

基于改进 YOLOv5s 的遥感目标检测

王悦炜¹, 焦良葆^{1,2}, 高 阳¹

(1. 南京工程学院 人工智能产业技术研究院, 南京 211167;

2. 江苏省智能感知技术与装备工程研究中心, 南京 211167)

摘要: 优化遥感目标检测对于军事民生方面都有着重要意义; 由于遥感数据中成像模糊、目标较小、检测对象数量多, 导致检测精度不高, 现提出一种新网络: 新网络在 YOLOv5s 原网络的基础上使用 Mish 激活函数代替 SiLU 激活函数; 为解决遥感图像中小目标的问题, 采用了对小目标、低分辨率友好的 SPD-Conv 模块; 考虑到使用耦合检测头会存在回归、分类两个任务之间的冲突问题, 采用 YOLOX 中解耦的检测头, 提高了模型查准率; 实验结果表明, 相比于原始 YOLOv5s, 新网络在 mAP 平均精度均值方面提升了 7%, 查全率 (recall) 提升了 10.9%, 检测速度 FPS 提升了 16.95%; 改进后的网络模型相对于原始模型具有明显优越性, 识别效果提升明显。

关键词: 深度学习; 遥感; 激活函数; 小目标; 解耦检测头

Remote Sensing Target Detection Based on Improved YOLOv5s

WANG Yuewei¹, JIAO Liangbao^{1,2}, GAO Yang¹

(1. AI Industrial Technology Research Institute, Nanjing Institute of Technology, Nanjing 211167, China;

2. Jiangsu Intelligent Perception Technology and Equipment Engineering Research Center, Nanjing 211167, China)

Abstract: Optimizing remote sensing target detection is of great significance to military and people's livelihood. Due to the blurred image, small target and large number of detected objects in the remote sensing data, the detection accuracy is not high. A new network is proposed; the new network uses the Mish activation function to replace the SiLU (Sigmoid weighted liner unit) activation function on the basis of the original YOLOv5s (you only look once v5s) network; In order to solve the problem of small targets in remote sensing images, SPD-Conv (Space-to-depth-Conv) module which is friendly to small targets and low resolution is adopted; Considering the conflict between regression and classification tasks when using coupling detector heads, the decoupled detector heads in YOLOX (you only look once X) are used to improve the model precision. The experimental results show that compared with the original YOLOv5s, the new network has improved the average accuracy of mAP (mean average precision) by 7%, the recall rate by 10.9% and the detection speed by 16.95%. The improved network model has obvious advantages over the original model, and the recognition effect is significantly improved.

Keywords: deep learning; remote sensing; activation function; small goals; decoupling detector

0 引言

军事民生方面的遥感目标检测具有越来越重要的意义。遥感目标检测要求从复杂的遥感图像中找到感兴趣的目标, 并精确地标注目标的位置和名称^[1]。同时, 由于遥感图像的拍摄角度特殊, 很大一部分遥感图像是从高空俯视的角度进行拍摄的, 造成一幅遥感图像中常会包含很多无用信息, 这大大增加了检测遥感目标的难度^[2]。

这就需要高精度和对遥感目标检测有针对性的优化方法, 但是遥感目标的检测除了拍摄角度问题, 还存在着其他一些难题: 由于小目标覆盖面积小、像素面积占比少、特征表达不充分、拍摄距离较远、图像覆盖范围较大而导致的形态畸变等特点, 使得基于遥感图像的小目标检

测面临更大挑战^[3]。

随着机器学习的快速发展, 深度学习越来越多地应用在目标识别方面, 图像目标识别技术也越来越先进。当前, 目标检测的主流方法大致分为两种: 以 R-CNN、Fast R-CNN、Faster R-CNN 网络为代表的两阶段目标检测以及以 YOLO、SSD 为代表的单阶段目标检测^[4]。由于两阶段目标检测在进行图像识别时需要先选出候选区域, 再对选出的候选区域进行调整分类, 这就使得双阶段比单阶段的检测精度高, 但同时检测速度 (FPS) 大打折扣, 并且随着单阶段检测算法 YOLO 的不断更新, 单阶段目标检测已经能在保证较高检测速度的同时, 兼顾良好的检测精度。

在单阶段目标检测中, 检测模型存在因目标重叠带来的漏检、模型计算量大导致检测速度较慢、损失函数 GIOU

收稿日期: 2023-02-01; 修回日期: 2023-03-04。

基金项目: 江苏省自然科学基金资助项目 (BK20201042); 江苏省政策引导类计划项目 (SZ-SQ2020007)。

作者简介: 王悦炜 (1999-), 男, 河北邯郸人, 硕士研究生, 主要从事目标检测、图像处理方向的研究。

焦良葆 (1972-), 男, 安徽安庆人, 硕士, 教授, 硕士研究生导师, 主要从事图像信号处理、视觉信息理解方面的研究。

引用格式: 王悦炜, 焦良葆, 高 阳. 基于改进 YOLOv5s 的遥感目标检测[J]. 计算机测量与控制, 2023, 31(8): 70-76.

