



基于全局与局部多尺度上下文的电表数据检测

马天磊 符俊 马琪 杨震 刘新浩

Electric meter data detection based on global and local multi-scale context

MA Tianlei, FU Jun, MA Qi, YANG Zhen, LIU Xinhao

引用本文:

马天磊, 符俊, 马琪, 等. 基于全局与局部多尺度上下文的电表数据检测[J]. 应用光学, 2024, 45(4): 804–811. DOI: 10.5768/JAO202445.0403006

MA Tianlei, FU Jun, MA Qi, et al. Electric meter data detection based on global and local multi-scale context[J]. *Journal of Applied Optics*, 2024, 45(4): 804–811. DOI: 10.5768/JAO202445.0403006

在线阅读 View online: <https://doi.org/10.5768/JAO202445.0403006>

您可能感兴趣的其他文章

Articles you may be interested in

基于自适应像素级注意力模型的场景深度估计

Depth estimation based on adaptive pixel-level attention model

应用光学. 2020, 41(3): 490–499 <https://doi.org/10.5768/JAO202041.0302002>

一种基于三维形状上下文特征的点云配准算法

Point cloud registration algorithm based on 3D shape context features

应用光学. 2023, 44(2): 330–336 <https://doi.org/10.5768/JAO202344.0202005>

多局部残差连接注意网络的图像去模糊

Image deblurring based on multiple local residual connection attention network

应用光学. 2023, 44(2): 337–344 <https://doi.org/10.5768/JAO202344.0202006>

基于NPU的实时深度学习跟踪算法实现

Implementation of real-time deep learning tracking algorithm based on NPU

应用光学. 2022, 43(4): 682–692 <https://doi.org/10.5768/JAO202243.0402003>

基于多尺度残差注意力网络的水下图像增强

Underwater image enhancement based on multiscale residual attention networks

应用光学. 2024, 45(1): 89–98 <https://doi.org/10.5768/JAO202445.0102003>

基于深度学习激光熔覆层树枝晶的形貌识别

Morphology identification of dendrites of laser cladding layer based on deep learning

应用光学. 2022, 43(3): 532–537 <https://doi.org/10.5768/JAO202243.0307001>



关注微信公众号，获得更多资讯信息

文章编号:1002-2082 (2024) 04-0804-08

基于全局与局部多尺度上下文的电表数据检测

马天磊,符俊,马琪,杨震,刘新浩

(郑州大学电气与信息工程学院,郑州 450001)

摘要:电力系统中配电箱的电表数据检测为电力管理和安全运行提供了重要的数据支持。传统的人工电表数据读取方法效率低下且易出错,而现有深度学习方法因模型参数量大限制了模型的应用。针对上述问题,提出了一种轻量化鲁棒的实时电表检测方法。通过减少特征提取网络的层数和通道数,减少模型的参数量,实现网络的轻量化。在减少网络参数量的同时,为了保证网络的特征表达能力和拟合能力,引入全局上下文和局部多尺度上下文丰富目标特征表达。全局上下文关注电表数据在电表箱中的位置,局部多尺度上下文适应不同尺寸的电表数据。实验结果表明,所提网络在参数量更小的情况下,仍能获得比其他检测方法更高的准确率和更快的检测速度。

关键词:电表数据检测;全局上下文;局部上下文;深度学习;注意力机制

中图分类号:TN201; TP391.41

文献标志码:A

DOI: 10.5768/JAO202445.0403006

Electric meter data detection based on global and local multi-scale context

MA Tianlei, FU Jun, MA Qi, YANG Zhen, LIU Xinhao

(School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract: The detection of electricity meter data in the distribution box of a power system provides important data support for power management and safe operation. Traditional manual methods for reading electricity meter data are inefficient and prone to errors, while existing deep learning methods are limited in model application due to large model parameter sizes. To address these issues, a lightweight and robust real-time electricity meter detection method was proposed. The parameter size of the model was reduced by reducing the number of layers and channels in the feature extraction network, and the lightweight of the network was achieved. While reducing the network parameter size, the global context and local multi-scale context were introduced to ensure the network feature representation and fitting capabilities. The global context focused on the position of electricity meter data in the meter box, while the local multi-scale context adapted to different sizes of meter data. Experimental results show that the proposed network achieves higher accuracy and faster detection speed than other detection methods, even with smaller parameter sizes.

