



# 面向电力通信配电屏数字化的人工智能识别关键方法

刘林, 杨志敏, 黄强, 杨经纬, 陈一童

(中国南方电网电力调度控制中心, 广东 广州 510525)

**摘要:** 为了提高电力通信网的运维管理效率, 研究利用人工智能技术的通信电源系统配电屏数字化方法, 通过目标检测和文字识别来实现配电屏中各下级支路供电状态的实时监测。首先, 提出了一个多层嵌套识别网络 (multi-layer nested recognition network, MLNRN) 架构, 结合轻量化策略, 降低该架构对终端设备的存算需求, 高效准确地实现了配电屏供电状态的结构化输出。其次, 针对图元检测任务提出了一个改进的YOLOv5网络, 通过引入ConvNext Block和双向特征金字塔网络 (bidirectional feature pyramid network, Bi-FPN) 结构, 提高了对状态灯等小目标的识别精度。最后, 基于卷积递归神经网络-连接时态分类 (convolutional recurrent neural network-connectionist temporal classification, CRNN-CTC) 网络构建针对配电屏下级支路标签的文本识别模型, 采用迁移学习和图像增广的策略, 提升模型对多字符种类、不规范配电屏图像中文本识别的准确率。仿真实验表明, 图像识别的平均准确率为97.2%, 文本方法的识别准确率为92%, 这验证了所提架构在配电屏数字化任务中的有效性和适用性, 为实现电力通信网的视频巡检和智慧运维提供了有效的解决方案。

**关键词:** 电力通信网; 配电屏数字化; 图像识别; 文字识别; YOLOv5; 轻量化; 卷积递归神经网络

**中图分类号:** TM73

**文献标志码:** A

**doi:** 10.11959/j.issn.1000-0801.2025041

## Key method of digitization of power distribution panel with artificial intelligence identification for power communication network

LIU Lin, YANG Zhimin, HUANG Qiang, YANG Jingwei, CHEN Yitong

Power Dispatching & Control Center of China Southern Grid, Guangzhou 510525, China

**Abstract:** To enhance the operational management efficiency of the power communication network, artificial intelligence technology was utilized to develop a digitalization method for distribution panels within the communication power system. By implementing object detection and text recognition, real-time monitoring of the power supply status for each subordinate branch in the distribution panel was achieved. Firstly, a multi-layer nested recognition network (MLNRN) architecture was proposed, which incorporated lightweight strategies to reduce the computational demands

收稿日期: 2024-11-14; 修回日期: 2024-12-17

通信作者: 刘林, liulin2@csg.cn

基金项目: 南方电网重点科技项目 (No.ZDKJXM20222029)

**Foundation Item:** The Key Technology Project of China Southern Power Grid (No.ZDKJXM20222029)

on terminal devices, allowing for efficient and accurate structured output of the distribution panel's power supply status. Secondly, an improved YOLOv5 network was introduced for the task of icon detection. By integrating ConvNext Block and bidirectional feature pyramid network (Bi-FPN) structures, the recognition accuracy for small targets, such as status lights, was significantly enhanced. Finally, a text recognition model targeting the labels of subordinate branches in the distribution panel was constructed on the basis of the convolutional recurrent neural network-connectionist temporal classification (CRNN-CTC). Transfer learning and image augmentation strategies were employed to improve recognition accuracy for text with multiple character types in non-standard distribution panel images. Simulation experiments demonstrate that the average accuracy of image recognition is 97.2%, whereas the accuracy of text recognition is 92%. These results validate the effectiveness and applicability of the proposed architecture in the digitalization of power distribution panel, which provides an effective solution for video inspection and intelligent maintenance in power communication networks.

**Key words:** power communication network, digitization of power distribution panel, image identification, character recognition, YOLOv5, lightweight, CRNN

## 0 引言

电力通信网承载着电力系统的监控、调度、数据传输等功能,保障电力设备安全高效地运行,但传统的电力通信网存在资源利用率低、故障率高等问题<sup>[1]</sup>。数字化技术的应用为运维管理的实时监控和自动化诊断提供了重要的实现途径,可以显著提高电力通信网的运维管理效率和精准度。通信电源系统是电力通信网的重要组成部分,确保变电站与调度中心之间的可靠数据传输,支撑电力设备的实时监测状态并及时下发远程控制指令,从而实现保护动作的有效通信<sup>[2]</sup>。通信电源系统通常包括交流配电屏、高频开关整流模块、直流配电屏等设备<sup>[3]</sup>,各部分设备需要定期巡检和维护。目前,配电设备的资产密集性和运维资源的匮乏性逐渐凸显<sup>[4]</sup>,考虑其数量多、运行状态复杂以及重复性高等特点,人工巡检和状态分析消耗大量成本,且可能造成误判,无法高效、准确地获取配电设备的运行状态信息。因此,通过数字化技术手段提高配电设备的巡检效率和准确度,成为运维管理的关键问题。

随着人工智能技术的快速发展,电力人工智能 (electric power artificial intelligence, EPAI) 融合了人工智能技术与电力系统技术,能够实现

从数据采集、设备感知到智能决策的全流程优化<sup>[5]</sup>。其中,感知智能通过多模态感知手段,借助计算机视觉技术对图像、视频等数据进行分析 and 理解。该技术在输电线路巡检、变电站智能巡视、电力违章监测等场景中已经得到了广泛的应用。配电屏作为通信电源系统的核心设备,承担着下级各支路电源的分配任务,利用感知智能技术对配电屏进行实时监控和状态识别,可以实现配电屏供电状态的数字化映射,从而大幅提升运维的效率和精准度。

配电屏的数字化映射主要分为3个方面:对各供电支路完整结构单元的定位,对开关、端子、状态灯等目标的图像识别和对各下级电源标号的文字识别。通过识别配电屏图像中各开关和状态灯的供电状态,以及每个开关所对应的文本标签,并将相同结构单元中“状态-标签”的关联匹配,实现各下级支路电源供电状态的结构化输出。其中,图像识别任务主要依赖于基于深度学习技术的目标检测方法<sup>[6-8]</sup>,YOLO系列网络<sup>[9-11]</sup>作为单阶段算法<sup>[12]</sup>的典型代表,能够同时满足信息图像识别任务中对可靠性和实时性的要求。在电力资产图像的目标识别领域,一些研究人员在YOLOv5网络的基础上分别引入深度展开超分网络<sup>[13]</sup>、GAM注意力模块和ASFF模块<sup>[14]</sup>、



图论分割 Grabcut<sup>[15]</sup>、元学习<sup>[16]</sup>等技术,从而降低复杂背景的干扰并提高绝缘子缺陷的检测精度。文献[17]引入 Ghost 卷积并结合自注意力改进的 C3 模块,提升了对变电设备红外图像的识别精度和速度。文献[18]利用窗口自注意力网络优化 YOLOv5 的骨干结构以扩大模型感受野,并改进自适应空间特征融合模块以提高输电线路通道中异物定位的准确性。文字识别任务主要包含卷积循环神经网络(convolutional recurrent neural network, CRNN)-光学字符识别(optical character recognition, OCR)和 Attention-OCR 两种主流技术<sup>[19]</sup>。其中,CRNN-OCR 无须分割图片中的文字即可识别文本,而且可以适应任意长度的文本序列。在电力领域,CRNN 被用于识别电网厂站接线图中的文本信息<sup>[19]</sup>、电表示数<sup>[20]</sup>、电力铭牌文本信息<sup>[21]</sup>等场景中,展现出了较好的识别效果。

