

# 基于状态空间模型的输电线路隐患目标检测方法

贺慧心<sup>1</sup>, 樊永生<sup>1</sup>, 崔文红<sup>2</sup>

(1. 中北大学 计算机科学与技术学院, 山西 太原 030051; 2. 太原理工大学 数学学院, 山西 太原 030600)

**摘要:** 输电线路巡检是保证电力稳定运输的重要手段, 现有的隐患检测方法难以做到在保障模型实时性的同时有效建模全局信息。本文在YOLOv8的基础上引入状态空间模型, 通过分块操作与低秩近似实现了空间长距离隐患的检测, 以较低的复杂度提取全局信息。在状态空间模型中, 使用交叉扫描和动态多路径激活机制来解决状态空间模型对图像数据的方向不敏感和非因果特性, 捕获隐患目标结构和模式信息, 并利用空间上下文信息识别出局部特征; 在目标分类和定位阶段, 设计了基于空间对齐的分离式检测头, 通过不同的表征方式对齐空间错位, 增强分类和定位的精确性。最后, 通过实验证明了新模型在平均精度和帧率方面优于主流的单阶段和双阶段目标检测模型, 精度提升了2.8%。

**关键词:** 状态空间模型; YOLO; Mamba; 目标检测; 输电线路隐患

**中图分类号:** TJ303.4 **文献标识码:** A **doi:** 10.62756/jnuc.issn.1673-3193.2024.07.0007

**引用格式:** 贺慧心, 樊永生, 崔文红. 基于状态空间模型的输电线路隐患目标检测方法[J]. 中北大学学报(自然科学版), 2025, 46(4): 508-516.

HE Huixin, FAN Yongsheng, CUI Wenhong. Target detection method of hidden danger targets in transmission lines based on state space models[J]. Journal of North University of China(Natural Science Edition), 2025, 46(4): 508-516.

## Target Detection Method of Hidden Danger Targets in Transmission Lines Based on State Space Models

HE Huixin<sup>1</sup>, FAN Yongsheng<sup>1</sup>, CUI Wenhong<sup>2</sup>

(1. School of Computer Science and Technology, North University of China, Taiyuan 030051, China;

2. School of Mathematics, Taiyuan University of Technology, Taiyuan 030600, China)

**Abstract:** Transmission line inspection is an important means to ensure the stable transportation of electricity, and existing hidden danger detection methods are difficult to effectively model global information while ensuring real-time model performance. This article introduced a state space model based on YOLOv8, and achieved long-distance hidden danger detection in space through block operation and low rank approximation, and global information was extracted with lower complexity. In the state space model, cross scanning and dynamic multi-path activation mechanisms were used to address the direction insensitivity and non-causal characteristics of the state space model to image data, capture hidden target structure and pattern information, and using spatial context information to identify local features; In the stage of target classification and localization, a separated detection head based on spatial alignment was designed to align spatial misalignment through different representation methods to enhance the accuracy of classification and localization. Finally, experiments show that

收稿日期: 2024-07-08

作者简介: 贺慧心(1999-), 女, 硕士生, 主要从事目标检测图像处理方面的研究。

通信作者: 樊永生(1967-), 男, 教授, 博士, 主要从事图形图像与信息处理等方面的研究。E-mail: fanys@nuc.edu.cn。

the new model outperforms mainstream single-stage and two-stage object detection models in terms of average accuracy and frame rate with an accuracy improvement of 2.8%.

**Key words:** state space models; YOLO; Mamba; target detection; hidden dangers in transmission lines

## 0 引言

输电线路是电力系统的基础设施,其安全稳定运行对于保障社会生产和居民生活至关重要<sup>[1]</sup>。然而,输电线路易受自然环境和人为因素的影响,出现各种隐患,如线路老化、杆塔倾斜、绝缘子损坏等,这些隐患可能导致大规模停电甚至引发安全事故<sup>[2-4]</sup>。因此,开发高效的输电线路隐患目标检测技术具有重要意义。采用先进的人工智能技术可以实现对输电线路隐患的自动识别和评估,显著提高检测效率和准确性,预防故障发生,确保电力系统稳定运行<sup>[5]</sup>,同时可以减少劳动力和时间成本。

输电线路隐患检测存在一定难点:1)目标尺度变化大;2)输电线路布置范围广,待检测目标背景复杂;3)图像质量容易受到环境干扰;4)部件之间容易互相遮挡。现有的基于深度学习的输电线路缺陷检测模型分为单阶段模型和双阶段模型两类。双阶段模型先生成候选区域,再对每个区域进行分类和定位,其准确率高,适用性强,可

解释性好,但训练速度慢,参数多,耗资源多,难以满足实时性要求<sup>[6]</sup>。赵振兵等<sup>[7]</sup>使用双阶段模型 Faster R-CNN 对输电线路中几种金具进行检测,取得了较高的精度;翟永杰等<sup>[8]</sup>采用迁移学习训练 Faster R-CNN,提升了模型对缺陷绝缘子的检测精度。单阶段模型能同时生成候选区域并预测出物体的类别和定位,网络结构简单,检测速度快,对小目标检测精度高,已成为当前主流算法。但该方法易受背景干扰,对大目标检测精度差。例如,李斌等<sup>[9]</sup>改进了一种单阶段无锚框目标检测模型 Single Shot Multi Box Detector (SSD),通过多尺度特征融合,提升了多尺度特征表达能力,对绝缘子及其缺陷检测取得了 94.3% 和 91.2% 的平均检测精度;徐昊<sup>[10]</sup>先对图像进行超分辨率重构,再用改进的 YOLOv5 模型检测绝缘子故障,获得了较高精度;为了追求更高精度和召回率,臧积业等<sup>[11]</sup>牺牲了一定的速度,结合 Deformable-DETR 和 YOLO 模型在多个关键目标的检测上获得了 96.8% 的精度。以上方法的优缺点对比如表 1 所示。