_LOSS 本身缺陷导致特殊情况下无法判断预测框的位置状态, 从而影响检测精度等问题。为解决上述存在的问题, 张云佐^[5]等人使用 DIoU-NMS, 以及在空洞空间金字塔池化 ASPP 基础上设置不同膨胀系数以获得不同的感受野, 得到了更好的信息增强检测效果, 提出了 ASPP+ 模块, 减少了传统 NMS 非极大值抑制对目标的漏检问题, 提高了精度。成倩^[6]提出了对 yolov5 检测端的优化, 对于舰船等小目标来说, 专门负责大目标的检测头并不能充分发挥作用, 甚至会对于小目标的检测产生一定的干扰, 所以将负责大目标的检测头进行移除, 减少了不必要的计算量, 在原网络基础上增加了检测速度。张宏群^[7]将 yolov5s (you only look once v5s) 中使用的 GIOU_LOSS 损失函数更换为 CI-OU_LOSS, 避免了当预测框在真实框内的时候 GIOU_LOSS 损失函数无法判断预测框的位置状态, 进而提升了检测的精度。

随着 YOLO 模型算法的不断更新, 对于遥感目标的检测不应单单局限于舰船、车辆等单一的目标。为了有效地对多种遥感目标进行检测, 本文通过对 DOTA 数据集中的游泳池、客机、直升机、战斗机、运输机、网球场共 6 种待检测目标进行学习并检测, 对 yolov5s 网络模型的激活函数、小目标检测能力、任务检测头三方面进行优化。

1 YOLOv5s-MSY 网络及改进策略

1.1 YOLOv5s-MSY 网络结构

基于 YOLOv56.0 版本改进之后的 YOLOv5s-MSY 结构如图 1 所示。对 YOLOv5 来说共有 5 个不同的版本, 根据模型的深度和宽度不同, 分别是: YOLOv5l、

YOLOv5m、YOLOv5n、YOLOv5s、YOLOv5x^[8]。在这 5 种 YOLOv5 模型中, YOLOv5s 模型相对于网络最小, 速度最快^[9]。

对于 YOLOv5s-MSY 网络来说, 结构分为 4 个部分, 分别是: 输入端、骨干网络 Backbone、颈部网络 Neck、预测网络 Prediction。首先, YOLOv5s-MSY 网络中的 CBS 模块更换为 CBM 模块, 二者的区别在于在卷积 Conv 以及 BN 之后的激活函数不同, 网络的激活函数由 SiLu (Sigmoid weighted liner unit) 函数更换为梯度图像更加平滑的 Mish 函数, 比原网络更加方便让大量遥感图像信息深入流动, 提升了检测性能、模型泛化能力。其次, 该网络将空间金字塔由 SPP 改进为 SPPF, 比 SPP 速度快上很多。在 SPPF 中, 数据首先经过 CBM 模块, 然后依次递减的经过 3 个 MaxPool 最大池化层, 最后将 4 个支路通过 concat 叠加, 最后再通过一个 CBM 模块。在骨干网络 Backbone 中取消了原本的 Focus 层, 取而代之的是一个卷积。同时对于 CSP1_X 以及 CSP2_X 进行了更新: CSP1_X 中的一个分支首先经过 CBM 模块然后通过残差网络, 另外一个分支仅通过一个 CBM 模块, 然后两个分支进行 concat 叠加, 最后再经过一个 CBM 模块, CSP1_X 中的参数 X 指的是残差网络 Resunit 的数量。CSP2_X 的两个分支中一个分支通过 CBM 模块, 另外一个模块首先经过 CBM 模块再经过 $2 * X$ 个 CBM 模块, 然后两个分支同样经过 concat 叠加, 最后再通过一个 CBM 模块。对于输入端来说, YOLOv5s-MSY 同样使用了 Mosaic (马赛克) 数据增强、自适应锚框计算、自适应图片缩放^[10]。Mosaic 数据增强: 通过随机缩