Key words: meter data detection; global context; local context; deep learning; attention mechanism

引言

电力系统是现代社会不可或缺的基础设施,而配电箱则是电力系统中一个非常重要的组成部分。在配电箱中,各种电表记录着配电箱内电力系统的实时状态,为电力管理和安全运行提供了重要的数据支持。然而,手动读取电表数据方法

效率低下且易出错,难以满足快速准确的读取需求。因此,开发一种高效准确的配电箱电表数据读取方法,已经成为电力系统管理中的一个重要内容^[1-5]。

传统的电表数据读取方法通常首先采用阈值分割或滤波的方法将数据从图像中分离出来,然

收稿日期:2023-07-07; 修回日期:2023-11-17

基金项目:国家自然科学基金 (62373330); 河南省本科高校青年骨干教师培养计划 (2023GGJS005)

作者简介:马天磊 (1989—),男,博士,教授,主要从事红外小目标检测、图像处理、模式识别等研究。E-mail: tlma@zzu.edu.cn

通信作者:符俊 (1995—),男,硕士,主要从事红外弱小目标检测和图像处理研究。E-mail: yunke120@gs.zzu.edu.cn

后使用分类模型对数据进行分类^[6]。例如, 沈美丽^[7]首先通过特有图像特征对单个数字进行分割, 再采用 K 最近邻分类算法判断数字类别; 尹建丰等人^[8]对图像进行滤波, 提取单个数字图像, 然后使用 SVM 对单个数字图像进行识别。这些传统方法相对于手动读取数据有着更快的速度, 但是并没有形成端到端的系统。同时, 由于采用阈值分割或者滤波提取数据, 在准确率上存在很大不足, 特别是当数据中存在小数点的情况下, 很难有效检测到小数点。

近年来, 深度学习技术在目标检测等领域取得了显著的进展, 其中有代表性的包括单阶段检测网络如 YOLO 系列^[9,10], SSD^[11] 和 EfficientDet^[12]; 两阶段检测网络如 R-CNN^[13], Faster R-CNN^[14] 和 Mask R-CNN^[15]。基于深度学习的目标检测模型可以有效地识别和提取配电箱内各电表的数据, 并对其进行快速准确地读取。高雪飞等人^[16]提出了一种基于 Faster R-CNN 的检测网络, 实现了电表读数的自动读取; 张俊等人^[17]在 YOLOv3 的基础上加入了注意力机制, 先检测图像中的表盘, 然后在所检测出的表盘图像部分检测指针和刻度值, 实现了电表数值的读取; SRIPANUSKUL N 等人^[18]使用一种数据扩充方法生成带标签信息的数据, 然后使用生成对抗网络进行数据的识别; LAROCAR^[19]等人设计了一种两阶段的方法, 使用 Fast-YOLO 检测器检测数据, 也取得了不错的效果。但是, 这些深度学习模型在应用过程中存在模型过大、计算量大、检测准确性低等问题, 限制了其在实际场景中的应用。因此, 轻量化的深度学习模型成为电力系统应用中的一个研究热点。

本文旨在提出一种轻量化电表数据检测网络, 该网络可以实现配电箱内各电表数据的实时检测和读取, 以便进行电力管理和安全运行的决策支持。本文的主要贡献在于提出了一种基于深度学习的轻量化鲁棒检测模型, 用于实现配电箱内各电表数据的实时识别和读取。该网络模型通过引入目标的全局上下文与局部多尺度上下文特征来丰富目标的特征表达, 由此来减少特征提取所需的网络深度, 从而减少模型拟合所需的参数量。并通过实验证明了其在检测准确性、运行速度和模型大小等方面的优越性。

1 轻量化电表数据检测网络

卷积神经网络通过大量的卷积、池化等操作

提取语义信息, 这些操作可以在一定程度上提高模型的表达能力, 但往往需要较深的层数才能提取到目标的语义信息。然而, 太深的卷积神经网络需要大量的计算和推理时间, 不利于实际应用。因此, 为了减少网络深度和模型所需的参数量、加快网络训练与推理速度, 提出了一种轻量化电表数据检测网络, 通过引入目标的全局上下文与局部多尺度上下文特征来丰富目标的特征表达, 从而增加网络的宽度和提高并行计算的效率, 加速模型的训练和推断过程。具体网络结构如图 1 所示, 主要包括特征提取, 全局与局部上下文特征融合和最终的检测头 3 个部分。

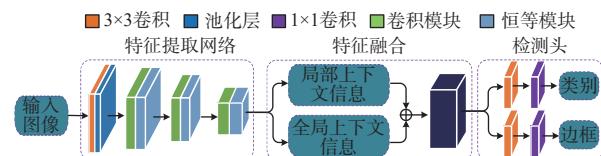


图 1 网络总体框架图

Fig. 1 Overall framework diagram of network

1.1 特征提取网络

特征提取网络是深度学习中的一种网络结构, 用于从输入数据中提取有用的特征。这些特征可以用于分类、回归、分割等任务。特征提取网络通常由多个卷积层和池化层组成, 可以自动学习输入数据中的特征。这些层会逐渐减小特征图的大小, 并增加特征的数量和复杂度。

特征提取网络对检测网络速度有重要的影响。一方面, 特征提取网络的深度和复杂度会影响特征的表达能力, 从而影响检测网络的精度。另一方面, 特征提取网络的计算量和参数量也会影响检测网络的速度。通常情况下, 特征提取网络的深度和复杂度越高, 其计算量和参数量也越大, 从而导致检测网络的速度变慢。而在电表读数检测中, 为了能够实时检测电力系统的状态, 对检测网络的速度有很高的要求, 因此具备一种轻量化的特征提取网络是十分必要的。具体网络结构数据如表 1 所示。

