

DOI: 10.19783/j.cnki.pspc.211176

基于贝塞尔曲线网络的电表用电信息识别算法

孙福明¹, 高严¹, 许蕊¹, 李明渊¹, 魏晓鸣²

(1. 大连民族大学信息与通信工程学院, 辽宁 大连 116600; 2. 大连民族大学网络中心, 辽宁 大连 116600)

摘要: 利用计算机视觉技术来快速、准确地获得用户的用电信息对于电力部门具有重要意义。针对传统算法精度低、速度慢等问题, 提出了一种基于自适应贝塞尔曲线网络的用电信息识别算法。该框架集检测、识别于一体, 实现了端到端的文本定位和预测。检测端结合了特征金字塔网络和残差网络, 对输入图像进行特征提取, 并通过四个控制点生成贝塞尔曲线, 能更好地拟合文本框。识别端采用了卷积循环神经网络, 引入门控循环单元替代长短期记忆网络, 再结合注意力机制对文本进行识别。最后在数据集上开展五组消融实验, 进行性能对比和评估分析。实验结果显示, 该算法识别精度高达99.08%, 且推理速度快, 可被用于用电信息检测与识别的实际应用中。
关键词: 机器视觉; 自适应贝塞尔曲线网络; 电表; 端到端; 注意力机制

Recognition algorithm of electricity consumption information of an electric energy meter based on an adaptive Bezier curve network

SUN Fuming¹, GAO Yan¹, XU Rui¹, LI Mingyuan¹, WEI Xiaoming²

(1. School of Information and Communication Engineering, Dalian Minzu University, Dalian 116600, China;
2. Network Center, Dalian Minzu University, Dalian 116600, China)

Abstract: The power sector can collect user electricity information quickly and accurately using computer vision technology. Given the problems of low precision and slow speed of traditional algorithms, a recognition algorithm of electricity consumption information based on an adaptive Bezier curve network is proposed. The framework integrates detection and recognition, and realizes end-to-end text location and prediction. The detection end extracts the feature from the input image by combining a feature pyramid and residual networks, and generates a Bezier curve through four control points, which can better fit the text box. A convolutional recurrent neural network is adopted at the recognition end, and the gate recurrent unit is introduced to replace the long short-term memory network, and is then combined with the attention mechanism to recognize the text. Finally, five ablation experiments are carried out on the data set for performance comparison and evaluation analysis. The results show that the recognition accuracy of the algorithm is up to 99.08%, and the reasoning speed is fast. It can be used in the practical application of electricity consumption information detection and recognition.

This work is supported by the National Natural Science Foundation of China (No. 61976042).

Key words: machine vision; adaptive Bezier curve network; electric energy meter; end-to-end; attention mechanism

0 引言

电表作为一种用电计量工具, 在日常生活中发挥着重要作用。采用人工抄表方法来获取传统的没有物联网功能的电表, 严重制约了电力企业的信息化服务水平。因此如何快速、准确获得此类用户电

表的用电信息对于电力企业具有重要意义。

计算机技术的发展推动了电力系统信息化, 如智能抄表系统^[1]、储能系统^[2]、竞价交易^[3]和电网巡检^[4]等。基于计算机技术的图像信息检测与识别方法, 可分为两类: 基于人工设计特征的方法和基于深度学习特征的方法。基于人工设计特征的方法是先通过手工方式获取图像特征, 再对文本信息进行检测和识别, 如文献[5-6]采用二值化算法和形态学操作, 结合电表示数的几何特征定位, 根据数字的

基金项目: 国家自然科学基金项目资助(61976042); 兴辽英才计划项目资助(XLYC2007023)

笔画特征进行识别。在光照充足时,这类方法具有较高的识别精度;而当环境不理想时,它们的识别精度会变差。因此,基于手工设计特征的方法不但消耗了大量的人力和时间,而且训练时间也比较长,实际检测效果也并不理想。

基于深度学习的方法,如文献[7]以YOLO算法为基础使用轻量级的YOLOv2-Tiny网络快速定位电表示数区域,然后分别使用CR-NET^[8]和卷积循环神经网络^[9]两种识别方法进行数字识别,它的识别准确率达到89.5%;文献[10]在此基础上将网络升级,提出了轻量级的YOLOv3-Tiny定位和识别网络模型,计数器识别准确率达到92.13%;文献[11]同样以YOLO为基础,采用多阈值软切分方法进行识别,识别准确率达到93.61%;文献[12]提出了ABCNet(Adaptive Bezier Curve Network),该网络集检测与识别于一体,检测端结合了残差网络^[13]和特征金字塔网络^[14]两种思想,使用贝塞尔曲线拟合文本框;识别端则采用了基于卷积循环神经网络(Convolutional Recurrent Neural Network, CRNN)的轻量级识别算法,准确率达到93.56%。显然,相比较于人工设计特征方法^[5-6],深度学习^[7-12]优势明显,显著提升了检测速度和检测精度。

本文针对电表用电信息开展智能检测与识别研究,所提出的网络框架检测准确率达到99.86%,识别准确率为99.08%。本文主要贡献如下。

1) 在文本检测方面,改进了特征提取网络,获得了更丰富的全局语义。在上采样过程中,先将前两级提取的特征作为输入,使高层的语义信息映射到每一级,与位置信息更好的融合,从而捕获丰富的多尺度上下文信息。然后将四个特征图以合并多个数组的方式连接起来,合成一个高级特征,这样不仅改善了梯度消失和网络退化问题,而且特征增强网络使该算法具备检测多尺度特征的能力。

