

# 基于改进 AdvancedEAST 的变电站仪表数显区域检测 \*

王嘉璇<sup>1</sup> 王天宁<sup>1</sup> 刘兵<sup>2</sup> 王祝<sup>1</sup>

(1. 华北电力大学自动化系 保定 071003; 2. 天津新松智能科技有限公司 天津 301800)

**摘要:** 为了提高数字仪表数显区域检测的实时性和鲁棒性,提出一种改进 AdvancedEAST 算法,以快速、准确地完成变电站数字仪表数显区域检测。首先,在 AdvancedEAST 模型框架下,构建一种注意力机制增强的轻量级主干网络 ES-MobilenetV3,通过考虑不同层提取特征对结果的影响,引入 ECA 和多维注意力机制(ECA-SE)对 MobileNetv3 的 Bneck 块进行改进,在保持性能和复杂性平衡的前提下,突出关键特征。在颈部网络中引入深度可分离卷积,降低网络计算的复杂度,提高检测速度。同时,使用迁移学习策略,提高小样本下模型的泛化能力。最后,在构建的变电站数字仪表数据集上进行试验验证,结果表明,所提算法在保证检测精度的情况下,模型的参数量相比 AdvancedEAST 算法减少 82%,检测速度提升近 2 倍。

**关键词:** 数字仪表; AdvancedEAST; 数显区域检测; 注意力机制增强; 深度可分离卷积; 迁移学习

**中图分类号:** TP391; TP216    **文献标识码:** A    **国家标准学科分类代码:** 510.1050

## Detection of digital display area for substation instrument based on improved AdvancedEAST

Wang Jiaxuan<sup>1</sup> Wang Tianning<sup>1</sup> Liu Bing<sup>2</sup> Wang Zhu<sup>1</sup>

(1. Department of Automation, North China Electric Power University, Baoding 071003, China;

2. Tianjin Xinsong Intelligent Technology Co., Ltd., Tianjin 301800, China)

**Abstract:** To enhance the real-time and robustness of digital instrument digital display area detection, an improved AdvancedEAST algorithm is proposed to complete the detection of substation digital instrument digital display area quickly and accurately. First, under the framework of the AdvancedEAST model, an ES-MobilenetV3 lightweight backbone network enhanced by the attention mechanism is constructed. By considering the influence of different layers of extracted features on the results, the ECA and multi-dimensional attention mechanism (ECA-SE) is introduced to The Bneck block of MobileNetv3 is improved to highlight key features while maintaining a balance between performance and complexity. A depthwise separable convolution is introduced in the neck network to reduce the computational complexity of the network and improve the detection speed. At the same time, the transfer learning strategy is used to improve the generalization ability of the model under small samples. Finally, the experimental verification was carried out on the constructed substation digital instrument dataset. The results showed that the proposed algorithm reduced the number of parameters of the model by 82% compared to the AdvancedEAST algorithm and increased the detection speed by nearly 2 times while ensuring detection accuracy.

**Keywords:** digital instrument; AdvancedEAST; digital display area detection; enhanced by the attention mechanism; depthwise separable convolution; transfer learning

## 0 引言

随着数字化转型的不断推进,变电站数字化建设逐渐成为电力行业发展的趋势。变电站数字仪表可以实时显示变电站中各种设备的运行情况,对保障电网运行安全、提高

电网运行效率至关重要<sup>[1]</sup>。传统的人工采集仪表数据方式已经无法满足变电站对数据实时性的需求,变电站仪表读数已经向自动读数方案转变<sup>[2]</sup>。

数显区域检测定位是读数检测识别的重要环节之一,其作用是在图像中准确地定位数显区域,并提取出感兴趣

的数字,以便进行后续的识别和处理。数显区域检测定位的重要性在于它直接影响到读数识别的准确性和效率。传统的数显区域检测定位的主要方法有双向投影法、色彩定位法和数字形态法等。肖佳<sup>[3]</sup>在图像预处理后,利用投影法粗定位仪表数显区域,再利用数学形态学精确定位,得到读数区域。苑玮琦等<sup>[4]</sup>利用改进曲率尺度空间角点检测原理和液晶边框顶点的位置关系检测仪表数显区域。陈刚等<sup>[5]</sup>通过列切分法对单个数码字符进行切分,使用七段检测和五线相交检测实现轻微斜体读数的快速检测和识别。但以上方法由于光照变化、尺度变化、噪声等因素的影响在复杂环境下鲁棒性较差。