首先使用一个卷积核大小为 3×3 的卷积层和一个池化层, 对输入的图片进行初步特征提取。然后将得到的特征进一步送入 3 个连续的卷积模块和恒等模块, 卷积模块的作用是改变网络的维度和特征层的大小, 恒等模块的作用是提高网络的拟合能力^[20]。值得注意的是, 第一个卷积模块卷积步长为 1, 仅仅改变网络的维度, 不改变特征层

的大小。卷积模块的结构如图 2(a)所示, 恒等模块的结构如图 2(b)所示。

表 1 特征提取网络的结构数据

Table 1 Structure data of feature extraction network

层	输入	输出
3×3 卷积层	(1280, 720, 3)	(640, 360, 8)
池化层	(640, 360, 8)	(320, 180, 8)
卷积模块1	(320, 180, 8)	(320, 180, 16)
恒等模块1	(320, 180, 16)	(320, 180, 16)
卷积模块2	(320, 180, 32)	(160, 90, 64)
恒等模块2	(160, 90, 64)	(160, 90, 64)
卷积模块3	(160, 90, 64)	(80, 45, 128)
恒等模块3	(80, 45, 128)	(80, 45, 128)

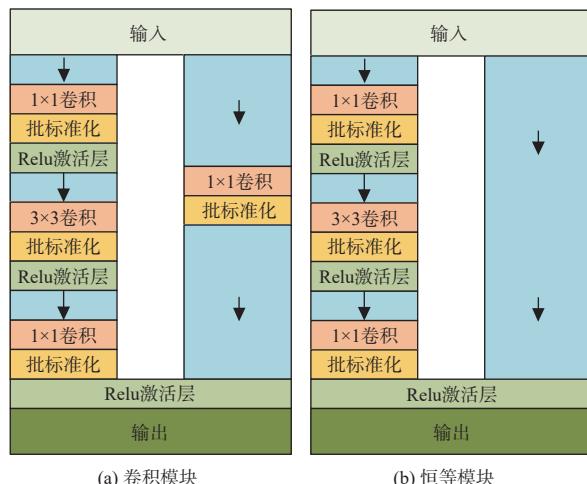


图 2 卷积模块与恒等模块结构图

Fig. 2 Structure diagram of convolutional module and identity module

1.2 全局与局部多尺度上下文特征融合

在特征提取网络对图片特征进行提取后, 为了提升网络拟合能力和对不同场景的适应性, 我们将使用空洞卷积获得的局部多尺度上下文信息与使用通道、空间双注意力机制获得的全局上下文信息进行融合。

空洞卷积(dilated convolution)^[21]可以扩大卷积核的感受野, 从而增加网络的感知范围。在传统的卷积操作中, 卷积核的每个权重只与相邻的像素进行卷积计算, 而空洞卷积通过在卷积核中插入空洞(即0值), 使得卷积核可以跨越更大的距离进行计算。空洞卷积的具体实现方式是, 在卷积核的权重之间插入一定数量的0值, 从而增加卷积核的感受野。例如, 将一个大小为3×3的卷积核插

入一个空洞, 就可以得到一个大小为5×5的卷积核。通过将不同空洞率的卷积核获得的特征进行融合, 便可以获得目标的局部多尺度上下文信息。局部多尺度上下文信息提取模块结构如图3所示。

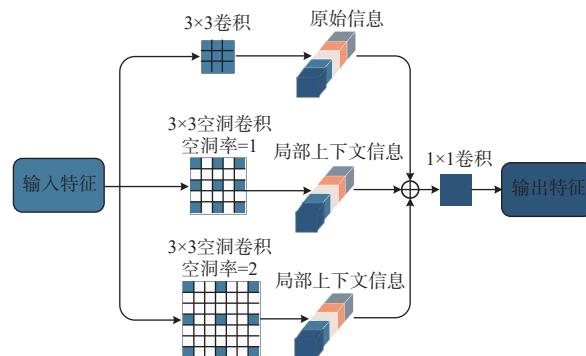


图 3 局部多尺度上下文信息提取模块结构图

Fig. 3 Structure diagram of local multi-scale context information extraction module

局部多尺度上下文信息提取计算如式(1)和式(2)所示:

$$F_{\text{out}} = \text{Conv}_{1 \times 1} \text{Concat}(F_{c1}, F_{d1}, F_{d2}) \quad (1)$$

$$\begin{cases} F_{c1} = \text{Conv}_{3 \times 3}(F_{\text{in}}) \\ F_{d1} = \text{DConv}_{(3 \times 3, r=1)}(F_{\text{in}}) \\ F_{d2} = \text{DConv}_{(3 \times 3, r=2)}(F_{\text{in}}) \end{cases} \quad (2)$$

