



2024 年（第 17 届） 中国大学生计算机设计大赛

人工智能实践赛作品报告

作品编号： 2024004375

作品名称： 慧眼智巡-基于旋转框提示 EdgeSAM 的无人机
航拍绝缘子语义分割系统

填写日期： 2024 年 4 月 8 日

目 录

- 第 1 章 作品概述 1
 - 1.1 背景..... 1
 - 1.2 功能与特色 2
 - 1.2.1 功能 2
 - 1.2.2 特色 2
 - 1.3 应用前景..... 3
- 第 2 章 问题分析 3
 - 2.1 问题来源..... 3
 - 2.2 现有解决方案..... 4
 - 2.3 本作品要解决的痛点问题 5
 - 2.4 解决问题的思路 6
 - 2.4.1 作品的功能和性能需求..... 6
 - 2.4.2 数据集 7
- 第 3 章 技术方案 8
 - 3.1 技术方案..... 8
 - 3.2 模块介绍..... 9
 - 3.2.1 数据集采集与制备..... 9
 - 3.2.2 模型训练与结合 9
 - 3.2.3 模型部署 10
 - 3.2.4 模型性能分析..... 10
- 第 4 章 系统实现 10
 - 4.1 数据来源..... 10
 - 4.2 模型训练和改进 11
 - 4.3 模型边缘端设备部署 12
 - 4.4 界面及成果展示 12
 - 4.5 系统部署方法..... 14
 - 4.6 问题及解决方法..... 14
- 第 5 章 测试分析 15
 - 5.1 运行结果分析..... 15
 - 5.2 准确率评估..... 16
 - 5.3 运行速度评估 17
- 第 6 章 作品总结 18
 - 6.1 作品特色与创新点 18
 - 6.2 应用推广..... 19
 - 6.3 作品展望..... 19
- 参考文献..... 20

第1章 作品概述

1.1 背景

在电力系统领域中，绝缘子作为高压输电线路中的关键部件，其稳定性和安全性对于电网的整体运行具有至关重要的影响。绝缘子主要承担支撑载流导体并隔绝电流泄露至地面的重要职责，确保电网在特高压、超高压远程输电以及日常供电过程中能够稳定运行。然而，由于长期工作于高压电场等复杂自然环境下，绝缘子容易受到老化、损坏等多种不利因素的影响，这对其性能的稳定性和安全性构成了严重威胁。

因此，对绝缘子进行日常的巡检与维护工作显得尤为关键。巡检工作不仅涉及因安装不当导致的断裂、恶劣天气造成的物理损伤、表面脏污引发的污闪事故等问题的检测，还需要对绝缘子的运行状态进行准确的评估，以预防潜在的安全隐患。然而，传统的巡检方式往往依赖于人工进行，这不仅效率低下，而且给巡检人员带来了高空作业和高压环境带来的安全风险。



图 1 常见的绝缘子故障示意图

近年来，随着计算机视觉和人工智能技术的快速发展，语义分割技术为绝缘子的巡检提供了新的解决方案。语义分割技术能够实现对图像中像素级信息的精确处理，从而实现对绝缘子的精确识别和状态评估。然而，基于深度学习的语义分割技术通常需要大规模的像素级数据集进行模型训练，而绝缘子复杂的结构特性使得标注数据的获取变得困难且成本高昂。

为应对这一难题，零样本分割模型 Segment-Anything（SAM）的提出激发了广泛研究。随后各种改进型 SAM、轻量化 SAM，使 SAM 的落地应用迈出重要一步。轻量化的 EdgeSAM，让零样本语义分割模型的快速部署和高效推理成为

可能。定向边界框检测算法 YOLOv8-OBB (OBB, Oriented Bounding Box) 打破传统水平边界框 (HBB, Horizontal Bounding Box) 检测的壁垒, 充分挖掘待测目标的位置信息和姿态信息, 为 EdgeSAM 的语义分割提供更加精准的提示。基于 YOLOv8-OBB 提供提示, EdgeSAM 承担语义分割任务的思想, 我们提出了基于旋转框提示 EdgeSAM 的无人机航拍绝缘子语义分割技术。

1.2 功能与特色

1.2.1 功能

本作品开发了一种基于旋转框目标检测提示 EdgeSAM 的绝缘子目标检测和语义分割方法及装置, 包括基于 YOLOv8-OBB 旋转框目标检测方法, 绝缘子目标姿态点求算方法, 基于 EdgeSAM 的连续点框结合提示的零样本学习语义分割方法, 优化的前后处理算法, 多线程、多 CUDA 流推理方法以及基于 Nvidia Jetson Orin Nano 边缘运行装置。

本作品实现了 YOLOv8-OBB 进行绝缘子旋转框目标检测, 将待测目标的位置信息和姿态信息作为提示下发给 EdgeSAM 进行语义分割。本作品创造性地提出了框+姿态点的提示模式, 可以提高绝缘子语义分割准确性和鲁棒性。依靠 EdgeSAM 零样本学习分割, 节省了语义分割数据集采集和标注的繁杂工作, 能够快速、高效率地投入绝缘子检测工作。

1.2.2 特色

(1) 旋转框目标检测算法, 对绝缘子目标进行更精确的定位, 充分挖掘待测目标的姿态信息, 利用定向边界框精确抓取待测目标, 提高了检测的准确性和鲁棒性, 为绝缘子目标姿态的确定提供了可能;

(2) 基于 θ 角的定向边界框, 快速求算绝缘子目标的姿态信息, 创造性地提出框+姿态点的提示模式, 为 EdgeSAM 提供更加精细的分割提示;

