

基于 U-Net 网络及 YOLOv4 的缺陷绝缘子智能检测

摘 要

高空架空线路作为电力输送网络建设中的基础设施,通常情况下都架设在环境极为复杂的自然场景中。而输电线路上的绝缘子,由于长时间暴露在自然环境下,容易产生自爆缺失故障,从而引起异常的能量供应,进而容易损害整个输电线路的使用安全和运行寿命,产生不可估量的经济损失。因此,为保障电网的安全,绝缘子的故障检测已成为电力系统维护的首要任务。近年来,随着无人机和智能机器人技术的迅速发展,可利用航拍对架空线路进行巡检,对其中的绝缘子进行航拍,但航拍图片数目多、图片尺寸大、存在人工标注时间长、标注效率低和易出错的弊端。为了能够良好的解决这个问题,通过将对这一棘手问题进行前期分析,本文提出了基于 U-Net 网络及 YOLOv4 的缺陷绝缘子智能检测级联模型。

针对任务一绝缘子串珠分割。首先对巡检所得 40 张图像进行压缩、旋转、剪裁等处理获得 300 张图像创建一个新的绝缘子缺失故障数据集,然后建立语义分割 U-Net 网络为图像中的每个像素指定类别标记,最后随机选取数据集的 85% 进行模型训练,15% 进行模型验证,并采用 Dice 系数评价模型从复杂的背景中分割出绝缘子串珠性能,最终网络训练得到绝缘子串珠分割准确率为 87.04%。

针对任务二绝缘子自爆识别和定位。首先对 CPLID 数据集进行初步处理,人工筛选出 60 张绝缘子缺失部位清晰的图像并与原始数据集中的 40 张掩膜图合并,创建一个新的绝缘子缺失故障数据集;通过进行压缩、旋转、剪裁、随机复制等处理手段,获得共 800 张绝缘子缺失掩膜图像。其次,对这数据集应用 labeling 标注工具进行标注。最后,基于 Keras 框架下构建 YOLOv4 目标识别网络,运用 Kmeans 聚类方法确定网络预选框的大小,并随机选取缺陷绝缘子数据集的 90% 用作训练集,共 720 张;10% 用作验证集,共 80 张。绝缘子自爆区域评价采用 IOU 系数计算网络识别效果,该模型 IOU 识别率达到了 79.4%。

关键词: 绝缘子 U-Net 图像分割 YOLOv4 目标识别

ABSTRACT

As an infrastructure in the construction of power transmission networks, high-altitude overhead lines are usually installed in natural scenes with extremely complex environments. Insulators on transmission lines are prone to self-explosive failure due to long-term exposure to the natural environment, which can cause abnormal energy supply, which can easily damage the safety and operating life of the entire transmission line, and cause inestimable economic losses. Therefore, to ensure the safety of the power grid, fault detection of insulators has become the primary task of power system maintenance. In recent years, with the rapid development of drone and intelligent robot technology, aerial photography can be used to inspect overhead lines and aerial photography of the insulators. However, there are many aerial pictures, large picture sizes, long manual annotation time, annotation The disadvantages of low efficiency and error-prone. In order to solve this problem well, through a preliminary analysis of this thorny problem, this paper proposes a cascade model for intelligent detection of defective insulators based on U-Net network and YOLOv4.

Bead segmentation of insulators for task one. First, the 40 images obtained by inspection are compressed, rotated, and cropped to obtain 300 images to create a new insulator missing fault data set. Then, a semantic segmentation U-Net network is established to assign category labels to each pixel in the image. Finally Randomly select 85% of the data set for model training and 15% for model verification, and use Dice coefficient to evaluate the model to segment the insulator bead performance from the complex background. Finally, the network training obtained the insulator bead segmentation accuracy rate of 87.04%.

Aiming at task two, the self-detonation identification and positioning of insulators. First, the CPLID data set is initially processed, and 60 clear images of the missing parts of the insulator are manually selected and merged with the 40 mask images in the original data set to create a new data set of the missing faults of the insulator; by compression, rotation, and cropping In other processing methods, a total of 400

insulator mask images were obtained. Secondly, labeling this data set using labeling tool. Finally, based on the Keras framework, YOLOv4 target recognition network was constructed, Kmeans clustering method was used to determine the size of the network preselection frame, and 90% of the defective insulator data set was randomly selected as the training set, a total of 360; 10% was used as the verification set. 40 photos in total. The evaluation of insulator self-detonation area uses the IOU coefficient to calculate the network identification effect. The IOU recognition rate of this model reached 79.4%.

Keywords: insulator U-Net image segmentation YOLOv4 target recognition

摘 要	1
ABSTRACT.....	2
1 绪论	5
1.1 背景、目的及意义	5
2 数据预处理.....	6
2.1 绝缘子自爆缺陷原图处理.....	6
2.1.1 椒盐噪声	6
3 基于 U-Net 的绝缘子串珠分割算法	8
3.1 语义分割 U-Net 网络.....	8
3.1.1 上采样和下采样.....	8
3.1.2 He 参数初始化	10
3.1.3 数据增强	11
3.2 Unet 网络模型训练.....	12
3.2.1 拆分训练集和验证集.....	13
3.2.2 测试时增强（TTA）	13
3.3 结果分析.....	15
3.3.1 网络训练损失及评价指标	15
3.3.2 测试结果对比.....	17
4 基于 YOLOV4 的绝缘子缺陷检测算法	19
4.1 网络技术背景	19
4.2 网络模型搭建	19
4.2.1 制作训练样本.....	20
4.2.2 YOLOV4 金字塔多级检测	21
4.3 YOLOV4 网络训练.....	21
4.3.1 迭代策略	21
4.3.2 先验框选择	22
4.3.3 训练结果分析.....	22
5 总结与展望.....	26
参考文献	27

1 绪论

1.1 背景、目的及意义

在电力系统之中，输送电路绝缘子是一种特殊的绝缘控件，它能够在架空输电线路中起到重要作用。早年间绝缘子多用于电线杆，慢慢发展于高型高压电线连接塔的一端挂了很多盘状的绝缘体，它是为了增加爬电距离的，通常由玻璃或陶瓷制成。绝缘子不应该由于环境和电负荷条件发生变化导致的各种机电应力而失效，否则绝缘子就不会产生重大的作用，就会损害整条线路的使用和运行寿命，在输送电路中，一旦绝缘子出现故障则会引起供电的中断和电力系统的瘫痪，动辄将造成巨大的经济损失。

为了避免不必要的财产损失，保障电力系统的稳定运行，及时准确的发现绝缘子故障并采取补救措施是维护系统正常运行的首要任务，所以定期对含有绝缘子部件的电路进行巡检是电路检测的必要工作。然而基于人工作业的绝缘子自爆巡检含有诸多复杂性，包括了劳动强度大，工作效率低，野外巡检环境恶劣，后期工作中对于自爆绝缘子数据人工录入数据量大，易产生人为失误，成本高昂。近年来随着无人机和智能机器人技术的高速发展，逐渐由机器代替人工巡检，通过拍摄大量电力设备和野外输电线路的图片交送于数据中心处理，具有作业灵活、成本低、数据获取快等优点。但同时，无人机所采集的海量复杂绝缘子图像数据也有着以下几点缺陷：(1)拍摄绝缘子种类繁多，并且有的类与类之间的差别较大；(2)不同绝缘子照片之间的尺度差别很大；(3)目标绝缘子的周围还存在强遮挡、背景干扰等问题；(4)采集的图像还存在光照变化、分辨率不统一等问题。若采取经由工作人员主观判断的方式进行绝缘子缺陷检测，极易造成误判和漏判，从而影响故障的排除，留下安全隐患。因此如何通过图像处理结合机器学习（深度学习）准确且快速的检测出绝缘子故障，是目前提高电力系统安全检测效率的重要任务。