式中: F_{out} 是局部多尺度上下文信息提取模块的输出; F_{in} 是特征提取模块的输出; $\text{Conv}_{1 \times 1}(\cdot)$ 和 $\text{Conv}_{3 \times 3}(\cdot)$ 分别是卷积核大小为1和3的卷积; $\text{DConv}_{(3 \times 3, r=1)}(\cdot)$ 是空洞率为1的空洞卷积; $\text{DConv}_{(3 \times 3, r=2)}(\cdot)$ 是空洞率为2的空洞卷积。

使用空洞卷积提取局部多尺度上下文信息, 可以很好地捕捉到目标的特征, 但是却没能捕捉到图像的全局信息, 导致模型的泛化能力不足。而电表数据在配电箱中的位置往往是有规律的, 因此为了帮助模型能够应对更复杂的场景, 进一步提高模型的泛化能力, 我们使用空间和通道双注意力机制来提取图像的全局上下文信息^[22]。具体结构如图4所示。

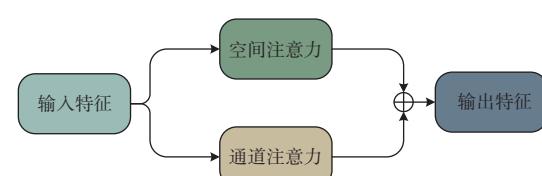


图 4 全局上下文信息提取模块结构图

Fig. 4 Structure diagram of global context information extraction module

其中, 空间注意力模块的结构如图 5 所示。 A 表示输入特征, 形状为 $C \times H \times W$, 首先对 A 进行卷积获得 3 个形状为 $C \times N$ 的新的特征 B, C, D , 其中 $N=H \times W$; 然后对 C 和 B 进行转置矩阵乘法, 并通过 Softmax 函数计算获得空间注意力图 S , 形状为 $N \times N$; 再将 D 和 S 2 个矩阵相乘, 并将结果形状转换为 $C \times H \times W$; 最后, 将获得的空间特征乘以可学习参数 α , 并与输入特征 A 相加获得输出特征。

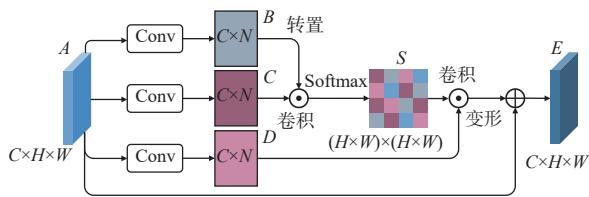


图 5 空间注意力模块结构图

Fig. 5 Structure diagram of spatial attention module

通道注意力模块的结构如图 6 所示。与空间注意力不同的是, 通道注意力直接通过输入特征 A 来计算通道注意力图 X 。首先将 A 的形状转换为 $C \times N$, 然后将 A 与 A 的转置相乘并进行 Softmax 函数计算, 获得形状为 $C \times C$ 的注意力图 X 。最终, 将 X 与 A 相乘, 并将结果形状转换为 $C \times H \times W$, 与可学习参数 β 相乘后再与 A 相加, 获得输出特征。

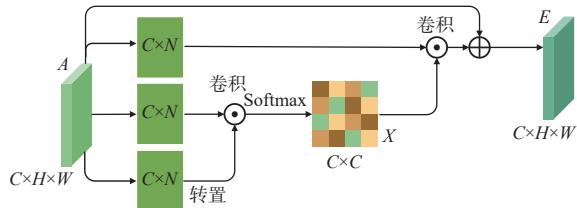


图 6 通道注意力模块结构图

Fig. 6 Structure diagram of channel attention module

在获得空间注意力特征 E_1 和通道注意力特征 E_2 后, 将两者相加得到全局上下文特征 E_{out} 。最终的全局与局部多尺度上下文特征 I 表示为

$$I = F_{\text{out}} + E_{\text{out}} \quad (3)$$

最终, 我们将全局与局部多尺度上下文特征图送入检测头进行下一步的预测。

1.3 检测头与损失函数

为了适应不同大小的电表数字, 我们在特征图上的每个点预设 3 个尺度的锚框进行回归。使用一个 3×3 的卷积和 1×1 的卷积将特征图转换成 $(80 \times 45 \times 3 \times 16)$ 大小的张量, 其中 80×45 为特征图的大小, 3 为每个特征点的锚框数量, 16 为 $4+1+11$, 其中 4 是预测框中心点坐标的横向偏移、纵向偏

移和宽高, 1 是预测框内包含目标的置信度, 11 为预测的类别。对这个张量进行解码, 便能得到预测结果。预测框进行非极大值抑制得到最终的边界框, 本文取非极大值抑制的阈值为 0.4。我们的损失函数包含三部分, 分别是类别损失, 置信度损失和位置损失。

$$L(o, c, O, C, t, g) = \lambda_1 L_{\text{conf}}(o, c) + \lambda_2 L_{\text{cla}}(O, C) + \lambda_3 L_{\text{loc}}(t, g) \quad (4)$$