(3) 依靠 EdgeSAM 零样本学习分割, 节省大量数据集收集和标注工作, 实现快速部署;

(4) 更加优化的图像预处理模式, 将语义分割预处理过程压缩至 15ms; 独特的掩膜后处理流程, 让语义分割结果的后处理过程压缩至 8ms。

(5) 基于 Nvidia Jetson Orin Nano 边缘端设备部署，充分优化模型推理流程和数据处理过程，利用边缘端设备有限的计算资源，实现多模型高速推理。

1.3 应用前景

本作品结合了目标检测和语义分割技术，针对绝缘子进行检测和分割，具有广泛的应用前景。其中，基于 YOLOv8-OBB 的绝缘子旋转框目标检测模型可以精准地检测绝缘子的位置和旋转角度，为电力线路维护和安全管理提供重要支持；基于 YOLOv8-OBB+EdgeSAM 的绝缘子语义分割模型可以将绝缘子与周围环境进行有效分离，有助于细化绝缘子的检测与分析，提高检测的准确性和效率。

通过部署在基于 Gradio 的 Web 端运行界面，用户可以方便地上传图像或视频进行绝缘子检测和分割，实现对电力设备的智能监测和管理。而基于 Nvidia Jetson Orin Nano 边缘端的绝缘子语义分割装置，则将这一智能检测技术推广到边缘端设备上，实现了在离线环境下对绝缘子进行高效准确的分割，为电力行业提供了更便捷、快速的解决方案。

我们创造性地提出了姿态点的概念，连续点框结合的提示模式以及数据处理方案，不断提高模型的精准度和执行效率。通过后续不断优化模型结合策略，本作品可以应用于电力设备巡检、故障诊断、安全监测等领域，将为电力行业的智能化发展和安全生产提供重要支持。

第2章 问题分析

2.1 问题来源

绝缘子在输电线路中，起到支撑导线和防止电流回地两大重要作用。绝缘子一般工作在户外，长期暴露大自然环境中，受强紫外线照射、雨水侵蚀、空气污染、气候变化等因素影响，存在断串、内部击穿等故障风险，对线路安全运行造成重大威胁。因此，输电线路绝缘子检测是电网巡检的重点。

鉴于传统人力巡检方式使得检测人员面临高空坠落以及高压击穿的风险，检测效率不高。我们考虑结合无人机技术、计算机视觉技术、深度学习等对绝缘子

进行视觉检测。然而航拍的输电线路图像信息复杂多样，借助现有的绝缘子目标检测算法，输出的结果存在大量背景像素，使得后续难以对绝缘子进行缺陷检测。

2.2 现有解决方案

目前用于绝缘子语义分割方法大致可以分为四类：基于阈值的分割方法、基于边缘的分割方法、基于区域的分割方法以及基于深度学习的分割方法。

其中，徐俊良等人提出一种将单一像素点属性判断和局部均值法相结合的绝缘子分割方法，实现了灰度非均匀的绝缘子图像的准确分割，但是计算量大，无法达到实时检测的需求^[1]；吴文海等人将绝缘子图像划分为具有平滑轮廓的子区域，采用灰度共生矩阵提取特征，并通过快速灰度共生集成算法进行计算，同时引入纹理描述符主动轮廓的非凸模型实现绝缘子的分割^[2]；杜晓亮等人首先使用 RPN 对绝缘子图像生成建议区域，再使用基于 Faster R-CNN 算法提取特征，然后利用基于 VGG16 的 FCN-8s 全卷积网络分割绝缘子与背景的基础上分别使用 VGGNet 和 AlexNet 作为特征提取网络对绝缘子图像进行语义分割^[3]，但是分割结果较粗糙，并未分割出具有边缘特征的绝缘子串。随着深度学习算法的不断发展，后来的研究者相继提出了基于改进的 U-Net 系列^[4]，YOLOv8 系列^[5]语义分割算法对绝缘子进行语义分割，在缺乏大规模公开数据集的情况下，难以对模型进行充分训练。

然而，上述方法需使用大量语义分割数据进行训练，但每个绝缘子包含多个绝缘薄片（如图 2 所示），导致绝缘子的高质量语义分割标签难以获取。因此，当前仍缺少绝缘子的高质量语义分割数据集。通用语义分割模型可以实现对绝缘子的零样本语义分割，但多数绝缘子在图像中的倾斜角度较大（如图 3 所示），导致绝缘子在边界框提示中的占比较小，使得通用语义分割模型根据边界框提示的分割结果不理想；同时，通用语义分割模型参数较多，导致其难以部署在无人机边缘端。鉴于此，本作品提出一个基于旋转框提示 EdgeSAM 的无人机航拍绝缘子语义分割系统，为通用语义分割模型设计了精确的绝缘子姿态信息提示方案，实现了通用语义分割模型在边缘端的部署。该模型减轻了基于深度学习的语义分割模型对高质量语义分割数据集的依赖，推动无人机全自动电力巡检的发展进程。



图 2 绝缘子形态示意图



图 3 绝缘子在边界框中占比较少示意图

2.3 本作品要解决的痛点问题

(1) 缺乏高质量绝缘子语义分割数据集

高质量绝缘子语义分割数据集的缺乏已成为制约相关研究的瓶颈。这主要体现在两个方面：一是缺乏公开大规模的数据集，使得研究者难以获取充足的训练样本；二是人工制备数据集困难重重。由于绝缘子结构复杂，每个绝缘子包含多个薄片，且边缘特征显著，这使得人工进行语义分割标注时难度极大。因此，现有的语义分割模型在缺乏足够数据支持的情况下，其训练效果往往难以达到理想状态。