2) 在文本识别方面,提出了基于改进CRNN的文本识别算法。首先,矫正网络选用贝塞尔对齐网络,采样网格不再受矩形框的条件限制,将采样点线性索引后进行双线性插值,实现图像矫正和文本对齐。然后,在识别算法中引入门控循环单元^[15](Gate Recurrent Unit, GRU)代替长短期记忆^[16](Long Short Term Memory, LSTM)网络对特征序列进行编码,使用结合注意力机制算法^[17]代替连续时序分类算法(Continuous Temporal Classification, CTC)对特征序列进行解码,得到最终的识别结果。消融实验表明:矫正网络可以提升模型性能;GRU的引入可以提高模型精度、降低模型复杂度;注意力机制对识别效果有明显提升。

3) 所提出的网络框架集检测与识别于一体,实现了端到端的文本定位和识别,具有网络结构简单、推理速度快、识别精度高,对阴暗、强光、灰尘等复杂背景具有抗干扰能力等特点。

1 相关工作

1.1 文字检测

文本检测算法可以分为4种类型:基于回归的方法、基于分割的方法、基于脑神经元(Giant Cerebral Neuron, GCN)的方法和基于混合的方法。为了检测具有不同纵横比且不规则形状的文本,RRD^[18]在SSD^[19]的基础上重新设定锚框,以适应不规则形状的文本,并减少纵横比变化所带来的误差。Textboxes++^[20]则是把卷积内核和锚框机制进行修改,以检测各种形状的文本。这类算法的局限性在于锚框机制比较复杂。经典的回归算法CTPN(Connectionist Text Proposal Network)^[21]的创新之处在于引入垂直锚框机制,将不同尺寸的文本简化成多个固定宽度的候选框,能够检测多尺度多语言的文本。该类算法的不足之处在于检测弯曲文本时精度不高。为了解决深度学习算法耗时长和弯曲文本检测精度不高的问题,文献[22]提出EAST算法,该算法在速度上获得提升,能直接预测任意形状文本,省去了候选方案、文本区域形成等交互步骤;文献[23]提出了基于像素分割的渐进尺度扩展网络,使得该方法可以很好地适用于曲线文本以及区分距离较近的文本行。但基于分割算法容易受到分割精度和内核数量的影响。文献[24]提出采用基于频谱的图卷积网络进行深度关系推理,适应更复杂情况下任意形状文本的检测,但是该方法的泛化能力不理想。

1.2 文字识别

第一种文字识别算法是连续时序分类算法^[25],主要解决输入序列和输出序列对应的问题,但是在进行去重操作时容易产生误差,把原本相邻的两个相同字符识别成一个。CTC的升级版GTC^[26],是利用CTC进行监督和引导,并加入GCN图卷积神经网络提高模型表达能力,但是前向后向算法实现比较复杂。第二种文字识别算法是基于注意力机制算法(Attention)^[27],提供了一种代替RNN的新思路。文献[28]在此基础上提出了融合语句间相互影响的注意力方案。文献[29]提出了一种基于CNN、RNN和注意力机制的神经网络模型^[30],这类算法识别的准确率高,但不足之处在于存储需求和计算消耗较大。第三种文字识别算法是聚合交叉熵^[31](Aggregation Cross-entropy, ACE),把每一个时间序

列预测的概率进行聚合再归一化, 具有计算简单、节省内存的优点。

2 算法原理

ABCNet 是一个可以实现端到端训练的网络框架, 不需要预定义锚设置, 并且计算开销小。它主要包含 3 个模块: 主干网络、贝塞尔检测网络和基于注意力机制的识别模块, 整体结构框架如图 1 所示。

它的主要步骤包括: 首先构建 50 层的残差神经网络作为骨干网络, 用特征金字塔网络^[14]来提取输入的图像特征; 然后通过基于贝塞尔曲线检测算法的单阶段无锚网络^[32]回归坐标点, 通过坐标点的参数信息将电表显示屏中的文字信息框出; 接着通过特征匹配算法, 将区域内文字特征进行对齐; 最后通过一个轻量级的文字识别模型进行文本识别, 并以图像的形式输出最终的可视化结果。