2 数据预处理

2.1 绝缘子自爆缺陷原图处理

数据集的好坏在很大程度上决定了网络训练结果的好坏，因此要对图像进行预处理，制作符合所选网络的数据集。基于原始数据集的数量有限会导致网络训练结果缺乏泛化性，图像分辨率过高影响网络运行速度和处理效率。因此在初始阶段要对数据集进行预处理，通过对数据集图像进行旋转，裁剪，压缩以及加椒盐噪声等处理扩充和优化数据集。

2.1.1 椒盐噪声

椒盐噪声也称为脉冲噪声，是图像中经常见到的一种噪声，它是根据图像的信噪比在图像上随机出现的白点或者黑点，可能是亮的区域有黑色像素或是在暗的区域有白色像素。椒盐噪声概率密度函数为：

$$P(z) = \begin{cases} P_a, & z = a \\ P_b, & z = b \\ 1 - P_a - P_b, & \text{其他} \end{cases} \quad (1)$$

给图像添加信噪比不同的椒盐噪声，经过对大量样本的分析发现，椒盐噪声会干扰模糊图像，添加椒盐噪声后的图像背景更加复杂，网络训练难度增加。如下为原图与添加椒盐噪声的图像对比：



图 2.1 绝缘子自爆缺陷原图

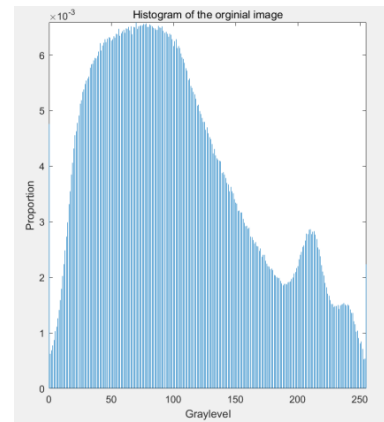


图 2.2 图 2.1 的像素分布图



图 2.3 信噪比为 0.005

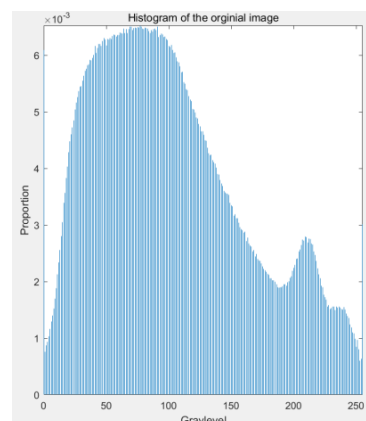


图 2.4 图 2.2 的像素分布图

从上面原图和像素分布图可以看出来，添加噪声之后和添加噪声之前没有明显的变化。从侧面猜测，添加噪声后的样本跟原始样本基本一致，对网络训练没有干扰性。

2.2 掩膜图像处理

掩膜图像主要用于绝缘子自爆识别和定位^[1]，绝缘子的种类繁多及大小不一会使有缺陷的绝缘子串珠存在难以识别和定位的问题。本文首先对 CPLID 数据集进行初步处理，人工筛选出 60 张绝缘子缺失部位清晰的图像并与原始数据集集中的 40 张掩膜图合并，创建一个新的绝缘子缺失故障数据集；通过进行压缩、旋转、剪裁、复制等处理手段，获得共 800 张绝缘子缺失掩膜图像。其次，对这数据集应用 labelimg 标注工具进行标注。

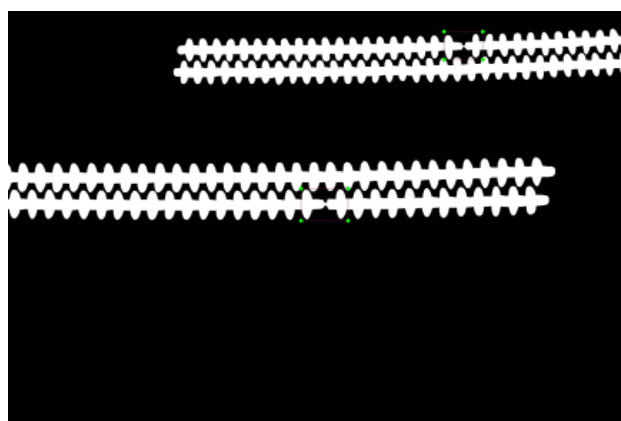


图 2.5 使用 labelimg 标注的掩膜图像

3 基于 U-Net 的绝缘子串珠分割算法

3.1 语义分割 U-Net 网络

U-Net 是一种基于深度学习的图像语义分割方法^[2]，它能够以较少的数据准确的预测绝缘子串珠所在位置。语义分割要从像素层次来识别图像，为图像中的每个像素指定类被标记，与图像分类或目标检测相比，语义分割使我们对图像有更加细致的了解。

本文使用 U-Net 处理绝缘子串珠图像分割问题，具体流程如下图所示。

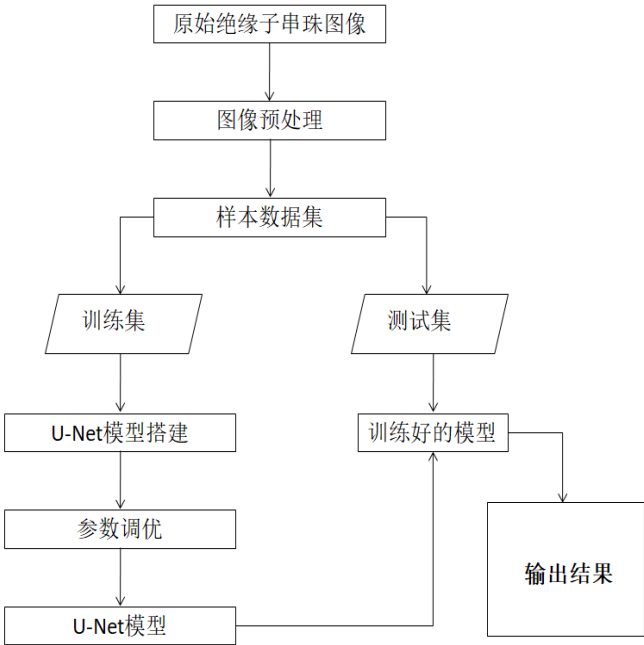


图 3.1 U-Net 框架算法流程图

3.1.1 上采样和下采样

本文是基于 U-Net 对绝缘子串珠图像进行分割，构建的网络如图 3.2 所示。该网络包含两个路径，一个是左侧用于特征提取的下采样路径，一个是右侧用于精准定位像素分类的上采样路径，这两个路径结合起来呈现一个英文字母 U，因此被命名为 U-Net。

下采样由许多收缩块组成，因此下采样也称为收缩，在每个收缩块之后，核或特征映射的数量会加倍，这样体系结构就可以有效地学习复杂的结构。每个收

缩块可以接受一个输入，应用两个 3×3 的卷积层^[3]，每个卷积层后面跟着一个修正线性单元 (RELU)，然后是一个 2×2 的最大池化，最后是为防止卷积过拟合的 Dropout。其中最大池化更像是做了特征选择，选出了分类辨识度更好的特征。