式中: $\lambda_1, \lambda_2, \lambda_3$ 为平衡系数; $L_{\text{conf}}(o, c)$, $L_{\text{cla}}(O, C)$, $L_{\text{loc}}(t, g)$ 分别为置信度损失、类别损失和位置损失。

$$L_{\text{conf}}(o, c) = -\frac{\sum_i (o_i \ln(\hat{c}_i) + (1 - o_i) \ln(1 - \hat{c}_i))}{N} \quad (5)$$

$$\hat{c}_i = \text{Sigmoid}(c_i)$$

式中: o 表示预测目标边界框与真实目标边界框的 IOU; c 为预测值; \hat{c}_i 为 c 通过函数得到的预测置信度; N 为正负样本总个数。

$$L_{\text{cla}}(O, C) = -\frac{\sum_{i \in \text{pos}} \sum_{j \in \text{cla}} (O_{ij} \ln(\hat{C}_{ij}) + (1 - O_{ij}) \ln(1 - \hat{C}_{ij}))}{N_{\text{pos}}} \quad (6)$$

$$\hat{C}_{ij} = \text{Sigmoid}(C_{ij}) \quad (6)$$

式中: O_{ij} 表示预测框 i 中是否含有第 j 类目标; C_{ij} 为预测值; N_{pos} 为正样本个数。

$$L_{\text{loc}}(t, g) = \frac{\sum_{i \in \text{pos}} (\sigma(t_x^i) - \hat{g}_x^i)^2 + (\sigma(t_y^i) - \hat{g}_y^i)^2 + (t_w^i - \hat{g}_w^i)^2 + (t_h^i - \hat{g}_h^i)^2}{N_{\text{pos}}} \quad (7)$$

$$\left\{ \begin{array}{l} \hat{g}_x^i = g_x^i - c_x^i \\ \hat{g}_y^i = g_y^i - c_y^i \\ \hat{g}_w^i = \ln(g_w^i / p_w^i) \\ \hat{g}_h^i = \ln(g_h^i / p_h^i) \end{array} \right.$$

式中: $t_x^i, t_y^i, t_w^i, t_h^i$ 为网络预测的回归参数; $g_x^i, g_y^i, g_w^i, g_h^i$ 为真实值的中心点坐标和宽高。

2 实验结果与分析

2.1 数据集与实验设置

由于本文针对的是变电站仪表检测, 场景较为特殊, 并没有大量公开可用的专业数据集, 因此在实验中采取了网上收集和自行制作的多种电表数据集, 四种不同的电表图像序列如图 7 所示。在经过筛选之后, 进行图像裁剪、添加噪声、对比度处理等基本图像扩充手段, 一共制作了 850 张图片, 利用 LabelImg 对其进行手工标注, 共标注 11 个类

别,包括 0 到 9 和小数点。我们将训练集和测试集分为 5:5, 6:4, 7:3, 8:2 和 9:1 这 5 个比例进行训练和测试,结果如表 2 所示。从表 2 可以看出,当训练集和测试集的比例为 8:2 时,训练得到的模型在测试集上的泛化能力最强。因此,本文将训练集和测试集的比例设置为 8:2。

表 2 不同训练比例模型 11 个类别的 AP 和 M_{AP}
Table 2 AP and M_{AP} for 11 categories of models with different training ratios

比例	类别											
	0	1	2	3	4	5	6	7	8	9	小数点	M_{AP}
5:5	0.93	0.98	0.93	0.96	0.91	0.95	0.92	0.96	0.91	0.97	0.91	0.94
6:4	0.92	0.91	0.94	0.94	0.97	0.90	0.98	0.99	0.93	0.99	0.93	0.95
7:3	0.95	0.98	0.96	0.93	0.94	0.94	0.99	0.95	0.96	0.97	0.94	0.96
8:2	0.99	0.99	0.99	0.99	0.99	0.95	0.99	0.99	0.99	0.99	0.95	0.98
9:1	0.96	0.98	0.98	0.97	0.98	0.96	0.97	0.96	0.96	0.97	0.96	0.97

我们的实验是在 NVIDIA 3080Ti GPU 上使用 Pytorch 1.7.1 进行的。在训练过程中,我们将批量大小设置为 8 个图像,并使用 Adam 优化器。我们对网络进行了总共 200 次迭代训练,前 50 次迭代的学习率设置为 0.001,然后在每次后续迭代中乘以 0.99。在预测过程中,我们首先使用类别的置信度得分阈值 0.001 来过滤检测结果,丢弃类别置信度得分小于 0.001 的任何检测框。然后,我们使用阈值为 0.7 的非最大值抑制(NMS)来确保每个目标只有一个检测框。