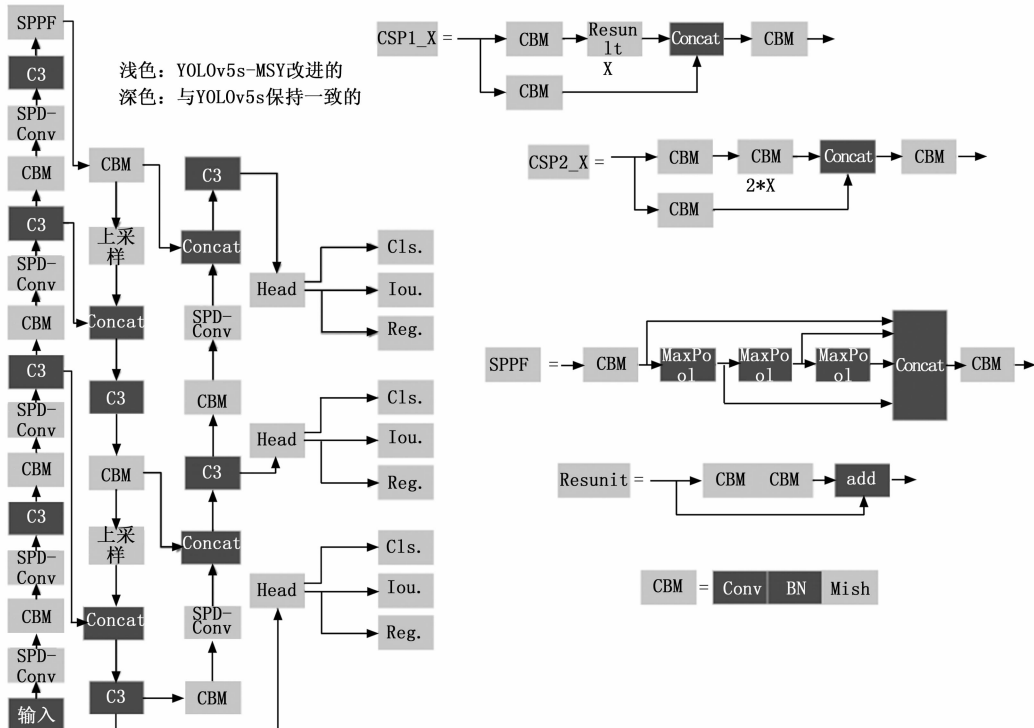


图 1 YOLOv5s-MSY 的网络结构图

放、随机裁剪、随机排布的方式来拼接 4 张图片^[11]，丰富了数据集，做到同时计算 4 张图片数据，减少了 GPU；自适应锚框计算：在设定了初始的预测框尺寸之后，与真实框的尺寸进行对比，再反向更新网络参数；自适应图片缩放：自动计算输入图像最适合的缩放比例，减少由图像填充带来的黑边干扰。在颈部网络 Neck 方面，依旧采用了自上而下的 FPN 结构和自底向上的 PAN 结构^[12]，PAN 结构首先是在 yolov4 中提出的，在之前的 yolov3 中只使用 FPN 结构。FPN+PAN 这种结构保证了在通过 FPN 将高层特征逐渐向底层传达强语义特征的同时，使用 PAN 将底层特征自下而上传达强定位特征。最后，通过预测网络 Prediction 中进行对指定对象的识别预测。

1.2 Mish 激活函数

为提升对复杂遥感目标的检测性能，在网络中加入激活函数：激活函数是神经网络的核心单元，给 CNN 网络加入了非线性特性，大大提升了 CNN 网络的学习能力^[13]。