表 1 文献中方法的优缺点

Tab. 1 Advantages and disadvantages of the methods in the literature

| 模型    | 方法  | 优点                        | 缺点                        |
|-------|---|---------------------------|---------------------------|
| 双阶段模型 | Faster R-CNN <sup>[7]</sup>                               | 精度较高                      | 模型复杂,无法实时巡检               |
|       | 基于迁移学习的 Faster R-CNN <sup>[8]</sup><br>SSD <sup>[9]</sup> | 降低训练成本和模型复杂度<br>速度快,复杂度很低 | 易受背景干扰,鲁棒性差<br>对多类目标识别效果差 |
| 单阶段模型 | YOLOV5 结合超分辨率重构 <sup>[10]</sup>                           | 精度较高                      | 超分辨率重构导致模型不是端到端,训练成本较高    |
|       | Deformable-DETR 结合 YOLO <sup>[11]</sup>                   | 能够很好的平衡精度和速度              | 缺乏对于全局信息的建模,对遮挡目标检测效果不佳   |

综上,现有模型通常基于 CNN 或 Transformer,Transformer 能够捕捉远程依赖关系,具备全局感受,但计算复杂度高且缺乏有益的归纳偏好。CNN 虽然能够进行轻量化结构设计并更容易收敛,但缺乏对于全局信息的捕捉。这些方法对图像质量较为敏感,依赖于特定的先验场景,对遮挡目标效果差。

输电线路图像中绝缘子体积较大但容易受到遮挡,夹具尺度较小且特征模糊,因此容易漏检或误检。本文在 YOLOv8 基础上将状态空间模型这一新方法引入输电线路隐患目标检测领域。首先通过交叉扫描机制对输电线路图像特征进行四种策略的扫描,缓解其对方向不敏感的特性;其次,设计动态多路径激活机制增强对非因果数据

的建模能力;最后通过空间对齐的分离式检测头提高模型检测精度和速度。

## 1 相关工作

### 1.1 单阶段目标检测模型 YOLOv8

由于输电线路的巡检需要在满足实时性检测的前提下保证较高的检测精度,因此需要合理平衡模型的精度和速度。单阶段模型 YOLOv8 具备上述特点,适用于输电线路隐患目标检测任务。具体的,YOLOv8 模型以 YOLOv5 为基础,在骨干网络中并行更多梯度流分支,获取更多梯度流信息,在轻量化的同时提高了精度。检测头从共

享头改为解耦头以解耦分类和定位任务,能够分区关注隐患目标的类别和位置信息。并且,其使用无锚框思想,该方法利用关键点或中心点以及目标的边界信息来表示物体,更适用于较为密集的输电线路隐患检测任务。

YOLOv8模型的结构如图1所示,分为骨干、颈部、头部三部分,首先将输电线路图像输入骨干网络,提取图像中待检测目标的结构信息。再将所提取的特征输入颈部网络,融合骨干网提取的多尺度特征,增强特征表达能力。最后,将融合后的特征输入检测头,计算并输出最终的预测结果。

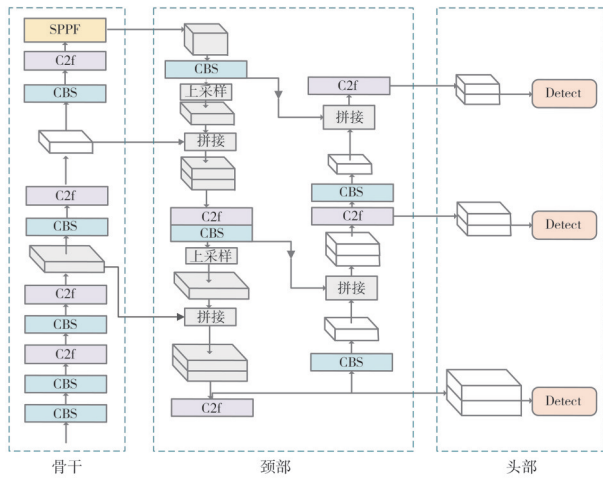


图1 YOLOv8模型结构

Fig. 1 YOLOv8 model structure

## 1.2 状态空间模型

状态空间方程是时序分析领域采用的特征提取方法,在因果数据处理中表示系统状态的动态演变。状态空间模型以较低的计算量进行全局建模,被逐渐引入目标检测领域。在输电线路缺陷检测任务中,状态空间模型表达了图像中缺陷目标的特征状态如何从细粒度转变为粗粒度,也就是纹理细节信息如何转换为空间语义信息。具体的,状态方程通常表达为线性时不变系统<sup>[12]</sup>

$$\begin{aligned} h'(t) &= Ah(t) + Bx(t), \\ y(t) &= Ch(t). \end{aligned} \quad (1)$$

