

# 面向电力作业的工作票分割与 作业信息提取方法

丘 浩<sup>1</sup>, 张 炜<sup>1</sup>, 林翔宇<sup>1</sup>, 蒙 亮<sup>2</sup>, 彭博雅<sup>1</sup>

(1. 广西电网有限责任公司电力科学研究院, 广西 南宁 530023; 2. 广西电网有限责任公司, 广西 南宁 530023)

**摘 要:** 电力工作票中包含的电力作业关键信息是进行作业前风险评估的重要基础, 为了快速且准确地从实际工作票文档中提取所需信息, 提出一种电力工作票分割与作业信息提取方法。首先, 采取二值化、膨胀和腐蚀等操作从电力工作票图像中提取表格框线; 然后, 基于框线检测结果对工作票进行分割操作得到单元格图片, 再使用光学字符识别方法(OCR)检测各单元格内对应的作业信息; 最后, 基于正则匹配方法对识别结果进行结构化处理, 实现电力作业信息的有效提取与匹配。基于实际多任务工作票的实验测试表明, 所提方法能将提取到的电力作业信息按所属单元格进行合乎上下文语义的组合, 关键信息识别效果优于通用商业 OCR 软件。

**关 键 词:** 电力工作票; 表格分割; 光学字符识别

DOI:10.19781/j.issn.1673-9140.2022.06.023 中图分类号: TM721 文章编号: 1673-9140(2022)06-0198-08

## Power operation oriented approach of work ticket segmentation and work information extraction method

QIU Hao<sup>1</sup>, ZHANG Wei<sup>1</sup>, LIN Xiangyu<sup>1</sup>, MENG Liang<sup>2</sup>, PENG Boya<sup>1</sup>

(1. Electric Power Science Research Institute, Guangxi Power Grid Co., Ltd., Nanning 530023, China;  
2. Guangxi Power Grid Co., Ltd., Nanning 530023, China)

**Abstract:** The key information of system operation contained in the power work ticket is an important basis for pre-operation risk assessment. In order to quickly and accurately extract the required information from the actual work ticket document, a method of power work ticket segmentation and information extraction is proposed. First, the binarization, expansion, and corrosion operations are utilized to extract the table frame lines from the power work ticket image. Then the segmentation operations are performed to obtain cell images of the work ticket based on the frame lines detection results. The optical character recognition (OCR) is further leveraged to detect the corresponding information in the cell. Finally, the recognition result is structured based on the regular matching method to realize the effective extraction and matching of power operation information. Experimental tests based on actual work tickets show that the proposed method can combine the extracted power operation information in a contextual and semantic combination, and the key information recognition effect is better than general commercial OCR software.

**Key words:** power work ticket; table segmentation; optical character recognition

电力作业中票证作业制度是保障危险性较大作业(如动火、动电、动土、进入受限空间、高处作业及检查维修作业等)安全性的基本制度。工作票的主要内容是明确工作职责、评估危险作业时可能发生的的问题以及应采取的应对措施<sup>[1-2]</sup>。基于工作票中的作业信息,可以有效引导工作人员规避作业风险。

为了便于传输与打印,电力工作票文件通常以 PDF(portable document format)格式存储,需要人为将其内容录入作业前风险评估系统。由于工作票表格结构复杂、内容繁多,人工提取表格信息不仅耗费时间,而且容易出错。光学字符识(optical character recognition, OCR)技术是近年来得到快速发展的一种自动文本提取技术。1929 年德国科学家 Tausheck 首次提出了 OCR 的概念<sup>[1]</sup>。1960 年后得益于计算机技术的发展,OCR 技术逐渐得以实现。近 10 年来,深度学习算法的发展成熟显著地推动了 OCR 技术的普及应用。1998 年 LeCun 等人提出了 LeNet 卷积神经网络框架<sup>[3]</sup>,并首次将深度学习算法用于识别手写数字;为了提高卷积神经网络的特征提取能力,Alex 于 2012 年提出了 AlexNet 深度卷积神经网络<sup>[4]</sup>。文献[5]在 AlexNet 的基础之上提出了 VGGNet,增加网络深度的同时使用了更多连续的小尺寸卷积核,可以减小模型的参数量、提高易用性;文献[6]将字符识别看成序列识别问题,提出卷积循环神经网络(convolutional recurrent neural network, CRNN),成为了当下的主流 OCR 模型之一。OCR 技术的识别准确率能较好地满足一般应用,已在交通和贸易等领域得到广泛应用<sup>[7-8]</sup>。