在 YOLOv5 之前的版本中，激活函数使用的是带泄露修正线性单元函数，即 Leaky Relu 激活函数^[14]，该函数表达式如式 (1) 所示，函数形状如图 2 所示，通常 $\alpha=0.01$ ，该函数在大于 0 的定义域上为 $y=x$ 直线，梯度为 1；在小于 0 的定义域上为非零数值。但是该激活函数存在着一些问题：LeakyRelu 激活函数在定义域大于 0、小于 0 两个区间上都是线性分布的^[15]，而采用线性的激活函数就会影响模型对于一些非线性数据的表达能力；另外，因为 LeakyRelu 函数没有上界，如果网络单元输出过大，或者网络是循环结构的，就会导致梯度累积超过计算机的数值上限，造成“梯度爆炸”现象，由于遥感目标数据量大的特点，这种情况极易发生。为改善上述情况，本文使用 Mish 激活函数，Mish 表达式如式 (2) 所示，函数形状如图 3 所示：

$$Leaky\ Relu = \max(\alpha x, x) \quad (1)$$

$$Mish = x * \frac{e^{\ln(1+e^x)} - e^{-\ln(1+e^x)}}{e^{\ln(1+e^x)} + e^{-\ln(1+e^x)}} \quad (2)$$

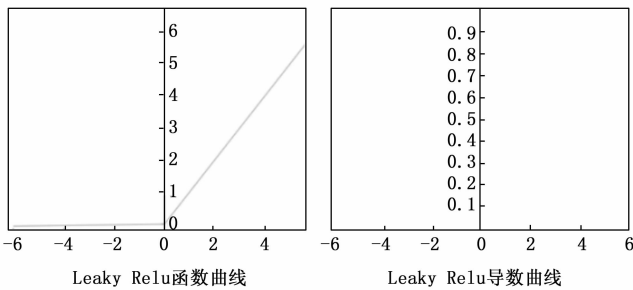


图 2 LeakyRelu 激活函数形状以及梯度

在 YOLOv5 网络当中，激活函数由 LeakyRelu 函数更换为 SiLU 函数，该函数可在负数域对数据进行拟合^[16]。此处改进的 Mish 激活函数与 SiLU 函数相似，SiLU 函数表达式如式 (3) 所示，函数形状如图 4 所示，但是 Mish 函数对外有着更好的表现效果，下面以 Mish 函数进行说明：

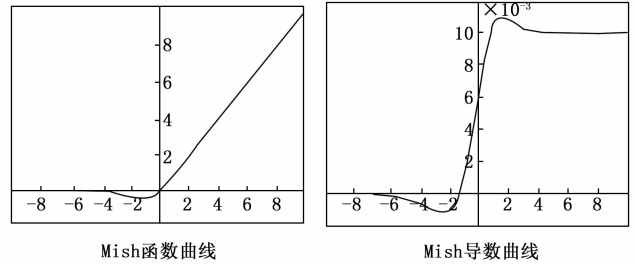


图 3 Mish 激活函数形状以及梯度

首先，Mish 函数非饱和，避免了因为饱和而导致的梯度消失现象；存在下界，在负半轴有着较小的权重，即在负半轴 Mish 的输出无限趋近于 0 但又不等于 0，保留了少量的负信息，防止了有些激活函数在定义域小于 0 的范围内函数输出为 0，导致一些“网络单元”始终不被激活，这种情况也被称为“神经元坏死”现象；函数自身具有正则化效果，对应的梯度图像更加平滑，这样就更加容易地对其进行优化。同时随着网络逐渐加深，信息可以更深入地流动，提升了对遥感目标检测的泛化能力。

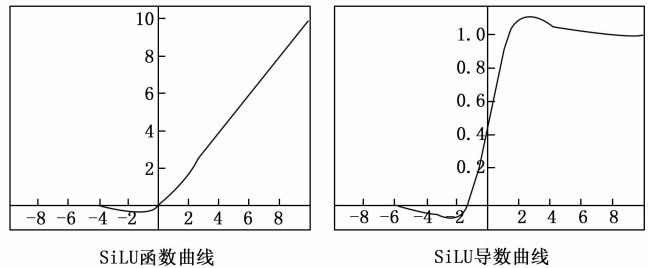


图 4 SiLU 激活函数形状以及梯度

$$SiLU = x * \frac{1}{1 + e^{-x}} \quad (3)$$

1.3 使用 SPD-Conv 模块

现在的 YOLOv5 模型归根到底使用的是 CNN 卷积神经网络，所以 YOLOv5 会因为 CNN 卷积神经网络本身的不足而存在一些弊端。卷积神经网络在对大目标以及高分辨率图像的识别已经较为成功，但是早期的遥感图像存在较多小尺寸、低分辨率的目标，这是因为遥感的定义就是一切无接触的远距离探测，导致遥感图像存在背景复杂、目标方向任意和目标尺寸小等问题^[17]。在待检测的遥感目标图像中，小目标与大目标常常一起出现。