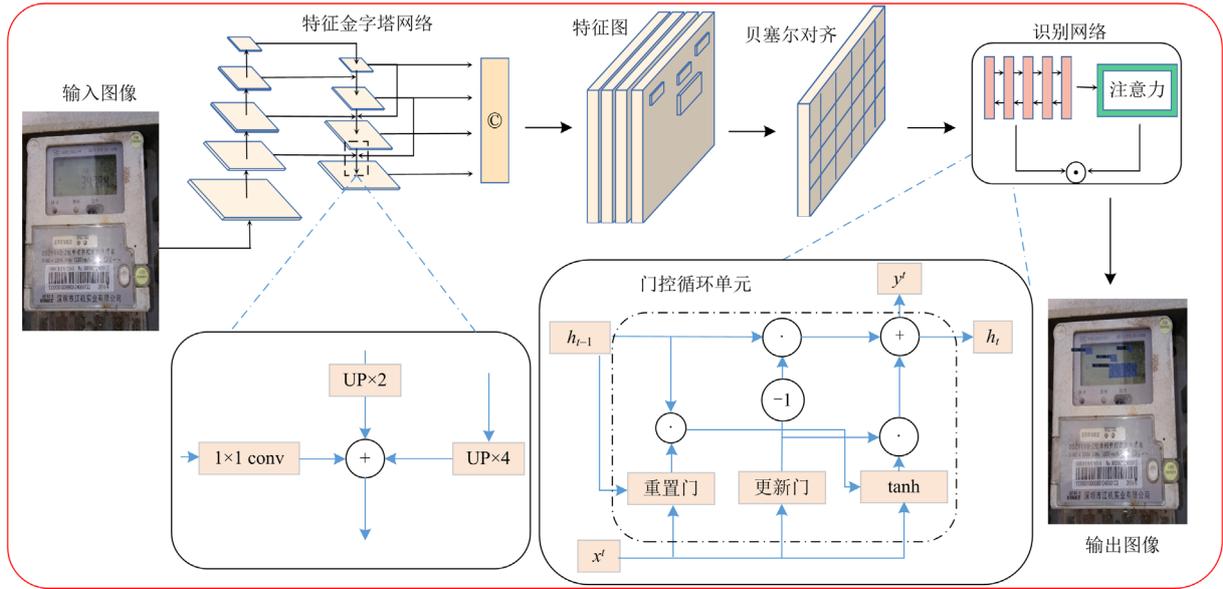


图 1 网络整体结构

Fig. 1 Overall network structure

2.1 骨干网络

骨干网络结构由残差网络构成, 如图 2 所示。首先, 利用残差网络提取特征, 接着对特征进行融合, 然后对文本的位置和语义进行预测。其中, 卷积层每进行一级卷积之后, 图像特征尺寸变成上一级的一半, 通道数变为原来 2 倍。该残差网络有 2 种瓶颈结构: BTNK1 和 BTNK2。当特征图尺寸不同时, 使用 BTNK1 结构进行连接; 当特征图尺寸相同时, 使用 BTNK2 结构进行连接。并采用 3×3 卷积核进行滤波, 减轻上采样过程带来的混叠影响。上采样采用最近邻插值, 将特征图尺寸扩大 2 倍。最后, 将每一级生成的特征以合并多个数组的方式进行特征融合。

2.2 贝塞尔算法相关理论

2.2.1 贝塞尔曲线

贝塞尔曲线^[33]可以用伯恩斯坦多项式^[34]表示成一条参数化曲线 $c(t)$, 其定义公式为

$$c(t) = \sum_{i=0}^n b_i B_{i,n}(t), \quad 0 \leq t \leq 1 \quad (1)$$

式中: n 表示贝塞尔曲线的阶数; b_i 表示第 i 个控制点; $B_{i,n}(\cdot)$ 表示伯恩斯坦多项式, 其定义公式为

$$B_{i,n}(t) = C_n^i t^i (1-t)^{n-i}, \quad i = 0, \dots, n \quad (2)$$

式中, C_n^i 代表多项式系数。根据经验以及实验验证得知, 当 $n=3$ 时, 可以较好地拟合任意形状文本框。

图 3 为电表用电信息的文本框示例, 文本框边界由 8 个控制点组成, 包括 4 个顶点及两条长边上的 4 个三等分点。利用长边上的 4 个控制点分别生成一条贝塞尔曲线。当标注点大于 4 个时, 可以通过最小二乘法来获取 4 个控制点的最优参数组合, 实现多边形标注的贝塞尔参数化。

2.2.2 贝塞尔对齐

为了实现端到端训练, 通常在连接识别网络分支前进行一步预处理操作, 即将候选文本区域的视觉特征对齐后输出一个新的特征图。特征对齐方法主要包括 3 种方法: 感兴趣区域对齐法、感兴趣区域池化法^[35]和文本对齐采样法^[36]。

本文采用贝塞尔特征匹配法构建了一个自适应贝塞尔对齐网络(Bezier Align Network, BAN)。它的

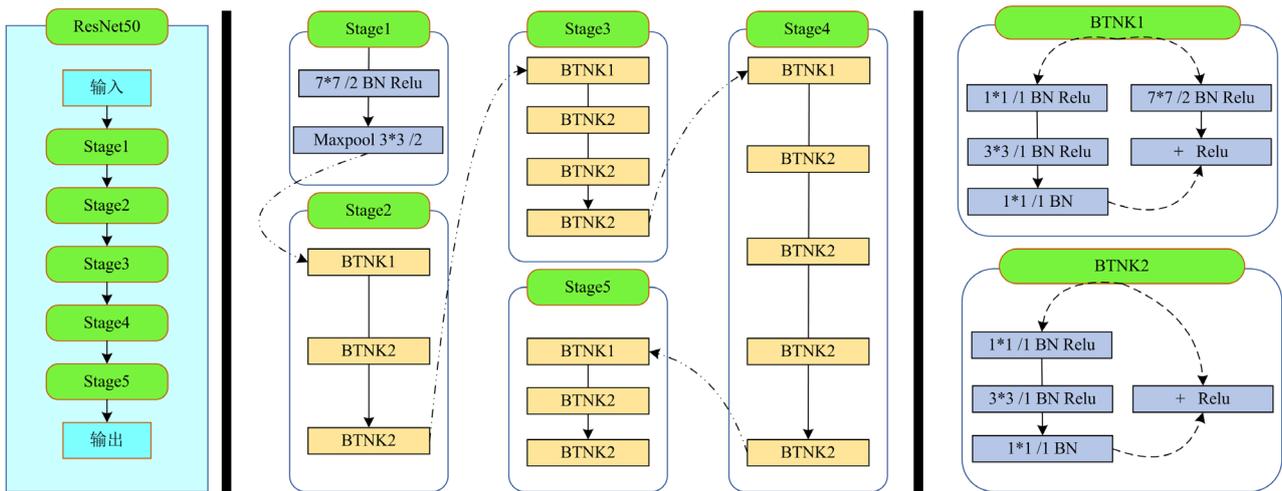


图 2 骨干网络结构

Fig. 2 Backbone network structure



图 3 文本框示例

Fig. 3 Textbox instance

创新之处在于：采样网格可以是任意形状，不受矩形框的条件限制，该方法可以同时处理输入特征图的所有像素点。以像素点位置坐标 $g_i(g_{iw}, g_{ih})$ 为例，采用比例式(3)和线性索引式(4)计算。

$$t = \frac{g_{iw}}{w_{out}} \tag{3}$$

$$o_p = b_p \cdot \frac{g_{ih}}{h_{out}} + t_p \cdot (1 - \frac{g_{ih}}{h_{out}}) \tag{4}$$