式(1)中包含了状态更新方程和观测方程,分别用于描述系统内部状态随时间的演化以及从状态获得观测值。状态更新方程的输出将作为系统在下一个时刻的内部状态,并影响接下来的观测生成。其中, $y(t) \in \mathbf{R}^N$ 代表输出的特征图,是通过输入特征图 $x(t) \in \mathbf{R}^N$ 和隐藏状态 $h(t) \in \mathbf{R}^N$ 计算得到的,参数 $A \in \mathbf{R}^{N \times N}$ 是状态转移矩阵, $B, C \in \mathbf{R}^N$ 代表投影矩阵,用于特征提取的可训练参数,可以将高维特征降维或变换到新的特征空间中。为了使其嵌入YOLOv8这一深度学习模型中,需要以离散化方程实现上述公式<sup>[12]</sup>。

$$h_k = \bar{A}h_{k-1} + \bar{B}x_k, y_k = \bar{C}h_k, \bar{A} = e^{\Delta A},$$

$$\bar{B} = (\Delta A)^{-1} (e^{\Delta A} - I) \cdot \Delta B. \quad (2)$$

因此,输出可表示为卷积的形式<sup>[12]</sup>

$$K = (\bar{C}\bar{B}, \bar{C}\bar{A}\bar{B}, \dots, \bar{C}\bar{A}^{L-1}\bar{B}), y = xK, \quad (3)$$

式中: $L$ 代表输入序列的长度; $K$ 代表卷积核。

## 2 改进后的模型

YOLOv8的整体流程如图2所示,从数据输入到检测结果输出经过了高效优化的4个步骤。1)如图3所示,输入图像经过Focus Layer分块后,通过Mamba改进骨干网络提取特征,获取具备长距离依赖性的全局信息。这一步整合了不同分辨率信息,以保证特征表达的丰富性。2)这些特征被传递到包含FPN和PANet的颈部网络,FPN自下而上提取不同尺度特征,PANet则自上而下增强特征交互,使模型在多尺度物体检测上更具鲁棒性。3)特征送入Anchor-free基于任务对齐的检测头,通过挖掘多种表征方式之间的联系,捕捉分类和定位任务之间的隐藏关系,对齐二者之间的空间错位,增强分类和定位特征,输出更精准的预测结果。4)模型通过改进的CIoU和分类损失函数优化边界框和分类结果,生成检测输出。在推理阶段,YOLOv8应用混合精度和量化,确保在资源受限设备上实现高效、实时检测。

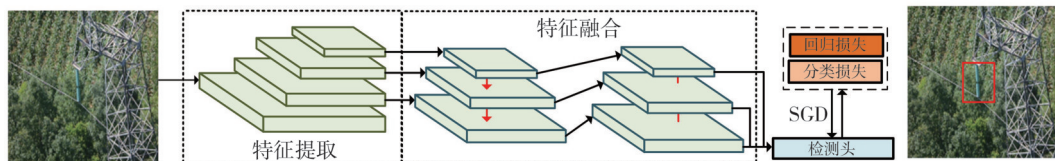


图2 模型整体流程

Fig. 2 Overall process of the model

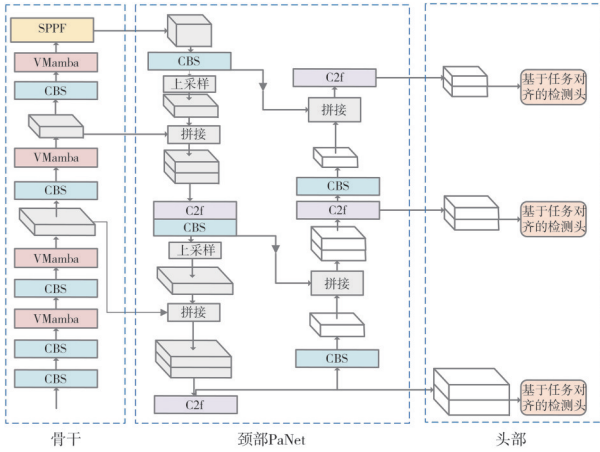


图 3 改进后的 YOLOv8 模型结构  
Fig. 3 Improved YOLOv8 model structure

### 2.1 基于状态空间模型的骨干网络

Mamba<sup>[13]</sup>模块结合了分块操作和低秩近似,进一步降低了长距离依赖建模的计算需求,具体结构如图 4 所示。分块操作将特征图划分为多个小块,分别建模局部关系并在合适位置进行特征融合。低秩近似则减少了特征图的维度和计算量,在降低模型复杂度的情况下仍能获得较好的长距离依赖建模效果。

对于 3 通道 RGB 输电线路图像  $x \in \mathbf{R}^{h \times w \times 3}$ ,先通过二维卷积模块(CBS)对图像编码,获取粗粒度特征。再次卷积操作会使图像特征采样降至原来的

1/4,通道数会增加至 32。表示为

$$\hat{x} = \phi_{C \times 2}(x), \quad (4)$$

式中:  $\phi_{C \times 2}$  代表两次卷积核大小为 3, 步长为 2 的卷积操作,用批量标准化和激活函数确保训练过程的梯度流稳定。

Mamba 这种因果数据分析模块是把图像分块后展开,获取  $n$  个一维向量,再将其依次输入模块内进行计算。这忽略了图像的空间结构,导致模型只能沿某个固定方向学习图像特征。交叉扫描模块是先将图片分块,再将每个块视为矩阵并展开为向量。为了使模型学习不同的空间分布,破坏模型方向不敏感的特性,通过 4 种排布方式排列这些向量,扫描的方向分别是: 1) 从左到右,从上到下; 2) 从右到左,从下到上; 3) 从上到下,从左到右; 4) 从下到上,从右到左。为了保证不丢失重要的位置信息,选择嵌入可学习位置编码,表示为

$$T = \text{Flatten}(\hat{x}) + P, \quad (5)$$

式中:  $T \in \mathbf{R}^{L \times D}$ ,  $P \in \mathbf{R}^{L \times D}$ , 代表将特征图展开为  $L$  个长度为  $D$  的序列,  $L$  是分块数,  $D$  是块内像素点数量,  $P$  是位置编码。将展开后的特征图分别沿 4 个方向排序,得到具备四种空间排序方式的序列,表示为