而 CNN 在图像分辨率较低或物体较小时，检测性能迅速下降的原因是：CNN 网络使用的是卷积步长层（strided convolution）或卷积步长层+池化层。这导致在每一次进行“卷积步长”操作之后，会伴随着一些细微的信息丢失，造成学习特征不足，在进行多次操作之后甚至会使得遥感图像中一些较小的待检测学习目标的特征消失，网络无法正常高效地对遥感目标进行学习，这也就是 CNN 网络对小尺寸、低分辨率遥感目标检测性能下降的原因。

在参考文献 [18] 中，作者提出了一种新构建的模块，

称为 SPD-Conv, 这是由一个空间到深度 (SPD) 层与一个无步长卷积层组成。SPD 层对输入的遥感图像特征进行下采样, 与卷积步长层不同的是: SPD 层保留了通道维度中的所有信息, 有效避免了细微信息的丢失。作者将这种新型的组方式成功应用在整个 CNN 网络。

SPD-Conv 原理:

1) SPD 模块: 对于大小为 $S * S * C1$ 的中间特征图 X 来说, 属于 X 的一系列子特征图的尺寸如公式 (4) 所示 (公式 (4)、图 5 均来自于文献 [18]):

$$f_{0,0} = X[0:S:scale, 0:S:scale], f_{1,0} = X[1:S:scale, 0:S:scale],$$

.....,

$$f_{scale-1,0} = X[scale-1:S:scale, 0:S:scale];$$

.....

$$f_{0,scale-1} = X[0:S:scale, scale-1:S:scale], f_{1,scale-1}, \dots, \dots,$$

$$f_{scale-1,scale-1} = X[scale-1:S:scale, scale-1:S:scale]. \tag{4}$$

子特征图一般来说可以认为是从特征图 X 中剥离出来的, 因此子特征图的 w 、 h 由可整除特征图 X 的 w 、 h 的所有元素组成。Raja Sunkara、Tie Luo 二人在文章中给出了当 $scale=2$ 时的例子, 这里也按照 $scale=2$ 的情况进行解释说明。例图如图 5 所示。

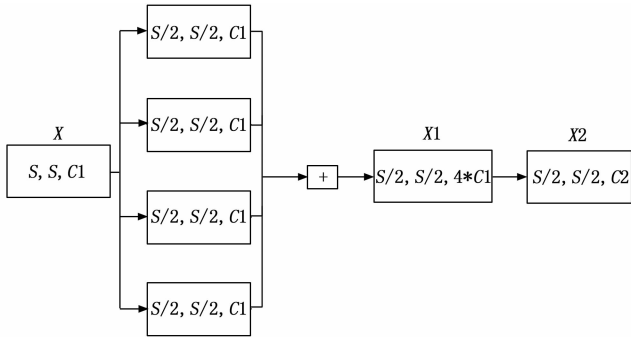


图 5 $scale=2$ 时的 SPD-Conv 示意图

当特征图 X 输入之后, 按照 $scale=2$ 对 X 进行处理, 由图可看出, 特征图 X 尺寸为 $(S, S, C1)$, 被分为了 $scale * scale=2 * 2=4$ 个子图, 每一个子特征图的尺寸为 $(S/2, S/2, C1)$ 。之后, 将这些子特征图沿着通道进行连接, 得到了另外一个特征图 $X1$, $X1$ 的尺寸为 $(S/scale, S/scale, scale * scale * C1) = (S/2, S/2, 4 * C1)$, 通道长度变为了原来的 4 倍。

2) 无步长卷积层 Non-strided Convolution:

在得到了中间特征图 $X1$ 之后, 作者添加了一个带有 $C2$ 滤波器的无步长卷积层 ($stride=1$), 其中 $C2$ 小于中间特征图 $X1$ 的通道数: $C2 < scale * scale * C1$ 。该模块将中间特征图 $X1$ 变换为特征图 $X2 (S/2, S/2, C2)$, 这一步对于特征图的 w 、 h 未进行改变, 只是改变了特征图的通道数。总的来看, 特征图 X 在经过了 SPD-Conv 层之后, 尺寸