式中： h_{out} 和 w_{out} 分别示输出的矩形特征图高度和宽度； t_p 、 b_p 分别表示贝塞尔曲线的上、下边界点； o_p 表示采样点。

2.3 基于 CRNN 的文本识别模块

为了减轻识别分支的负担，文本识别选用轻量级的网络结构。该网络将提取出的图像深度特征送入循环，经网络预测每帧序列的标签分布，然后使用 CTC 进行解码，将特征序列转化为标签序列。改进的 CRNN 模型结构如图 4 所示。它引入门控循环单元代替长短期记忆网络，并使用基于注意力机制

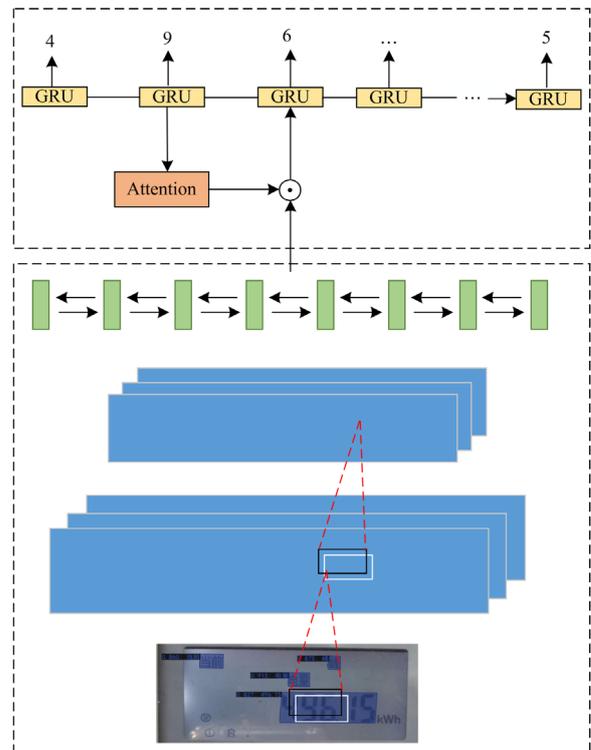


图 4 改进的 CRNN 模型

Fig. 4 Improved CRNN model

的解码器代替 CTC，这样不仅可以降低网络的复杂度，还可以提升网络的性能。

2.3.1 序列建模

经过贝塞尔对齐处理的图像已经被调整成统一的高度和宽度，所以在构建卷积特征提取网络时，只选取卷积层、池化层和激活函数的叠加组合。输入图像特征可以比较完整地映射到隐藏层，因而可以舍弃全连接层。最终生成的卷积特征图，可以用

于提取特征序列。特征序列的划分标准则是以列为单位, 从左至右每一个像素点生成一个特征向量。也就是说, 特征序列的长度就是特征图的列长, 特征序列的宽度则为固定数值, 即一个像素点的大小。

假设一共输入 T 条特征序列, 理论上每一条序列都会对应一个标签, 所以循环层发挥的主要作用就是对输入的特征序列进行标签标注和标签预测。一般采用 \tanh 函数或者 Logistic 函数作为激活函数, 在进行反向传播时极易发生梯度消失或者梯度爆炸现象, 影响训练时网络的权重参数。GRU 与 LSTM 一样都用于解决长期记忆和反向传播中的梯度消失和梯度爆炸问题, 主要通过引入重置门和更新门来修改循环神经网络中隐藏状态的计算方式。

考虑到单向的 GRU 网络只能通过一个方向来捕捉其上下文的信息, 因此本文采用双向 GRU, 从特征序列的两个方向进行序列建模, 然后再拼接得到一个更好的特征序列。双向 GRU 网络结构如图 5 所示。这样的特殊设计不仅能够缓解 RNN 的梯度消失或者梯度爆炸问题, 而且可以更好地捕捉到序列上下文之间的内部联系。

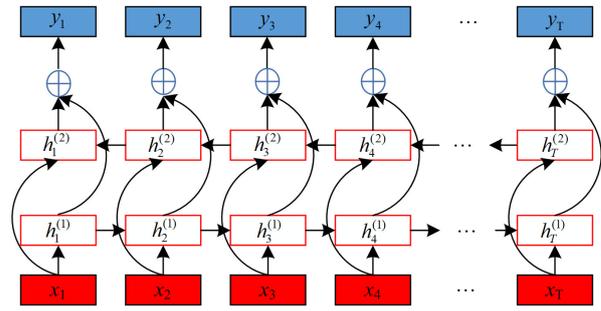


图 5 双向 GRU 结构

Fig. 5 Bidirectional GRU structure

2.3.2 解码网络

解码网络是由引入了注意力机制的双向 GRU 网络构成。它的核心思想是将序列建模所得的特征序列转换为字符序列, 得到识别的最终结果。注意力机制是一种模仿人类视觉注意力的算法, 起初将其加入序列到序列模型^[37]用于处理图像中视觉特征对齐和词嵌入, 增强算法的可读性。注意力处理机制如图 6 所示。