U-Net 架构的核心在于上采样即扩张部分，可以将图片恢复到全分辨率的分割图。与收缩层类似，它也包含几个扩张块。首先是 2×2 的反卷积上采样，然后是与下采样进行跳跃连接，最后进行反卷积。每个块将输入传递到两个 3×3 CNN 层和每个卷积层后面的一个 RELU。由于每一个卷积中的边界像素的缺失，故需要进行裁剪。在最后一层，一个 1×1 的卷积被用来将每一个特征向量映射到需要的类的数量。上采样使用 2×2 上卷积 (up-convolution)，输入图像的每个像素，产生 4 个输出像素。在每次上卷积之后，图像的深度是减半 (即通道数减半)，但分辨率增大 4 倍，意味着采用低分辨率图像并将其转换为更高分辨率的图像。因此可以堆叠多个上卷积以创建更高深度的输出图像。

由图 3.2 可以看出，U-Net 是一个全卷积神经网络，输入和输出都是图像，没有全连接层。较浅的高分辨率层用来解决像素定位的问题，较深的层用来解决像素分类的问题，并将特征图从收缩路径的每个级别传递到扩展路径中的相应级别。

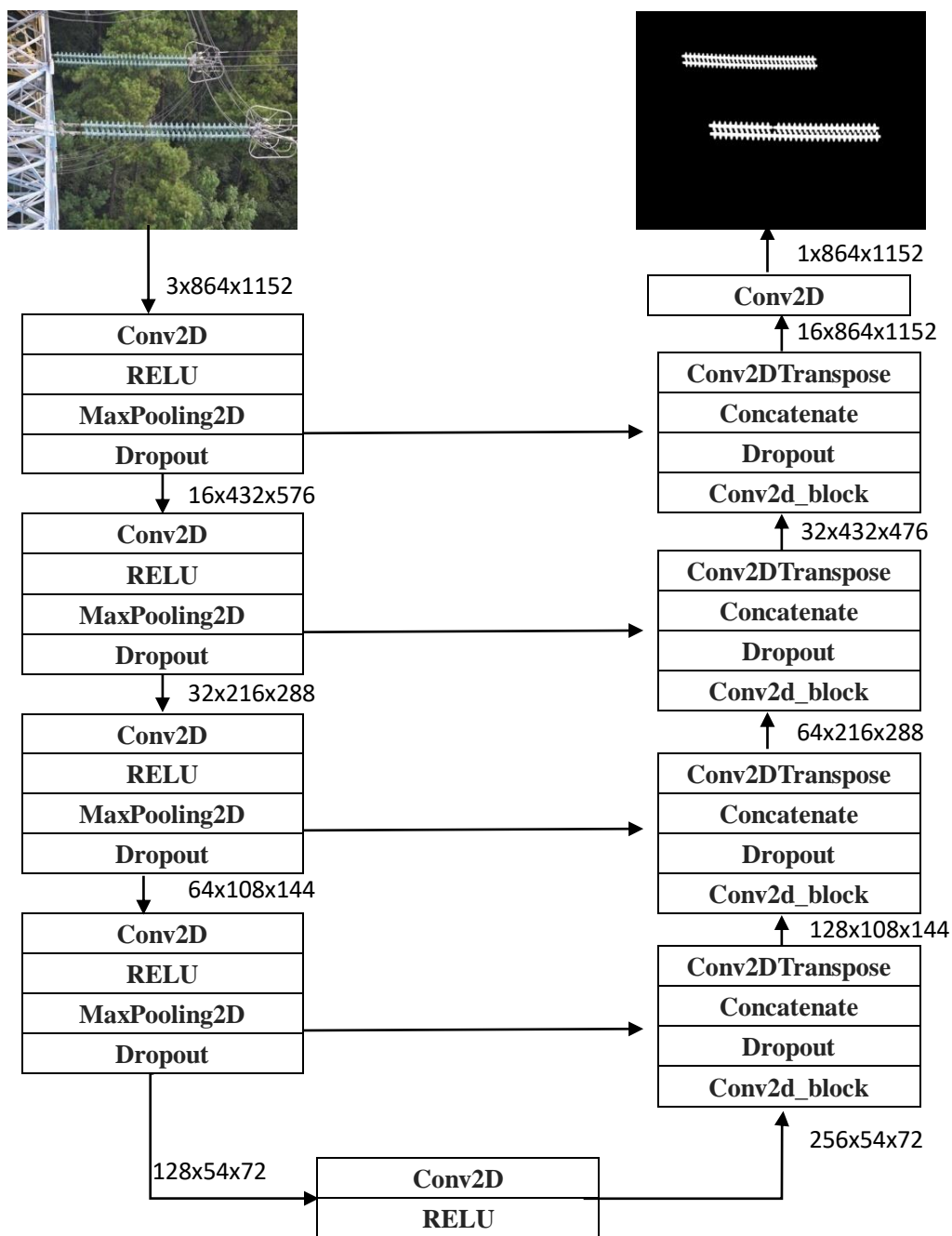


图 3.2 U-Net 网络框架

3.1.2 He 参数初始化

在网络模型建立的初期，实验在网络训练前对其各个节点进行参数的初始化，用以提高网络的运行速度和训练结果的准确率。本实验采用了“正态化的kaiming 初始化”方法中的 he_normal 方法^[4]。

为确保试验参数初始化的正确运行，实验中参数调整将满足 Glorot 条件，

即：

(1) 网络各层的激活值 h 的方差要保持一致，计算公式如下：

$$\forall(i, j): \text{Var}(h^i) = \text{Var}(h^j) \quad (2)$$

(2) 网络中各个层对状态 z 的梯度的方差要保持一致，计算公式如下：

$$\forall(i, j): \text{Var}\left(\frac{\partial \text{cost}}{\partial z^i}\right) = \text{Var}\left(\frac{\partial \text{cost}}{\partial z^j}\right) \quad (3)$$

以上公式计算中， h 代表激活值，它是经过网络激活函数运算后的网络层输出值； z 代表状态值，它位于网络层内部且担任了激活函数的输入。具体神经网络单层的运行图如下所示：

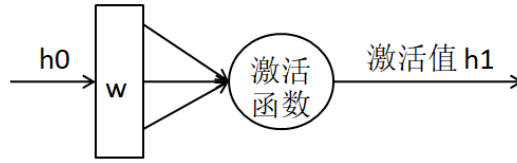


图 3.3 神经网络单层运行图

本实验采用的 `he_normal` 初始化方式又称为 **He** 正态分布初始化器，这种方式使得数据在网络的正向传播时，状态值的方差保持不变；反向传播时，关于激活值的梯度的方差保持不变。同时考虑了不同激活函数对输出数据分布的影响，使得输入和输出数据方差一致，加快了网络运行中参数的收敛速度。

3.1.3 数据增强

因为可用于训练的数据非常少，所以要对可用的训练图像应用弹性变形达到数据增强^[5]。数据增强可以增加训练的数据量，提高模型的泛化能力，增加噪声数据，提升模型的鲁棒性。本文是对数据预处理后再进行的数据增强，利用 `keras` 中 `ImageDataGenerator` 生成器的 `flow` 来输出增强后的图像。