解决方案：探索半自动标注数据集的方法，减轻数据集制备的负担。同时，考虑零样本学习语义分割，寻找零样本分割模型。

(2) 传统目标检测绝缘子在水平边界框中的占比较小的问题

由于绝缘子通常具有较长的形态，并可能存在一定的旋转角度，传统目标检

测算法在处理绝缘子时，往往会出现绝缘子在水平边界框中占比较小的情况。这不仅无法充分挖掘绝缘子的旋转信息，还可能导致预测锚框过大，进而使得检测结果中包含大量的冗余背景信息。这些冗余信息对后续的绝缘子检测工作产生了极大的干扰。

解决方案：制定复杂背景下的绝缘子图像采集计划，通过半自动标注和人工矫正的方式，制备绝缘子定向边界框检测数据集。训练基于定向边界框绝缘子目标检测模型，充分挖掘绝缘子的位置信息和姿态信息，排除环境像素的干扰。

（3）框提示对 EdgeSAM 提示不足的问题

在利用 EdgeSAM 进行绝缘子分割时，我们发现仅提供绝缘子的位置信息作为提示，往往得到的分割掩码并不完整，甚至可能不产生任何掩码。此外，EdgeSAM 目前仅支持单一类型的连续提示，无法满足多类型混合连续提示的实际需求。

解决方案：我们计划研究如何增强对 EdgeSAM 的提示效果。在已有的框提示基础上，我们将设计“框+X”的多元化提示模式。具体来说，我们将尝试“框+点”、“框+掩膜”、“框+文本”等多种不同的组合方式，并通过对比分析，找出易于实现且能有效提升分割效果的提示组合方式。

（4）边缘端计算资源有限，模型部署困难问题

由于使用了 YOLOv8-OBB 和 EdgeSAM 两个模型，对计算资源的需求比较大。若按照模型串行推理，将占用大量的系统资源，且在推理过程中，存在资源因为无效等待，造成的资源空等数据或任务指令的现象，无法满足系统的实时性要求。

解决方案：利用 TensorRT 深度学习推理引擎，加速模型推理。同时，利用 CUDA 流机制，切分推理任务，细粒度管理推理流程，实现并行推理图片，串行处理结果。

2.4 解决问题的思路

2.4.1 作品的功能和性能需求

基于 YOLOv8-OBB 和 EdgeSAM 的无人机视角下的绝缘子语义分割模型，实现对无人机采集的绝缘子图像进行语义分割，减少对绝缘子后续检测的干扰。

实现快速的绝缘子语义分割。在满足推理精准度的同时，提高模型的运行速度，减少模型的计算成本和资源占用。

2.4.2 数据集

(1) 数据来源和规模

采用大疆无人机在徐州本地拍摄 4K 绝缘子视频共计 275 个，大小 78.7G。每个视频按照 30 帧/张抽取图像，从中筛选出 4016 张含绝缘子目标的图像。利用半自动标注工具制备标签，通过人工修正制成旋转框目标检测数据集。将数据集按照 3:1:1 划分为训练集，测试集和验证集，对模型进行训练。