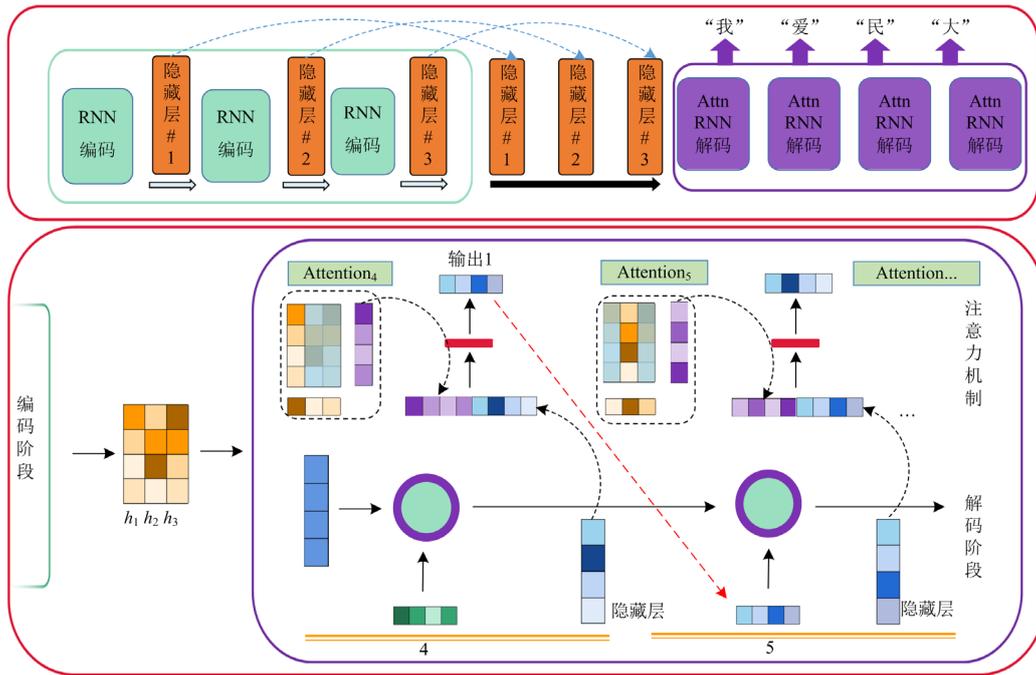


图 6 注意力处理机制

Fig. 6 Attention processing mechanism

实际上, 注意力机制就是 QKV 模型, 其中 Q 代表查询(Query), 即要处理的序列; K 代表关键信息(Key), 表示每一种可能存在的对应输出序列; 而每一序列对应一个权重值, 即为 V(Value)。因此 Q 与 K 构成了一对相关变量, 相关性由数值 V 来决定,

所有的权重值进行加权求和, 便可以得到 Attention Value(注意力权重)。注意力向量的分数 $e_{t,i}$ 及权重值 $\alpha_{t,i}$ 分别由式(5)和式(6)获得。

$$e_{t,i} = \mathbf{w}^T \tanh(\mathbf{W}s_{t-1} + \mathbf{V}h_i + \mathbf{b}) \quad (5)$$

$$\alpha_{t,i} = \frac{\exp(e_{t,i})}{\sum_{i'=1}^n \exp(e_{t,i'})} \quad (6)$$

式中： w 、 W 和 V 为可训练的权重值； s_{t-1} 表示注意力向量； h_t 表示输出序列； b 表示偏置。权重值 $\alpha_{t,i}$ 作为系数，将编码器输出序列线性组合成向量 g_t 。

$$g_t = \sum_{i=1}^n \alpha_{t,i} h_i \quad (7)$$

作为译码器循环单元的输入，产生一个输出向量 x_t 和一个新的注意力向量 s_t 。

$$(x_t, s_t) = rnn(s_{t-1}, (g_t, f(y_{t-1}))) \quad (8)$$

式中： $(g_t, f(y_{t-1}))$ 为 g_t 与 y_{t-1} 的one-hot嵌入； $rnn(\cdot)$ 表示循环单元GRU的阶跃函数，其输出和新的内部状态分别用 x_t 和 s_t 表示。采用 x_t 来预测当前步的字符。

$$p(y_t) = \text{soft max}(W_o x_t + b_o) \quad (9)$$

$$y_t \sim p(y_t) \quad (10)$$

这是一种隐式的语义模型，它可以帮助识别之前所学的语义信息。

3 实验过程及结果分析

为了验证所提出的电表检测与识别模型的有效性，在电表数据集上进行了5组消融实验，分别是：(1)选取近5年提出的文本检测算法在公开数据集和自定义数据集上分别进行对比；(2)探究残差网络的结构和深度对模型的影响；(3)对比不同Epoch对训练结果的影响；(4)对比文本矫正网络对识别准确度的影响；(5)对比CTC和ATTN两种不同机制的识别准确度。

3.1 数据集

实验收集并使用2200张coco格式的低压电表数据集。其中，2000张用于训练，剩余的200张用于测试，目标区域为液晶屏的文字区域。在数据集的预处理阶段，采用数据增强方法，选择随机尺度训练，最小训练尺寸在640~896随机选取，最大训练尺寸小于1600；在确保文字部分不被裁剪的情况下进行0.1比例随机尺度裁剪。该自定义的数据集具有通用性。