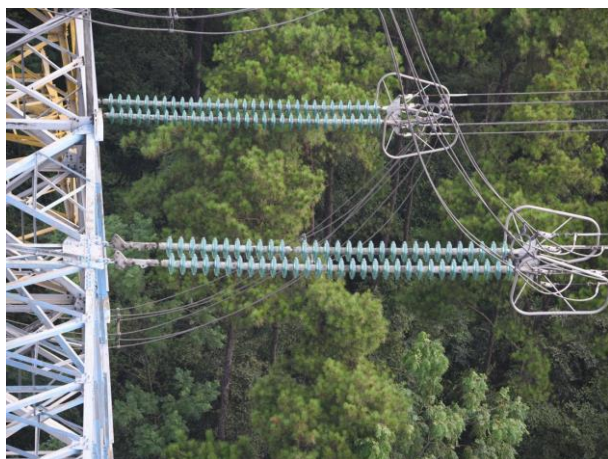


图 3.4 绝缘子原图

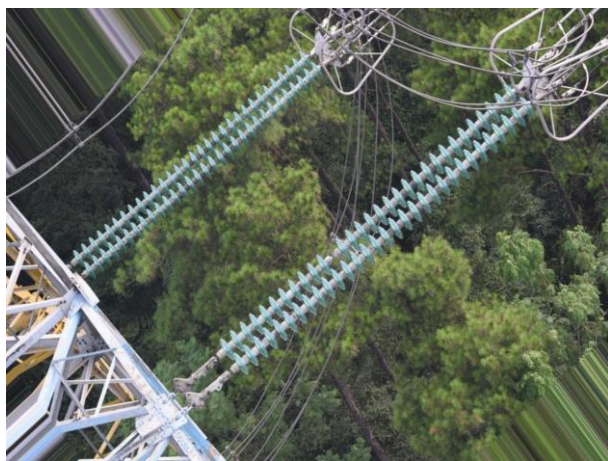


图 3.5 增强后图像

3.2 Unet 网络模型训练

该 U-Net 所需的训练样本较少，速度快,对比滑动卷积网络效果更好。使用绝缘子串珠图像及其对应的分割图谱来训练网络,进行样本训练 100 次。该网路能量函数通过最终特征图上的逐像素 softmax 值与交叉熵损失函数的组合来计算。

soft-max 函数的定义为:

$$p_k = \exp(a_k(x)) / (\sum_{k'=1}^K \exp(a_{k'}(x))) \quad (4)$$

其中 $a_k(x)$ 表示像素位置 x 在特征通道 k 上的激活值； k 表示的是类别的数量； p_k 是近似的最大函数。通过 softmax 得到 p_k 值之后再进行加权相乘求交叉

熵，加权可以让一些像素更重要的训练。

交叉熵损失函数定义为：

$$E = \sum_{x \in \Omega} \omega(x) \log(p_{l(x)}(x)) \quad (5)$$

其中： $\Omega \rightarrow \{1, \dots, k\}$ 是每个像素的真实标签; $\omega: \Omega \rightarrow \mathbb{R}$ 。

3.2.1 拆分训练集和验证集

对于一个监督模型来说，通过训练数据来调整模型参数，在调参的过程中要进行学习和验证，而用于验证的数据就是验证集，而真正的预测是在测试上面去进行。本文将 300 张图片的 85% 共 255 张作为训练集，15% 共 45 张作为验证集。

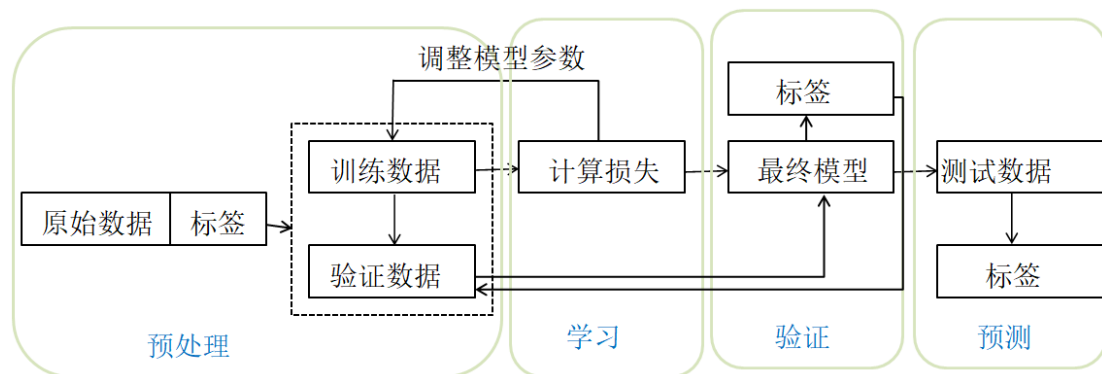


图 3.6 监督学习过程

3.2.2 测试时增强（TTA）

测试时增强技术（Test Time Augmentation）又称 TTA，这种技术会在图片测试时将原始图片的不同区域进行裁剪，尺寸缩放操作等^[6]，并将处理后的图片输入进模型，在随后的测试工作中，网络将对多个版本的处理图片进行计算得到平均输出，作为图片的最终输出分数。在预测工作中，对数据进行测试时增强技术（TTA）的原因是：若在测试阶段仅仅对图像做一种增强,这一种增强所采用的变换可能会使图像关键信息(即特征)丢失,比如在对图像做剪切变换时,可能会把关键特征丢掉，最终导致样本的预测错误。因此，一般情况下此类技术可以进一步提升结果的稳定性和精准度。

然而在本实验中，绝缘子图像数据集的预测在经过 TTA 技术后增强效果并不

明显。如下图为对训练集的预测结果。



图 3.7 绝缘子原始掩膜图



图 3.8 网络模型预测



图 3.9 TTA 模型预测

不难发现，经过 TTA 处理的绝缘子预测图像上出现了噪声点，而且在放大图像以后还可以发现，绝缘子串边缘的轮廓细节并不光滑，一部分噪声使得轮廓看起来凹凸不平。实验初步分析产生这种现象的原因是根源于数据集的局限性，后

续网络预测和实验第二步所需的图像将主要采用不经 TTA 处理的预测图像。

3.3 结果分析

本章是实验处理工作的第一阶段，在这一阶段里，实验旨在通过深度学习网络来提取和学习图像的特征形成映射以便于减少无用信息干扰，提高目标检测的效率和准确率。U-net 网络针对实验第一阶段中图像分类问题有十分显著的成效，它将图像信息及特征信息转化成为向量数据信息，再由这些向量数据信息重建图像，利用 U-net 网络的优势可以有效的保持重建图像的完整性，降低失真率。

在数据集的预处理阶段，由于实验提供的数据集有限，为了防止模型产生欠拟合或过拟合的现象，首先对数据集采取了 `resize` 和加噪声等方式，将 40 张数据集扩充至 300 张。

3.3.1 网络训练损失及评价指标

在网络正式运行之前，为方便网络训练时数据集的读取，将统一输入的数据集尺寸调整为 1152×864 ，将 300 张数据集拆分为 255 张训练集和 45 张测试集，通过增强数据处理之后放入网络运行训练 100 次，U-net 网络训练的 loss 曲线如图 3.10 所示。

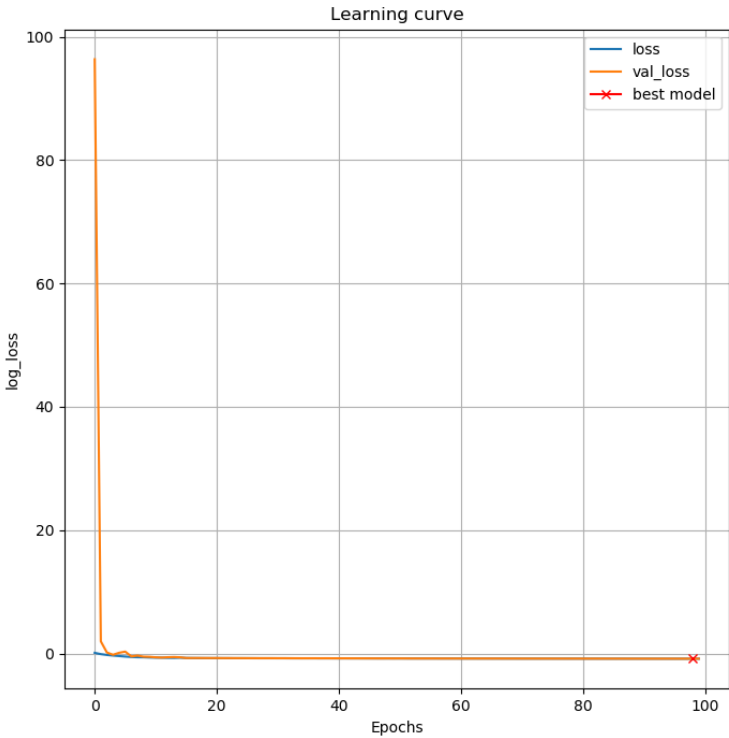


图 3.10 U-net 网络训练的 loss 曲线

由网络的学习曲线可以看出，在 epoch 增加的过程中，训练集和测试集的正确率稳步上升，随着 epoch 增加趋近于 100，正确率的增长速率逐渐减慢并趋于稳定。