(2) 数据样例

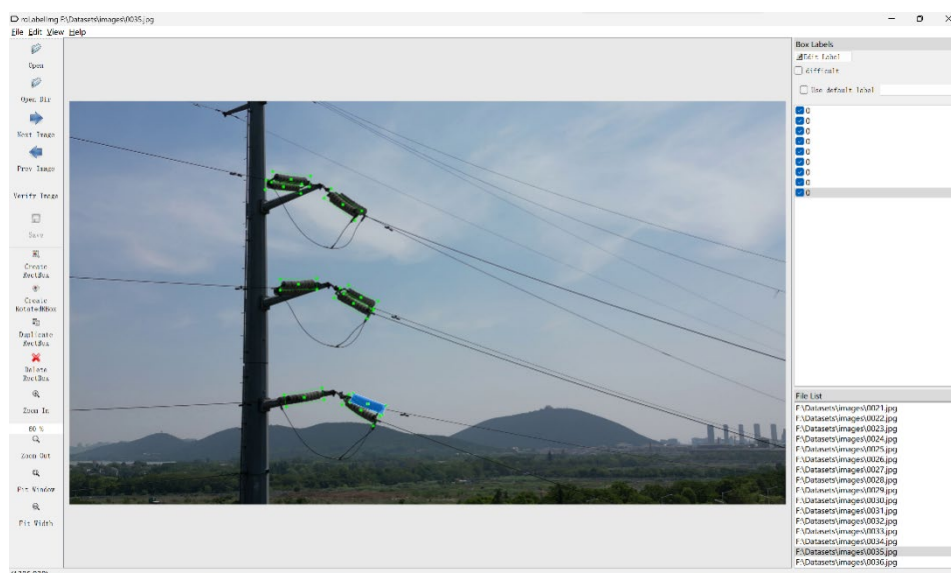


图 4 RolabelImg界面下的数据集图像

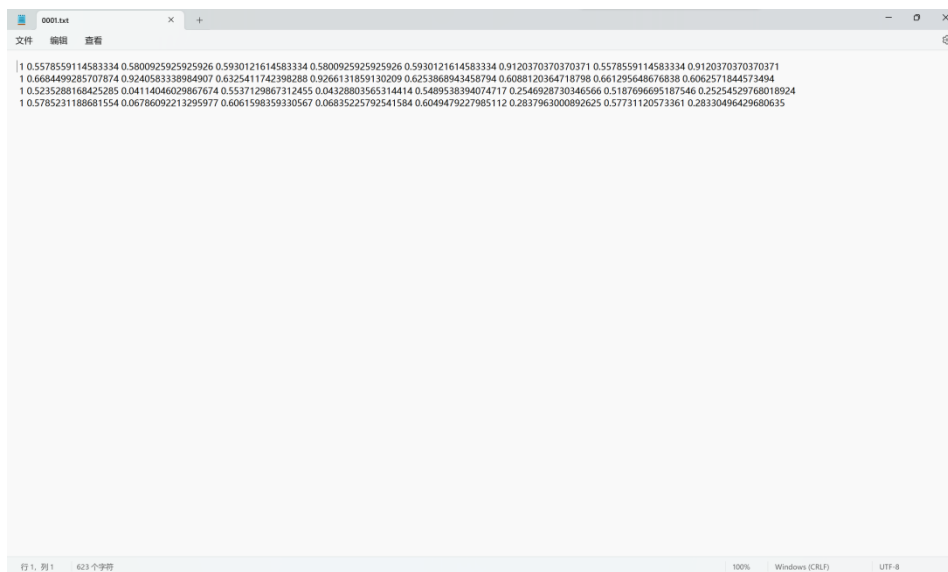


图 5 YOLOv8-OBBox可使用的训练集标签格式

第3章 技术方案

3.1 技术方案

本作品可以分为四个部分执行：数据集采集与制备，模型训练与结合，模型性能分析和模型部署。

我们将利用航拍绝缘子照片制备数据集，训练 YOLOv8-OBBox 模型，实现对绝缘子进行旋转框目标检测。对 YOLOv8-OBBox 的预测结果，我们将通过算法映射为框提示和姿态点提示，用以指导 EdgeSAM 进行语义分割。我们将从准确度和速度两个方面，对比不同语义分割模型，以验证我们方法的有效性和价值。最后我们会开发简单易用的 UI 界面，并将模型部署到 Nvidia Jetson Orin Nano 边缘的开发板上。

本作品涉及的开源代码是 YOLOv8 系列代码（包括 YOLOv8-OBBox 和 YOLOv8-SEG）和 EdgeSAM 代码。为了实现作品功能，我们在代码中做出了改进措施。本作品涉及的硬件主要包括，大疆无人机，Nvidia Jetson Orin Nano 开发者套件以及显示设备。

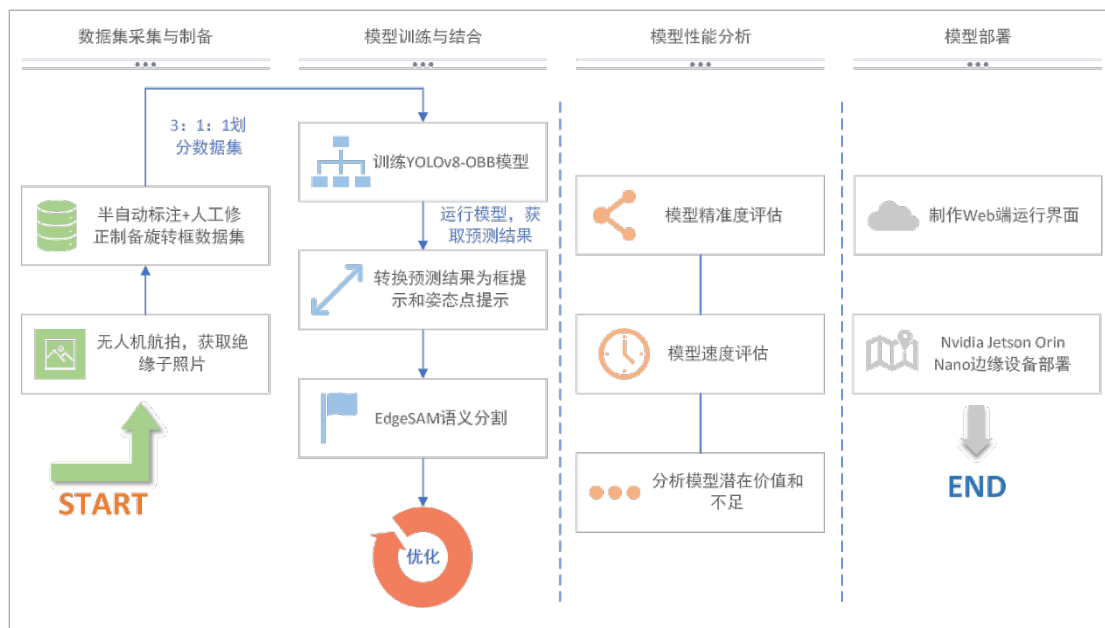


图 6 作品开发流程图

3.2 模块介绍

3.2.1 数据集采集与制备

我们将采用以下四步，完成绝缘子数据集的采集与制备：

- 1) 使用大疆无人机在徐州本地拍摄 4K 分辨率的绝缘子视频
- 2) 从每个视频中按照 30 帧/张抽取图像，从中筛选出 4016 张有绝缘子目标的图片
- 3) 采用半自动标注和人工矫正的方法制作标签
- 4) 将图片和标签按照 3:1:1 的比例划分成训练集，验证集和测试集

3.2.2 模型训练与结合

我们将训练 YOLOv8-OBb 模型，实现对绝缘子进行旋转框目标检测。对 YOLOv8-OBb 的预测结果，我们将通过算法映射为框提示和姿态点提示，用以指导 EdgeSAM 进行语义分割。