近年来,深度学习算法在图像处理领域表现出色,相比传统的检测和识别算法,有较强的鲁棒性和泛化能力。目前,主流检测算法分为以 YOLO<sup>[6]</sup>(you only look once)、SSD<sup>[7]</sup>(single multibox detector)系列为代表的单阶段检测算法和以 RCNN<sup>[8]</sup>(regions with CNN features)系列为代表的两阶段检测算法。两阶段目标检测算法基于候选区域思想,利用卷积神经网络对提取区域的目标进行分类和定位,但存在模型规模大、检测速度慢的不足<sup>[9]</sup>。单阶段目标检测算法采用回归分析的思想,省略了候选区域生成阶段,直接得到目标分类和位置信息,加快了检测速度,适合用于嵌入式设备。Waqar 等<sup>[10]</sup>提出了一种基于 Faster RCNN 的电表数字提取方法,通过对单个数字进行标注与检测,从而实现数字示数的检测。华泽玺等<sup>[11]</sup>提出一种基于轻量级 YOLOv4 模型的数字仪表检测识别方法,该方法在仪表检测的基础上对数字字符逐个进行检测以获取数字示数。以上两种方法当数字字符较小或相邻数字字符间距较小时容易导致漏检或误检。侯卓成等<sup>[12]</sup>提出了深度学习与图像处理相结合的识别方法,利用 YOLOv4 对数字仪表进行定位,采用双向投影法对数字字符进行分割,使用模板匹配法获取识别结果,并在机器设备上得到了验证。电表数显区域检测与水表数显区域检测是比较接近的应用,二者在技术上有不少可以相互借鉴之处。周度<sup>[13]</sup>将旋转变量引入到 YOLO 检测模型中,训练出的 ROYOLO 水表检测模型可以很好地解决水表数显区域检测的多方向问题。

数显区域存在倾斜、边缘轮廓不明显、尺寸大小不一等情况,一般的目标检测算法应用于数显区域检测时精确率不高。由于数显区域具有明显的序列特征,更适用于文本检测算法。主流文本检测算法有 DBNet<sup>[14]</sup>(detection with differentiable binarization)、CTPN<sup>[15]</sup>(connectionist text proposal network)、EAST<sup>[16]</sup>(efficient and accurate scene text)等。鲁乾鹏<sup>[17]</sup>设计了基于 EAST 的工业数字仪表读数检测模型,有效地处理了长读数检测,但计算量相对较大。AdvancedEAST 算法在 EAST 基础上增加了后置处理,提升了对长文本的处理能力和整体的检测精度,但仍具有较大的参数量,在处理高分辨率图像或复杂场景时,算法

实时性较差。AdvancedEAST 算法在车牌识别领域进行了应用, Yin 等<sup>[18]</sup>通过迁移学习在车牌图像上训练 AdvancedEAST 文本检测模型,该算法能够很好地适应光照、低分辨率、车牌倾斜等复杂条件,在自然场景中有较高的精度。

为了解决变电站数字仪表数显区域的检测问题,特别是针对现有算法对数字仪表数显区域检测效果不佳、网络参数量大和实时检测效率不高的问题,提出基于改进 AdvancedEAST 算法的变电站数字仪表数显区域检测方法,在保证检测精度的情况下有效的缩减了模型计算复杂度和参数量。本文主要包含以下 3 个方面工作:

1) 以 AdvancedEAST 为基础框架,基于 MobileNetv3<sup>[19]</sup>改进网络主干,提高模型的内存使用效率和推理速度。

2) 引入 ECANet<sup>[20]</sup>(efficient channel attention network)和多维注意力机制(ECA-SE)对 MobileNetv3 的 Bneck 块进行改进,突出关键特征,提高模型性能。

3) 在特征融合中引入深度可分离卷积,进一步压缩模型大小,提升模型的计算效率。

## 1 AdvancedEAST 算法原理

AdvancedEAST 算法<sup>[21]</sup>是 2018 年提出的一种场景文本检测算法,其将长文本检测转化为检测文本头部和尾部边界区的问题,使得长文本检测更加准确。AdvancedEAST 的网络架构如图 1 所示。

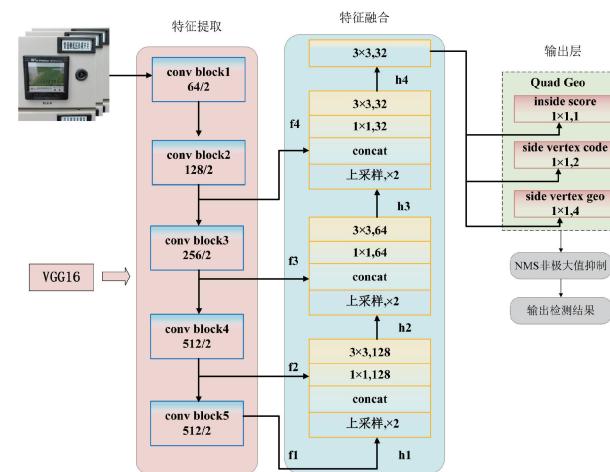


图 1 AdvancedEAST 网络架构

在模型结构上,AdvancedEAST 沿用 EAST 的三大组成部分:特征提取、特征融合、输出。其中,特征提取选用 VGG16 作为主干网络,分别提取 4 个阶段的特征图,其大小分别为输入图像的 1/32、1/16、1/8 和 1/4。不同尺度的特征图融合能够适用于多尺度目标检测任务,通过对输出的 4 个阶段的特征图上采样并与前一层的特征图进行通道上的拼接(concat)和卷积操作后,将得到的特征图送到输出层。特征融合过程:

$$h_i = \begin{cases} f_i, & i = 1 \\ conv_{3 \times 3}(conv_{1 \times 1}([g_{i-1}; f_i])), & \text{其他} \end{cases} \quad (1)$$

$$g_i = \begin{cases} unpool(h_i), & i \leq 3 \\ conv_{3 \times 3}(h_i), & i = 4 \end{cases}$$