绝缘子串珠分割采用 Dice 系数进行评价，他是一种集合相似度量度的函数，通常用于计算两个样本呢的相似度：

$$\text{dice}(A, B) = \frac{2|A \cap B|}{|A| + |B|} \quad (6)$$

其中 A 为 Ground Truth 区域即专业人士标注的区域, B 为算法 U-Net 算法分割所得到的区域。Dice 系数取值范围是[0, 1]，取值越接近 1 则越表明预测的结果与专业人士标注的结果相符合。本文利用 U-Net 训练网络 dice 系数曲线如图 3.11 所示。

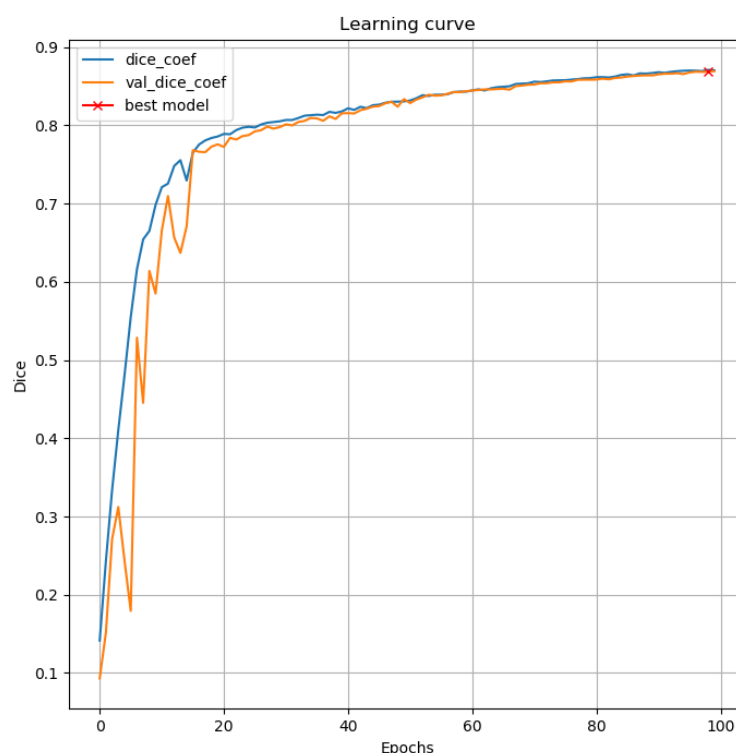


图 3.11 U-net 网络训练的 Dice 曲线

3.3.2 测试结果对比

网络训练结束后，加载训练好的网络模型，随意选取训练集和验证集里的一张绝缘子原图对他们进行性能评估，验证集性能评估准确率达到 85.9%。训练集和验证集数据再经 TTA 处理显著增强绝缘子串辨识度，方便在实验第二阶段进行目标检测。图 3.13，图 3.14 为测试结果对比图。

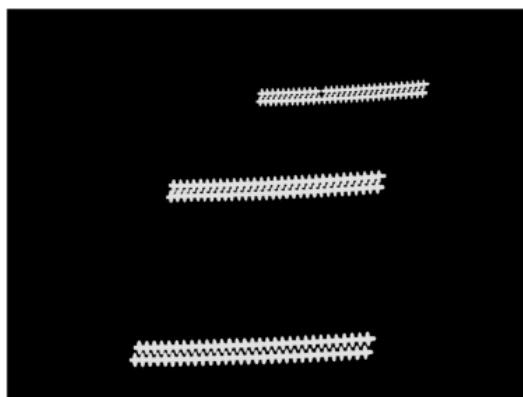


图 3.12 原始掩膜图像

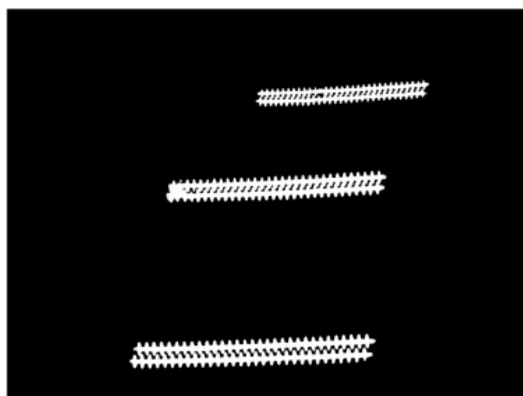


图 3.13 预测掩膜图像

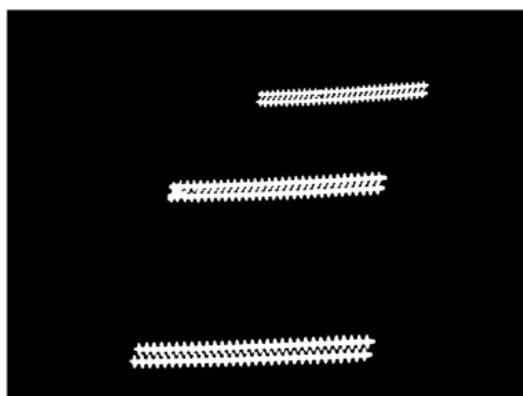


图 3.14 TTA 预测掩膜图像

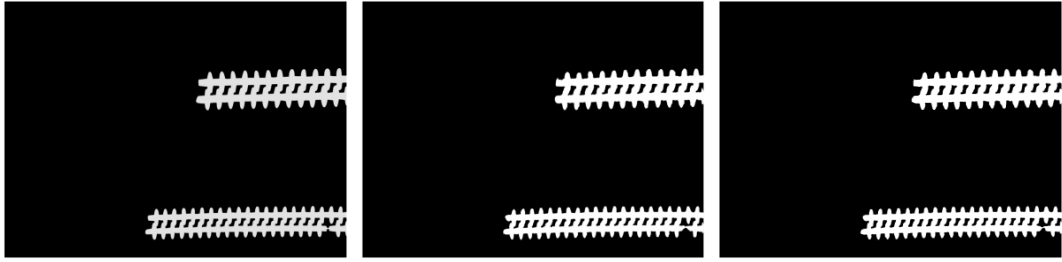


图 3.15 验证集测试结果对比图

（从左到右为：原始掩膜图像，预测掩膜图像，TTA 预测掩膜图像）

对验证集进行评估之后，对未放入网络训练的测试集（40 张）进行网络模型测试，并用 dice 系数来进行准确率的评估，网络模型测试结果准确率达到 90% 及以上的占 15%；准确率达到 80% 及以上的占 87.5%；准确率达到 70% 及以上的占 100%。