3.2 实验细节

网络在一个GTX 2080TiGPU显卡上进行训练，训练批次为2。初始学习率为0.01，分别在迭代次数为160000和220000时乘以0.1，最大迭代次数为400000，训练时长大约为1天。

3.3 消融实验

3.3.1 文本检测算法

为了验证模型对更复杂情况的鲁棒性，选取近

5年提出的5种文本检测算法，在尽量保证公平公正的条件下进行对比实验。评价指标有准确率、召回率和综合评价指标(F 值)。衡量检测效果采用ICDAR 2015数据集和自定义的电表数据集。表1比较了本文算法和其他算法^[20-23]。从表1中可以看出，本文算法在检测精度上具有一定的优越性，因此在后面的实验中，选择该算法对电表信息进行定位和预测。

表1 ICDAR2015和电表数据集的检测结果比较

Table 1 Comparison of detection results in ICDAR2015 and electric meter dataset

算法	数据集	准确率/%	召回率/%	F 值/%
EAST	ICDAR 2015 数据集	83.27	78.33	80.72
PSENet		87.26	85.37	86.30
DRRG		89.24	84.69	86.91
ABCNet		93.56	89.60	91.54
本文方法		95.28	90.86	93.02
EAST	电表 数据集	92.62	94.31	93.46
PSENet		97.25	95.49	96.36
DRRG		98.36	97.64	98.00
ABCNet		99.02	99.78	99.40
本文方法		99.86	99.63	99.75

3.3.2 残差网络结构和深度对模型的影响

为了避免深度神经网络的梯度消失、梯度爆炸和网络退化问题，选取残差网络作为特征提取的骨干网络。消融实验选取ResNet34、ResNet50和ResNet101残差网络进行对比，实验结果如表2所示，可以看出，ResNet50得益于BottleNeck结构的特殊性，在参数量和精度之间取得了平衡。

3.3.3 Epoch对准确率的影响

表3为Epoch对准确率的影响。可以看出，在相同的权重衰减作用下，模型在Epoch=400，IoU大于或等于0.5和0.75两个标准下，训练结果都达到基本稳定。

表2 不同深度残差网络的对比结果

Table 2 Comparison results of residual network with different depths

网络层数	准确率/%	召回率/%	F 值/%	参数
34	91.30	93.76	92.51	63.47×10^6
50	99.08	99.30	99.19	76.16×10^6
101	99.12	99.45	99.28	85.21×10^6

表3 Epoch对准确率的影响

Table 3 Influence of Epoch on accuracy

Epoch	权重衰减	$IoU \geq 0.5$	$IoU \geq 0.75$
100	0.0001	0.684	0.563
200	0.0001	0.906	0.832
400	0.0001	0.991	0.947
1000	0.0001	0.995	0.951

3.3.4 矫正网络的有效性研究

为了验证 BezierAlign 矫正网络对识别性能所起的作用, 分别使用带有 BAN 网络和没有 BAN 网络的模型在数据集上进行消融实验, 对比结果如表 4 所示。可以看出, 经过矫正网络处理的识别算法, 不仅识别准确率提升了一个百分点, 而且在推理速度方面也没有引入太多的额外消耗。由此可见, BAN 对识别性能具有一定的提升作用。

表 4 矫正网络有效性的对比结果

Table 4 Comparison results of the effectiveness of the corrective network

算法	速度/fps	准确率/%	召回率/%	F 值/%
无 BAN	22.8	98.06	99.15	98.60
有 BAN	22.5	99.08	99.30	99.19

3.3.5 基于不同机制的识别准确度分析

识别准确度的实验结果如表 5 所示。从表 5 中可知: GRU 和 LSTM 的识别准确率相差不大, GRU 微微领先, 且 GRU 模型大幅度减少了参数量、降低了耗时; 引入注意力机制后, 模型可以更好地获取特征信息, 准确率提升了 3.33%, 在耗时、参数量和精度之间取得了平衡。

表 5 不同机制的识别准确度对比结果

Table 5 Comparison results of the recognition accuracy of different mechanisms

识别机制	准确率/%	耗时/ms	参数
LSTM+CTC	95.69	13.2	52.72×10 ⁶
GRU+CTC	95.75	9.8	42.58×10 ⁶
GRU+ATTN	99.08	16.5	46.16×10⁶

3.4 实验结论

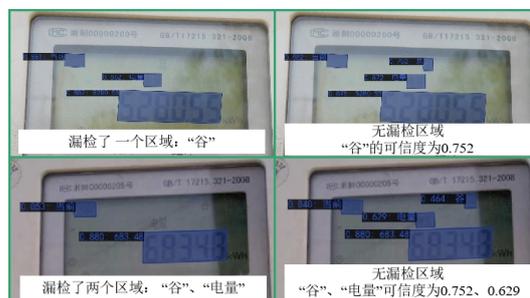
两种识别算法的可视化结果对比如图 7 所示, 输出结果前面的数字代表在序列结果中预测正确的概率, 数值越接近于 1, 表明可信度越高。其中, 左侧为 CTC 算法的检测与识别结果, 右侧为本文提出的结合了注意力机制算法的检测识别结果。通过对比可以看出, 本文算法具有以下 4 点优势。

1) 本文所提出的算法, 文字区域检测的成功率更高, 即便受图片模糊、拍照时相机抖动等因素影响, 也能成功定位目标文本区域; 相比之下, 左图漏检了多个文字区域, 如图 7(a)所示。