其中,  $h_i$  表示局部信息融合后的输出,  $g_i$  表示特征融合后的输出。

AdvancedEAST 的输出层共有 7 位输出, 即 1 位 score map、2 位 vertex code 和 4 位 vertex geo, 分别表示点在文本框内的置信度(内部像素)、是否属于文本框边界像素以及是头还是尾和边界像素可以预测的 2 个顶点坐标。

## 2 改进 AdvancedEAST 算法

本文提出的改进 AdvancedEAST 算法整体框架如图 2

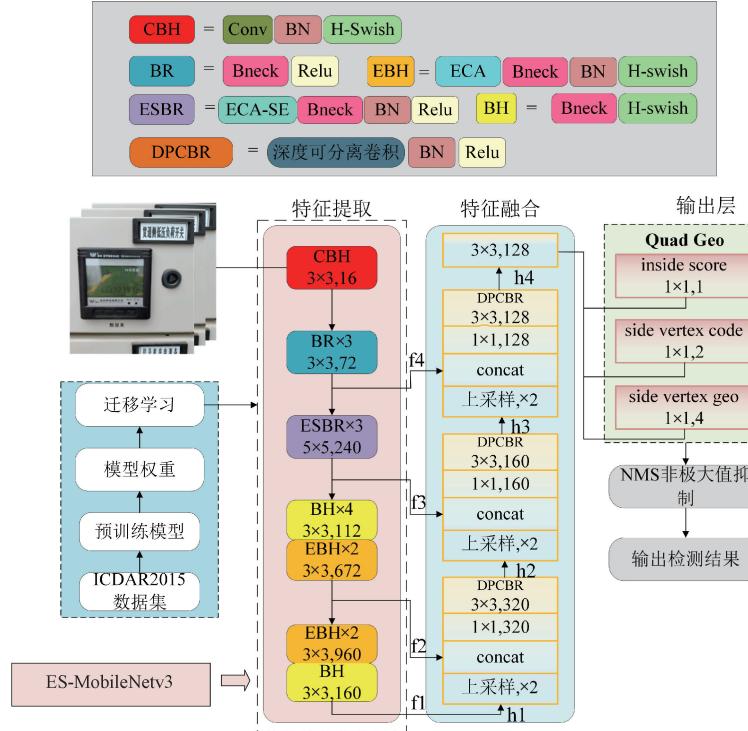


图 2 改进 AdvancedEAST 算法框架

预训练并初始化网络参数。

### 2.1 主干网络结构改进

变电站数字仪表检测具有实时性和准确性的应用要求, 且巡检机器人携带的嵌入式设备性能有限。AdvancedEAST 原主干网络 VGG16 模型参数量相对较大, 难以满足实时检测的问题。为了减少网络参数量, 使模型在嵌入式设备部署时满足实时性需求, 对 AdvancedEAST 主干网络进行改进, 设计基于 ES-MobileNetv3 的主干网络。改进主要包括下述两个方面。

1) 使用 MobileNetv3 large 作为基础网络, 保留其深度可分离卷积结构、线性瓶颈的倒残差结构, 并在深层网络中使用 h-swish 激活函数。在 MobileNetv3 large 的 Bneck

所示。主干网络 ES-MobileNetv3 以 MobileNetv3 为基准, 在 Bneck 块基础上引入多维注意力模块构建 ESBR 块, 加入 ECANet 构建 EBH 块, 以有效减少模型复杂性和提高网络在检测过程中对关键特征的提取能力。经主干网络特征提取后, 将输出的 4 个阶段的特征图送入颈部网络中进行特征融合。为了进一步缩减模型参数量, 在特征融合网络中引入深度可分离卷积结构, 提升模型的推理速度。最后, 将提取的特征图送入到输出层进行检测, 经非极大值抑制后输出最终的检测结果。

相较原算法, 引入 ES-MobileNetv3 代替 VGG16 作为主干网络, 并将特征融合部分的  $3 \times 3$  卷积层使用深度可分离卷积(DPCBR)代替。采用迁移学习的训练策略, 使用 ICDAR2015 作为源数据域对改进的 AdvancedEAST 进行

块中, 引入更快速的 ECANet 代替 SENet<sup>[22]</sup> (squeeze and excitation), ECANet 通过一维卷积来完成跨通道间的信息交互, 卷积核的大小通过一个函数来自适应变化, 使得通道数较大的层可以更多地进行跨通道交互。相比较 SENet, 结构复杂度更低, 推理速度更快, 并有效捕获了跨通道交互。ECANet 的结构图如图 3 所示。自适应函数为:

$$k = \left| \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right| \quad (2)$$