本部分的重点在于：

- 1) 将 YOLOv8-OBb 的预测结果处理成“框+姿态点”的形式
- 2) 降低 EdgeSAM 图像预处理成本，避免二次预处理

3) 优化语义分割结果的后处理算法，快速保存语义分割结果

3.2.3 模型部署

我们将制作基于 Gradio 的 UI 界面，力求降低模型的使用难度，打造简单易用的绝缘子语义分割平台。

同时我们将模型部署到 Nvidia Jetson Orin Nano 边缘端设备上。充分利用边缘端设备有限的计算资源，实现绝缘子实时语义分割边缘设备平台。

3.2.4 模型性能分析

我们将会从准确性和运行速度两个方面对模型的性能进行对比。在准确性方面，我们将纵向比较多个 SAM 以及 YOLOv8-SEG 模型的准确性，同时我们将横向对比点框结合提示相较于传统框提示对语义分割准确度的影响。在速度方面，我们将对比多个 SAM 和 YOLOv8-SEG 的运行速度，包括前处理速度，推理速度和后处理速度，得出模型在一张图像上执行全流程平均速度。最后，我们将总结模型潜在的利用价值，并分析模型尚存在的不足，指导后续的模型优化和模型部署。

第4章 系统实现

4.1 数据来源

作品面临的挑战之一是绝缘子公开数据集稀缺，增加了数据采集和标注的工作量。我们使用大疆 mini3 pro 无人机在徐州市拍摄当地绝缘子视频，历经时长 6 天，飞行总距离约 18.23 公里，获得绝缘子 4K 视频 275 个，总大小为 78.7 GB，视频总时长约 2 小时。我们从视频中按照 30 帧/秒的频率提取图像，筛选出 4016 张含有绝缘子目标的图像。通过半自动标注和人工矫正的方式，制备旋转框目标检测的数据集。

4.2 模型训练和改进

我们将使用制作好的数据集，训练 YOLOv8-OBb 模型，实现对绝缘子目标的旋转框检测。根据 YOLOv8-OBb 的预测结果我们提出了基于 θ 角定向边界框的姿态点求算算法。其原理如下图所示：

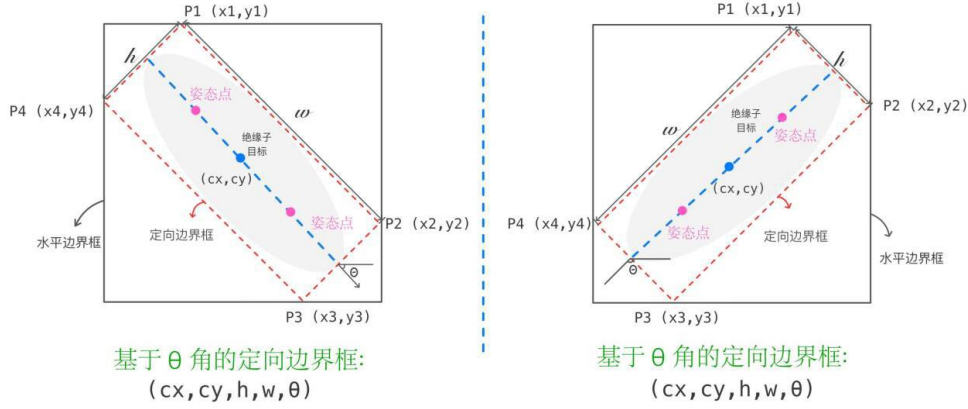


图 7 基于 θ 角的定向边界框示意图

针对图中两种旋转情况，我们给出姿态点的求算公式：

$$\theta = \begin{cases} \theta, & \theta \leq \frac{\pi}{2} \\ \theta - \frac{\pi}{2}, & \theta > \frac{\pi}{2} \end{cases} \quad \begin{aligned} vector &= [\frac{w}{4} \cdot \cos\theta, \frac{w}{4} \cdot \sin\theta] \\ \text{Pose - Point1} &= [cx, cy] + vector \\ \text{Pose - Point2} &= [cx, cy] - vector \end{aligned}$$

图 8 姿态点计算公式

通过改进 EdgeSAM 的提示方法，我们将框提示和姿态点提示进行分组、合并，形成 $(1, n, 4, 256)$ 的提示张量。其含义为：在 1 个批次中，对 n 个绝缘子目标给予 4 个提示，分别代表 1 个框提示的两个定界点和 2 个姿态点提示，每个提示包含 1 个长度为 256 的子提示向量。

另外，为了提高模型效率，我们提出了数据处理的三大策略：

- 1) 输入缓冲区共享：我们通过为 YOLOv8-OBb 模型和 EdgeSAM 模型分配相同的缓冲区，实现模型输入共享。其好处在于节约系统显存开销的同时大大节约了图像预处理时间和传输时间。
- 2) 掩码选择与合并：对于 EdgeSAM 解码器为每个绝缘子目标生成的多张掩码，我们取得分最高者作为每个目标的最终掩码。通过将掩码合并成一张，可以大大节约 GPU 向 CPU 拷贝的数据量，提高效率。

- 3) 图层按权值叠加：我们摒弃在图片上画一层掩码的方式，改为待测图像层和分割掩码层按照 1.0:0.8 的权值叠加。

4.3 模型边缘端设备部署

在边缘端设备部署方面，我们将模型从 Pytorch 框架转换为 TensorRT 的推理引擎，在模型部署过程中，我们继承了模型结合提出的三大策略：输入缓冲区共享、掩码选择与合并以及图层按权值叠加。同时，为了进一步提高效率，节约系统资源，我们再次创造性地提出了基于 CUDA 流的模型推理策略，包括多 CUDA 流串行并行结合推理和多线程优化。