2.2 评价指标

为了验证我们所提出的检测方法的优越性,我们使用 R (召回率)、 P (准确率)和 M_{AP} 来验证模型的准确度,使用模型每秒钟处理的图像帧数(FPS)来验证模型的速度。这些评价指标的定义如下:

$$P = \frac{TP}{TP + FP} \quad (8)$$

$$R = \frac{TP}{TP + FN} \quad (9)$$

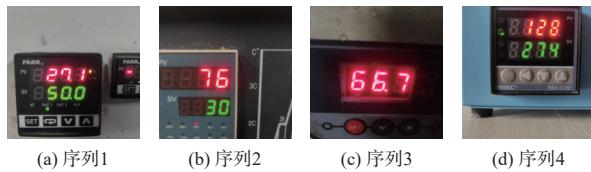


图 7 数据集中部分图像

Fig. 7 Partial images in dataset

$$P_{AP} = \int_0^1 PR \, dr \quad (10)$$

$$M_{AP} = \frac{\sum_{q=1}^Q AP(q)}{Q} \quad (11)$$

式中: TP 表示将正类预测为正类的数目; FP 表示将正类预测为负类的数目; FN 表示将正类预测为负类的数目; P_{AP} 表示通过改变判断结果是否为正类的阈值获得的一系列 P 、 R 值连成的曲线与 X 轴围成的面积; Q 表示检测的类别数量; M_{AP} 表示所有类别 P_{AP} 值的平均值。

2.3 实验分析

为了验证本文提出方法相对于其他方法的优越性,我们分别与传统电表数据检测方法(ITOIS^[8]),经典目标检测算法(YOLO-v3^[9], YOLO-v5^[10] 和 Faster-RCNN^[14])以及基于深度学习的电表数据检测算法(ESPA-YOLOv3^[17])进行了对比试验,实验参数设置如表 3 所示。

表 3 对比实验参数设置
Table 3 Parameter settings of comparison experiment

模型	参数
ITOIS ^[8]	滤波方法(均值滤波), 模板窗口(3, 3), 输入尺寸(30, 0)
YOLO-v3 ^[9]	输入尺寸(416, 416), 主干(Darknet53), 优化器(SGD), 批次大小(8)
Faster-RCNN ^[14]	输入尺寸(600, 480), 主干(Resnet50), 优化器(Adam), 批次大小(2)
YOLO-v5 ^[10]	输入尺寸(640, 640), 主干(CSPDarknet), 优化器(SGD), 批次大小(8)
ESPA-YOLOv3 ^[17]	输入尺寸(416, 416), 主干(ESPAnet), 优化器(SGD), 批次大小(8)

图 8 给出了部分实验结果, 模型得到的数据如图 9 所示。

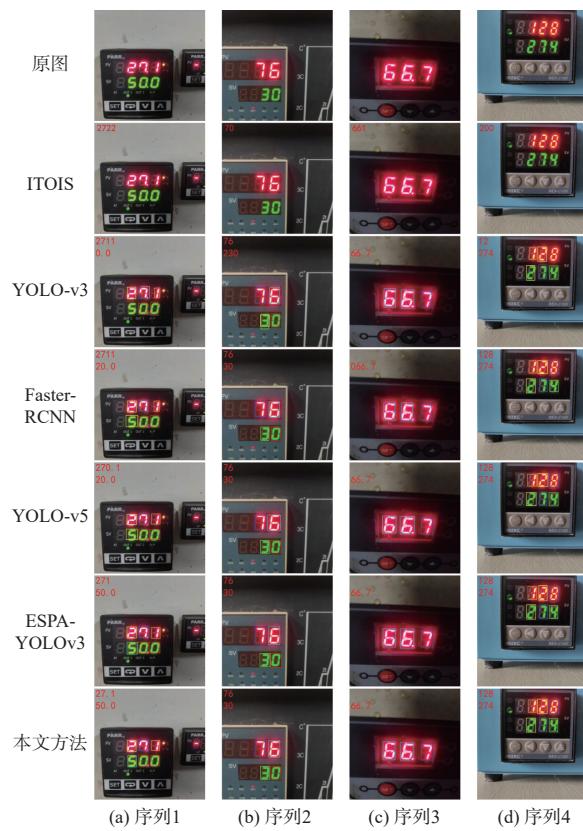


图 8 检测结果图像

Fig. 8 Detection result images

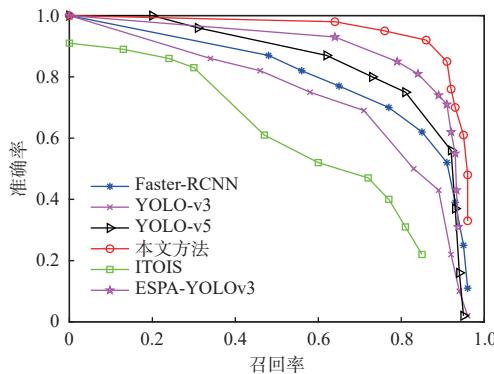


图 9 P-R 曲线

Fig. 9 Curves of P-R values

图 9 中可以看出, 传统的电表数据在多行数据的情况下难以进行有效检测, 这是由于传统检测方法缺少对数字的定位能力。另外, 基于阈值分割的数据提取方法在小数点的识别上准确率也很低。而其他基于深度学习的方法由于网络过深, 对小数点的识别也效果不佳, 并且由于网络缺乏对全局与局部上下文语义信息的提取能力, 对电