准确率	90% 及以上	80% 及以上	70% 及以上
网络模型测试	15%	87.5%	100%

表 3.1：准确率

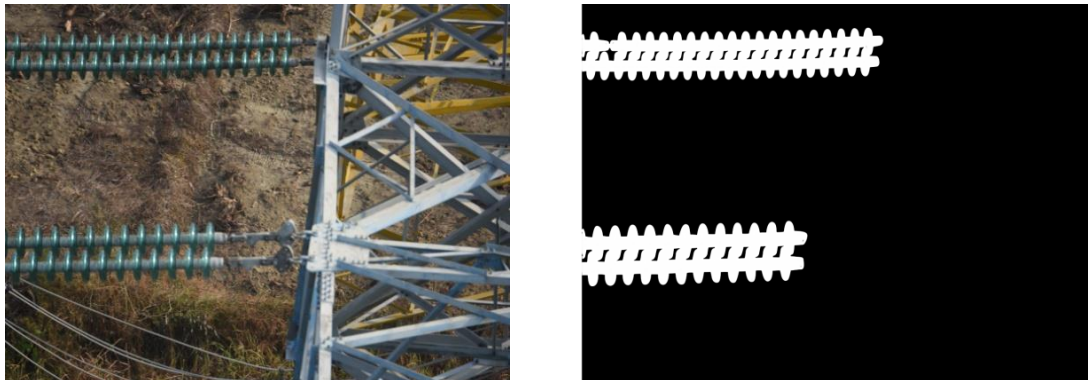


图 3.16 准确率为 90% 以上的原图和掩模图

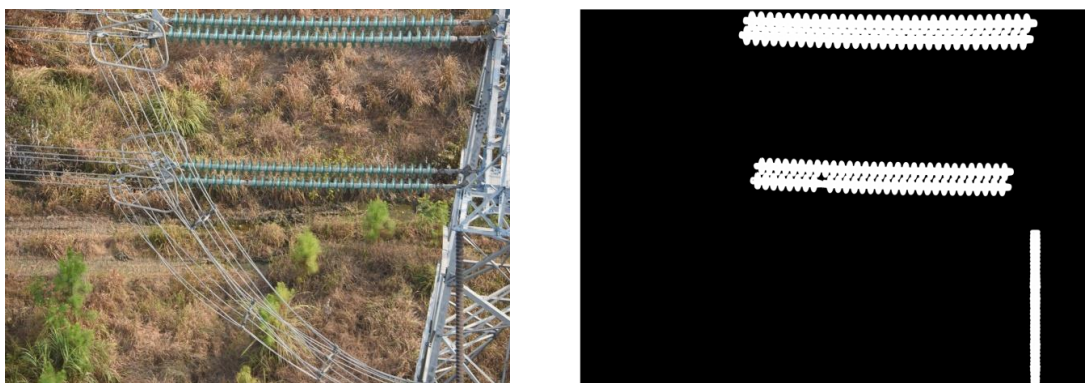


图 3.17 准确率为 80% 以上的原图和掩模图



图 3.18 准确率为 70% 以上的原图和掩模图

从上图可以看出绝缘子串珠分割准确率在 90% 以上的掩膜图像形状光滑，与原图基本完全重合；准确率在 80% 以上的掩膜图像边上存在粗齿；准确率在 70% 以上的掩膜图像存在空心情况。

4 基于 YOLOV4 的绝缘子缺陷检测算法

4.1 网络技术背景

随着深度学习的发展，可用于绝缘子自爆识别和定位的网络算法也越来越多，如：王万国等提出的基于 RCNN 利用 Faster R-CNN 的无人机巡检图像电力小部件识别研究^[7]；陈涛等人提出的轻量级 YOLOV3 的绝缘子缺陷检测方法^[8]等。以上这些算法模型均可在一定程度上识别检测出绝缘子缺陷的位置，但会存在无法在复杂背景下准确检测、无法满足实时要求、抗干扰能力差等缺点。而 YOLOV4^[9]可以针对不同的检测器水平，从不同的主干水平中选择更多的块以增加感受野和最佳参数聚集方法，从而在输入网络分辨率，卷积层数，参数之间找到最佳平衡。因此本文选择 YOLOV4 网络模型来对绝缘子自爆识别和定位无疑是最佳的选择。

4.2 网络模型搭建

Joe Redmon 提出的 YOLO 模型是目标检测领域一种著名的端到端学习模型。其特点是与 RCNN 级数的两步模型相比，YOLO 模型的执行速度快得多，在回归检测中表现良好，最新一代的 YOLO v4 在继承 YOLOv3 的骨干特征提取网络

(darknet-53) 后，并添加了 SPP 模块，不仅能显著增加感受野，分离出标注图片上最重要的特征，并且几乎不会降低网络运行速度，其网络模型简图如下图所示。

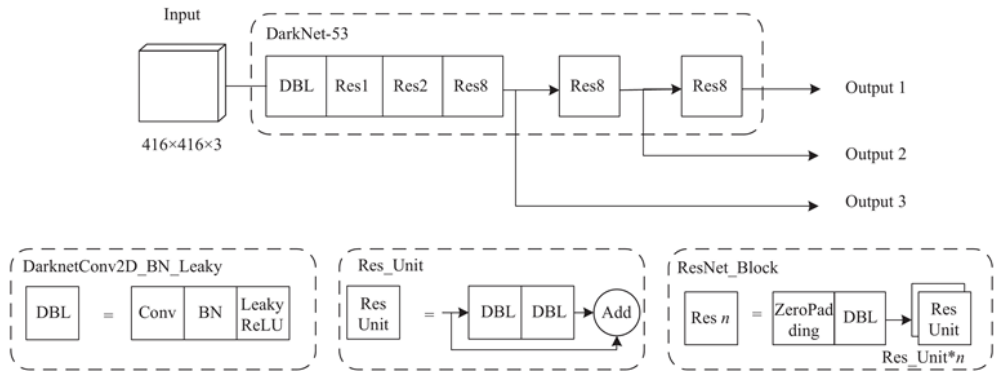


图 4.1 YOLOV4 网络训练流程简图

本文提出了基于 Kreas 框架下的一种 YOLOv4 的绝缘子缺陷检测方法，处理的主要步骤分为以下几步，首先利用 labelImg 工具精确标注分割后的绝缘子掩膜数据集，其次将标注文件里的数据格式转换为符合 YOLOV4 网络模型的数据格式；最后使用调整后的 YOLOV4 网络模型检测目标绝缘子缺陷所在位置。

4.2.1 制作训练样本

首先利用 labelImg 标注工具精确标注前期制作的绝缘子掩膜数据集中的 400 张缺陷绝缘子图片，并将标注文件里的数据 XML 格式利用脚本文件归一化转换为符合 YOLOV4 网络模型的 TXT 数据格式；通过 K-Means 算法获取九类预选框大小，使其符合缺陷绝缘子大小特征；最后对 YOLOv4 模型进行训练，训练结束后，将其用于检测目标绝缘子的缺陷位置。



图 4.2 绝缘子原图

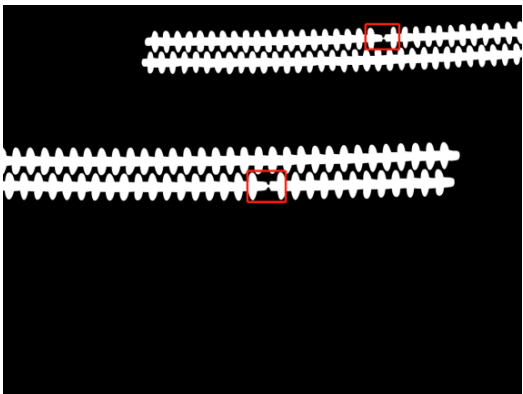


图 4.3 绝缘子自爆定位

4.2.2 YOLOV4 金字塔多级检测

绝缘子缺陷块的尺寸通常非常小，仅占原图的 1%~5%，仅用最后一层的特征图预测根本无法有效检测出绝缘子的缺陷部分。因此，YOLOV4 使用特征金字塔（FPN）^[10]结构来融合高层特征语义和低层位置信息，更精准的定位缺陷绝缘子。FPN 特征金字塔如图 4.4 所示。