如图 9 所示，在视频流中，我们在主推理线程之外建立子线程，用以降低摄像头采集图像，抽取关键帧的延迟。在主推理线程的处理流中，我们继承了模型结合时采用的三大策略。在推理流中，我们为 YOLOv8-OBB、EdgeSAM encoder、EdgeSAM decoder 分别建立了独立的 CUDA 流和推理上下文。推理过程为 YOLOv8-OBB 流和 EdgeSAM encoder 流并行推理，分别完成绝缘子位置检测任务和待测图片编码任务，同时阻塞主线程，等待 GPU 任务完成并同步。对推理结果进行相应处理后，送入 EdgeSAM decoder 流解码，完成语义分割。由此，形成了“两并一串”的推理流程，大大提高了系统的运行效率。

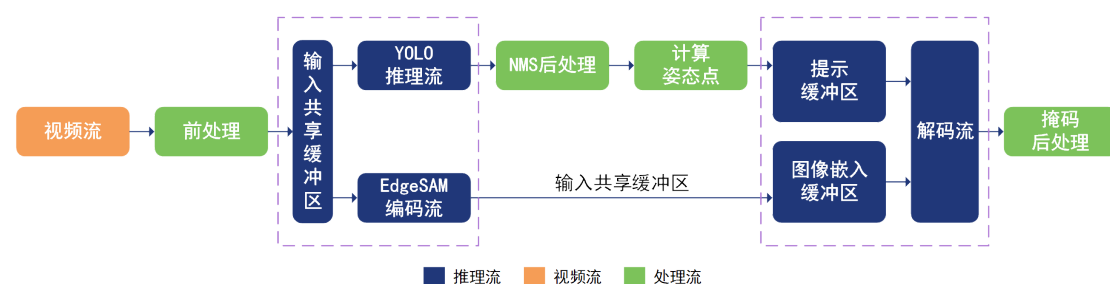


图 9 多CUDA流推理示意图

4.4 界面及成果展示

制作基于 Gradio 的运行界面，实现上传图片或视频即可进行语义分割。在 UI 界面下，用户可以选择 EdgeSAM 模型和提示模式。

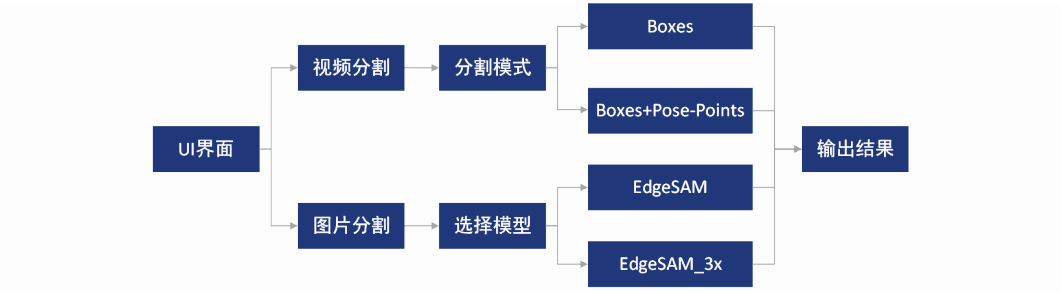


图 10 UI界面功能示意图



图 11 UI界面初始状态示意图

YOLOv8-OBb+EdgeSAM



图 12 UI界面图片语义分割结果预览图

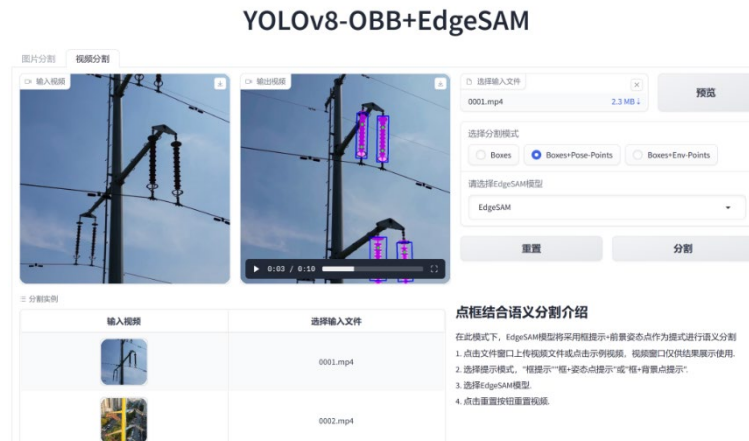


图 13 UI界面视频语义分割结果预览



图 14 基于Nvidia Jetson Orin Nano边缘端设备运行结果

4.5 系统部署方法

详见《用户手册》

4.6 问题及解决方法

(1) 目前开源的数据集虽然丰富，但存在明显的局限性。部分数据集基于3D建模，无法真实还原航拍环境中复杂多变的光线干扰，使得模型训练与实际应用之间存在较大差距。此外，有些数据集中的绝缘子材质和外观与实际应用中的绝缘子存在显著差异，这也影响了模型的准确性和泛化能力。因此，即便有开源数据集可用，也无法直接应用于绝缘子检测模型的训练。人工制备绝缘子语义分割数据集成本消耗巨大。由于绝缘子目标边缘起伏变化明显，边缘轮廓细密复