与简单的字符识别不同,实际业务文档往往包含多个相对复杂的结构性模块,这些模块之间具有相对独立而又关联的复杂关系,从具有复杂结构的文档中提取有效信息仍然是一项具有挑战性的工作。除识别文字外,准确识别复杂表格结构以及单元格信息之间的匹配关系是在实际业务文档中应用 OCR 技术的瓶颈。

文献[9]提出了一种基于深度学习的表格结构识别方案,该方案可处理结构简单的表格文件,但失败案例也将随结构复杂度的提升而明显增多;文献[10]设计了一个票据字符识别平台,对市场交易中常见票据的采取票据版面分析后,由基于 YOLO-

v3 的目标检测算法提取定位票据中表格区域信息,再采取 OCR 引擎识别。该模型能泛化不同格式的票据图像,但由于模型参数量庞大,需要有海量的训练数据,对硬件性能的要求也较为苛刻。文献[11]提出使用语义识别的方法进行电力工作票的文字识别与语义分割,但 OCR 字符检测效果受表格框线影响较大并会导致识别精度低下、影响语义识别的效果。以上方法均基于“深度学习+OCR”的方法实现表格分割,但普遍存在模型训练量大、工作票分割效果受 OCR 检测准确率影响较大等问题。

本文利用电力作业工作票的复杂结构具有框线分割的特点,提出一种电力作业工作票分割与作业信息自动提取方法。该方法首先提取框线来分割单元格,然后使用 OCR 技术识别各单元格的工作票作业信息,最后基于正则匹配方法对识别结果进行结构化处理。与传统方法相比,该方法不需要额外的模型训练,运算量小,同时可有效避免表格框线对 OCR 检测精度的影响,提高电力工作票的识别效率。

## 1 基于 CRNN 的光学字符识别算法

光学字符识别是模式识别领域的重要分支在图像文档的信息提取中有广泛应用。传统 OCR 方法采取人工设计的特征,在复杂场景中表现出识别率低、泛化能力弱等缺点<sup>[12-13]</sup>。因此,对基于深度学习的 OCR 方法的研究成为主流。在基于深度学习框架的 OCR 技术中,人工神经网络主要用于提取字符特征以及根据特征进行分类的功能。人工神经网络的加入省去了人工设计特征和训练分类器的步骤,极大地简化了人工设计的过程。

在 CRNN 应用于字符识别技术之前,文字识别过程分为单字切割和分类任务两步。识别过程高度依赖于字符切分结果,误切分对识别准确率有突出影响。基于 CRNN 的字符识别模型将文字识别转化为序列学习问题,不再需要字符切割环节,在图像的序列识别上具有良好的表现。CRNN 由卷积层、循环层和转录层 3 个部分组成,其结构如图 1 所示。

CRNN 从输入图像提取文字信息的流程可以分为两部分:①特征序列提取,即在 CRNN 底部,由连续分布的多层卷积神经网络自动从输入图像中提取特征,得到特征图,然后转化成特征序列;②特征

序列识别,即循环神经网络对特征序列中的每个特征向量进行学习后输出预测标签分布,由转录层从预测的标签分布中找到概率最高的标签序列,之后转换成最终的识别结果。

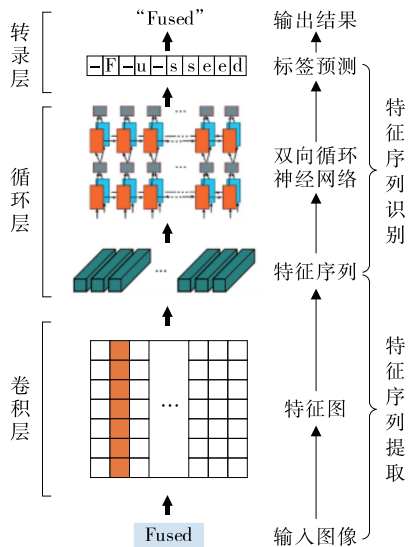


图 1 CRNN 结构

Figure 1 The network architecture of CRNN

1.1 特征序列提取

CRNN 的卷积层采用卷积计算层和最大池化层的组合。卷积计算层的核心是一个二维滤波器(卷积核)。卷积操作如图 2 所示,卷积核中每个元素都代表一个权重值,通过滑动滤波的方式,将输入图像与卷积核中的对应元素依次相乘并求和,进而确定特征图中的元素值<sup>[14]</sup>。