$$T^k = \text{CSM}_k(T), \quad (6)$$

式中:  $k=1, 2, 3, 4$ , 代表 4 种排序方式,  $\text{CSM}()$  代表用于排序的函数<sup>[14]</sup>。

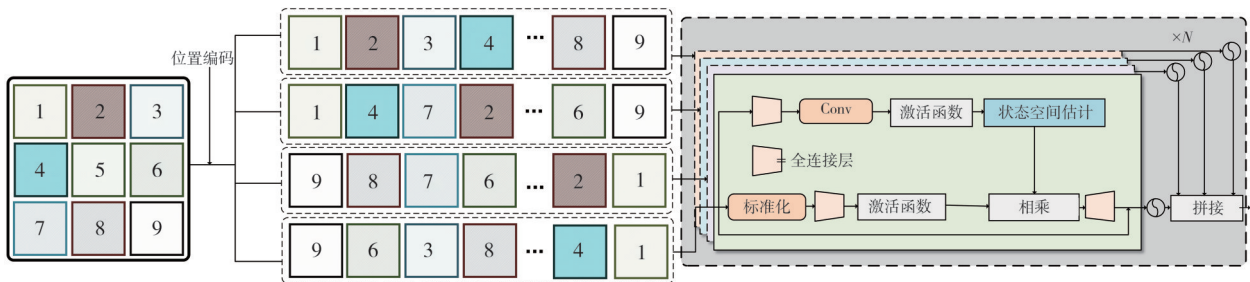


图 4 Mamba 模块的结构

Fig. 4 The specific structure of the Mamba module

多路径激活机制是聚合不同空间分布的结果,将 4 种排列方式输入 4 个并行的状态空间模型中,分别计算后得到 4 个结果并聚合。该机制可以增强状态空间模型对输电线路隐患图像这一非因果数据的建模能力,捕捉输电线路隐患图像中的全局依赖性,聚合更为丰富的细粒度特征。单个模块的状态空间估计表示为

$$y = \varphi_{sm}(T), \quad (7)$$

式中:  $y$  表示状态空间估计的输出,  $T^0 =$

$T \in \mathbf{R}^{L \times D}$ ,  $\varphi_{sm}()$  代表多路径动态激活的状态空间估计模块,每个阶段共有  $N$  块。 $\varphi_{sm}()$  函数来源于一维因果序列建模,因此不具备捕捉空间位置信息的能力,限制了对图像特征的表达。为此,本文将经过四个方向排序的序列输入四个并行的状态空间估计模块,表示为

$$y^k = \varphi_{sm}^k(T^k). \quad (8)$$

再将几个并行的模块输出进行门控聚合,整合 4 种排序序列的信息,获得最终结果,表示为

$$\hat{y} = \sum_{k=1}^4 y^k g_k, \quad (9)$$

$$g_k = \text{Softmax}(\varphi_{\text{pro}}(\varphi_{\text{mean}}(\varphi_{\text{concat}}(y^k))))), \quad (9)$$

式中： $\varphi_{\text{concat}}()$ 用于拼接4个并行状态空间估计模块的输出， $\varphi_{\text{mean}}()$ 用于沿长度维度求均值操作， $\varphi_{\text{pro}}()$ 将求均值后的投影压缩为长度为4的向量，Softmax 激活函数用于获得每个并行模块的最终归一化权重。

## 2.2 基于空间对齐的分离式检测头

第2.1节中提取了模型的高阶特征，并通过颈部网络PANet(Path Aggregation Network)进行特征融合，增强了多尺度特征的表达能力。在目标检测模型中，PANet通过自下而上和自上而下两条路径，将不同层次的多尺度特征进行融合，有助于提升对不同大小目标的检测能力，从而提高模型整体检测的准确性。在多尺度目标检测任务中，PANet会输出3个不同尺度的特征，分别对应于输入图像的1/8、1/16和1/32的尺度。这些特征将被分别送入不同的检测头中执行最终的分类和回归计算，以输出检测结果。

然而，YOLOv8中使用的解耦检测头虽然能使分类和定位任务共享特征映射的不同部分，但忽略了分类和定位任务之间理应具备的某种潜在关系。在检测头中增加用于增强分类和定位任务之前的空间对齐模块，利用中心点信息增强预测框，实现分类与定位任务间的空间对齐，从而提升定位精度。

空间对齐模块受Transformer<sup>[15]</sup>的启发，将捕捉二者之间的关系作为Key，用Key增强定位任务Query，提升定位任务的精度<sup>[16]</sup>。Query和Key和注意力机制的含义完全相符，Query代表查询，Key代表键。由于输电线路隐患图像中目标特征具有差异显著，分类任务在已有特征基础上已具备较高鲁棒性，因此仅通过空间对齐模块单向增强定位任务精度，进一步提升边界框的准确性，具体的融合方式表示为

$$f_i^q = f_i^q + \sum_j T(f_i^q, f_j^k, g_i^q, g_j^k) \text{Linear}_v(f_j^k), \quad (10)$$

式中： $f_i^q, f_i^q, g_i^q$ 分别为Query中像素点*i*的输入、输出和几何矢量； $f_j^k, g_j^k$ 为Key中像素点*j*的输入和几何矢量； $T(\cdot)$ 为计算相似度的函数； $\text{Linear}(\cdot)$ 为线性投影，用于获得Value。