在目标识别过程中,配电屏状态灯的目标尺寸较小,包含的像素点有限,使得在逐层深入的特征提取与融合过程中,小目标的原始信息容易丢失,从而降低目标检测的准确率。而且配电屏图像中包含开关、端子、状态灯等多种目标,目标尺寸的不一致导致特征提取困难,从而影响识别性能。文本识别任务的主要目标是识别序列长度不一致的字母、数字和文字混合的开关标签信息,但由于采集的配电屏图像种类有限且拍摄角度单一,文本图像相对规范,模型在实际应用中面临局限性,难以适应多样化场景。综合考虑以上挑战,本文研究面向电力通信配电屏数字化的人工智能识别关键方法,具体贡献如下。

(1) 提出了一个多层嵌套识别网络(multi-layer nested recognition network, MLNRN)架构,含结构单元检测、图元检测和文字识别模块,并结合轻量化策略,高效准确地实现了配电屏信息获取和结构化报文输出。

(2) 为实现配电屏图像中状态灯等小目标的准确、快速检测,在 YOLOv5 骨干特征提取网络

中引入了 ConvNext Block 和双向特征金字塔网络(bidirectional feature pyramid network, Bi-FPN)结构。ConvNext Block 模块通过倒瓶颈和大卷积核的设计有效地提取高维度特征信息,避免小目标特征的缺失,Bi-FPN 结构增强了浅层特征与深层特征的融合。

(3) 构建了一个基于卷积递归神经网络-连接时态分类(convolutional recurrent neural network-connectionist temporal classification, CRNN-CTC)网络的配电屏文字识别模块。该模块采用迁移学习的策略,解决了实际采集图像中字符种类不足的问题,有效地提升了配电屏图像中文本识别的准确性和适应性。

## 1 MLNRN 总体架构

### 1.1 实现框架

电力通信配电屏数字化的本质是识别配电屏图像中各开关和状态灯的供电状态,以及每个开关所对应的文本标签,通过“状态-标签”的关联匹配,并与系统中设定的各下级电源的名称、类型等数据相结合,得到对下级电源的供电状态。针对这一识别任务,提出一个多层嵌套识别网络架构,该架构由结构单元检测模块、图元检测模块和文字识别模块构成。其中,结构单元检测模块采用轻量化的 YOLOv5 (YOLOv5-tiny) 网络,主要负责快速检测由状态灯、开关和标签等多个图元对象组成的结构单元。考虑图元检测模块仍以 YOLOv5 架构为网络基础结构,两层 YOLOv5 网络产生的大量参数会降低算法在终端设备上的运行速度。因此,在结构单元模块中对原始 YOLOv5 网络轻量化,使该网络在保证检测精度的同时,具有较低的计算复杂性和快速的推理速度,在动态环境中高效地识别并定位关键对象,为后续精细化识别操作提供位置特征。定位后的结构单元输入图元检测模块,该模块使用 Imp-YOLOv5 网络,考虑检测目标的尺寸不一致

及小目标检测问题, 在 YOLOv5 网络中引入 ConvNext Block 模块和 Bi-FPN 结构以通过增强特征提取能力来优化检测效果。在获取各支路供电状态后, 需要确定该支路的身份信息, 将该结构单元中的文本标签部分输入文字识别模块, 该模块采用改进的 CRNN-CTC。因为配电屏的数据样本有限, 所以利用大型中文识别数据集, 采用迁移学习的方法提高对标签和名称两类图元的文本识别精度, 最后输出结构化的配电屏数字化报文。电力通信配电屏数字化方法如图 1 所示。

本文配电屏数字化过程主要分为如下 4 个阶段。

(1) 数据准备阶段。将捕获的配电屏的正面彩色图像进行预处理操作, 并标注为对应的数据集。

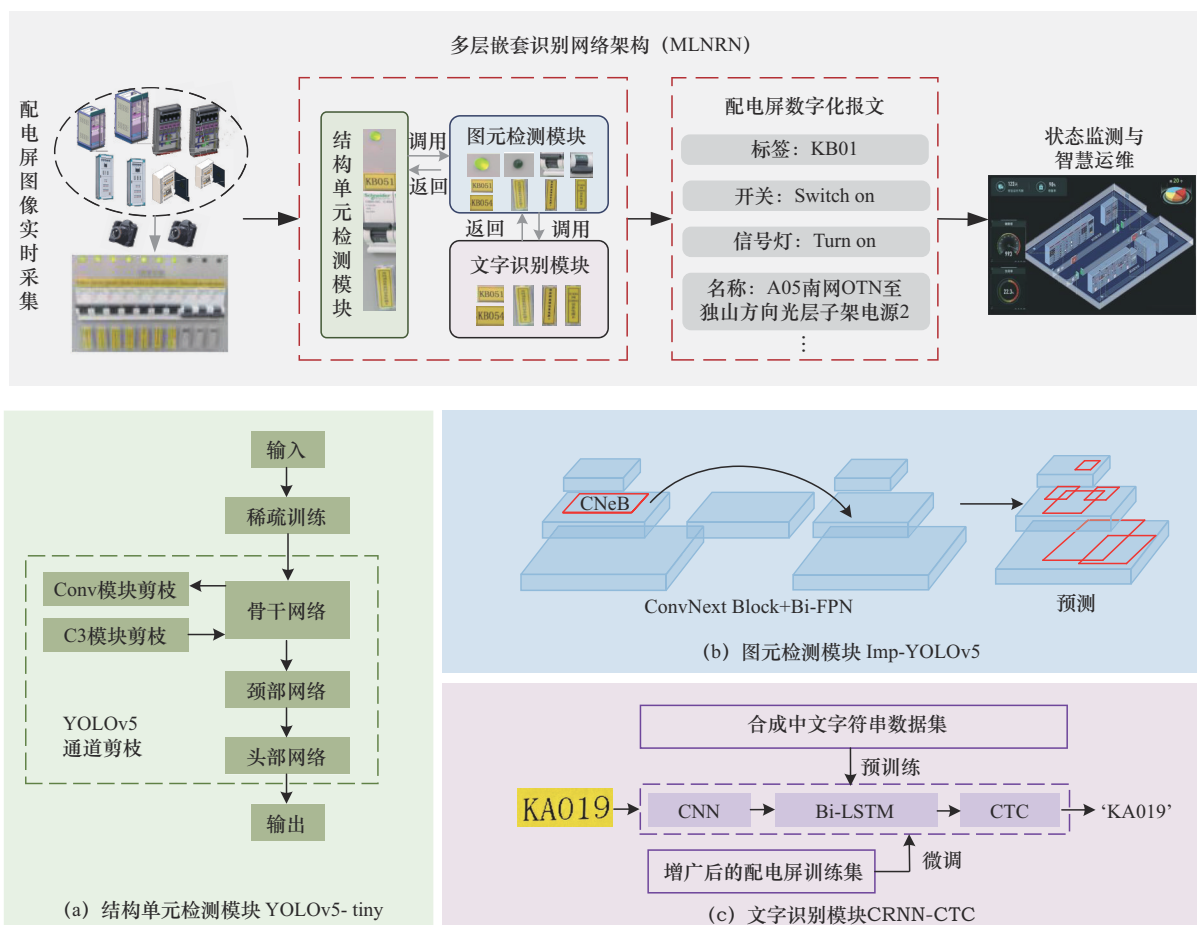
(2) 模型训练阶段。本阶段通过使用不同的标注数据集, 分别对结构单元检测模块中的 YOLOv5-tiny 网络、图元检测模块中的 Imp-YOLOv5 网络和文字识别模块中的 CRNN-CTC 进行预训练与微调。