其中,  $C$  表示通道数,  $\gamma = 2, b = 1$  以此自适应函数确定卷积核的大小。

引入 ECANet 后 MobileNetv3 的 Bneck 块的结构图如图 4 所示。

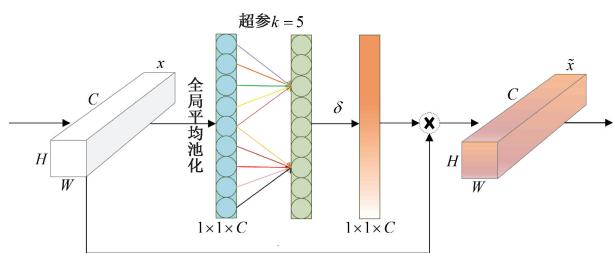


图 3 ECA-Net

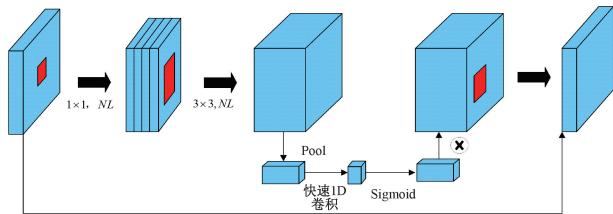


图 4 MobileNetv3 的 Bneck 块

2)为了更有效的捕捉数显区域的语义信息,对低维特征提取块进行改进,设计基于 ECA-SE 的多维注意力模块。低层特征包含丰富的小目标的语义信息,适用于小目标的检测,数显区域在仪表图像中占据面积较小,通过 SENet 获取全局通道信息和 ECANet 获取局部通道信息,将 ECANet 与 SENet 有效结合,使网络能够关注更加有用的通道信息赋予每一通道不同的权重,以此来捕捉数显区域的语义信息。ECA-SE 多维注意力模块如图 5 所示。本文改进后的 ES-MobileNetv3 的主干如表 1 所示。

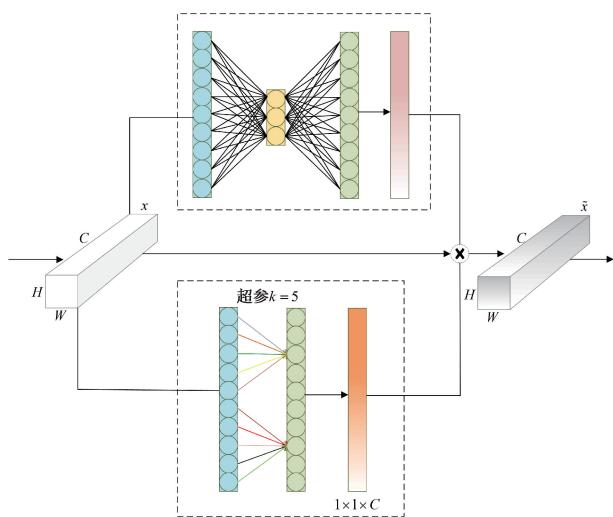


图 5 ECA-SE 多维注意力模块

## 2.2 基于深度可分离卷积的特征融合网络结构改进

深度可分离卷积<sup>[23]</sup>将传统卷积分解为用于空间滤波的轻量级深度卷积和用于特征生成的  $1 \times 1$  逐点卷积。这个分解过程能够极大地减少参数量和模型大小。在 AdvancedEAST 的特征融合网络中引入深度可分离卷积,

表 1 ES-MobileNetv3 主干

特征层	输入尺度	基本单元	注意力机制	激活函数
1	$1600 \times 1600 \times 3$	Conv2d	不施加	h-swish
2	$800 \times 800 \times 16$	bneck, $3 \times 3$	不施加	ReLU
3	$800 \times 800 \times 16$	bneck, $3 \times 3$	不施加	ReLU
4	$400 \times 400 \times 24$	bneck, $3 \times 3$	不施加	ReLU
5	$400 \times 400 \times 24$	bneck, $5 \times 5$	ECA-SE	ReLU
6	$200 \times 200 \times 40$	bneck, $5 \times 5$	ECA-SE	ReLU
7	$200 \times 200 \times 40$	bneck, $5 \times 5$	ECA-SE	ReLU
8	$200 \times 200 \times 40$	bneck, $3 \times 3$	不施加	h-swish
9	$100 \times 100 \times 80$	bneck, $3 \times 3$	不施加	h-swish
10	$100 \times 100 \times 80$	bneck, $3 \times 3$	不施加	h-swish
11	$100 \times 100 \times 80$	bneck, $3 \times 3$	不施加	h-swish
12	$100 \times 100 \times 80$	bneck, $3 \times 3$	ECA	h-swish
13	$100 \times 100 \times 112$	bneck, $3 \times 3$	ECA	h-swish
14	$100 \times 100 \times 112$	bneck, $3 \times 3$	ECA	h-swish
15	$50 \times 50 \times 160$	bneck, $3 \times 3$	ECA	h-swish
16	$50 \times 50 \times 160$	bneck, $3 \times 3$	ECA	h-swish
17	$50 \times 50 \times 160$	bneck, $3 \times 3$	不施加	h-swish