从 $X (S, S, C1)$ 转变为了 $X2 (S/scale, S/scale, C2)$ 。

无步长卷积层的作用是减少 CNN 网络对遥感特征图中特征的丢失, 尽可能保留遥感图片的特征信息, 改善了网络对遥感目标检测的性能。例如: 使用 $stride=4$ 的 $4 * 4$ 卷积, 特征图在经过该卷积操作之后尺寸会缩小, 但是该卷积只会对原输入特征图中的每一个像素点信息只进行一次采样, 这样就会导致原始特征图的特征丢失。由于遥感目标中小目标居多, 小目标中的图像特征本身就要比大目标的图像特征要少, 所以这种特征丢失相比于大目标, 对遥感小目标的影响更大。总的来说, 只要卷积层的步长 $stride$ 大于 1 就会导致图片的特征丢失, 这也是 CNN 网络中的一种缺陷, 使用 SPD-Conv 的无步长卷积层就是为了尽可能地减少特征丢失, 这也就保证了网络对遥感小尺寸、低分辨率目标识别能力的提升。

1.4 将 YOLOX 解耦头应用在 YOLOv5

在通常使用的目标检测网络中执行回归和分类任务时, 系统常常共用一个检测头, 即: 二者对应的检测头是耦合的。导致二者会被迫使用一些相同的数据参数值, 造成准确度下降以及收敛速度变慢。同时, 遥感目标具有数量众多、目标小、训练参数大的特点, 共用检测头容易导致数据“混乱”, 造成“数据拥堵”。所以将 YOLOX (you only look once X) 中的解耦头^[19]应用在 YOLOv5 的模型中, 以此来增加遥感目标检测的准确度, YOLOX 的解耦头以及耦合头的收敛速度如图 6 所示 (图 6、图 7 均来自于文献 [20])。由图可看出 (Decoupledhead 曲线位于 YOLO head 曲线上方), 将 YOLO 的耦合头部换成 YOLOX 中的解耦头, 大大提升了收敛速度。在 YOLOX 中回归和分类任务分别使用不同的检测头, 避免了由耦合检测头带来的检测性能下降。