(3) 目标检测及识别阶段。本阶段利用 MLNRN 架构中的各个模块, 对目标进行图像检测和文字识别。

(4) 结构化输出阶段。将各图元的识别结果按网络层级向上返回, 输出结构化报文信息, 实现配电屏数字化映射。

## 1.2 结构单元检测模块

结构单元检测模块由一层 YOLOv5-tiny 网络构成, 该模块的任务是快速检测配电屏图像中由



注: CNeB (ConvNext-based Block) 是一种目标检测模块

图 1 电力通信配电屏数字化方法



多个图元对象构成的结构单元,为后续模块的图元检测和文字识别任务提供基础的位置特征。

### 1.2.1 YOLOv5 网络结构

YOLOv5是一种高效的单目标检测方法,分为输入、骨干网络、颈部网络和头部预测4个部分。

#### (1) 输入部分

作为YOLOv5目标检测模型的输入端,该部分对输入的待检测图像进行缩放以适应骨干网络的输入尺寸,并执行归一化及图像去噪等预处理步骤。

#### (2) 骨干网络部分

该部分主要由Focus模块、Conv模块、C3模块和SPPF模块组成。Focus模块对图像进行切片,以降低图像的维度并增加通道数;C3模块包含多个Conv和Bottleneck模块,利用 $1 \times 1$ 的卷积核实现通道数减半以降低计算量,利用 $3 \times 3$ 的卷积核扩大网络的感受野;SPPF模块对图像做进一步的特征提取。

#### (3) 颈部网络部分

颈部网络用来衔接骨干网络和输出,利用C3模块和Conv模块进一步加强网络特征提取能力,基于双层金字塔结构和Concat实现多特征层间的特征融合输出预测特征图。

#### (4) 头部预测部分

YOLOv5系列利用3个检测层对目标进行识别,设置基于3个不同大小的锚框进行目标检测,计算每个检测框的置信度和得分,得分最高的类别为检测结果。

### 1.2.2 YOLOv5 轻量化

因为MLNRN架构中包含两层YOLOv5网络,这会产生大量的参数,从而导致在终端设备上运行速度变慢。因此,在结构单元模块中采用稀疏训练和通道剪枝的方式将原始YOLOv5网络轻量化,以提升整体MLNRN架构的整体性能。

#### (1) 稀疏训练

YOLOv5网络利用批归一化(batch normal-

ization, BN)层对每一层的输入进行规范化,从而减少内部协变量的偏移,使网络训练更稳定,计算式为:

$$Z_{out} = \gamma \hat{Z} + \beta \quad (1)$$

其中, $Z_{out}$ 为BN层的输出, $\hat{Z}$ 为标准化后的输入值, $\gamma$ 和 $\beta$ 是可训练的缩放因子和平移参数。通过对缩放因子施加L1正则化,可以引导不重要的通道值接近于零,从而实现通道的稀疏化。这种稀疏训练的损失函数表示为:

$$L = \sum_{(x,y)} l(f(x,W),y) + \lambda \sum_{\gamma \in \Gamma} \gamma \quad (2)$$

式(2)由两项加和得到,左边的求和项是YOLOv5网络训练损失函数, $(x,y)$ 表示训练的输入和输出, $W$ 代表可训练的权重, $l(\cdot)$ 代表正常训练的损失;右边是正则化项,稀疏率 $\lambda$ 用来平衡损失函数和正则化项, $\lambda$ 值越大,网络稀疏程度则越大, $\Gamma$ 为所有缩放因子构成的集合。

#### (2) 通道剪枝

稀疏训练结束后网络基本收敛,大多数 $\gamma$ 因子都接近于0。采用全局阈值策略设定全局剪枝率,对所有稀疏训练后的缩放因子 $\gamma$ 进行统计和排序,根据设定的全局剪枝率来剪除排序靠后的通道权重参数。通道剪枝过程如图2所示。

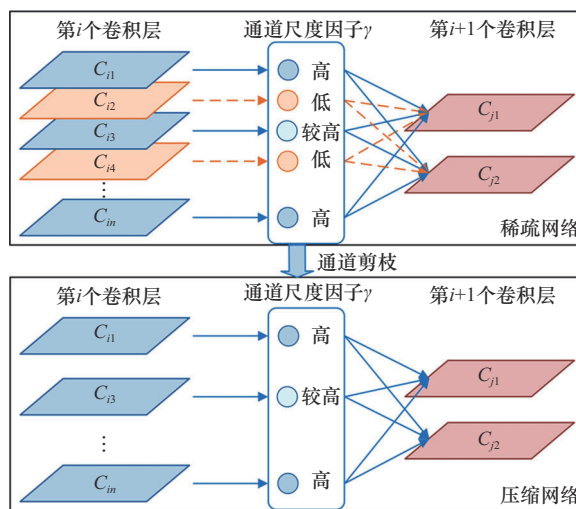


图2 通道剪枝过程

图2中,上层为稀疏后的网络, $n$ 是第 $i$ 层的卷积通道数,其中一些通道的批归一化层 $\gamma$ 值接近于0,删去后得到下层压缩网络。

综合考虑模型整体性能和剪枝后对特征提取能力的影响,选择对原始YOLOv5网络中Conv模块、C3模块和SPPF模块的部分卷积块进行通道剪枝操作。此外,单一阈值控制剪枝的方式会使网络中某层的所有通道都被剔除,从而破坏原始网络的规整结构。因此,引入局部安全阈值 $\tau$ ,即逐层减除同时小于全局阈值 $\hat{\gamma}$ 与局部阈值 $\tau$ 的通道,以防止过度剪枝。

### 1.3 图元检测模块

图元检测模块的任务是结构单元目标中典型图元的检测和识别,包括状态灯、开关、标签和名称4种图元对象的种类、状态、方向识别和图元定位。然而,这些图元的大小、清晰度和规范差别较大,且部分图元之间具有一定的相似度,这给识别模型带来了巨大的挑战。

针对这一问题,设计了一种基于改进YOLOv5网络以提升图元的精准识别和定位。具体而言,在YOLOv5骨干特征提取网络中引入ConvNext Block

模块,其结构能有效地提取骨干网络中图片高维度的特征信息,避免状态灯等小目标特征信息的缺失。此外,引入Bi-FPN结构替换原有的特征金字塔-路径聚合网络(feature pyramid network-path aggregation network, FPN-PAN)结构,加强接近输入的浅层特征和接近输出的深层特征的融合。改进YOLOv5的结构框架如图3所示。

#### 1.3.1 基于ConvNext Block的小目标检测优化

ConvNext Block模块是一种结合卷积神经网络(convolutional neural network, CNN)和Transformer的深度学习模块,去除了CNN的冗余设计,同时保持高效性能,并利用Transformer结构实现类注意力机制以提取精细特征,在微调精度和减少训练时间方面具有优势<sup>[22]</sup>。在原始YOLOv5的骨干网络结构中引入ConvNext模块<sup>[23]</sup>,可以提高骨干网络提取特征的能力。

CNeB模块分为两个分支,结构如图3中所示。设CNeB模块的输入特征为 $F_1$ ,输入骨干网络分支后,经过尺寸为 $7 \times 7$ 的卷积核进行深度卷积并进行层归一化,利用两个 $1 \times 1$ 的卷积核进行特征增强,其中包含GeLU函数的激活过程,如式(3)所示。