函数 $T(\cdot)$ 表示为

$$T(f_i^q, f_j^k, g_i^q, g_j^k) =$$

$$\text{Softmax}[T^A(f_i^q, f_j^k) + T^G(g_i^q, g_j^k)], \quad (11)$$

式中： $T^A()$ 代表Query中像素点*i*和Key中像素点*j*的外观相似度； $T^G()$ 代表Query中像素点*i*和Key中像素点*j*的相对位置几何项。

如图5所示，Query代表预测框，形状为 $H \times W \times 4$ ，因此几何向量为4，4代表通过每个像素点预测一个目标边界框。Key是中心点表征，形状为 $H \times W \times 2$ ，几何向量为2，显然Query和Key的形状不符，无法进行相似度计算，因此通过卷积核为1的卷积将Query压缩为 $H \times W \times 2$ 再进行后续计算。

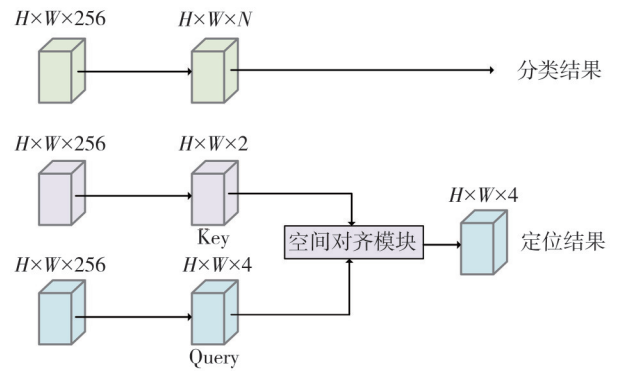


图5 检测头的具体结构

Fig. 5 The specific structure of the detection head

## 3 实验结果与分析

### 3.1 数据集

本文数据集来源于国家电网和某省电业局实际巡检中采集到的绝缘子数据集；在其基础上对绝缘子，夹具，屏蔽环共3类目标，2721张图片重新进行了标注。

实际巡检过程中无人机拍摄可能面临各种角度和光照等变化，因此本文随机抽取了部分样本，分别进行高斯噪声、颜色抖动及仿射变换处理，模拟拍摄过程中可能受到的干扰、环境变化和角度变化，增强模型泛化能力。扩充后的数据集共包含6230张图片。从构建的数据集中挑选80%作为训练集，10%作为测试集，10%作为验证集，采用分层抽样对原始数据进行随机划分，同时保留训练和测试集中不同类别内的样本分布，在此基础上进行了模型训练，超参数设定如表2所示。

表 2 超参数设置

Tab. 2 Hyperparameter settings

| 超参数     | 选取值     |
|---------|---------|
| 批量大小    | 32      |
| 训练轮次    | 200     |
| 初始学习率   | 0.01    |
| 优化器     | SGD     |
| 输入图像分辨率 | 640×640 |

### 3.2 评测指标

本文使用常用的指标均值平均精度 mAP 来评估模型的性能, 评估值越高说明性能越高, 反之亦然。

$$P = \frac{N_{TP}}{N_{TP} + N_{FN}}, R = \frac{N_{TP}}{N_{TP} + N_{FP}},$$

$$AP = \int_0^1 p(r) dr, mAP = \frac{\sum_{i=1}^n AP_i}{n}, \quad (12)$$

式中: 真阳性(TP)、真阴性(TN)、假阳性(FP)和假阴性(FN)为目标检测模型检测结果的 4 种类型(公式中用到了 3 种),  $N_{TP}$  为真正的正样本或目标的数量,  $N_{FP}$  为将负样本或非目标错误地标记为正样本或目标的数量,  $N_{FN}$  为错误的未能检测到的正样本或目标的数量,  $N_{TN}$  为被正确识别为负样本的数量;  $P$  为精确率;  $R$  为召回率;  $AP$  为平均精度;  $mAP$  为均值平均精度, 表示所有类别的  $AP$  的平均值,  $mAP$  的值越大, 模型的综合性能越好;  $n$  为类别数。

### 3.3 模型的性能验证

模型损失分为三部分, 分别是边界框损失、分类损失和关键点损失。边界框损失用于衡量预测边界框与实际边界框之间的差异。分类损失用于衡量预测目标类别与实际类别之间的差异。关键点损失通过对标签附近的数值聚焦, 使标签处的概率密度变大。通过观察损失函数的变化能有效判断模型是否收敛, 是否出现过拟合, 从而使模型性能变化直观可视化。

图 6 所示, 在模型训练的 11 个轮次内, 损失迅速下降, 预测值与真实值差异很大, 因此反向传播所求得的梯度值就越大, 此时模型精度迅速提升。在 11 到 21 个训练轮次中, 模型损失下降速度逐渐放缓, 但下降趋势仍明显。21 到 111 轮, 损失仍在下降, 但基本难以通过肉眼判断, 此时模型向收敛靠拢, 仅对参数微调。111 到 200 轮, 模型损失在最后 20 个轮次内基本保持不变, 代表模型已收敛, 此时再训