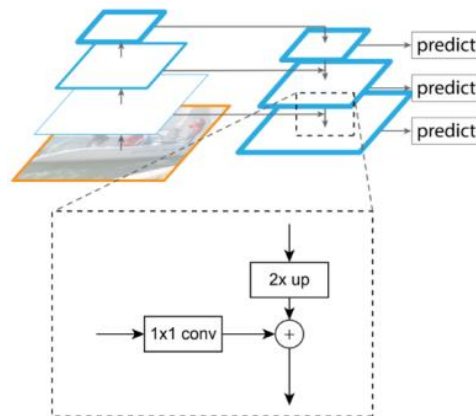


图 4.4 FPN 特征金字塔

4.3 YOLOV4 网络训练

本文在训练时，通过预设参数，网络接收 1152×864 分辨率的绝缘子图像，生成不同尺度的特征，用于 YOLOv4 的检测任务。绝缘子自爆区域用 IOU 系数来评价，IOU 表示产生的候选框（Candidate Bound）与原标记框（Ground Truth Bound）的交叠率或者说重叠度，也就是它们的交集与并集的比值。相关度越高该值。最理想情况是完全重叠，即比值为 1。

$$IOU = \frac{\text{area}(C) \cap \text{area}(G)}{\text{area}(C) \cup \text{area}(G)} \quad (7)$$

4.3.1 迭代策略

实验中使用随机梯度下降(Stochastic Gradient Descent, SGD)作为基础迭代器，相较于其他类型迭代器，SGD 可以在每轮迭代中，随机优化某一条训练数据上的损失函数，这样每一轮参数的更新速率会大大加快。实验设置参数中，训练 100 轮，其中冻结层前期训练共 10 轮，初始学习率为 0.001，动量为 0.949，正则化参

数为 0.0005, 批量大小为 2, 便于防止训练初期网络发散。以达到训练进行到解冻层后, 网络拥有一个稳定的模型。解冻层训练共 10 轮, 初始学习率为 0.0001, 动量为 0.949, 正则化参数为 0.0005, 批量大小为 1。为防止训练后期 SGD 优化器训练带来的不稳定性, 望江将会灵活调整学习率, 通过将学习率降低 10 倍, 进而继续训练。为保证参与损失计算的边界框数量, 将 ignore_thresh 设为 0.5。

4.3.2 先验框选择

先验框是在网络自身调节 box 大小之前, 提前给定一个或者多个大小尺寸的目标框帮助网络学习备选。本实验中采用 k-means 聚类方法在训练集所标注好的 bounding boxes 上来寻找先验框的尺寸。

k-means 算法中, k 代表了聚类的个数: 训练集为 $X^{(1)}, X^{(2)}, \dots, X^{(m)}$; 随机初始化 K 个聚类中心 $\mu_1, \mu_2, \dots, \mu_k$, $C^{(i)}$ 表示从 1 到 k 的所有聚类中心索引 (index) 中最接近于 $X^{(i)}$ 的索引, 公式表示为:

$$c^{(i)} = \min_k \|x^{(i)} - \mu_k\|^2 \quad (8)$$

其中 μ_k 表示接近于聚类 k 点处平均值。

通过交叉验证计算损失函数, 最终可以选择最优的 k 值。标准的 k-means 聚类方法所采用的是欧氏距离, 实验计算出了 9 个先验框的尺寸。本实验将绝缘子图像输入至网络提取 3 个尺度特征, 采用 NMS (非极大抑制) [11] 的概念, 即抑制不是极大值的元素, 搜索局部的极大值。通过 NMS 的方法可以有效剔除概率过低的框, 保留 IOU 更大的边界框, 提高运行效率。

4.3.3 训练结果分析

通过分析所给掩膜图数据集及其 XML 文件可得, 40 张掩膜图中, 有 8 张是无缺陷的, 另外 32 张掩膜图均有 1 个或 1 个以上的绝缘子缺陷。同时, 掩膜图图片的分辨率大小不一致, 掩膜图上绝缘子面积占图片面积的相对值不一, 而 YOLOv4 网络在测试时, 统一将图片 resize 到 1152 * 864 大小, 在这个过程中, 图片信息可能会有一定的损失, 容易造成不良的检测结果; 同时 YOLOv4 预选框的大小仅取决于训练数据集, 如果测试图片中的绝缘子过大或者过小, 并不能完美的契合我们的结果模型, 极易造成误检出多个或漏检的情况因此, 我们在测试过程中, 对测试数据集中的部分文件进行了二次处理。针对不同尺寸的绝缘

子我们采取的方法有放缩法（大型绝缘子），截取法（小型绝缘子）。放缩法是将原图重录到原来的 1/4 大小，位置在图片左上方，示意图如图 4.5 及图 4.6 所示，经 `resize` 后，使得原来相对较大的图片最终变得相对较小，其检测出的坐标可以直接适用在 `IOU` 计算中。截取法是将图片的绝缘子部分单独截取出来，重新生成一个与原图分辨率一致的图片，示意图如图 4.7 及图 4.8 所示，经 `resize` 后，相对较小的绝缘子最终变得相对较大，从而便于检测,其检测出的坐标需要与截取时的参数相加减之后，就能应用于 `IOU` 计算中。

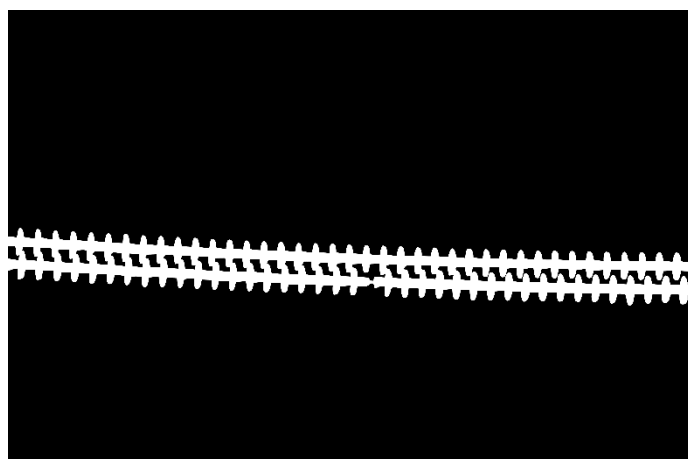


图 4.5 放缩法原图



图 4.6 经放缩法后的图片

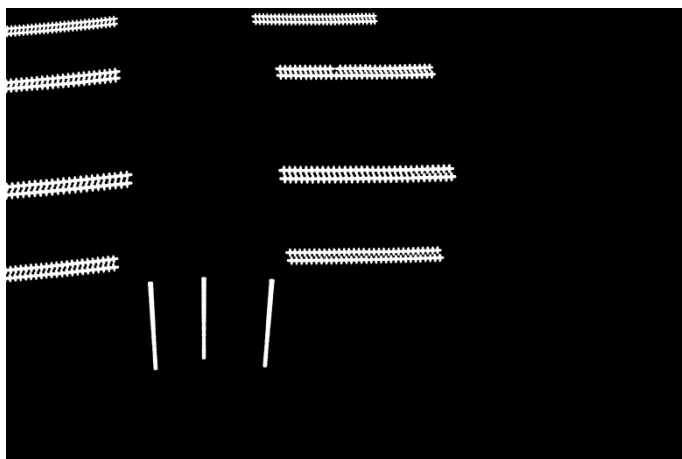


图 4.7 截取法原图



图 4.8 经放缩法后的图片

最终，我们成功的检测出了其中的 29 张缺陷掩膜图，部分检测结果及对比图如下所示，其整体 IOU 均值达到了 79.4%；对于 8 张无缺陷掩膜图没有任何一处误检；同时，对于多个绝缘子检测，也有着良好的结果，如图 4.11。