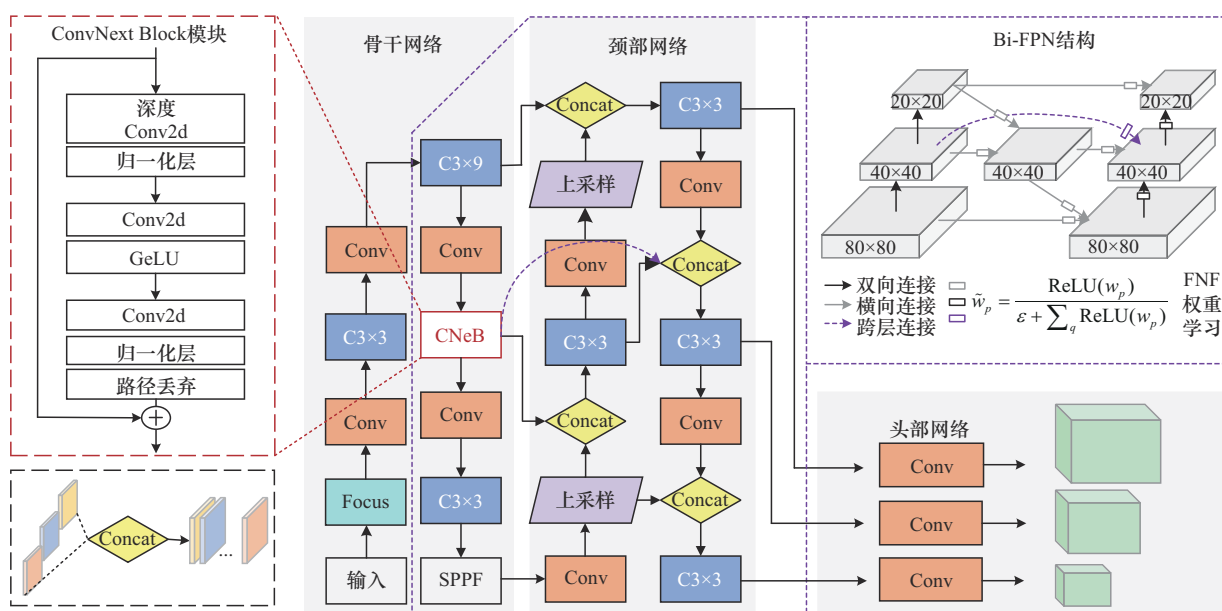


图3 改进YOLOv5的结构框架



$$F_2 = \text{Conv}_{1 \times 1} \left( \text{GeLU} \left( \text{Conv}_{1 \times 1} \left( \text{LN} \left( \text{Conv}_{7 \times 7} (F_1) \right) \right) \right) \right) \quad (3)$$

然后, 将增强后的特征  $F_2$  进行层缩放归一和路径丢弃, 通过对特征值进行缩放和正则化得到骨干网络分支的结果。最后与原始输入特征聚合, 多分支的卷积结构使 CNeB 模块通过不同感受野和参数的组合提取的特征包含更多原始图像的细节信息。

对比原始 C3 结构, 设计倒瓶颈的结构可以降低网络的运算量, 并通过大卷积核的形式扩大特征提取的感受野以提取图片中的小目标精细特征, 减少特征提取过程中对状态灯图元特征的丢失。同时, CNeB 将 BN 替换为层归一化 (layer normalization, LN), 减少归一化和使用激活函数的次数, 降低网络运算的复杂度。

### 1.3.2 基于 Bi-FPN 结构的特征融合增强

在传统 YOLOv5 算法中, 通常用 FPN-PAN 来完成在不同尺度和分辨率下的目标识别, 在特征提取过程中对每个尺度的特征图逐级进行采样和融合, 构建特征金字塔完成更准确的目标检测。但是在多尺度采样与逐级融合过程中存在细节信息丢失的风险以及信息损失与信息模糊的问题。

为解决上述问题并提高检测准确性, 利用 Bi-FPN 结构替换传统结构。在配电屏目标对象的检测过程中, 对于不同尺度和分辨率的照片, 在双向信息流动和横向连接的基础上引入跨层连接, 有效地提升了目标检测的准确性和稳定性, Bi-FPN 结构如图 4 所示。

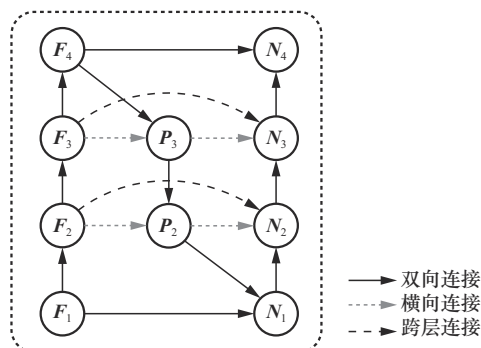


图4 Bi-FPN 结构

Bi-FPN 旨在用高效的方式融合不同分辨率的特征, 不局限于单个输入节点, 允许多个特征层的加权输入, 采用双向连接、跨层连接和横向连接机制来促进输入照片特征之间的信息传递和特征融合。其中, “跨层连接”是指若输入和输出节点是同一尺度的特征层, 则添加一条额外的通道跨过中间的  $P$  层, 在不增加计算量的同时加强该尺度的特征融合, 如图 4 中的虚线连接线, 以此来全面捕捉不同尺度和分辨率照片的特征信息。Bi-FPN 中不同尺度特征的 FPN 层的表达式如下:

$$P_3 = \text{Concat}(F_3, \text{upsample}(F_4)) \quad (4)$$

$$P_2 = \text{Concat}(F_2, \text{upsample}(P_3)) \quad (5)$$

$$N_2 = \text{Concat}(P_2, F_2, \text{Conv}(N_1)) \quad (6)$$

$$N_3 = \text{Concat}(P_3, F_3, \text{Conv}(N_2)) \quad (7)$$

其中, Conv 代表卷积层, upsample 代表上采样操作, 使不同尺度的特征可以融合, Concat 对多输入进行合并, 代表通道数的增加, 如图 3 所示。

另一方面, 因为具有不同分辨率的多尺度特征图对目标检测所提供的信息量不同, 在特征金字塔网络中进行由上至下和由下至上的双向信息流动的过程中, 引入权重系数以量化不同输入特征的重要程度。权重系数需要通过归一化来约束其取值范围, 但 softmax 中的指数计算部分会增加运算量, 因此, 采用快速归一化融合 (fast normalized fusion, FNF) 的方法实现权重系数的归一化。具体为取消 softmax 函数的指数部分, 并利用 ReLU 激活函数保证权重的非负性, 可以提高训练的速度、训练效率和训练稳定性, 权重系数的计算过程为:

$$\tilde{w}_p = \frac{\text{ReLU}(w_p)}{\varepsilon + \sum_q \text{ReLU}(w_q)}, q = 1, 2, \dots, Q \quad (8)$$

其中,  $w_p$  和  $w_q$  分别代表第  $p$  和第  $q$  个输入特征图

的权重,  $\tilde{w}_p$  代表归一化后的结果,  $Q$  代表融合特征图的总数。ReLU 可以保证权重的非负性,  $\varepsilon$  是一个微小的常量, 用来保证数值的稳定。

以  $N_2$  为例, 式 (6) 的求解变为式 (9)。

$$N_2 = \text{Concat}(\tilde{w}_1 \cdot P_2, \tilde{w}_2 \cdot F_2, \tilde{w}_3 \cdot \text{Conv}(N_1)) \quad (9)$$

Bi-FPN 的引入使每个层级的特征可以相互影响和融合, 提升特征的多样性和丰富性, 进而提高了小目标检测的准确性和鲁棒性。

#### 1.4 文字识别模块

经结构单元检测模块后, 原始图片中的状态灯、标签、开关和名称等图元分别被标注和定位。基于 CRNN-CTC 结构, 利用迁移学习的方法训练文字识别模块, 用于对标签和名称两类图元的文本识别, 以实现配电屏图像的数字化的。

##### 1.4.1 CRNN-CTC 结构

CRNN-CTC 能够处理长度不固定的文本识别任务, 将识别转化为一个时序依赖的序列学习问题。引入 CRNN-CTC 构建配电屏文字识别模块, 无须进行字符分割<sup>[24]</sup>, 具有速度快、性能好和参数少的优势。

