48.2	71.1
Faster RCNN	测试集	72.1	73.9	56.1	80.9	41.8	64.9
Mask RCNN	测试集	70.2	69.8	65.5	84.7	45.3	68.7
Unet	测试集	71.2	70.4	64.9	81.1	42.5	66.0
FRFTLR (Ours)	测试集	74.3	69.9	71.2	86.4	48.7	70.2

4.4. 消融实验

考虑本文提出的模型所利用的信息主要包含区域特征和表格线特征提取两大部分。为探究这两部分特征对于提高医学化验单布局分析的重要性，本节对上述两大特征在医学化验单数据集上进行了消融实验。实验结果如表 2 所示。

Table 2. Results of ablation experiments
表 2. 消融实验结果

特征	数据集	Hospital mAP	Personal mAP	Test item mAP	mAP ₅₀	mAP ₇₅	mAP
区域特征	训练集	69.5	70.5	65.7	83.9	44.9	68.1
表格线特征	训练集	73.4	69.7	67.4	85.1	45.3	69.7
特征融合	训练集	75.1	70.4	70.8	87.4	48.2	71.1
区域特征	测试集	70.2	69.8	65.5	84.7	45.3	68.7
表格线特征	测试集	73.5	69.1	66.9	85.3	45.7	70.0
特征融合	测试集	74.3	69.9	71.2	86.4	48.7	70.6

其中, 区域特征提取仅使用了 Mask R-CNN 模型, 表格线特征提取仅采用了 Unet 模型, 特征融合为完成的模型结构。实验结果表明, 表格线特征对于医学化验单布局识别的影响因素更大, 这与医学化验单属于少线表, 表格线有较强的引导功能相符合。特征融合实验也表明, 融合上述两种特征, 对于实验效果提升有较大帮助。

5. 结论

本文提出了融合区域特征和表格线特征用于医学化验单布局识别的模型。首先, 利用 Mask R-CNN 卷积网络提取医学化验单布局的区域特征; 其次, 针对医学化验单属于少线表, 线条具有强引导性的特征, 使用 Unet 网络提取医学化验单表格的表格线信息; 最后, 对上述两种信息进行特征融合, 并输出最终结果。与先前的模型对比, 本文提出模型的性能有显著提升; 消融实验也表明表格线对于布局识别有较为重要的作用。在未来, 将研究本文的模型, 并将其扩展到其他任务, 如文档布局识别等。

基金项目

北京市自然科学基金项目, 项目资助号: L192026。

参考文献

- [1] Li, X., Yin, F., Xue, T., Liu, L., Ogier, J. and Liu, C. (2019) Instance Aware Document Image Segmentation Using Label Pyramid Networks and Deep Watershed Transformation. 2019 *International Conference on Document Analysis and Recognition (ICDAR)*, Sydney, 20-25 September 2019, 514-519. <https://doi.org/10.1109/ICDAR.2019.00088>
- [2] Li, K., Wigington, C., Tensmeyer, C., Zhao, H., Barmpalios, N., Morariu, V.I., Manjunatha, V., Sun, T. and Fu, Y. (2020) Cross-Domain Document Object Detection: Bench-Mark Suite and Method. *Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 12915-12924. <https://doi.org/10.1109/CVPR42600.2020.01293>
- [3] Chen, K., Seuret, M., Liwicki, M., Hennebert, J. and Ingold, R. (2015) Page Segmentation of Historical Document Images with Convolutional Autoencoders. 2015 *13th International Conference on Document Analysis and Recognition (ICDAR)*, Tunis, 23-26 August 2015, 1011-1015. <https://doi.org/10.1109/ICDAR.2015.7333914>
- [4] Wick, C. and Puppe, F. (2018) Fully Convolutional Neural Networks for Page Segmentation of Historical Document Images. *DAS 2018: 13th IAPR International Workshop on Document Analysis Systems*, Vienna, 24-27 April 2018, 287-292. <https://doi.org/10.1109/DAS.2018.39>
- [5] Gatos, B., Louloudis, G. and Stamatopoulos, N. (2014) Segmentation of Historical Handwritten Documents into Text Zones and Text Lines. 2014 *14th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, Herso-nissos, 1-4 September 2014, 464-469. <https://doi.org/10.1109/ICFHR.2014.84>

-
- [6] Lee, J., Hayashi, H., Ohyama, W. and Uchida, S. (2019) Page Segmentation Using a Convolutional Neural Network with Trainable Co-Occurrence Features. 2019 *International Conference on Document Analysis and Recognition (ICDAR)*, Sydney, 20-25 September 2019, 1023-1028. <https://doi.org/10.1109/ICDAR.2019.00167>
- [7] He, D., Cohen, S., Price, B.L., Kifer, D. and Giles, C.L. (2017) Multi-Scale Multi-Task FCN for Semantic Page Segmentation and Table Detection. 2017 *14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, Kyoto, 9-15 November 2017, 254-261. <https://doi.org/10.1109/ICDAR.2017.50>
- [8] Krishnamoorthy, M.S., Nagy, G., Seth, S.C. and Viswanathan, M. (1993) Syntactic Segmentation and Labeling of Digitized Pages from Technical Journals. *The IEEE Transactions on Pattern Analysis and Machine Intelligence*, **15**, 737-747. <https://doi.org/10.1109/34.221173>
- [9] Shilman, M., Liang, P. and Viola, P.A. (2005) Learning Non-Generative Grammatical Models for Document Analysis. *10th IEEE International Conference on Computer Vision (ICCV 2005)*, Beijing, 17-20 October 2005, 962-969.
- [10] Zhong, X., Tang, J. and Yepes, A.J. (2020) PubLayNet: Largest Dataset Ever for Document Layout Analysis. 2019 *International Conference on Document Analysis and Recognition (ICDAR) IEEE*, Sydney, 20-25 September 2019, 1015-1022. <https://doi.org/10.1109/ICDAR.2019.00166>
- [11] Aggarwal, M., Sarkar, M., Gupta, H. and Krishnamurthy, B. (2020) Multi-Modal Association-Based Grouping for form Structure Extraction. 2020 *IEEE Winter Conference on Applications of Computer Vision (WACV)*, Snowmass Village, 1-5 March 2020, 2064-2073. <https://doi.org/10.1109/WACV45572.2020.9093376>
- [12] Yang, X., Yumer, E., Asente, P., Kralej, M., Kifer, D. and Giles, C.L. (2017) Learning to Extract Semantic Structure from Documents Using Multimodal Fully Convolutional Neural Networks. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 4342-4351. <https://doi.org/10.1109/CVPR.2017.462>
- [13] Technicolor, T., Related, S., Technicolor, T., *et al.* (2017) ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, **60**, 84-90.
- [14] He, K., Zhang, X., Ren, S., *et al.* (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>