杂。这将耗费大量的人力和时间，制备有效的数据集。解决方法：利用目标检测算法，对绝缘子进行目标检测。通过转换算法将目标检测结果转换为 EdgeSAM 所需的位置提示。将语义分割数据集制作转化为目标检测数据集制作，大大节省了数据集制作的人力和时间成本。经过我们小组的实际测算，在组员完全熟悉数据集制备方法的情况下，要制作 200 张能够训练绝缘子语义分割模型的的数据集，需要 30 个小时。而人工制备 200 目标检测数据集，只需要 5 小时，且目前存在许多半自动标注工具辅助，大大节约了人工制备目标检测数据集的成本。

(2) 只提供绝缘子位置信息对 EdgeSAM 进行提示，得到的掩码不全甚至不产生掩码。解决方法：用 YOLOv8-OBb 目标检测算法，对绝缘子目标进行旋转框目标检测。我们将绝缘子的旋转信息映射为姿态点，与绝缘子位置信息组合成新提示，指导 EdgeSAM 语义分割。

(3) YOLOv8n-OBb 的结果是基于 θ 角的定向边界框，与 EdgeSAM 的提示模式不完全契合。解决方法：将绝缘子的姿态抽象化，简化为其中心轴线上的两点表示，并引入公式来计算这两个姿态点的位置。

(4) EdgeSAM 推理结果难处理。解决办法：尽量在 GPU 中将图像的处理工作等效成张量并行运算，然后将处理结果，写入文件。

(5) 边缘设备计算资源有限。解决办法：分析模型推理过程中的数据流向，减少数据的重复处理和无效拷贝。充分利用 TensorRT 的批处理机制，多 CUDA 流推理机制，优化模型推理流程。

第5章 测试分析

5.1 运行结果分析

以下三组运行结果图片直观地展示了我们所提出的点框结合提示方法的有效性。其中，左侧为点框结合提示结果，右侧为传统框提示结果。可以看出点框结合提示对语义分割结果完整性的提升。



图 15 点框结合提示结果与框提示结果对比图

5.2 准确率评估

通过生成掩膜图像,我们完成了不同语义分割模型结果和人工抠取结果的交并比实验,测试指标表包括平均交并比 $mIoU$, 平均召回率 $mRecall$ 以及平均像素准确率 mPA 。以下为不同模型对 210 张绝缘子照片的进行语义分割形成的掩膜二值图像。

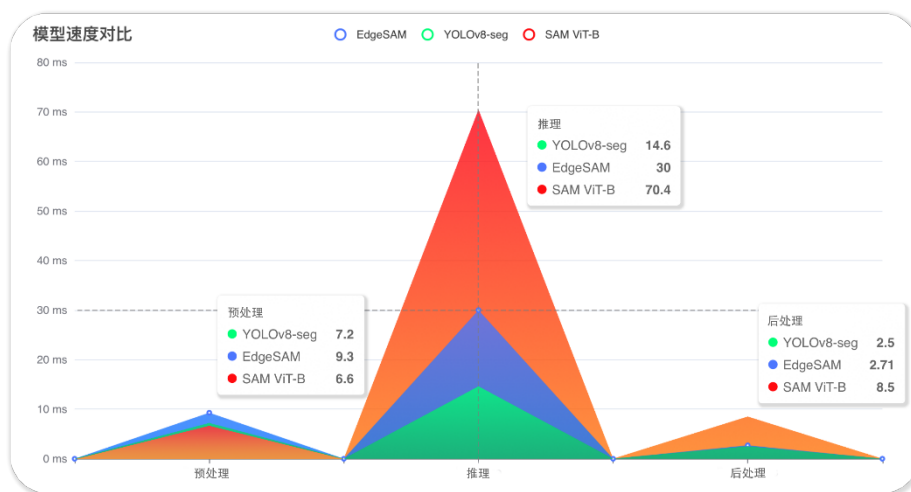


图 17 不同模型速度对比图

第6章 作品总结

6.1 作品特色与创新点

(1) 使用旋转框检测。绝缘子目标多为长条状，长度可达数米且具有一定的旋转角度。基于传统水平边界框的目标检测方法生成的检测框十分庞大，且其中包含大量的背景像素，对 EdgeSAM 语义分割产生巨大干扰。另外，基于传统 HBB 的无法确定绝缘子在边界框中的旋转信息和姿态信息。在后处理过程中，基于 NMS 的处理方法会导致绝缘子边角信息的缺失。旋转框目标检测算法，可以充分挖掘待测目标的姿态信息，利用定向边界框精确抓取待测目标，提高了检测的准确性和鲁棒性，为绝缘子目标姿态的确定提供了可能。

(2) 确定绝缘子的姿态。YOLOv8-OBb 的推理结果是基于 θ 角的定向边界框，为了适应 EdgeSAM 的提示模式，将绝缘子的姿态抽象成其中心轴线上的两点，对于不同的旋转情况，我们设计出对应的计算公式求姿态点。为了节约计算的时间成本，将张量引入到 GPU，对姿态点进行并行运算。

(3) 点框结合提示。点框结合提示是指将待测目标最小边界框以及姿态前景点作为 EdgeSAM 的提示进行语义分割。改进后的 EdgeSAM 允许多个点框坐标组合批量对图片进行连续分割，获取分割掩码。

(4) 结合 YOLOv8-OBb 和 EdgeSAM。输入缓冲区共享：通过为模型分配

相同的缓冲区，实现模型输入共享，节约图像预处理时间和传输时间。掩码选择与合并：我们选取绝缘子目标得分最高的掩码作为最终掩码，并将掩码合并成一张，可以大大节约 GPU 向 CPU 拷贝的数据量，提高效率。图层按权值叠加：我们待测图像层和分割掩码层按照 1.0:0.8 的权值叠加。