练会导致模型过拟合, 因此停止训练。

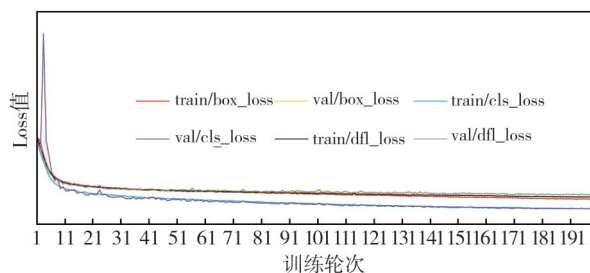


图 6 模型的收敛曲线

Fig. 6 Convergence curve of the model

通过模型的混淆矩阵和准确率-召回率曲线(Precision-Recall Curve, PR 曲线)分析模型对不同类别目标的检测精度和漏检程度, 如图 7 和图 8 所示, 1、2、3 分别代表夹具、绝缘子和屏蔽环。显然, 模型对夹具、绝缘子和屏蔽环都有不同程度的漏检和误检。本文模型的漏检度相对较低, 其中绝缘子和屏蔽环的平均精度达到了 96.8% 和 95.1%, 夹具属于较小且模糊的目标, 平均精度较低, 为 82.1%。

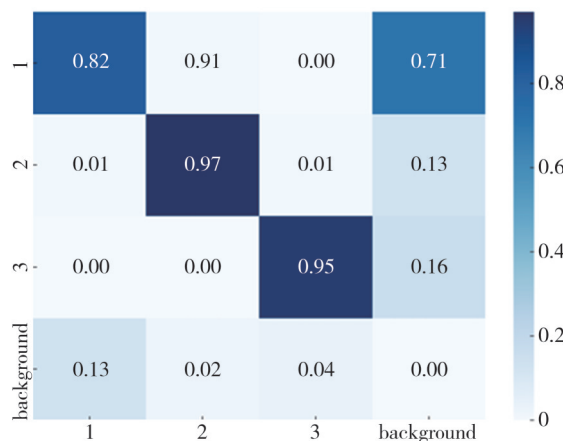


图 7 模型的混淆矩阵

Fig. 7 Confusion matrix of the model

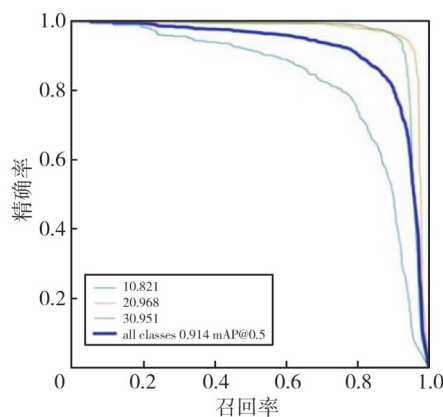


图 8 各类别缺陷的 PR 曲线

Fig. 8 PR curves for various types of defects

模型对小目标的检测性能如图9所示,对于尺寸较小的夹具类目标,模型的检测置信度相对较低,但几乎未出现漏检或误检现象。此外,模型能够有效检测到小尺度和模糊的金具,这得益于状态空间模型所提供的全局感受野,使得模型能够精确地关注目标区域。特别是左上角绿色框标记的检测结果中,一个被严重遮挡的绝缘子仍然被成功识别。尽管数据集中未对遮挡目标进行专门标注,但模型借助状态空间特征分布,减弱了因遮挡导致的信息缺失。与此同时,基于空间对齐的分离式检测头能够有效捕捉分类与定位任务间的关联性,显著提升边界框的精度,从而使边界框与目标区域的契合度更高。

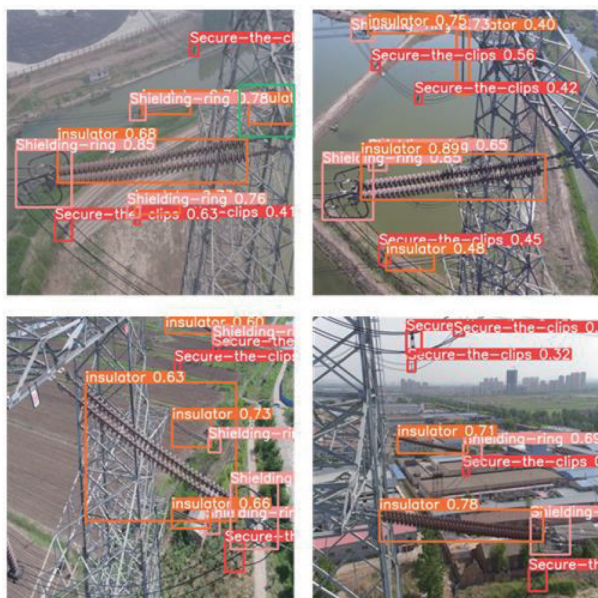


图9 模型输出的检测结果

Fig. 9 Detection results of model output

### 3.4 消融实验

消融实验用于定量分析改进的有效性。本文以YOLOv8为基准,先分析基于状态空间模型的骨干网络。为了缓解Mamba的因果特性和空间不敏感,采用了交叉扫描机制和动态多路径激活机制策略改进Mamba。因此,需要对这两点改进进行消融实验。

如表3所示,本文首先将基础Mamba模块替换C2f模块,由于Mamba原本为处理因果数据的模型,因此模型的mAP略有降低(减少1%)。随后通过引入交叉扫描机制对Mamba模块进行优化,使其在精度上相较C2f模块提升3.2%。进一步地,利用动态多路径激活机制整合多个并行

Mamba模块的输出,显著提升mAP表现。

表3 骨干网络的消融实验

Tab. 3 Ablation experiments of backbone networks

| 状态空间估计<br>(基础Mamba) | 交叉扫描机制 | 动态的多路径<br>激活机制 | mAP@0.5 |
|---------------------|--------|----------------|---------|
|                     |        |                | 0.882   |
| ✓                   |        |                | 0.872   |
| ✓                   | ✓      |                | 0.904   |
| ✓                   | ✓      | ✓              | 0.910   |