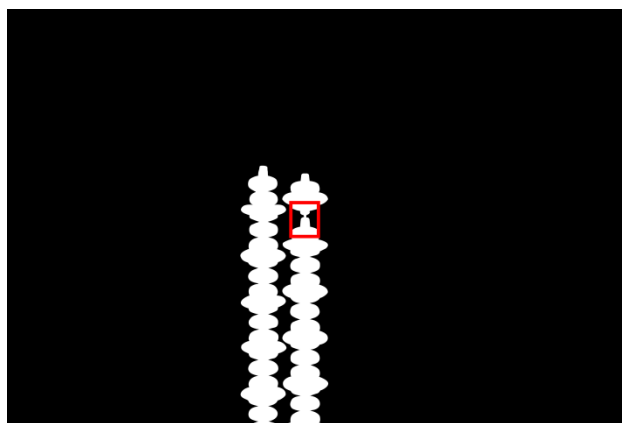


图 4.9 YOLOV4 检测结果

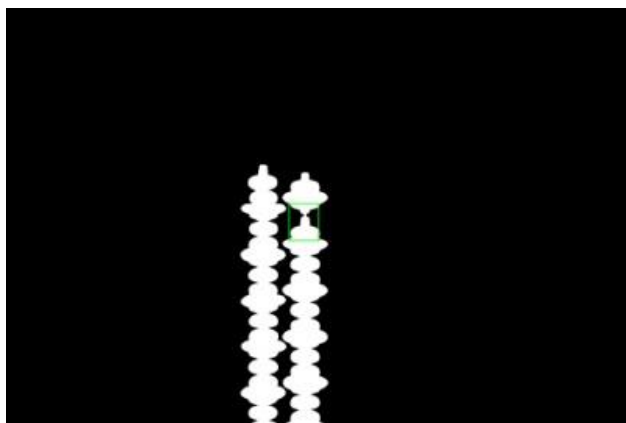


图 4.10 真实结果

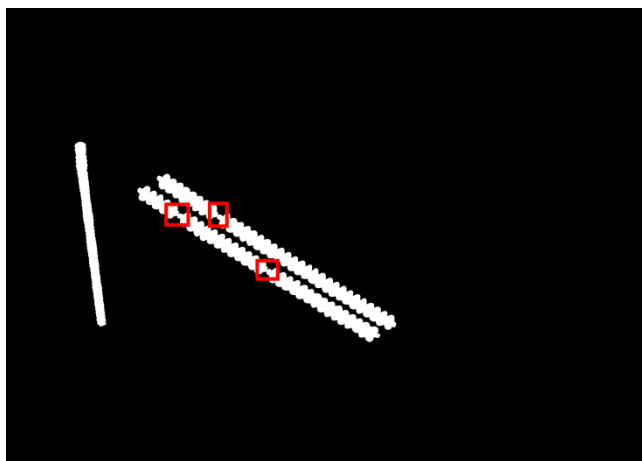


图 4.11 YOLOV4 检测结果(多个绝缘子)

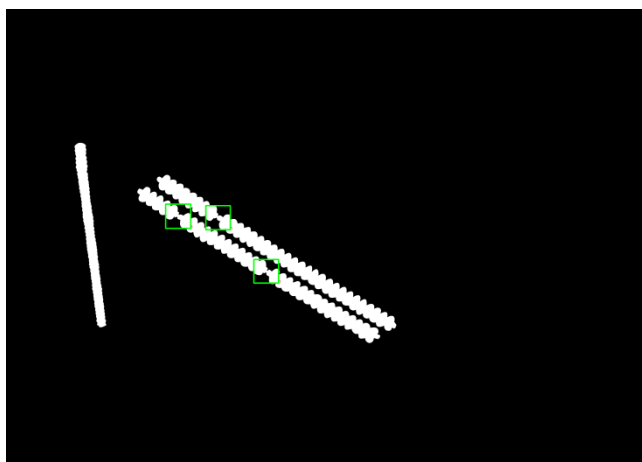


图 4.12 真实结果

5 总结与展望

对于绝缘子串珠分割任务，本文提出了一种 U-Net 语义分割网络，由于所建立模型的缺陷及硬件设备的不完善，在对绝缘子串珠分割时会对图片进行压缩导致部分信息的丢失，虽然模型训练准确率能够达到 85% 左右，但由于缺少样本数据集，训练好的模型会缺乏泛化性，即对与训练样本相差较大的绝缘子串珠无法分割；对绝缘子自爆识别和定位任务，本文提出基于 YOLOv4 网络的缺陷检测方法，利用训练好的网络直接检测绝缘子缺陷位置，由于样本数据集的不足，YOLOv4 网络模型的预选框具有一定的局限性，模型初期对部分绝缘子缺陷定位效果不好，特别是尺寸异于通常绝缘子的部分图片，但通过对测试图片进行简单的图像处理手段，我们仍然完成了 90% 的测试图片检测任务，IOU 值能达到 79.4%。无论是 U-Net 语义分割网络还是 YOLOV4 网络，本文结果均表明网络模型在保持绝缘子检测性能的前提下，大幅提升了网络的检测速度，能够满足无人机电力巡检任务要求。若拥有更多类型的缺陷绝缘子数据集，本文所提出的两类网络将能够进一步提升检测准确率及泛化性；我们也将继续深入挖掘网络架构，使其能达到更高的准确率及更短的训练时间，满足实际生产问题的需要。

参考文献

- [1] 陈文浩, 姚利娜, 李丰哲. 无人机电网巡检中的绝缘子缺陷检测与定位[J]. 计算机应用, 2019(S1).
- [2] 周鲁科, 朱信忠. 基于 U-net 网络的肺部肿瘤图像分割算法研究[J]. 信息与电脑(理论版), 2018(5).
- [3] 袁甜, 程红阳, 陈云虹, 张海荣, 王文军. 基于 U-Net 网络的肺部 CT 图像分割算法[J]. 自动化与仪器仪表, 2017(06):59-61.
- [4] He K, Zhang X, Ren S, et al. Delving Deep into Rectifiers: Surpassing Human-Level Performance on, ImageNet Classification[J]. 2015.
- [5] Zhang H, Cisse M, Dauphin Y N, et al. mixup: Beyond Empirical Risk Minimization[J]. arXiv: Learning, 2017.
- [6] A. Krizhevsky, I. Sutskever, and G. Hinton. Imagenet classification with deep convolutional neural networks. In NIPS, 2012.
- [7] 王万国, 田兵, 刘越, 等. 基于 RCNN 的无人机巡检图像电力小部件识别研究[J]. 地球信息科学学报, 2017, 19(2): 256-263
- [8] 吴涛, 王伟斌, 于力, 等. 轻量级 YOLOV3 的绝缘子缺陷检测方法[J]. 计算机工程, 2019, 45(8), 275-280. DOI: 10.19678/j.issn.1000-3428.0053695.
- [9] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: Optimal Speed and Accuracy of Object Detection[J]. arXiv preprint arXiv:2004.10934, 2020.
- [10] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2117-2125.
- [11] Neubeck A, Van Gool L. Efficient non-maximum suppression[C]//18th International Conference on Pattern Recognition (ICPR'06). IEEE, 2006, 3: 850-855.