可以有效的进行模型压缩,提高模型的计算效率。文中将输入通道设为  $M$ ,输出通道数为  $N$ ,深度可分离卷积的具体过程如图 6 所示。

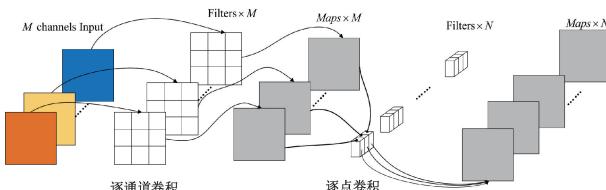


图 6 深度可分离卷积过程

逐通道卷积使用  $M$  个  $3 \times 3 \times 1$  大小的卷积核与输入的  $M$  通道的数据做卷积,得到  $M$  通道的特征图;逐点卷积使用  $N$  个  $1 \times 1 \times M$  的卷积核与输出的  $M$  个特征图做卷积操作,将  $M$  个通道的空间信息进行融合,最终得到  $N$  通道的特征图。其中,逐通道卷积的计算量为  $3 \times 3 \times M \times D_w \times D_h$ ,参数量为  $3 \times 3 \times M$ ;逐点卷积的计算量为  $M \times N \times D_w \times D_h$ ,参数量为  $M \times (N+M)$ ;故一次深度可分离卷积总的计算量为  $M \times D_w \times D_h \times (N+M)$ ,参数量为  $M \times (N+M)$ 。

标准卷积使用与输入数据具有相同通道数的卷积核在整个特征通道上进行加法和累加运算。结构图如图 7 所示。

标准卷积的单次计算量为  $3 \times 3 \times M \times N \times D_w \times D_h$ ,参数量为  $3 \times 3 \times M \times N$ 。深度可分离卷积计算量与普通卷积计算量之比为:

$$\frac{M \times D_w \times D_h \times (N+M)}{9 \times M \times N \times D_w \times D_h} = \frac{1}{N} + \frac{1}{9} \quad (3)$$

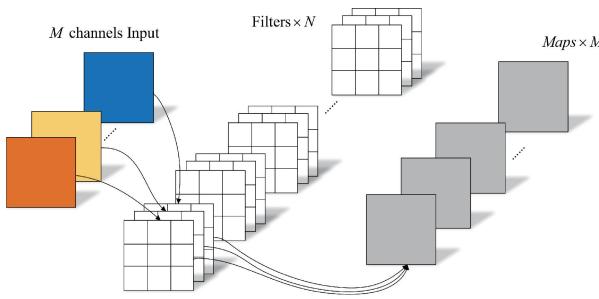


图 7 标准卷积过程

由此可得出,当通道数增加时,使用深度可分离卷积代替普通卷积可显著降低计算量。

### 2.3 迁移学习策略

针对电表的数据样本量较少的情况,且考虑到深度网络中存在大量的参数,模型从 0 开始训练,模型权值分布过于随机,训练难以收敛,特征提取效果不明显。对此,采用基于模型的迁移学习<sup>[24]</sup>策略,从已有的源数据域和目标数据域中找到可以共享的参数信息。本文所用的源数据域为 ICDAR2015,目标数据域为变电站数字仪表。

迁移学习策略能够有效的利用与当前任务相近的领域数据,与未使用迁移学习的模型相比较,预训练模型已经在源数据域上学习了丰富的特征,可以在小样本的条件下快速的训练出结果,且提升小样本条件下电表读数识别的准确率。使用 ICDAR2015 数据集作为源数据域对本文提出的改进 AdvancedEAST 模型进行预训练,将预训练好的模型迁移到本文模型中。迁移学习策略的训练过程分为冻结和解冻两个阶段。在冻结阶段,使用预训练权重初始化特征提取的主干网络参数并将其冻结,仅对主干网络以外的其他网络参数进行微调,这个过程可以加快训练效率和防止权值被破坏。在解冻阶段,将模型的主干网络参数解冻后继续训练,该过程对模型的所有参数进行微调,最终得到最优模型。

### 2.4 损失函数

损失函数由 3 部分组成,分别为 scope map、vertex code 和 vertex coord,Scope map 和 vertex code 损失为置信度损失,采用交叉熵损失函数<sup>[25]</sup>计算,vertex code 损失为回归损失,采用 Smooth L1 损失函数计算。

scope map 和 vertex code 损失函数公式为:

$$L_{scope} = -\beta Y^* \log \hat{Y} - (1 - \beta)(1 - Y^*) \log(1 - \hat{Y})$$

$$L_{code} = \frac{-\sum_i^N (\beta y_i^* \log \hat{y}_i + (1 - \beta)(1 - y_i^*) \log(1 - \hat{y}_i)) w_i}{\sum_i^N w_i} \quad (4)$$

其中,  $L_{scope}$  表示 Scope map 损失,  $L_{code}$  表示 vertex code 损失。 $Y^*$  表示正确标注,  $\hat{Y}$  表示预测值,  $N$  表示样本

数量。 $\beta$  是先验值,在标签生成过程中得到的,表示所有训练图像中为 1 的像素点数量占总像素点数量的比例,具体定义为:

$$\beta = 1 - \frac{\sum_{y^* \in Y^*} y^*}{|Y^*|} \quad (5)$$

其中,  $w$  表示归属权重:

$$w = \begin{cases} 0, & y_i^* = 0 \\ 1, & y_i^* = 1 \end{cases} \quad (6)$$

vertex coord 的损失函数定义如下:

$$L_{coord} = \frac{\sum_i^N S_i w_i}{\sum_i^N w_i} \quad (7)$$

其中,  $S$  表示 Smooth L1 损失, Smooth L1 函数定义如下:

$$SmoothL_1 = \begin{cases} |x| - 0.5, & |x| > 1 \\ 0.5x^2, & |x| \leq 1 \end{cases} \quad (8)$$