基于 CRNN-CTC 的文本识别模块包含 3 个部分: 首先, 利用卷积神经网络层对图像进行特征提取, 获取包含字符信息的局部特征; 然后, 通过特征序列映射层将字符特征序列输入长短期记忆网络, 捕捉字符间时序关系并预测每个时间步的标签分布; 最后, CTC 转录层根据每个时间步的概率分布进行解码, 生成输出的文本标签。

##### 1.4.2 迁移学习

在配电屏图像的文本识别任务中, 由于实际配电屏图片的采集范围有限, 训练集中图片包含的字符种类较少。为了使模型能够实现多种配电屏的文本识别, 采用迁移学习的方法提高模型的识别性能。选用 Synthetic Chinese String Dataset 作为预训练的源数据集, 利用大量的中文数据预

训练模型。随后, 利用配电屏图像中的标签和名称两类图元作为微调训练集, 通过迁移学习策略对模型进行微调, 以实现高效的文本识别, 主要由以下步骤组成。

(1) 利用源数据集对 CRNN-CTC 模型进行预训练, 得到一个基础的文字识别模型。

(2) 在微调过程中, 首先固定卷积层, 以保留在预训练阶段学到的特征。随后逐步解冻双向长短期记忆网络 (BiLSTM), 使模型能够学习配电屏文本特征。

(3) 使用配电屏文本图像的微调训练集对模型进行训练, 调整学习率以适应新的数据分布。

## 2 实验分析

### 2.1 图像预处理与数据集制作

#### 2.1.1 图像预处理

在对配电屏图片进行拍摄、读取和传输的过程中随机产生的噪声会导致画面模糊, 不利于图像解析。为提高图像质量, 采用二维高斯卷积核对整幅图像进行加权平均, 基于高斯滤波的方法实现图像去噪。利用二维高斯卷积核遍历图像中所有的像素点, 像素点的值可以通过自身以及周围邻域内其他像素值的加权平均计算得出。高斯去噪结果示意图如图 5 所示, 得到去噪后的图像如图 5 (b) 所示, 可以看出, 经过滤波后图像中的噪点被去除, 图像清晰度得到了有效提高。

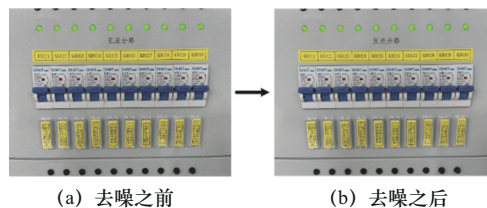


图 5 高斯去噪结果示意图

#### 2.1.2 数据集制作

采用的配电屏图像数据集来源于中国南方电网的实际运行场景, 共有 408 张不同的配电屏图



像。配电屏图像拍摄环境、角度、分辨率等差异较大,可以有效验证所提 MLNRN 架构的有效性与鲁棒性。

在结构单元检测模块中,使用 Labelimg 工具箱对结构单元对象进行标注,标注为“Unit”标签,共约 4 488 个标注框,划分 70% 作为轻量化 YOLOv5 网络的训练集,剩余 30% 作为测试集。

在图元检测模块中,使用 Labelimg 工具箱对状态灯、标签、开关和名称 4 个图元对象进行标注,分别标注为“Led on/Led off”“Label”“Switch on/Switch off”“Name”标签,共标注约 17 952 个标注框,划分 70% 作为 Imp-YOLOv5 网络的训练集,剩余 30% 作为测试集。

在实际的通信运维过程中,会采用摄像头定时自主拍摄并上传系统的方式,并非人为采集,这会导致获取的图片存在角度偏移、亮度的变化。为了使模型能够对上述图片均具有较高的识别准确率,对现有数据集进行样本扩充,分别采用调整亮度、旋转和添加噪声的方式对训练样本做一系列的随机改变,生成相似但不同的训练样本,以提高模型的泛化能力。其中,旋转变换是以正面拍摄图片为基础,以图像中心点为基准,随机将图像左旋或右旋 5° 以内。以其中一个文本切片为例示意样本扩充过程,样本扩充示意图如图 6 所示。



图6 样本扩充示意图

## 2.2 图像识别实验

### 2.2.1 实验设置

本实验于特定实验环境下展开,硬件包含

NVIDIA GeForce RTX 3090 图形处理单元,软件包含 Ubuntu20.04 LTS 操作系统、Python3.8 编程语言、PyTorch1.10 深度学习框架、CUDA11.3 和 cuDNN8.2 计算加速库。训练过程利用余弦退火算法对学习率进行调整,学习率的初始值设定为 0.01,批次大小设为 64,动量参数为 0.937,权重衰减设为 0.000 5。对于模型,anchor boxes 的尺寸经聚类计算确定,在颈部网络部分调整特征融合通道数以优化模型结构。

### 2.2.2 评估指标

#### (1) 平均精度 (mAP@0.5)

mAP@0.5 侧重于评估模型对目标的粗粒度检测能力,即当预测框与真实框的交并比 (intersection over union, IoU) 达到 0.5 及以上时,认为检测正确。

#### (2) 召回率

召回率是指模型正确检测到的正样本数与实际正样本数的比值,召回率衡量了模型对所有正样本的检测能力,反映了模型的查全率。计算式为:

$$R = \frac{TP}{TP + FN} \quad (10)$$

其中,TP 表示真正例,即被模型正确检测为正样本的数量;FN 表示假反例,即实际为正样本但未被模型检测到的数量。

#### (3) 精确率

精确率定义为模型正确检测到的正样本数与模型预测为正样本的总数的比值,用于衡量模型预测结果的准确性。计算式为:

$$P = \frac{TP}{TP + FP} \quad (11)$$

其中,FP 表示假正例,即模型错误地将负样本预测为正样本的数量。

### 2.2.3 结构单元检测结果与分析

#### (1) 结果分析

不同剪枝率下参数数量和精度指标对比见表 1。由表 1 可知,原始 YOLOv5 网络的参数量为

$7.022\ 3\times10^6$ , 随着剪枝率的增加, 模型的参数量可降低 $4.177\ 4\times10^6\sim5.211\ 5\times10^6$ , 验证了采用的通道剪枝方法有效地压缩了YOLOv5网络模型, 节省了网络资源。当剪枝率为0.5时, 当参数量大幅降低的同时精度仍维持在0.986, 当剪枝率为0.7时, 虽然参数量降低至 $1.810\ 8\times10^6$ , 但精度损失太多。因此, 综合考虑参数量及检测精度, 实验设置剪枝率为0.5。

| 表 1 不同剪枝率下参数量和精度指标对比 |                      |       |       |         |
|----------------------|----------------------|-------|-------|---------|
| 剪枝率                  | 参数量/( $\times10^6$ ) | P     | R     | mAP@0.5 |
| 未剪枝                  | 7.022 3              | 1.000 | 1.000 | 0.995   |
| 0.5                  | 2.844 9              | 0.986 | 0.986 | 0.986   |
| 0.6                  | 2.290 1              | 0.979 | 0.984 | 0.995   |
| 0.7                  | 1.810 8              | 0.956 | 0.956 | 0.977   |

不同剪枝率下精确率随迭代次数变化曲线如图7所示, 显示了训练过程中不同剪枝率下的精确率变化。从图7可以看出, 剪枝率为0.5时, 模型没有损失太多的精确率, 却有更快的收敛速度, 进一步证明了通道剪枝方法的有效性。