表数据识别的准确率不如本文提出的方法。具体性能指标如表 4 所示, 可以看出本文提出的方法在所有类别上的准确率均为最高, 特别是对小数点的检测远高于其他模型。

表 4 不同模型 11 个类别的 AP 和 M_{AP}

Table 4 AP and M_{AP} for 11 categories of different models

模型	类别					
	0	1	2	3	4	5
ITOIS ^[8]	0.73	0.68	0.53	0.76	0.51	0.75
YOLO-v3 ^[9]	0.82	0.71	0.64	0.61	0.87	0.80
Faster-RCNN ^[14]	0.74	0.88	0.76	0.83	0.84	0.94
YOLO-v5 ^[10]	0.89	0.90	0.83	0.86	0.92	0.93
ESPA-YOLOv3 ^[17]	0.94	0.91	0.96	0.97	0.93	0.89
本文方法	0.99	0.99	0.99	0.99	0.99	0.95

模型	类别					
	6	7	8	9	小数点	M_{AP}
ITOIS ^[8]	0.62	0.46	0.71	0.67	0.39	0.62
YOLO-v3 ^[9]	0.78	0.69	0.81	0.69	0.50	0.72
Faster-RCNN ^[14]	0.89	0.85	0.86	0.87	0.56	0.83
YOLO-v5 ^[10]	0.86	0.96	0.93	0.95	0.71	0.89
ESPA-YOLOv3 ^[17]	0.97	0.93	0.95	0.96	0.78	0.92
本文方法	0.99	0.99	0.99	0.99	0.95	0.98

表 5 展示了不同检测模型的模型大小和检测速度。其中, 传统方法不同于深度学习方法, 不存在模型大小, 表中仅给出了运行速度。从表 5 可以看出, 本文所提出的方法具有极其轻量化的特点, 凭借模型的轻量化, 我们的模型相对于其他模型具有更高的检测速度。传统方法由于步骤复杂, 因此检测速度低于深度学习方法。同时还可以看出, 单阶段的深度学习方法(YOLO-v3, YOLO-v5, SPA-YOLOv3 和本文模型)要快于两阶段的深度学习方法(Faster-RCNN)。

表 5 不同检测模型参数量大小和运行速度

Table 5 Parameter size and running speed of different detection models

模型	参数量/Mb	FPS
ITOIS ^[8]	—	0.5
YOLO-v3 ^[9]	130	22
Faster-RCNN ^[14]	108	5
YOLO-v5 ^[10]	122	31
ESPA-YOLOv3 ^[17]	135	20
本文模型	11.03	43

为了验证引入全局上下文、局部多尺度上下文的有效性,我们将本文方法与单独去掉全局上下文的网络、单独去掉局部上下文的网络和同时去掉全局和局部上下文的网络进行消融实验,实验结果如表 6 所示。从表 6 中可以看出,在单独加入全局上下文和局部上下文后,模型的准确率得到了很大的提升,但是检测速度略有下降,局部上下文对准确率的提升要优于全局上下文,但是速度上要比全局上下文慢。

表 6 消融实验结果

Table 6 Results of ablation experiment

全局上下文	局部上下文	M_{AP}	FPS
—	—	0.87	48
√	—	0.92	46
—	√	0.94	45
√	√	0.98	43

尽管本文提出的算法在彩色图像上的检测效果令人满意,但在灰度图像上的表现仍有待提升,如检测结果图 10 所示。这是由于灰度图相对于彩色图像缺乏了颜色特征,导致特征提取不足,因此检测效果差。在实际应用中,由于采集设备、环境光照等多种因素的影响,图像的颜色信息可能会失真或缺失。特别是在光线暗淡或不均匀的环境下,图像的颜色信息往往难以准确捕捉。此外,低质量的图像采集设备也可能导致颜色细节的丢失。在这种情形下,如何有效维持检测算法的性能,将是我们未来研究的重要方向。我们期待通过进一步的研究和改进,使得算法在各种图像环境下都能发挥出色的检测效果。



图 10 灰度图检测结果

Fig. 10 Detection results of grayscale image

3 结论

本文提出了一种轻量化鲁棒的实时电表检测网络。该网络的特征提取部分具有更少的层数和通道数,减少了网络的参数量,提高了网络的速度和可部署性。另外,该网络通过引入全局上下文