总的损失函数为:

$$L_{all} = \lambda_1 L_{scope} + \lambda_2 L_{code} + \lambda_3 L_{coord} \quad (9)$$

其中,  $\lambda_1$ 、 $\lambda_2$  和  $\lambda_3$  分别为 score map、vertex code 和 vertex coord 权重。

## 3 试验结果及分析

### 3.1 数据集与试验平台

试验中采用的数据集包括百度飞桨平台公开数据集和配电室自采集数据集,共计 1 045 张。使用 rolabelImg 对其进行手工标注,并进行预处理以满足 AdvancedEAST 网络的输入要求。首先,将每一幅图像标注后保存的 XML 文件转为 txt 文件,其中 txt 文件内容为标注区域的四点坐标,分别表示左下、右下、右上和左上。其次,由于试验所用数据集分辨率过大且分辨率不一致,将数据集归一化至同一分辨率,修改标签内容并保存。根据标注的 4 个顶点,确定电表数显部分的像素区域,以此构建新的训练数据标签并保存。数据集中的部分图像如图 8 所示。



图 8 数字仪表数据集示例

左侧图像为原始图像,右侧图像为处理后的标签图像。采用表 2 所示的试验环境配置进行模型的训练。

表 2 试验环境配置

操作系统	Ubuntu18.04
CPU	Intel(R) Xeon(R) Platinum 8255C CPU@2.50 GHz
GPU	NVIDIA GeForce RTX 3090
运行内存	43 GB
编程语言	python3.8
深度学习框架	Pytorch1.11.0
开发环境	pycharm

### 3.2 模型评价指标和训练设置

#### 1) 模型评价指标

采用精确率(precision,  $P$ )、召回率(recall,  $R$ )、 $F1$  值作为模型的评估指标,  $P$  表示正确预测为正样本的占全部预测为正样本的比例,  $R$  表示正确预测为正的正样本占全部实际预测为正样本的比例,  $F1$  值是精确率和召回率的加权调和平均。 $P$ 、 $R$  和  $F1$  值的计算公式如下:

$$\begin{aligned} P &= \frac{T_p}{T_p + F_p} \\ R &= \frac{T_p}{T_p + F_N} \\ F1 &= 2 \times \frac{PR}{R + P} \end{aligned} \quad (10)$$

其中,  $T_p$  为检测正确的正样本,  $F_p$  为检测错误的正样本,  $F_N$  为检测错误的负样本,  $T_N$  为检测正确的负样本。

#### 2) 模型训练设置

在模型训练前,将数据集按照 9:1 的比例划分为训练集和验证集,并将图片大小调整为  $1600 \times 1600$  进行训练。损失函数参数  $\lambda_1 = 4, \lambda_2 = 1, \lambda_3 = 1$ 。总的训练轮数(epoch)设置为 300, 将 ICDAR2015 数据集预训练网络参数作为预训练权重参数, 训练使用的优化器是 Adam, 参数为默认值。前 50 轮为冻结训练, 冻结主干网络, 只对除主干网络的其他网络层数进行训练, 初始学习率设为  $1e-4$ , 批量大小(batch\_size)设为 4。解冻之后, 初始学习率设为  $1e-5$ , 批量大小设为 2。设定边界点和内部像素阈值为 0.9, 长边收缩比例为 0.6, 短边收缩比例为 0.3, 交并比(intersection over union, IOU)大于 0.5 时, 将预测边界框定义为正; 否则, 定义为负。

### 3.3 试验结果与对比分析

#### 1) 计算量与模型参数对比

为了说明本文算法在改进过程中轻量化效果, 在硬件平台相同的情况下, 比较以下 4 种模型的计算量和参数量。

(1) AdvancedEAST 模型;

(2) M-AdvancedEAST 模型: Mobilenetv3 作为主干的

AdvancedEAST 模型;

(3) ME-AdvancedEAST 模型: ES-Mobilenetv3 作为主干的 AdvancedEAST 模型;

(4) MED-AdvancedEAST 模型: ES-Mobilenetv3 作为主干且特征融合中引入深度可分离卷积的 AdvancedEAST 模型。

表 3 给出了 4 种模型的计算量和参数量对比结果。由表 3 可知, M-AdvancedEAST 模型相对于 AdvancedEAST 模型, 浮点数计算量降低 91%, 内存占用量降低 73.7%。结合 ECANet 且引入多维注意力机制后, ME-AdvancedEAST 模型的参数量和计算量进一步降低。将特征融合网络中引入深度可分离卷积, 此时轻量化后的 MED-AdvancedEAST 模型的内存占用量仅为 10.5 M, 相较于 AdvancedEAST 模型, 模型计算量降低 95.8%, 参数量降低 82.6%, 内存占用量降低 82.6%。

表 3 不同网络模型大小对比结果

模型	计算量/	参数量/	模型
	$10^{12}$	$10^6$	大小/M
AdvancedEAST	1.590	15.09	60.5
M-AdvancedEAST	0.141	3.905	15.9
ME-AdvancedEAST	0.141	3.849	15.7
MED-AdvancedEAST	0.066	2.700	10.5