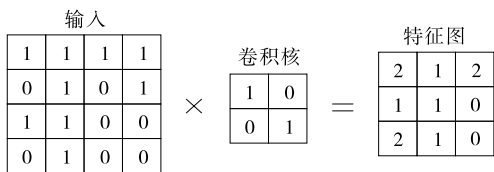


图 2 卷积操作

Figure 2 Convolution operation

池化层分布在卷积计算层之间,常用的池化方式有平均池化和最大池化,平均池化选取区域均值作为该区域值,而最大池化则选取区域最大值作为该区域值。池化操作可以去除冗余信息,压缩数据和参数的量,对邻域内特征进行聚合<sup>[15]</sup>。

通过卷积与池化操作会将输入图像所包含的原始数据映射到特征空间,形成特征图。从卷积层输出的特征图无法直接送入循环层进行训练,需要转

化为特征序列。因此,在 CRNN 的结构中加入了一个转换层,在该层将特征图中的某列像素顺序连接就形成了一个特征向量,所有的特征向量就构成一个序列。

1.2 特征序列识别

区别于传统的目标检测方法,字符识别的目标通常是一个序列,将字符识别看成序列识别问题,需要引入循环神经网络(recurrent neural network, RNN),RNN 以序列数据为输入,具有一定记忆能力,能够存储短时间内输入数据的信息,这是因为在 RNN 中,当前时刻的输出  $h_t$  由当前输入  $x_t$  和前一时刻输出  $h_{t-1}$  共同决定,即

$$h_t = f(Ux_t + Wh_{t-1} + b) \quad (1)$$

其中,  $U$ 、 $W$  是权重系数,  $b$  为偏差量,这些值都是通过训练得到的。

在传统的 RNN 中,时间累积会最终导致梯度消失现象。为克服该问题,有学者提出了一种改进后的 RNN 结构—长短期记忆网络(long short-term memory, LSTM)<sup>[16]</sup>,LSTM 单元引入门控机制,通过遗忘门、输入门和输出门来控制细胞状态,实现新信息的加入和无用信息的遗忘,其结构如图 3 所示,具体步骤如下。

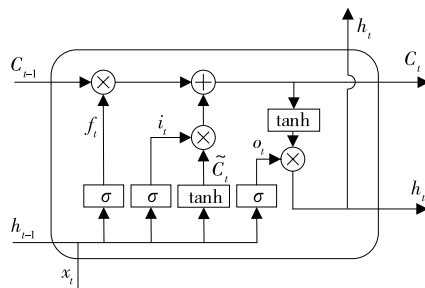


图 3 LSTM 单元结构

Figure 3 The structure of LSTM

1)由遗忘门决定细胞状态,需要遗忘上一个状态  $h_{t-1}$  的部分信息,其表达式为

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2)$$

2)经过输入门决定细胞状态添加哪些新的信息。输入门的输出可以表示为

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (3)$$

3)细胞的临时状态  $\tilde{C}_t$  以  $\tanh$  函数作为激活函数,即

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (4)$$

然后细胞状态由  $\tilde{C}_{i-1}$  更新为  $\tilde{C}_i$  :

$$C_i = f_i \cdot C_{i-1} + i_i \cdot \tilde{C}_i \quad (5)$$

4)由输出门输出更新后的状态  $o_i$ ,并将其与处理后的细胞状态相乘,获得序列的识别结果  $h_i$ 。

$$o_i = \sigma(W_o[h_{i-1}, x_i] + b_o) \quad (6)$$

$$h_i = \tanh C_i \quad (7)$$

在 LSTM 单元中,信息的传递是一个单向过程,因此,单个 LSTM 单元只能利用文字序列中某一个方向的信息。将一个 LSTM 向前另一个向后组合,形成一个双向 LSTM,就能同时利用上下文信息。当 LSTM 输出的序列分布输入转录层后,通过去冗余操作,解码即可输出概率最大的文字识别结果。

## 2 电力作业工作票分割与作业信息提取方法

本文设计的电力作业工作票分割与作业信息提取方法具体工作流程如图 4 所示。该方法主要包括表格分割和作业信息提取与结构化两部分。工作票表格分割部分包括 3 个模块:工作票二值化处理、表格框线检测和单元格分割模块。作业信息提取与结构化部分包括:作业信息提取和信息结构化处理模块。