（5）TensorRT 加速与多 CUDA 流推理：我们创造性地提出了基于 CUDA 流的模型推理策略，包括多 CUDA 流串行并行结合推理和多线程优化。

6.2 应用推广

随着无人机技术的不断创新和应用场景的不断拓宽，无人机视角下的绝缘子语义分割技术将在电力行业展现出更广阔的前景。我们深信，在未来，该技术将实现更大的突破。我们将继续研发高效的算法和检测设备，提升检测精度和速度，以满足电力行业对高效、精准检测的不断追求。我们期待与更多的电力企业、科研机构等合作伙伴开展深度合作，共同推动无人机绝缘子故障检测应用的发展和创新。我们设想，无人机视角下的语义分割技术将在电力线路巡检，应急抢修，新建线路验收等场景下大放异彩。

6.3 作品展望

随着智能电网和无人机巡检等应用场景的需求，绝缘子语义分割作品的实时性和效率将成为重要关注点。在接下来的过程中，我们将争取在以下几个方面对作品进行改进和完善：

优化 EdgeSAM 的图像编码器的处理时序，实现图像预处理过程中的并行编码，降低时间成本，实现实时语义分割。

增加故障检测模块，对绝缘子进行语义分割的同时将绝缘子破损，击穿等问题实时预测报警。

优化边缘端部署，降低运算成本与资源占用，实现边缘端检测，减少设备体积，实现无人机搭载。

优化模型结构，提高语义分割的速度和实时性能。使其能够应对各种复杂环境和条件下的绝缘子检测任务。

以与其他领域进行跨界融合，如计算机视觉、遥感技术、电力工程等。通过引入其他领域的知识和技术，为绝缘子的缺陷检测带来更多的创新点和可能性。

本作品 YOLOv8-OBB+EdgeSAM 的绝缘子语义分割技术, 为绝缘子检测提供高效精准的技术, 为推动电力行业的智能化发展和安全生产提供重要支持。

积小流以汇江河, 集细沙以聚巨塔, 不断学习与改进, 逐步推动技术的进步与完善。我们期待在未来的日子里, 能够将更多的精力投入到电力巡检的实际应用中, 使绝缘子语义分割技术更加精准、高效, 为电力行业的安全与发展贡献更多力量。

参考文献

- [1] 徐俊良,曹建,杨凯斌.基于局部均值的绝缘子图像分割算法设计[J].计算机工程, 2016.DOI:10.3969/j.issn.1000-3428.2016.09.046.
- [2] 于兰英,姚波,吴文海,等.一种基于多特征的绝缘子识别方法[J].电瓷避雷器, 2016(3):5.DOI:10.16188/j.isa.1003-8337.2016.03.015
- [3] 杜晓亮,姜香菊. 基于改进 Faster R-CNN 的铁路接触网绝缘子识别[J]. 兰州交通大学学报,2021,40(6):62-68. DOI:10.3969/j.issn.1001-4373.2021.06.010.
- [4] 唐小煜,熊浩良,黄锐珊,等. 基于改进的 U-Net 和 YOLOv5 的绝缘子掩模获取与缺陷检测 [J]. 数据采集与处理 ,2021,36(5):1041-1049. DOI:10.16337/j.1004-9037.2021.05.019.
- [5] Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics YOLO (Version 8.0.0) [Computer software]. <https://github.com/ultralytics/ultralytics>
- [6] Dillon Reis, Jordan Kupec, et al. Real-Time Flying Object Detection with YOLOv8. DOI:<https://doi.org/10.48550/arXiv.2305.09972>
- [7] 叶力鸣,陈蔚文. 一种结合语义分割和目标检测的级联式绝缘子缺陷检测方法 [J]. 计算机与现代化 ,2023(6):82-88. DOI:10.3969/j.issn.1006-2475.2023.06.014.
- [8] Chong Zhou, Xiangtai Li, et al. EdgeSAM: Prompt-In-the-Loop Distillation for On-Device Deployment of SAM. DOI: <https://doi.org/10.48550/arXiv.2312.06660>
- [9] 倪祺,崔宇铭,王方鑫,等. 基于改进 SECNet 的绝缘子图像分割研究[C]//中国自动化学会 .2023 中国自动化大会论文集 .[出版者不详],2023:6.DOI:10.26914/c.cnkihy.2023.092624.
- [10] Xu Zhao, Wenchao Ding, Yongqi An, Yinglong Du, Tao Yu, Min Li, Ming Tang,

- and Jinqiao Wang. Fast segment any thing. arXiv preprint, 2023. 2, 9
- [11]HanCai, Junyan Li, MuyanHu, ChuangGan, andSongHan. Efficientvit: Lightweight multi-scale attention for on-device semantic segmentation. In ICCV, 2023. 3, 7
- [12]Mazurowski M A, Dong H, Gu H, et al. Segment anything model for medical image analysis: an experimental study[J]. Medical Image Analysis, 2023, 89: 102918.
- [13]Kirillov A, Mintun E, Ravi N, et al. Segment anything[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 4015-4026.
- [14]Zhang C, Liu L, Cui Y, et al. A comprehensive survey on segment anything model for vision and beyond[J]. arxiv preprint arxiv:2305.08196, 2023.
- [15]韩谷静,何敏,雷宇航,等.基于改进 U-Net 的输电线路绝缘子图像分割方法研究[J].智慧电力,2022,50(03):93-99.