#### 2) 消融试验结果

为了验证改进模块对整体网络性能的影响, 消融试验结果如表 4 所示。由表可看出, 虽然所提方法相较于 AdvancedEAST 算法在  $F1$  值方面有较小的降低, 但网络的检测速度提升了近 2 倍。使用 MobileNetv3 作为主干网络后, M-AdvancedEAST 网络由于特征提取能力变弱, 不能有效的提取到深层特征, 导致  $F1$  值有所降低, 但由于其轻量化特点, 帧率提高明显, 由 13.6FPS 升至 34.2FPS。在此基础上, 结合 ECANet 进行特征增强, 且在低层特征层使用多维注意力机制, ME-AdvancedEAST 模型  $F1$  值提升 0.9%。在特征融合网络中引入深度可分离卷积(DP 模块), MED-AdvancedEAST 模型精确率相较于 ME-AdvancedEAST 模型降低了 2.1%,  $F1$  值降低 2.8%, FPS 提升了 5.1%。使用 ICDAR2015 数据集训练预训练模型, 初始化 MED-AdvancedEAST 模型的主干网络权重, 使用迁移学习后在精确率、召回率和  $F1$  值均有了提升, 精确率提升了 1.4%, 召回率提升 6.8%,  $F1$  值提升了 4.1%。改进前后损失值变化曲线如图 9 所示, 模型经过迁移学习后, 模型收敛速度更快, 且模型精度更高。该结果证明本文提出算法在保证精度的前提下, 有效提升了检测效率。

#### 3) 不同检测模型对比

数字仪表数显区域具有明显的文本特点, 而 EAST、DBNet 算法在文本检测领域有着广泛的应用, 为了验证本

表 4 消融试验

Mobilenetv3	注意力机制	DP 模块	迁移学习	精确率/%	召回率/%	F1 值/%	FPS
×	×	×	×	89.6	96.0	92.7	13.6
√	×	×	×	87.3	90.0	88.6	34.2
√	√	×	×	89.1	90.0	89.5	35.8
√	√	√	×	87.0	86.5	86.7	40.9
√	√	√	√	88.4	93.3	90.8	40.5

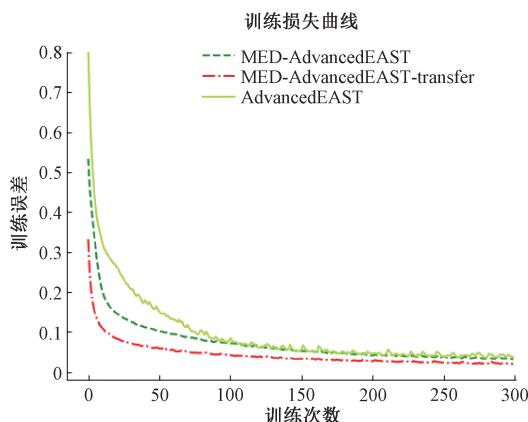


图 9 损失值变化曲线

文所提算法的性能,本文选用了基于分割方法的文本检测方法 DBNet 和基于目标检测的文本检测方法 EAST 分别在公开数据集和自制数字仪表数据集在相同的硬件环境下进行了试验对比。

采用 ICDAR2015 开源数据集作为提出模型的验证数据,ICDAR2015 数据集中包含路标、书籍封面和广告牌等场景文本图片。其中,训练集有 1 000 张,测试集有 500 张,每个图像中包含多个文字框,在标签文件中将文本框以坐标形式表示。主流文本检测算法与本文提出算法在 ICDAR2015 数据集的性能如表 5 所示。

在自制的数字电表数据集上对主流文本检测算法和本文所提算法进行训练,具体试验结果如表 6 所示。MED-AdvancedEAST 由于轻量化处理,模型参数相较其他模型明显减少,虽然特征提取能力有所减弱,但通过迁移学习策略训练,MED-AdvancedEAST 的调和平均值 F1 为 90.8%,相对于 AdvancedEAST 有所减小,但仍优于其

表 5 不同模型在公共数据集检测结果 %

模型	主干网络	精确率	召回率	F1 值
DBNet	Resnet18	87.0	75.1	80.6
EAST	VGG16	81.3	84.4	82.3
Advanced-EAST	VGG16	87.8	81.2	84.4
本文方案	ES-Mobilenetv3	86.4	78.8	82.4

他检测算法。本文算法检测精确率相比较 EAST 和 DBNet 分别提升了 12.5% 和 1.4%,参数量分别减少 82.1% 和 44.6%,计算量分别减少 95.8% 和 94.3%。试验结果验证了本文算法在保证了模型的精度前提下,有效降低运算量和内存占用量,符合变电站仪表实时检测需求。

表 6 不同模型的电表数显区域检测结果

模型	计算量/参数量/精确率/召回率/F1 值/			
	$10^{12}$	$10^6$	%	%
EAST	1.590	15.09	75.9	88.6
DBNet	1.157	3.905	87.0	92.3
AdvancedEAST	1.591	15.09	89.6	96.0
本文方案	0.066	2.700	88.4	93.3

改进网络 MED-AdvancedEAST 的检测效果图如图 10 所示,从图可以看出,对于多类别数字仪表数显区域,改进网络 MED-AdvancedEAST 均可对数显区域进行较好的检测。面对拍摄仪表存在倾斜、光照不同的复杂环境条件下,改进网络有较好的鲁棒性,能够准确检测出数显区域,检测能力较强。