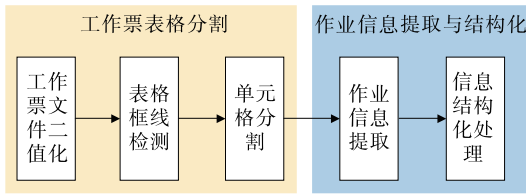


图 4 电力作业工作票分割与作业信息提取流程

Figure 4 The workflow of proposed method

### 2.1 工作票表格分割

#### 2.1.1 工作票二值化处理模块

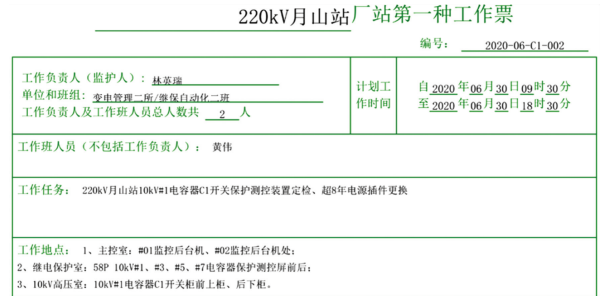
输入的工作票文件为彩色图像,为去除图像中的冗余信息、提高工作效率,需要对输入的工作票图像进行二值化处理,把大于某个临界灰度值的像素灰度设为灰度极大值,把小于这个值的像素灰度设为灰度极小值,其公式如下:

$$g(x, y) = \begin{cases} 0, & g(x, y) < T \\ 255, & g(x, y) > T \end{cases} \quad (8)$$

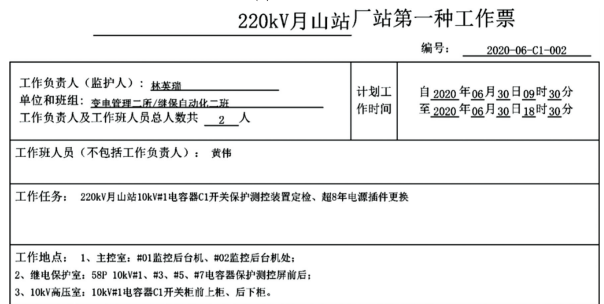
式中  $g(x, y)$  为该点的灰度值,范围为  $0 \sim 255$ ;  $T$  为临界灰度值(阈值),工作票文件存在的色系固定

且有限,因此阈值  $T$  可设为固定值。

对工作票进行二值化处理后的效果如图 5 所示。二值化仅去除对文字识别无用的部分颜色信息,不会改变图像中的作业信息与表格结构信息。



(a) 输入图像



(b) 二值化图像

图 5 工作票经二值化处理后的效果

Figure 5 Result of binarization of work ticket

#### 2.1.2 表格框线提取模块

表格框线提取实际上是将表格框线像素与非框线像素分离的过程,可以采取基于数学形态学的方法来实现。数学形态学中有 2 种最基本的运算:腐蚀和膨胀<sup>[17]</sup>。数学形态学的运算过程需要 2 个对象:一个是输入图像,记其代表的所有像素集合为  $A$ ;另一个是结构元素,记其代表的所有像素集合为  $B$ 。腐蚀和膨胀运算的过程如图 6 所示。

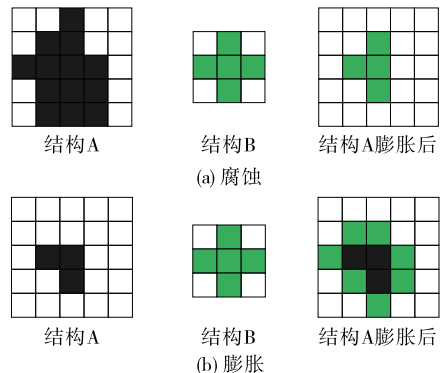


图 6 数学形态学基本运算

Figure 6 Basic operations of mathematical morphology

腐蚀是一种消除边界点,使边界向内收缩的过程。膨胀是腐蚀的对偶运算,是一种使边界向外扩展的过程。形态学运算可以看成是在目标图像中检测由结构元素所描述的特征形状的过程,这一特性可用于检测图像中的表格框线。采用形态学运算的表格框线提取步骤如下:

1)使用大小为  $l_1 \times 1$  的水平结构和  $1 \times l_2$  的垂直结构元素,对表格区域进行水平和垂直方向的腐蚀,分别获得水平和垂直框线;

2)使用大小为  $l_1 \times 1$  的水平结构和  $1 \times l_2$  的垂直结构元素,对腐蚀过的图像进行膨胀,更好地恢复水平和垂直表格框线;