在骨干网络的基础上,本文在检测头上进行了消融实验,如表4所示,基于空间对齐的分离式检测头通过异构表征显著提升边界框精度,使模型整体精度提高0.4%。

表4 检测头的消融实验

Tab. 4 Ablation experiment of detection head

| 基于空间对齐的分离式检测头 | mAP@0.5 |
|---------------|---------|
|               | 0.910   |
| ✓             | 0.914   |

### 3.5 对比实验

本研究将提出的新模型与当前主流的五种目标检测模型进行了系统性对比,包括带有锚框的单阶段模型YOLOv5,以及无锚框的单阶段模型SSD、YOLOv8、YOLOv10。此外,对比了两阶段目标检测模型Cascade R-CNN及端到端模型Deformable-DETR。根据表5的结果,相同测试条件下,各检测模型均存在不同程度的漏检现象,但所提出模型的漏检数量明显较少,表现出较优的检测性能。相比其他检测模型,新模型展现了更高的平均精度,体现了其在目标检测任务上的优越性。

表5 对比试验

Tab. 5 Comparative test

| 模型                              | 平均精度         | 帧率        |
|---------------------------------|--------------|-----------|
| YOLOv5                          | 0.872        | 17        |
| Deformable-DETR <sup>[17]</sup> | 0.895        | 5         |
| SSD <sup>[18]</sup>             | 0.709        | 5         |
| YOLOv8                          | 0.882        | 9         |
| Cascade-RCNN <sup>[20]</sup>    | 0.885        | 4         |
| YOLOv10                         | 0.905        | 23        |
| Dino-DETR                       | 0.913        | 10        |
| <b>本文模型</b>                     | <b>0.914</b> | <b>21</b> |

各模型的检测结果如图10所示,SSD模型存在明显的漏检和误检现象,而Deformable-DETR和Cascade R-CNN在检测夹具目标时也表现出不同程度的漏检。YOLOv8模型的检测性能与本文提出的方法相近。YOLOv10和Dino-DETR各具特点,

YOLOv10 的检测速度比本文提出的模型快,但其精度相对较低,在小目标检测方面容易出现漏检。Dino-DETR 的检测精度与本模型相当,但检测速度较慢。

本文所提出的模型在边界框的生成上精度更高,与目标的贴合度更好,具有更高的帧率,体现了该模型在精度与速度之间的良好平衡。

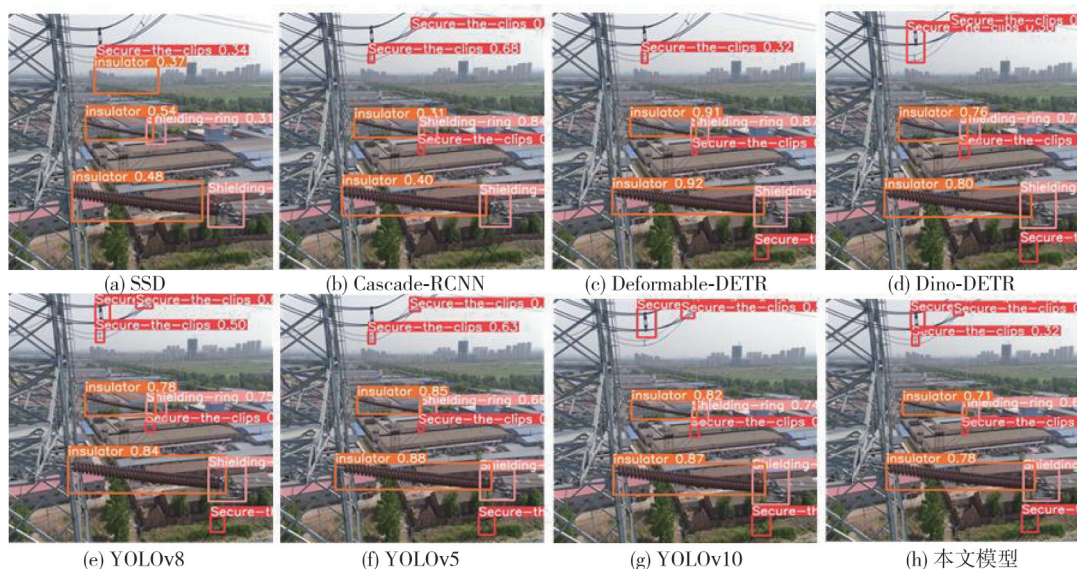


图 10 不同模型的输出结果

Fig. 10 Output results of different models

### 4 结 论

本文提出了一种基于状态空间模型的输电线路隐患目标检测方法,首次将状态空间模型技术引入 YOLOv8。用 Mamba 代替 C2f 改善长距离信息捕获和全局信息建模能力,针对 Mamba 擅长处理因果数据且对图片方向空间位置不敏感的问题,本文采用交叉扫描机制,使模型更好地学习特征中的空间结构,缓解对方向的不敏且感性;用动态多路径激活机制门控聚合不同扫描方向的状态空间估计后的输出结果,增强对非因果数据的建模能力,捕获图像的全局依赖性。对于分类和定位任务的特征分布差异,采用空间对齐的分离式检测头捕捉分类和定位之间的关系,通过增强定位任务输出的预测框,实现输电线路隐患目标的精准定位。实验证明了模型能够更好地平衡精度和速度,对比主流模型具备一定优势。

#### 参考文献:

[ 1 ] 董庆州. 面向电网输电线路部件识别的目标检测算法研究及实现[D]. 广州: 华南理工大学, 2021.