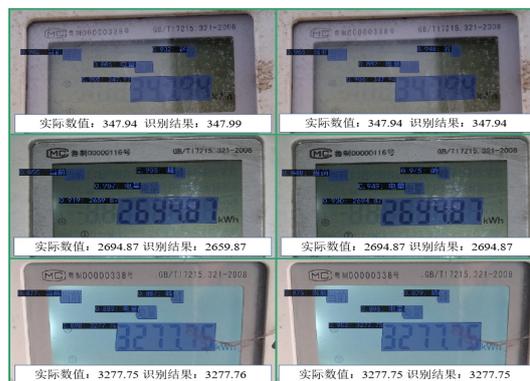
2) 受图片模糊、划痕、污渍、异物遮挡等因素影响时, Attention 算法可将电表用电信息相对准确地检测并识别出来, 且可信度很高, 如图 7(b)所示。

3) 处理相邻的两个相同字符时, Attention 算法可以避免 CTC 算法中去重操作所带来的误差, 识别精度更高, 如图 7(c)所示。

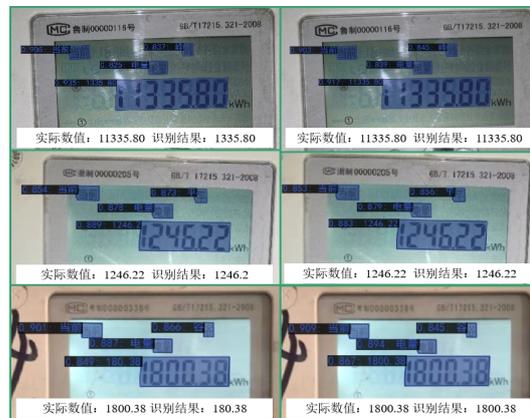
4) 处理 5/6/8 等相似字符时, 传统的 CTC 算法在字符预测时, 容易出错, 而结合 Attention 机制的识别算法错误率更低, 如图 7(d)所示。



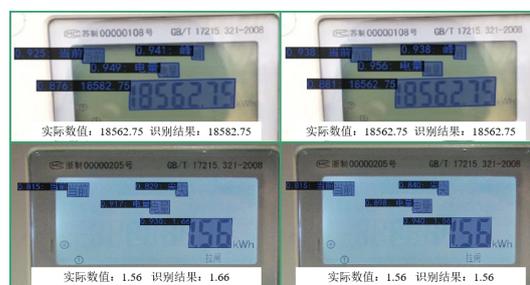
(a) 检测效果对比



(b) 受外界因素影响的识别效果对比



(c) 相同字符的识别效果对比



(d) 相似字符的识别效果对比

图 7 可视化结果对比

Fig. 7 Comparison of visual results

参考文献

- [1] 王鸿玺, 唐如意, 吴一敌, 等. 基于 HPLC 的智能抄表技术在客户侧泛在电力物联网中的研究及应用[J]. 电力系统保护与控制, 2020, 48(3): 92-98.
WANG Hongxi, TANG Ruyi, WU Yidi, et al. Research and application of smart meter reading technology based on HPLC in customer side universal power internet of things[J]. Power System Protection and Control, 2020, 48(3): 92-98.
- [2] 孙立明, 杨博. 蓄电池/超导混合储能系统非线性鲁棒分数阶控制[J]. 电力系统保护与控制, 2020, 48(22): 76-83.
SUN Liming, YANG Bo. Nonlinear robust fractional order control of battery/SMES hybrid energy storage systems[J]. Power System Protection and Control, 2020, 48(22): 76-83.
- [3] 何黎君, 程杉, 陈梓铭. 考虑交互功率控制和双边竞价交易的多微电网双层优化调度[J]. 电力系统保护与控制, 2020, 48(11): 10-17.
HE Lijun, CHENG Shan, CHEN Ziming. A scheduling model of a multi-microgrid system based on bi-layer optimization with consideration of PCC power control and bilateral bidding[J]. Power System Protection and Control, 2020, 48(11): 10-17.
- [4] HU Zhuangli, HE Tong, ZENG Yihui, et al. Fast image recognition of transmission tower based on big data[J]. Protection and Control of Modern Power Systems, 2018, 3(2): 149-158.
- [5] ANIS A, KHALILUZZAMAN M, YAKUB M, et al. Digital electric meter reading recognition based on horizontal and vertical binary pattern[C] // 3rd International Conference on Electrical Information and Communication Technology (EICT), December 7-9, 2017, Khulna, Bangladesh: 1-6.
- [6] ZHANG Yichen, YANG Xinxin, HONG Tao, et al. Digital meter recognition method based on topological features of image contour[C] // IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), June 5-7, 2019, Jeju, South Korea: 1-4.
- [7] LAROCA R, BARROSO V A, DINIZ M A, et al. Convolutional neural networks for automatic meter reading[J]. Journal of Electronic Imaging, 2019, 28(1).
- [8] SILVA S M, JUNG C R. Real-time Brazilian license plate detection and recognition using deep convolutional neural networks[C] // 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images, October 17-20, 2017, Niteroi, Brazil: 55-62.
- [9] SHI Baoguang, BAI Xiang, YAO Cong. 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.
- [10] 龚安, 张洋, 唐永红. 基于 YOLOv3 网络的电表示数识别方法[J]. 计算机系统应用, 2020, 29(1): 196-202.
GONG An, ZHANG Yang, TANG Yonghong. Automatic reading method of electric energy meter based on YOLOv3[J]. Computer Systems & Applications, 2020, 29(1): 196-202.
- [11] 厉建宾, 张旭东, 吴彬彬. 结合深度神经网络和多阈值软切分的电表示数自动识别方法[J]. 计算机系统应用, 2019, 39(1): 223-227.
LI Jianbin, ZHANG Xudong, WU Binbin. Automatic display number recognition method for electricity energy meter based on deep neural network and multi-threshold soft segmentation[J]. Journal of Computer Applications, 2019, 39(1): 223-227.
- [12] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA: 770-778.
- [13] LIN T Y, DOLLÁR P, GIRSHICK R B, et al. Feature pyramid networks for object detection[C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA: 936-944.
- [14] HE Kaiming, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 386-397.
- [15] DEY R, SALEM F. Gate-variants of gated recurrent unit (GRU) neural networks[C] // IEEE 60th International Midwest Symposium on Circuits and Systems, August 6-9, 2017, Boston, MA, USA: 1597-1600.
- [16] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [17] FUKUI H, HIRAKAWA T, YAMASHITA T, et al. Attention branch network: learning of attention mechanism for visual explanation[C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA: 10697-10706.
- [18] LIAO Minghui, ZHU Zhen, SHI Baoguang, et al. Rotation-sensitive regression for oriented scene text detection[C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA: 5909-5918.
- [19] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: single