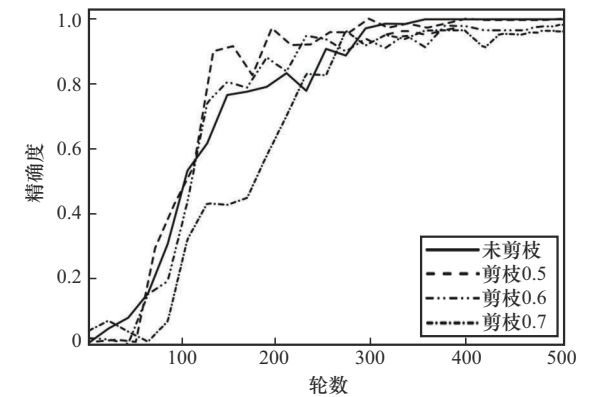


图7 不同剪枝率下精确率随迭代次数变化曲线

(2) 结果可视化

将待检测配电屏图像输入轻量化YOLOv5网络对结构单元对象进行检测, 结构单元识别结果如图8所示。图8中虚线框为每个“Unit”对象的检测结果, 置信度相对较高。

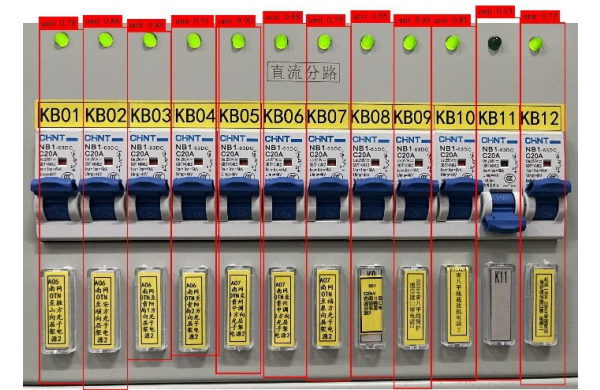


图8 结构单元识别结果

2.2.4 图元检测结果与分析

(1) 结果分析

为验证所提的改进YOLOv5算法的可行性与有效性, 进行了多组消融实验, YOLOv5消融实验结果见表2, 其中, 下降箭头表示下降幅值, 上升箭头表示上升幅值, 加粗表示效果最好。由表2可知, 单独添加Bi-FPN模块时, 精确率有很大幅度的提升, 但召回率和平均精度有所下降; 单独添加ConvNext Block时, 3项指标都有所提升; 两个模块同时添加时效果最好, 相对原始的YOLOv5网络, 精确率提升了10.5%、召回率提升了9.8%、平均精度提升了4.1%, 证明所提方法具有可行性与有效性。

将原始YOLOv5网络与引入ConvNext Block

| 表 2 YOLOv5消融实验结果 |                |        |                      |                     |                     |
|------------------|----------------|--------|----------------------|---------------------|---------------------|
| 算法               | 改进             |        | 指标                   |                     |                     |
|                  | ConvNext Block | Bi-FPN | P                    | R                   | mAP@0.5             |
| YOLOv5           | ×              | ×      | 86.7%(基准)            | 81.1%(基准)           | 83.8%(基准)           |
| YOLOv5-v1        | ×              | ✓      | 94.4%(↑ 7.7)         | 80.4%(↓ 0.7)        | 76.5%(↓ 7.3)        |
| YOLOv5-v2        | ✓              | ×      | 90.2%(↑ 3.5)         | 84.8%(↑ 3.7)        | 87.5%(↑ 3.7)        |
| Imp-YOLOv5       | ✓              | ✓      | <b>97.2%(↑ 10.5)</b> | <b>90.9%(↑ 9.8)</b> | <b>87.9%(↑ 4.1)</b> |



模块、Bi-FPN模块进行对比, 算法精确度对比曲线如图9所示。由图9可知, 原始YOLOv5模型较难收敛且精度低, 引入ConvNext Block结构和Bi-FPN结构后的3种模型收敛速度较快, 其中, Imp-YOLOv5达到的精确率最高。

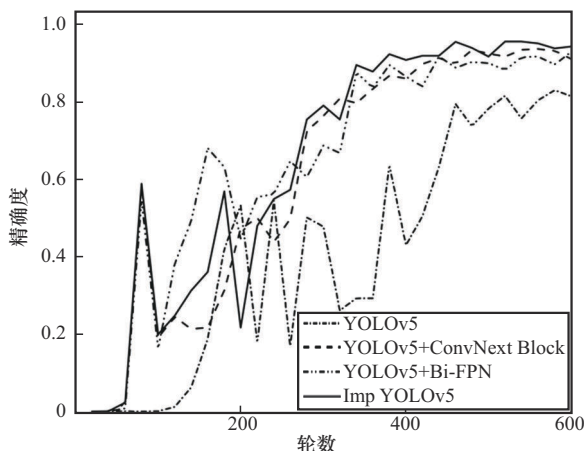


图9 算法精确度对比曲线

此外, 将所提改进YOLOv5算法与其他常见的目标检测算法进行对比, 不同算法的测试结果见表3。由表3可知, 对比YOLOv3、Faster-RCNN、SSD这3种目标检测算法, 本文改进的

YOLOv5算法能获得更高的目标检测精度, 具有更优越的性能。

表3 不同算法的测试结果

| 算法          | P            | R            | mAP@0.5      |
|-------------|--------------|--------------|--------------|
| YOLOv3      | 82.3%        | 76.9%        | 79.5%        |
| Faster-RCNN | 85.6%        | 79.2%        | 82.6%        |
| SSD         | 88.2%        | 82.9%        | 86.0%        |
| YOLOv5      | 86.7%        | 81.1%        | 83.8%        |
| Imp-YOLOv5  | <b>97.2%</b> | <b>90.9%</b> | <b>87.9%</b> |

## (2) 结果可视化

将结构单元模块所检测的结果进行图像切片并输入图元检测模块中进行检测, 得到各图元检测结果, 目标检测的图像示意图如图10所示。图10中黄色检测框是信号灯点亮, 棕色检测框为信号灯熄灭; 紫色检测框是“Label”图元检测结果; 蓝色检测框是开关打开, 深蓝色检测框是开关关闭; 绿色检测框是“Name”图元检测结果。由图10可知, 信号灯和开关图元的状

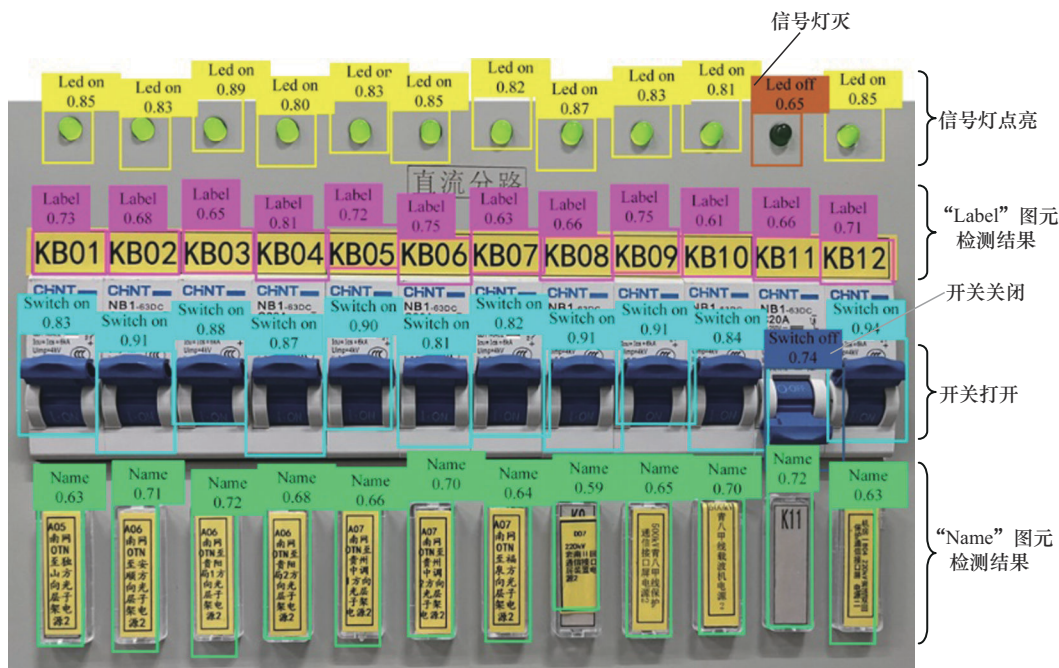


图10 目标检测的图像示意图

态检测及定位结果全部正确无误,且置信度相对较高;而“Label”和“Name”标签的置信度相对较低,这是因为两者视觉上呈现为“黑色外框”“黄色背景”“黑色字体”的相似特征,但定位信息基本准确。