与局部多尺度上下文特征来丰富目标的特征表达,增加网络宽度并提高并行计算效率,加速模型的训练和推断过程。实验结果表明,该网络在保证检测精度的同时提高了网络模型的可部署性和模型的检测速度,可以满足电力系统实时状态监控的需求。

参考文献:

- [1] 张慧. 基于轮廓提取的电能表视觉识别技术[J]. 自动化应用, 2011(8): 52-54.
ZHANG Hui. Watt-hour meter visual identity technology based on contour extraction[J]. Automation Application, 2011(8): 52-54.
- [2] 王舒憬, 陈凯. 基于图像处理的智能电表显示缺陷自动检测系统设计[J]. 电测与仪表, 2016, 53(4): 63-68.
WANG Shujing, CHEN Kai. Automatic testing system design of smart meter based on computer vision[J]. Electrical Measurement & Instrumentation, 2016, 53(4): 63-68.
- [3] 陈英, 蒋文峰, 杨丰玉, 等. 智能电能表数字识别算法研究[J]. 信息通信, 2020, 33(1): 17-21.
CHEN Ying, JIANG Wenfeng, YANG Fengyu, et al. Research on digits recognition algorithm of intelligent electric meter[J]. Information & Communications, 2020, 33(1): 17-21.
- [4] 杨娟. 基于数字图像处理的电能表图像识别技术研究与实现[D]. 南京: 南京理工大学, 2012.
YANG Juan. Research and implementation of image recognition technology of watt-hour meter based on digital image processing[D]. Nanjing: Nanjing University of Science and Technology, 2012.
- [5] 曾伟铭. 基于数字图像处理技术的智能电表读数识别方法研究[J]. 黑龙江电力, 2022, 44(5): 467-470.
ZENG Weiming. Research on reading recognition method of intelligent electric meter based on digital image processing technology[J]. Heilongjiang Electric Power, 2022, 44(5): 467-470.
- [6] 潘毓生. 基于图像处理的智能电表读数识别方法的研究[D]. 天津: 天津理工大学, 2021.
PAN Yusheng. Research on the reading recognition method of smart meter based on image processing[D]. Tianjin: Tianjin University of Technology, 2021.
- [7] 沈美丽. 数字电表图像的检测与识别[J]. 现代电子技术, 2022, 45(16): 110-114.

- SHEN Meili. Detection and recognition of digital ammeter image[J]. Modern Electronics Technique, 2022, 45(16): 110-114.
- [8] 尹建丰, 卫鑫, 顾雄伟, 等. 基于图像阈值优化及改进 SVM 的电表数字识别[J]. *计算机与现代化*, 2023(5): 106-110.
- YIN Jianfeng, WEI Xin, GU Xiongwei, et al. Digital identification of electric meter based on image threshold optimization and improved SVM[J]. *Computer and Modernization*, 2023(5): 106-110.
- [9] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08) [2023-07-05]. <https://pjreddie.com/media/files/papers/YOLOv3.pdf>.
- [10] WU W T, LIU H, LI L L, et al. Application of local fully convolutional neural network combined with YOLOv5 algorithm in small target detection of remote sensing image[J]. *PLoS One*, 2021, 16(10): e0259283.
- [11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[M]//Computer Vision–ECCV 2016. Cham: Springer International Publishing, 2016: 21-37.
- [12] TAN M X, PANG R M, LE Q V. EfficientDet: scalable and efficient object detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020: 10781-10790.
- [13] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014: 580-587.
- [14] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [15] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE, 2017: 2961-2969.
- [16] 高雪飞, 喻伟, 程飞, 等. 基于卷积神经网络的指针电表读数检测方法研究[J]. *电子乐园*, 2021(12): 116-117.
- GAO Xuefei, YU Wei, CHENG Fei, et al. Research on pointer electric meter reading detection method based on convolutional neural network[J]. *Electronics Garden*, 2021(12): 116-117.
- [17] 张俊, 杨光, 胡东升, 等. 基于注意力机制与 YOLOv3 的电表自动读数算法[J]. *电力信息与通信技术*, 2021, 19(12): 82-87.
- ZHANG Jun, YANG Guang, HU Dongsheng, et al. An automatic reading algorithm of power meter based on YOLOv3 and attention mechanism[J]. *Electric Power Information and Communication Technology*, 2021, 19(12): 82-87.
- [18] SRIPANUSKUL N, BUAYAI P, MAO X Y. Generative data augmentation for automatic meter reading using CNNs[J]. *IEEE Access*, 2022, 10: 28471-28486.
- [19] LAROCA R, BARROSO V, DINIZ M A, et al. Convolutional neural networks for automatic meter reading[J]. *Journal of Electronic Imaging*, 2019, 28(1): 1-14.
- [20] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 770-778.
- [21] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[EB/OL]. (2016-04-30) [2023-07-05]. <https://arxiv.org/pdf/1511.07122.pdf>.
- [22] FU J, LIU J, TIAN H J, et al. Dual attention network for scene segmentation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019: 3146-3154.