图 10 数字仪表数显区域检测结果

获取数显区域检测结果后,采用裁剪的方式将数显区域分离出来,并利用预先训练好的文本识别 CRNN 模型进行数字字符的识别。试验结果如图 11 所示,可以看出所

有检测到的数显区域示数均能被准确识别,这进一步验证了本文所提出的算法的有效性和实用性,可以满足数字仪表数显区域检测的应用需求。



图 11 数字仪表检测和识别结果

## 4 结 论

针对当前主流目标检测算法对于数显区域检测不佳,基于轻量化网络 MobileNetv3 进行改进,结合 ECANet 进行特征增强,并在低维特征层引入多维注意力机制,构建 ES-MobileNetv3 网络作为 AdvancedEAST 的主干网络,降低模型复杂度和提高模型性能。在特征融合网络中引入深度可分离卷积,使得模型的计算量和参数量明显减小,提升模型的检测速率。将 MED-AdvancedEAST 在 ICDAR2015 数据集上训练的权重作为模型预训练权重,增强了训练效果,提升了小样本条件下电表读数识别的准确率。试验结果表明,所提算法在保证检测精度的情况下,能够实时、准确的实现数显区域检测,具有很好的实际应用价值。

## 参考文献

- [1] 蔡焕青,邵瑰玮,胡霁,等.变电站巡检机器人应用现状和主要性能指标分析[J].电测与仪表,2017,54(14):117-123.
- [2] 童枭军,陈洋,路浩,等.面向表计读数的变电站巡检机器人路径规划[J].电子测量与仪器学报,2022,36(8):167-177.
- [3] 肖佳.基于机器视觉的数字仪表自动读数方法研究[D].重庆:重庆大学,2017.
- [4] 苑玮琦,江海燕.基于 ACSS 角点定位的液晶显示仪表示值识别算法[J].计算机应用系统,2015,24(7):111-116.
- [5] 陈刚,胡子峰,郑超.基于特征检测的数字仪表数码快速识别算法[J].中国测试,2019,45(4):146-150.
- [6] 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: 779-788.
- [7] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[M]. Springer: Springer International Publishing, 2016: 21-37.
- [8] ZHOU C, ZHOU J, YU C, et al. Multi-channel sliced deep RCNN with residual network for text classification [J]. Chinese Journal of Electronics, 2020, 29(5): 880-886.
- [9] 刘华玲,皮常鹏,赵晨宇,等.基于深度域适应的跨域目标检测算法综述[J].计算机工程与应用,2023,59(8):1-12.
- [10] WAQAR M, WARIS M A, RASHID E, et al. Meter digit recognition via Faster R-CNN [C]. 2019, International Conference on Robotics and Automation in Industry(ICRAI), IEEE, 2019: 1-5.
- [11] 华泽玺,施会斌,罗彦,等.基于轻量级 YOLO-v4 模型的变电站数字仪表检测识别[J].西南交通大学学报,

- 2024, 59(1): 70-80.
- [12] 侯卓成, 欧阳华, 胡鑫, 等. 基于改进的 YOLOv4 彩色数字仪表读数识别方法[J]. 电子测量技术, 2022, 45(6): 124-129.
- [13] 周度. 基于深度学习的水表读数检测及 iOS 应用的设计与实现[D]. 广州: 华南理工大学, 2020.
- [14] LIAO M, WAN Z, YAO C, et al. Real-time scene text detection with differentiable binarization [C]. Proceedings of the AAAI Conference on Artificial Intelligence, New York, 2020, 34(7): 11474-11481.
- [15] TIAN Z, HUANG W L, HE T, et al. Detecting text in natural image with connectionist text proposal network[C]. European Conference on Computer Vision, Springer, Cham, 2016: 56-72.
- [16] ZHOU X Y, YAO C, WEN H, et al. EAST: An efficient and accurate scene text detector[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2017: 2642-2651.
- [17] 鲁乾鹏. 面向工业巡检的数字仪表智能识别技术的研究与应用[D]. 杭州: 浙江大学, 2021.
- [18] YIN F F, WANG J X, XIONG W, et al. Detection optimization of license plate targets based on AdvancedEAST [C]. Journal of Computational Methods in Sciences and Engineering, 2021, 1-9.
- [19] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3[C]. In Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, 1314-1324.
- [20] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020. 11531-11539.
- [21] HUO Y J. AdvancedEAST[EB/OL]. [2023-05-09]. <https://github.com/huoyijie/AdvancedEAST>.
- [22] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2018: 7132-7141.
- [23] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1251-1258.
- [24] 庄福振, 罗平, 何清, 等. 迁移学习研究进展[J]. 软件学报, 2015, 26(1): 26-39.
- [25] LIN T Y, GOYAL P, GIRHICK R, et al. Focal loss for dense object detection. Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy. 2017. 2999-3007.

### 作者简介

王嘉璇, 硕士研究生, 主要研究方向为电力设备自主巡检。

E-mail: 18339181029@163.com

王天宁(通信作者), 博士, 讲师, 主要研究方向为机器人智能感知、自主控制。

E-mail: wangtianning1989@126.com