## 2.3 文本识别实验

### 2.3.1 实验设置

实验的软硬件环境与图像识别实验一致。在模型预训练阶段使用 Synthetic Chinese String Dataset 数据集,按照 9:1 的比例划分为训练集和测试集,使用 RMSprop 优化器,批大小为 256,共训练 20 个轮次;在模型微调阶段,利用预训练好的模型对参数进行初始化,使用约 8 976 个的文本切片图像训练模型,批大小设为 256,使用 Adam 优化方法,初始学习率设置为  $1.0 \times 10^{-3}$ ,共设置 30 000 个轮次,经过 14 000 次迭代时学习率调整为原来的 0.1 倍。

### 2.3.2 评估指标

文字识别实验采用识别准确率(正确识别的文字图片数量占识别文字图片的总数)作为评价指标。

### 2.3.3 文字识别结果与分析

#### (1) 结果分析

设置了 3 个对比组,对比组 1 仅用配电屏图像中的文本切片进行训练,对比组 2 仅用源数据集进行训练,对比组 3 采用源训练集进行训练,配电屏文本切片进行参数微调,不同对比组的识别准确率见表 3。

对比 3 组的识别准确率可以发现,利用迁移学习后的模型的识别准确率最高。对比组 1 难以

满足对含有其他字符的样本的精准识别;对比组 2 可以通过源数据集的训练,使 CRNN-CTC 模型学习到常见的文本特征,识别准确率有所提升;对比组 3 在此基础上,利用配电屏样本使模型能够兼顾配电屏文本常见的字符排列模式,进一步提升识别准确率,证明了训练的文字识别模块的有效性。

#### (2) 结果可视化

将待检测图像输入微调后的 CRCC-CTC 文字识别模块中,任取 4 张待识别文本切片为例,文本识别的图像示意图如图 11 所示。

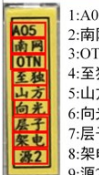
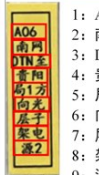
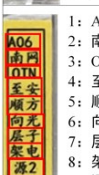
|                                                                                     |                                                                                                                              |                                                                                       |                                                                                                                                |
|-------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------|
|    | 1:KB01 0.991                                                                                                                 |    | 1:KA001 0.991                                                                                                                  |
|    | 1:KB02 0.994                                                                                                                 |    | 1:KA002 0.994                                                                                                                  |
|   | 1:A05 0.935<br>2:南网 0.942<br>3:OTN 0.987<br>4:至独 0.911<br>5:山方 0.942<br>6:向光 0.913<br>7:层子 0.921<br>8:架电 0.917<br>9:源2 0.889 |   | 1:A06 0.874<br>2:南网 0.931<br>3:DTN至 0.811<br>4:贵阳 0.894<br>5:局1方 0.927<br>6:向光 0.885<br>7:层子 0.944<br>8:架电 0.922<br>9:源2 0.983 |
|  | 1:A06 0.913<br>2:南网 0.927<br>3:OTN 0.932<br>4:至安 0.891<br>5:顺方 0.874<br>6:向光 0.867<br>7:层子 0.933<br>8:架电 0.949<br>9:源2 0.912 |  | 1:A06 0.892<br>2:南网 0.934<br>3:DTN至 0.876<br>4:贵阳 0.919<br>5:局2方 0.928<br>6:向光 0.953<br>7:层子 0.921<br>8:架电 0.857<br>9:源2 0.972 |

图 11 文本识别的图像示意图

图 11 中左侧是电源分支的标签和名称,右侧是文本识别结果和置信度。由结果可知,该模块经迁移学习后可以实现配电屏文本切片中的文字识别。

## 2.4 结构化输出结果

配电屏的数字化结果包括设备名称、设备类型、运行状态等信息的数字化呈现。所提方法通过对配电屏的开关、状态灯等目标对象进行状态和文本识别,输出各支路的运行状态信息,并与提前设定的设备名称、类型等数据相结合,形成配电屏运行状态的数字化报文,有效地支撑了配电屏数字化的实现,配电屏数字化结果示意图如

表 4 不同对比组的识别准确率

| 对比组 | 配电屏文本切片 | 源数据集 | 准确率   |
|-----|---------|------|-------|
| 1   | √       | ×    | 73.1% |
| 2   | ×       | √    | 90.7% |
| 3   | √       | √    | 92.0% |



图12所示。

### 3 结束语

本文深入探讨了面向电力通信配电屏数字化的人工智能识别关键方法,针对电力通信配电屏在供电状态监测与标签识别中面临的核心问题,提出了多层嵌套识别网络(MLNRN)架构。该架构通过整合结构单元检测、图元检测和文字识别模块,有效地实现了配电屏信息的结构化报文输出,提升了监测效率和准确性。

在结构单元检测模块中,为了提升模型在终端设备上的运行速度,采用了轻量化策略,

包括稀疏训练和通道剪枝,这些方法有效地减少了模型的参数量,同时保持了特征提取能力。在图元检测模块中,改进的YOLOv5网络通过引入ConvNext Block和Bi-FPN结构,显著地增强了对状态灯等小目标的识别性能。在文字识别模块中,结合了迁移学习和图像增广策略,增强了对多样化字符种类和不规范配电屏图像的文本识别能力。

实验结果表明,图像识别的平均准确率为97.2%,文本识别的准确率达到了92%,所提MLNRN架构为配电屏的数字化识别提供了有效的解决方案,提升了电力通信网的可靠性和运维管理