3)将前两步得到的框线在同一坐标系下相加,即得到完整的表格区域以及单元格顶点位置信息。

表格框线提取效果主要取决于结构元素的长度  $l_1, l_2$ 。若  $l_1, l_2$  太长,在步骤 1)中会腐蚀掉一些较短的表格框线,得到的表格结构不完整;若  $l_1, l_2$  太短,则会保留一些不必要的文字线段,对之后的识别造成干扰。因此,结构元素的长度应满足 2 个条件:小于最短表格线的长度和大于单个字符的尺寸。

使用本文提出的框线提取方法处理电力作业工作票的结果如图 7 所示,可以看出,2 种表格框线均被完整提取,而在水平表格框线图中存在部分非框线像素,这些像素来自于工作票文档中的文下划线,由于这些下划线不会与垂直表格框线相交,因此不会影响对表格结构的判断。

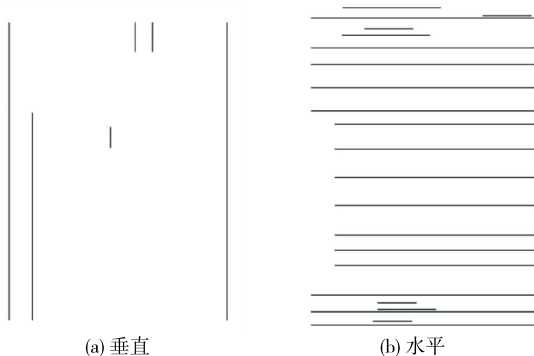


图 7 表格框线提取

Figure 7 Table frame lines extraction

### 2.1.3 单元格分割模块

该模块接收来自表格框线检测模块输出的表格框线和交点位置,然后对工作票表格进行单元格分割,其流程如图 8 所示。具体步骤如下:

1)从图片最上方开始,记录下第 1 条水平表格框线上所有点的位置,从第 1 个点开始,以它为一个单元格的左上角顶点,然后判断该点下方是否存在构成直线的连续像素;若存在,则该点为一个单元格的左上角顶点;

2)在同一水平位置上向右寻找该单元格的右上角顶点,找到右上角顶点后,判断该顶点下是否存在构成直线的连续像素;

3)在与左上角顶点同一垂直位置的直线上寻找该单元格的左下角顶点,最后一个顶点可以根据已知的 3 个顶点位置推断确定,只需验证在对应位置是否存在第 4 个顶点即可;

4)通过 4 个顶点的位置信息,可确定对应单元格所在区域;而后可将该单元格所在区域裁剪、保存并进行编号,保留其所在的行信息与列信息;

5)对处在同一水平位置的顶点重复步骤 1)~4)中的操作,可分割出行所有的单元格;