[ 2 ] 刘传洋, 吴一全. 基于深度学习的输电线路视觉检测方法研究进展[J]. 中国电机工程学报, 2023, 43(19): 7423-7446.

LIU Chuanyang, WU Yiquan. Research progress of

vision detection methods based on deep learning for transmission lines [J]. Proceedings of the CESS, 2023, 43(19): 7423-7446. (in Chinese)

[ 3 ] 刘传洋, 吴一全, 刘景景. 基于视觉的输电线路金具锈蚀缺陷检测方法研究进展[J]. 仪器仪表学报, 2024, 45(3): 286-305.

LIU Chuanyang, WU Yiquan, LIU Jingjing. Research progress of vision-based rust defect detection methods for metal fittings in transmission lines[J]. Chinese Journal of Scientific Instrument, 2024, 45(3): 286-305. (in Chinese)

[ 4 ] 陆佳政, 周特军, 吴传平, 等. 某省级电网 220 kV 及以上输电线路故障统计与分析[J]. 高电压技术, 2016, 42(1): 200-207.

LU Jiazheng, ZHOU Tejun, WU Chuanping, et al. Fault statistics and analysis of 220 kV and above power transmission line in province-level power grid[J]. High Voltage Engineering, 2016, 42(1): 200-207. (in Chinese)

[ 5 ] 蒲天骄, 乔骥, 韩笑, 等. 人工智能技术在电力设备运维检修中的研究及应用[J]. 高电压技术, 2020, 46(2): 369-383.

PU Tianjiao, QIAO Ji, HAN Xiao, et al. Research and application of artificial intelligence in operation and maintenance for power equipment [J]. High Voltage Engineering, 2020, 46(2): 369-383. (in Chinese)

[ 6 ] 林一峰, 王增平, 王彤, 等. 构建更加坚强电网安全“第一道防线”的探讨[J]. 电力系统保护与控制,

- 2022, 50(16): 170-178.
- LIN Yifeng, WANG Zengping, WANG Tong, et al. Discussion on building a stronger first line of defense for grid security [J]. *Power System Protection and Control*, 2022, 50(16): 170-178. (in Chinese)
- [7] 赵振兵, 李延旭, 甄珍, 等. 结合KL散度和形状约束的Faster R-CNN典型金具检测方法[J]. *高电压技术*, 2020, 46(9): 3018-3026.
- ZHAO Zhenbing, LI Yanxu, ZHEN Zhen, et al. Typical fittings detection method with faster R-CNN combining KL divergence and shape constraints [J]. *High Voltage Engineering*, 2020, 46(9): 3018-3026. (in Chinese)
- [8] 翟永杰, 杨珂, 王乾铭, 等. 基于混合样本迁移学习的盘型绝缘子缺陷检测[J]. *中国电机工程学报*, 2023, 43(7): 2867-2877.
- ZHAI Yongjie, YANG Ke, WANG Qianming, et al. Disc insulator defect detection based on mixed sample transfer learning [J]. *Proceedings of the CSEE*, 2023, 43(7): 2867-2877. (in Chinese)
- [9] 李斌, 屈璐瑶, 朱新山, 等. 基于多尺度特征融合的绝缘子缺陷检测[J]. *电工技术学报*, 2023, 38(1): 60-70.
- LI Bin, QU Luyao, ZHU Xinshan, et al. Insulator defect detection based on multi-scale feature fusion [J]. *Transactions of China Electrotechnical Society*, 2023, 38(1): 60-70. (in Chinese)
- [10] 徐昊. 基于机器视觉的输电线路小目标检测和缺陷识别方法研究[D]. 杭州: 浙江大学, 2022.
- [11] 臧积业, 曲朝阳, 董运昌, 等. 基于仿生视觉骨干的级联蒸馏输电线路目标检测模型[J]. *高电压技术*, 2024, 50(8): 3757-3768.
- ZANG Jiye, QU Zhaoyang, DONG Yunchang, et al. Target detection model of cascade distillation transmission line based on bionic visual backbone [J]. *High Voltage Engineering*, 2024, 50(8): 3757-3768. (in Chinese)
- [12] LI K C, LI X H, WANG Y, et al. VideoMamba: State space model for efficient video understanding [DB/OL]. (2024-03-12)[2024-07-07]. <https://arxiv.org/abs/2403.06977>.
- [13] GU A, DAO T. Mamba: Linear-time sequence modeling with selective state spaces [DB/OL]. (2024-05-31)[2024-07-07]. <https://arxiv.org/abs/2312.00752>.
- [14] LIU Y, TIAN Y, ZHAO Y, et al. Vmamba: Visual state space model [DB/OL]. (2024-01-18)[2024-07-07]. <https://arxiv.org/abs/2401.10166>.
- [15] KHAN S, NASEER M, HAYAT M, et al. Transformers in vision: A survey [DB/OL]. (2022-01-19)[2024-07-07]. <https://arxiv.org/abs/2101.01169>.
- [16] 赵丽娟. 输电线路组件及其缺陷检测的关键技术研究[D]. 北京: 华北电力大学, 2023.
- [17] ZHU X, SU W, LU L, et al. Deformable DETR: Deformable transformers for end-to-end object detection [DB/OL]. (2021-03-18)[2024-07-07]. <https://arxiv.org/abs/2010.04159>.
- [18] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]//*Computer Vision-ECCV*, 2016: 21-37.
- [19] VU T, JANG H, PHAM T X, et al. Cascade RPN: Delving into high-quality region proposal network with adaptive convolution [DB/OL]. (2019-12-04)[2024-07-07]. <https://arxiv.org/abs/1909.06720>.