图12 配电屏数字化结果示意图

的效率,为智能电网的发展奠定了坚实的基础。

## 参考文献:

- [1] 唐大钧. 基于数字化通信技术的电力通信网的运维探究[J]. 中国宽带, 2024(10): 164-166.  
TANG D J. Research on operation and maintenance of electric power communication network based on digital communication technology[J]. China BroadBand, 2024(10): 164-166.
- [2] 郝青云. 电力通信网中通信电源的典型故障案例分析[J]. 内蒙古科技与经济, 2024(15): 134-137.  
HAO Q Y. Analysis of typical fault cases of communication power supply in power communication network[J]. Inner Mongolia Science Technology & Economy, 2024(15): 134-137.
- [3] 李鑫. 智能变电站数字化站用电源交直流系统分析[J]. 光源与照明, 2024(10): 192-194.  
LI X. Analysis of AC/DC system of digital power supply for smart substation[J]. Lamps & Lighting, 2024(10): 192-194.
- [4] 尚宇炜, 周莉梅, 马钊, 等. 数字化主动配电系统初探[J]. 中国电机工程学报, 2022, 42(5): 1760-1773.  
SHANG Y W, ZHOU L M, MA Z, et al. Preliminary study of digitalized active distribution systems[J]. Proceedings of the CSEE, 2022, 42(5): 1760-1773.
- [5] 蒲天骄, 赵琦, 王新迎. 电力人工智能技术研究框架、应用现状及展望[J]. 电网技术, 2025: 1-21.  
PU T J, ZHAO Q, WANG X Y. Technology framework, application status, and prospects on electric power artificial intelligence[J]. Power System Technology, 2025: 1-21.
- [6] 李雪峰, 刘海莹, 刘高华, 等. 基于深度学习的输电线路销钉缺陷检测[J]. 电网技术, 2021, 45(8): 2988-2995.  
LI X F, LIU H Y, LIU G H, et al. Transmission line pin defect detection based on deep learning[J]. Power System Technology, 2021, 45(8): 2988-2995.
- [7] 颜宏文, 陈金鑫. 基于改进YOLOv3的绝缘子串定位与状态识别方法[J]. 高电压技术, 2020, 46(2): 423-432.  
YAN H W, CHEN J X. Insulator string positioning and state recognition method based on improved YOLOv3 algorithm[J]. High Voltage Engineering, 2020, 46(2): 423-432.
- [8] 刘思言, 王博, 高昆仑, 等. 基于R-FCN的航拍巡检图像目标检测方法[J]. 电力系统自动化, 2019, 43(13): 162-168.  
LIU S Y, WANG B, GAO K L, et al. Object detection method for aerial inspection image based on region-based fully convolutional network[J]. Automation of Electric Power Systems, 2019, 43(13): 162-168.
- [9] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2016: 779-788.
- [10] 陈嘉琛, 俞曜辰, 陈中, 等. 基于改进YOLOv3的输电线路缺陷识别方法[J]. 南方电网技术, 2021, 15(1): 114-120.  
CHEN J C, YU Y C, CHEN Z, et al. An improved method for defect identification of transmission lines based on YOLOv3[J]. Southern Power System Technology, 2021, 15(1): 114-120.
- [11] 缪希仁, 林志成, 江灏, 等. 基于深度卷积神经网络的输电线路防鸟刺部件识别与故障检测[J]. 电网技术, 2021, 45(1): 126-133.  
MIAO X R, LIN Z C, JIANG H, et al. Fault detection of power tower anti-bird spurs based on deep convolutional neural network[J]. Power System Technology, 2021, 45(1): 126-133.
- [12] 蔡正奕, 赵杰煜, 朱峰. 融合图像特征的单阶段点云目标检测[J]. 计算机工程与应用, 2023, 59(9): 140-149.  
CAI Z Y, ZHAO J Y, ZHU F. Single-stage object detection with fusion of point cloud and image feature[J]. Computer Engineering and Applications, 2023, 59(9): 140-149.
- [13] 黄悦华, 刘恒冲, 陈庆, 等. 基于USRNet与改进YOLOv5x的输电线路绝缘子故障检测方法[J]. 高电压技术, 2022, 48(9): 3437-3446.  
HUANG Y H, LIU H C, CHEN Q, et al. Transmission line insulator fault detection method based on USRNet and improved YOLOv5x[J]. High Voltage Engineering, 2022, 48(9): 3437-3446.
- [14] 肖俊杰, 潘睿志, 李超, 等. 基于改进YOLOv5s绝缘子缺陷检测技术研究[J]. 电子测量技术, 2022, 45(24): 137-144.  
XIAO C J, PAN R Z, LI C, et al. Research on defect detection technology based on improved YOLOv5s insulator[J]. Electronic Measurement Technology, 2022, 45(24): 137-144.
- [15] 何锦强, 李锐海, 李昊, 等. 基于Yolo v5与Grabcut的架空线路绝缘子可见光图像自动识别与分割方法[J]. 南方电网技术, 2023, 17(6): 128-135.  
HE J Q, LI R H, LI H, et al. Visible light image automatic recognition and segmentation method for overhead power line insulators based on Yolo v5 and Grabcut[J]. Southern Power System Technology, 2023, 17(6): 128-135.
- [16] 李新海, 曾令诚, 卢泳茵, 等. 一种小样本条件下的绝缘子外套裂纹缺陷图像识别方法[J]. 南方电网技术, 2021, 15(12): 86-94.  
LI X H, ZENG L C, LU Y Y, et al. An image recognition



- method for insulator jacket crack defects under small sample conditions[J]. Southern Power System Technology, 2021, 15(12): 86-94.
- [17] 王媛彬, 李媛媛, 段誉, 等. 基于轻量骨干网络和注意力结构的变电设备红外图像识别[J]. 电网技术, 2023, 47(10): 4358-4369.
- WANG Y B, LI Y Y, DUAN Y, et al. Infrared image recognition of substation equipment based on lightweight backbone network and attention mechanism[J]. Power System Technology, 2023, 47(10): 4358-4369.
- [18] 薛昂, 姜恩宇, 张文涛, 等. 基于窗口自注意力网络与YOLOv5融合的输电线路通道异物检测[J]. 上海交通大学学报, 2025: 1-22.
- XUE A, JIANG E Y, ZHANG W T, et al. Detection of foreign bodies in transmission line channels based on the fusion of swin transformer and YOLOv5[J]. Journal of Shanghai Jiao Tong University, 2025: 1-22.
- [19] 王梓耀, 罗庆全, 萧文聪, 等. 电网厂站接线图人工智能识别关键方法[J]. 电力系统自动化, 2023, 47(2): 115-124.
- WANG Z Y, LUO Q Q, XIAO W C, et al. Key method of wiring diagram recognition with artificial intelligence for power plant and substation[J]. Automation of Electric Power Systems, 2023, 47(2): 115-124.
- [20] 张彦龙, 许丹, 曹晖, 等. 基于多阶段协同的电网厂站接线图智能识别[J]. 电网技术, 2024, 48(10): 4341-4350.
- ZHANG Y L, XU D, CAO H, et al. Intelligent recognition of power grid station wiring diagrams based on multi-stage collaboration[J]. Power System Technology, 2024, 48(10): 4341-4350.
- [21] 曹倩倩. 电力铭牌图像中的文本检测与识别方法研究[D]. 南京: 南京理工大学, 2020.
- CAO Q Q. Research on text detection and recognition method in power nameplate image[D]. Nanjing: Nanjing University of Science and Technology, 2020.
- [22] WOO S, DEBNATH S, HU R H, et al. ConvNeXt V2: co-designing and scaling ConvNets with masked autoencoders[C]// Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2023: 16133-16142.
- [23] LIU Z, MAO H Z, WU C Y, et al. A ConvNet for the 2020s[C]// Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2022: 11966-11976.
- [24] 周乐. 基于YOLO v4以及CRNN+CTC算法的中文车牌识别系统[D]. 南京: 南京邮电大学, 2022.
- ZHOU L. Chinese license plate recognition system based on YOLO v4 and CRNN+CTC algorithm[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2022.

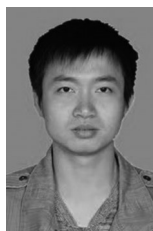
## [作者简介]



刘林 (1994-), 男, 博士, 中国南方电网电力调度控制中心工程师, 主要研究方向为人工智能在电力系统通信中的应用。



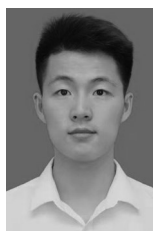
杨志敏 (1982-), 男, 博士, 中国南方电网电力调度控制中心正高级工程师, 主要研究方向为电力通信运行及技术支持系统、人工智能在电力通信中的应用。



黄强 (1982-), 男, 中国南方电网电力调度控制中心高级工程师, 主要研究方向为电力通信运行及技术支持系统、电力通信电源管理。



杨经纬 (1997-), 男, 中国南方电网电力调度控制中心工程师, 主要研究方向为电力通信运行及人工智能在电力通信中的应用。



陈一童 (1995-), 男, 中国南方电网电力调度控制中心工程师, 主要研究方向为电力通信运行及技术支持系统。