6)跳转到下一行重复上述所有操作,直至所有单元格分割完毕。

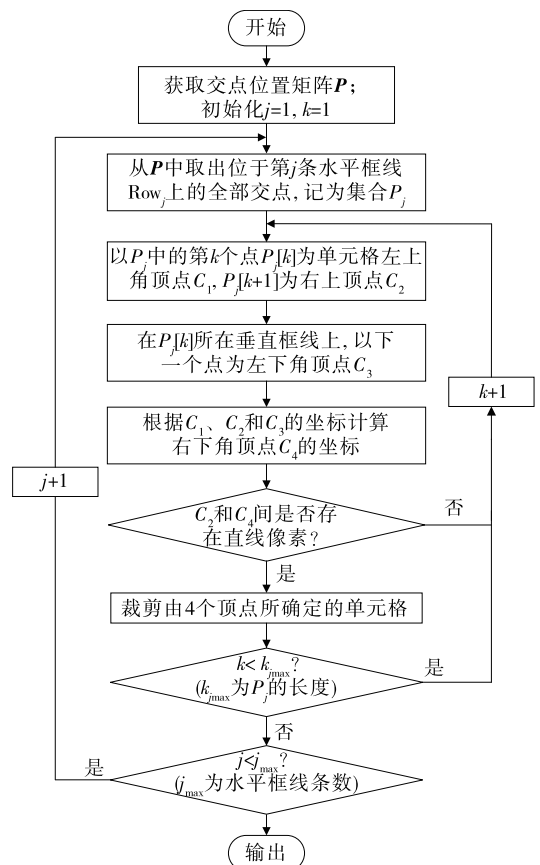


图 8 单元格分割流程

Figure 8 Flow chart of cell segmentation

单元格分割结果如图 9 所示。对单元格进行编号时应当保留单元格在原表格中的行列信息,统一

命名格式为“ $0i_0j_0k$ ”,表示该单元格在工作票文件的第  $i$  页,第  $j$  行,第  $k$  列。

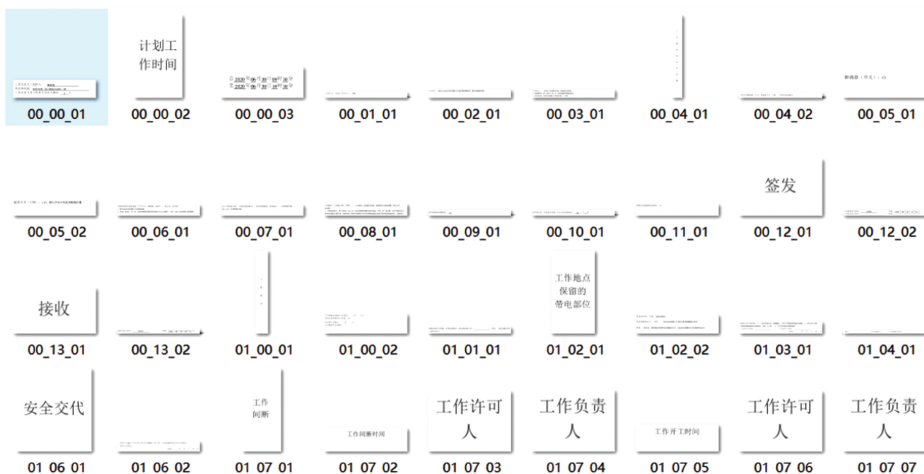


图 9 单元格分割结果

Figure 9 Result of cell segmentation

### 2.2 作业信息提取与结构化处理

分割单元格可将复杂表格的识别任务分解为多个单元格图像识别子任务。对所获得的单元格图像只需要通过 CRNN 模型就能准确提取其中的作业信息。进行本次电力作业风险评估只需要全部作业信息中的部分关键信息(如工作任务,工作地点,单位和班组等)。为此,需要对文字识别结果进行结构化处理,通过正则表达式检索和匹配符合所需模式(规则)的文本以提取所需信息。

## 3 实验验证

为评估所提方法的性能,训练一个卷积循环神经网络,并对 100 份实际的工作票文档进行信息自动提取。

### 3.1 训练过程

基于 MXNet 深度学习框架搭建 CRNN 模型,具体结构如表 1 所示。在特征提取阶段共包含 6 层卷积计算,均采用  $3 \times 3$  大小的卷积核,0~5 层的卷积神经网络滤波器数量分别为 64、128、256、512、512、512。循环层为 2 层各 256 个隐藏单元的双向 LSTM。转录层使用 CTC(connectionist temporal classifier)作为损失函数,对序列进行解码。表中  $k$  表示卷积核大小, $s$  表示卷积核滑动时的步长, $p$  表示可用于填充的零元素列数。

表 1 CRNN 网络参数

Table 1 Network parameters of CRNN

Layers	Configuration
Transcription(CTC)	—
Bidirectional-LSTM	# hidden units:256
Bidirectional-LSTM	# hidden units:256
Map-to-Sequence	—
Convolution	# maps:512, $k:3 \times 3, s:1, p:1$
MaxPooling	Window:1 $\times 2, s:2$
BatchNormalization	—
Convolution	# maps:512, $k:3 \times 3, s:1, p:1$
BatchNormalization	—
Convolution	# maps:512, $k:3 \times 3, s:1, p:1$
MaxPooling	Window:1 $\times 2, s:2$
Convolution	# maps:256, $k:3 \times 3, s:1, p:1$
MaxPooling	Window:2 $\times 2, s:2$
Convolution	# maps:128, $k:3 \times 3, s:1, p:1$
MaxPooling	Window:2 $\times 2, s:2$
Convolution	# maps:64, $k:3 \times 3, s:1, p:1$
Input Image	$W \times 32$ gray-scale image

从中文语料库抽取文字后,通过计算机合成文本图像样本作为数据集。该数据集共包括约 200 万张图片,涵盖了汉字、英文、数字和标点 4 大类共 3 870 个字符。按 9 : 1 将数据集分成训练集和测试集。在搭载 Intel Xeon Gold 6240 CPU(2.50 GHz)和 NVIDIA GeForce GTX 1650 显卡(4 G)的计算机上完成训练,训练准确率如图 10 所示。