- shot multibox detector[C] // European Conference on Computer Vision, October 8-16, 2016, Amsterdam, Holland: 21-37.
- [20] LIAO Minghui, SHI Baoguang, BAI Xiang. TextBoxes++: a single-shot oriented scene text detector[J]. IEEE Transactions on Image Processing, 2018, 27(8): 3676-3690.
- [21] TIAN Zhi, HUANG Weilin, HE Tong, et al. Detecting text in natural image with connectionist text proposal network[C] // European Conference on Computer Vision, October 8-16, 2016, Amsterdam, Holland: 56-72.
- [22] ZHOU Xinyu, YAO Cong, WEN He, et al. EAST: an efficient and accurate scene text detector[C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition, July 22-25, 2017, Hawaii, USA: 2642-2651.
- [23] LI Xiang, WANG Wenhai, HOU Wenbo, et al. Shape robust text detection with progressive scale expansion network[C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 16-19, 2019, CA, USA: 9328-9337.
- [24] ZHANG Shixue, ZHU Xiaobin, HOU Jiebo, et al. Deep relational reasoning graph network for arbitrary shape text detection[C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 14-19, 2020, CA, USA: 9696-9705.
- [25] GRAVES A, FERNÁNDEZ S, GOMEZ F, et al. Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks[C] // Proceedings of the 23rd International Conference on Machine Learning, June 25-29, 2006, PA, USA.
- [26] HU Wenyang, CAI Xiaocong, HOU Jun, et al. GTC: guided training of CTC towards efficient and accurate scene text recognition[C] // Association for the Advancement of Artificial Intelligence, February 7-20, 2020, CA, USA.
- [27] VASWANI A, SHAZEER NM, PARMAR N. Attention is all you need[J]. ArXiv, abs/1706.03762, 2017.
- [28] YIN Wenpeng, SCHÜTZE H, XIANG Bing, et al. ABCNN: attention-based convolutional neural network for modeling sentence[J]. Transactions of the Association for Computational Linguistics, 2016, 4: 259-272.
- [29] WOJNA Z, GORBAN A N, LEE D, et al. Attention-based extraction of structured information from street view imagery[C] // 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), November 13-15, 2017, Kyoto, Japan: 844-850.
- [30] SMITH R, GU C, LEE D, et al. End-to-end interpretation of the French street name signs dataset[C] // European Conference on Computer Vision, Part I, LNCS 9913, October 8-16, 2016, Amsterdam, Holland: 411-426.
- [31] XIE Zecheng, HUANG Yaoxiong, ZHU Yuanzhi, et al. Aggregation cross-entropy for sequence recognition[C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 16-19, 2019, CA, USA: 6531-6540.
- [32] TIAN Zhi, SHEN Chunhua, CHEN Hao, et al. FCOS: fully convolutional one-stage object detection[C] // IEEE/CVF International Conference on Computer Vision, October 27-November 2, 2019, Seoul, Korea (South): 9626-9635.
- [33] LIU Yuliang, CHEN Hao, SHEN Chunhua, et al. ABCNet: real-time scene text spotting with adaptive Bezier-curve network[C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 14-19, 2020, CA, USA: 9806-9815.
- [34] LORENTZ G G. Bernstein polynomials[M]. American Mathematical Soc., 2013.
- [35] LI Hui, WANG Peng, SHEN Chunhua. Towards end-to-end text spotting with convolutional recurrent neural networks[C] // IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy: 5238-5246.
- [36] HE Tong, TIAN Zhi, HUANG Weilin, et al. An end-to-end textspotter with explicit alignment and attention[C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-22, 2018, Salt Lake City, USA: 5020-5029.
- [37] SUTSKEVER I, VINYALS O, LE Q V. Sequence to sequence learning with neural networks[C] // Neural Information Processing Systems, December 8-13, 2014, Montreal, Canada.

收稿日期: 2021-08-28; 修回日期: 2022-01-26

作者简介:

孙福明(1972—), 男, 博士, 教授, 研究生导师, 研究方向为智能检测与模式识别; E-mail: sunfuming@dlnu.edu.cn

高 严(1992—), 男, 硕士研究生, 研究方向为智能检测与模式识别; E-mail: 863450662@qq.com

魏晓鸣(1963—), 男, 通信作者, 博士, 教授, 研究方向为人工智能与机器学习。E-mail: xmwei@dlnu.edu.cn

(编辑 许 威)