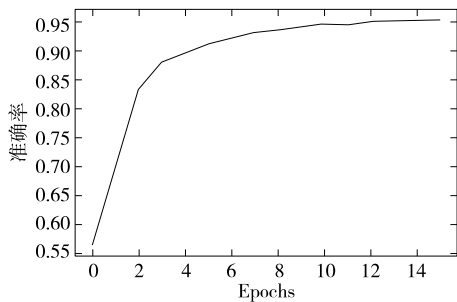


图 10 识别准确率

Figure 10 Training accuracy

### 3.2 实验结果与分析

分别采用本文所提方法与市面常见的商业文字识别软件对工作票文档进行信息提取,如图 11 所示。

工作负责人(监护人): 林英瑞 单位和班组: 变电管理二所/继保自动化二班 工作负责人及工作班人员总人数共 2 人	计划工作时间 自 2020 年 06 月 30 日 09 时 30 分 至 2020 年 06 月 30 日 18 时 30 分
工作班人员(不包括工作负责人): 黄伟	
工作任务: 220kV 月山站 10kV#1 电容器 C1 开关保护测控装置定检、超 8 年电源插件更换	

(a) 工作票局部

全文还原 json 返回

工作负责人(监护人): 林英瑞 计划工 自2020年06月30日09时30分 单位和班组: 变电管理二所/继保自动化二班 作时间 至 2020年06月30日18时30分 工作负责人及工作班人员总人数共 2 人 工作班人员(不包括工作负责人): 黄伟 工作任务: 220kV月山站10kV#1电容器C1开关保护测控装置定检、超8年电源插件更换 工作地点: 1、主控室: #01监控后台机、#02监控后台机处; 2、继电保护室: 58P10kV#1、#3、#5、#7电容器保护测控屏前后; 3、10kV高压室: 10kV#1

(b) 某商业软件文件识别结果

result\_save - 记事本  
 文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)  
 工作负责人(监护人): 林英瑞  
 单位和班组: 变电管理二所/继保自动化二班  
 工作负责人及工作班人员总人数共2人  
 计划工作时间  
 自2020年06月30日09时30分至2020年06月30日18时30分  
 工作班人员(不包括工作负责人): 黄伟  
 工作任务: 220kV月山站10kV#1电容器C1开关保护测控装置定检、超8年电源插件更换  
 2、继电保护室: 58P10kV#1、#3、#5、#7电容器保护测控屏前后;  
 B、10kV高压室: 10kV#1电容器C1开关柜前上柜、后下柜。

(c) 本文所提方法文字识别结果

图 11 文字识别结果

Figure 11 Results of character recognition

在图 11(b)中,当商业软件对标题“计划工作时间”及其对应字段值“自 2020 年 06 月 30 日 09 时 30 分至 2020 年 06 月 30 日 18 时 30 分”进行识别时,由于没有正确处理表格结构出现了识别错误,标题与字段值无法保持正确的匹配关系,导致后续数据结构化时提取“工作时间”项信息失败。本文所提方法采用单元格分割算法将复杂表格文档的文字识别

拆分为多个单元格图像的文字识别,识别结果正确保留了原表格中各个标题及对应字段值的逻辑关系。对该结果进行结构化处理得到的效果如图 12 所示。

工作任务	工作负责人	单位和班组	工作时间	工作班人员	工作地点	应拉开的断路器	应拉开的隔离开关	应拉切的相关画流源、低压及二次回路	应合上的接地刀闸、安装的接地线、应设绝缘挡板	是否需线路对侧接地	
220kV月山站10kV#1电容器C1开关保护测控装置定检、超8年电源插件更换	林英瑞	变电管理二所/继保自动化二班	2020年06月30日09时30分至2020年06月30日18时30分	黄伟	1、主控室: #01监控后台机、#02监控后台机处; 2、继电保护室: 58P10kV#1、#3、#5、#7电容器保护测控屏位置	C14、将C1开关小车拉至检修位置	C1	1、断开10kV#1电容器C1开关控制电源; 2、在58P10kV#1、#3、#5、#7电容器...	1、断开10kV#1电容器C1开关控制电源; 2、在58P10kV#1、#3、#5、#7电容器...	C138、C147 (共2把接地刀闸)	否

图 12 数据结构化处理

Figure 12 Data structure processing

对 100 份实际工作票文档进行信息提取的实验统计结果表明,本文所提方法文字识别准确率达 95.18%,略低于文字识别模型在测试集上的识别准确率。单次信息提取过程平均耗时为 22.41 s。

## 4 结语

本文着眼于实际应用场景,提出了一种电力作业工作票分割与作业信息提取方法,能够从工作票文档复杂的表格结构中快速且准确的提取电力作业的关键信息,提升了电力作业风险评估效率。在框线提取过程中,由于有少量下划线像素保留,可能会造成文字识别错误,因此,本文的研究工作还可以进一步完善。之后的研究也将关注自然语言处理领域<sup>[18]</sup>,对识别结果进行查错和纠错,以进一步提升输出的工作票信息的正确程度。

### 参考文献:

[1] 惠军.浅谈电力安全生产中的“两票”管理制度[J].湖北水力发电,2007(1):54-56.  
 HU Jun."Two-Card" management system in electrical production[J].Hubei Water Power,2007(1):54-56.  
 [2] 孙小龙,郝玉凤,樊宏亮,等.新能源场站“两票”管理及移动运维平台应用[J].电力安全技术,2020,22(4):49-53.  
 SUN Xiaolong, HAO Yufeng, FAN Hongliang, et al."Two Tickets" management of new energy station and

- application of mobile operation and maintenance platform[J]. Electric Safety Technology, 2020, 22(4): 49-53.
- [3] LECUN Y, CORTES C, BURGESS J C. The MNIST database of handwritten digits[EB/OL]. <http://yann.lecun.com/exdb/mnist>, 1998.
- [4] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [5] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]// 3rd International Conference on Learning Representations (ICLR), San Diego, CA, US, 2015.
- [6] SHI B G, BAI X, YAO C. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(11): 2298-2304.
- [7] 史建伟. 基于深度学习的车牌识别系统的设计与实现[D]. 南京: 南京邮电大学, 2020.
- [8] 张振宇, 姜贺云, 樊明宇. 一种面向银行票据文字自动化识别的高效人工智能方法[J]. 温州大学学报(自然科学版), 2020, 41(3): 47-56.
- ZHANG Zhenyu, JIANG Heyun, FAN Mingyu. An efficient artificial intelligence method for automatic recognition of bank bill text[J]. Journal of Wenzhou University (Natural Science Edition), 2020, 41(3): 47-56.
- [9] SCHREIBER S, AGNE S, WOLF I, et al. DeepDeSRT: deep learning for detection and structure recognition of tables in document images[C]// 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 2017.
- [10] 陈科峻. 票据字符识别平台研究与实现[D]. 长春: 中国科学院大学(中国科学院长春光学精密机械与物理研究所), 2020.
- [11] 徐楠楠, 王东风, 韩璞. 基于正向最大匹配算法的电力两票安全识别[J]. 计算机仿真, 2014, 31(1): 145-148+355.
- XU Nannan, WANG Dongfeng, HAN Pu. Recognition of two-tickets system in power station based on forward maximum matching algorithm[J]. Computer Simulation, 2014, 31(1): 145-148+355.
- [12] 李新良. 基于模板匹配法的字符识别算法研究[J]. 计算机技术与自动化, 2012, 31(2): 90-93.
- LI Xinliang. Research on character recognition algorithm based on template matching method[J]. Computing Technology and Automation, 2012, 31(2): 90-93.
- [13] 刘莉. 光学字符识别系统中关键算法的设计与研究[D]. 成都: 电子科技大学, 2011.
- [14] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
- ZHOU Feiyan, JIN Linpeng, DONG Jun. Review of convolutional neural network[J]. Chinese Journal of Computers, 2017, 40(6): 1229-1251.
- [15] BOUREAU Y L, PONCE J, LECUN Y. A theoretical analysis of feature pooling in visual recognition[C]// 27th international conference on machine learning (ICML), Haifa, Israel, 2010.
- [16] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [17] 周余涵, 张哲, 韩杰祥, 等. 基于开-闭滤波和多分辨率形态梯度的直流线路纵联保护新原理[J]. 中国电机工程学报, 2020, 40(6): 1897-1906.
- ZHOU Yuhuan, ZHANG Zhe, HAN Jiexiang, et al. A novel pilot protection based on open-close filtering and multi-resolution morphological gradient operators for HVDC transmission line[J]. Proceedings of the CSEE, 2020, 40(6): 1897-1906.
- [18] 奚雪峰, 周国栋. 面向自然语言处理的深度学习研究[J]. 自动化学报, 2016, 42(10): 1445-1465.
- XI Xuefeng, ZHOU Guodong. A survey on deep learning for natural language processing[J]. Acta Automatica Sinica, 2016, 42(10): 1445-1465.