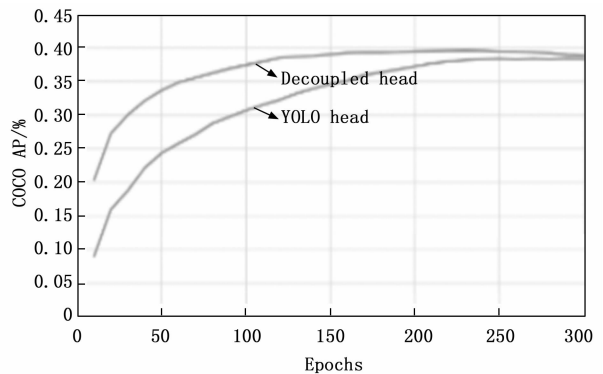


图 6 耦合头以及 YOLOX 解耦头的收敛速度示意图

同时, 由图 7 可看出, 分类、回归分别使用解耦的头部, 可减少头部耦合时带来的任务冲突, 防止“拥堵”, 减轻了模型训练遥感目标时的负担, 优化效果较好。

2 实验与分析

2.1 数据集的构建

本实验所使用的数据集来自于 DOTA 遥感数据集,

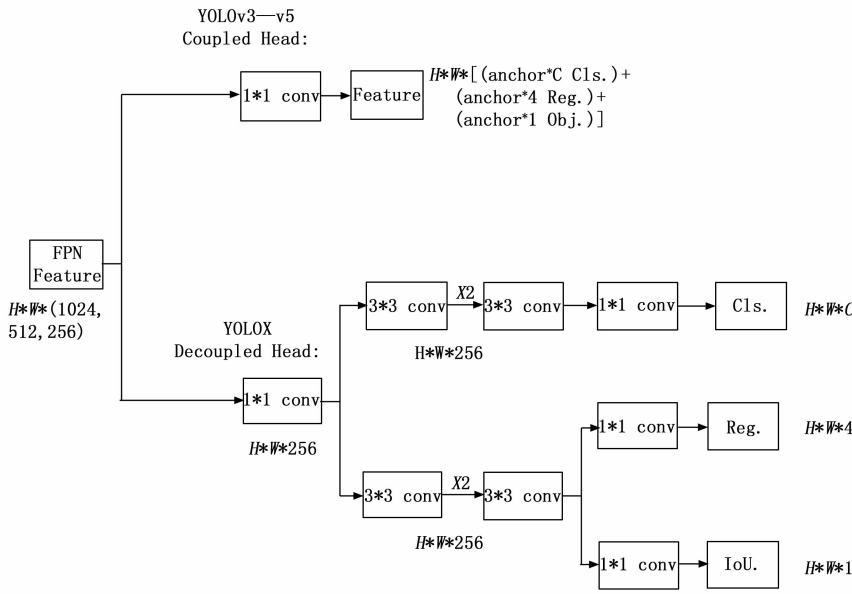


图 7 耦合头以及 YOLOX 解耦头示意图

DOTA 数据集是用于对航拍图像进行目标检测的大型图像数据集，其中的目标涵盖了各种各样的规模、位置、形状^[21]。它可用于发现和评估航拍图像中的物体。DOTA 数据集无论从数量还是质量上来说，比同类型数据集具有很大优势。为了使训练后的模型具有较好的泛化能力，共使用 1 162 张遥感数据图，包含 6 种分类：客机 (airplane)、直升机 (helicopter)、运输机 (transport plane)、战斗机 (warcraft)、游泳池 (swimmingpool)、网球场 (tennis-court)。数据集原图如图 8 所示，共包含上述提及的 6 种待检测目标。



图 8 实验中所用数据集原图

2.2 数据集的处理

将已经准备好的数据集图片放在 images 目录中，通过 Labellmg 软件进行这 6 种类别的标注，并将软件标注模式更改为 YOLO，使标图产生的文件是 YOLO 网络模型训练可以直接使用的 txt 格式，并保存在数据集 images 目录

相同路径的 labels 目录中。然后将数据集按照训练集、验证集和测试集对应数量为 670、245、247 进行划分。

2.3 实验平台搭建

本次实验的模型训练的平台搭建在服务器上，相关配置如下：实验操作平台为 Ubuntu18.04、CPU 为 AMD Ryzen Threadripper 3970X、GPU 为 NVIDIA GeForce RTX 3090 显卡 * 3，显存 24 G，内存 128 G、软件框架使用 Pytorch、编程环境为 Python 语言。

2.4 参数设置及评价指标

模型训练的参数设置如下：输入图片大小 (imgsz) 为 640 * 640、迭代轮数 (epochs) 为 300 轮、批次大小 (batch-size) 为 48，本文中的所有实验均是在不加载相应模型的预训练权重的情况下进行。

实验采用的参数指标：查全率、查准率 (precision)、平均精度均值、训练所用的权重文件大小 (MB)、检测帧数 (FPS)。部分指标的计算公式如下所示：

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

$$AP = \sum_{i=1}^{n-1} (r_{i+1} - r_i) P_{inter}(r_i + 1) \quad (7)$$

$$mAP = \frac{\sum_{i=1}^k AP_i}{k} \quad (8)$$

上述公式中：TP (true positives) 表示正确地将目标判断为正例；FP (false positives) 表示错误地将目标判断为正例；FN (false negatives) 表示错误地将目标判断为负例。其中 P 称为查准率，表示被检测出的目标当中真实目标所占的比例；R 称为查全率，表示被正确检测出的目标占有所有正例的比例。因为查准率、查全率二者通常有高有低、所以在二者变化不大的情况之下，以 mAP (mean average precision) 平均精度均值作为评判网络精度的标准。其中， $r_1, r_2 \dots r_n$ 是按升序排列的 Precision 插值段第一个插值处对应的 Recall 值。

2.5 实验结果评估分析

2.5.1 目标检测算法对比

本阶段，在已发表的有关遥感论文 (文献 [22]、[23]) 的实验数据中，选取一些主流的目标检测算法进行对比，两阶段：Faster-RCNN，单阶段：SSD、YOLOv3、YOLOv4、YOLOv5s、YOLOX、YOLOv7。论文的实验数据如表 1 所示：不论是在平均精度均值 mAP、查准率，还是在网络参数量、权重文件大小方面，使用 YOLOv5s 网络

检测遥感数据集比其他网络具有明显的优势。

表 1 主流目标检测算法对比

算法	权重/MB	P/%	R/%	mAP/%
Faster-RCNN ^[22]	108	63.5	90.8	76.47
SDD ^[22]	90.6	92.3	78.2	74.12
YOLOv3 ^[22]	235	88.7	86.1	84.35
YOLOv4 ^[22]	244	87.6	89.7	86.76
YOLOv5s ^[22]	13.7	89.9	88.2	92.02
YOLOX ^[23]	68.5	—	—	81.8
YOLOv7 ^[23]	74.9	85.9	74.3	79.3

2.5.2 激活函数实验

为了证明更换 Mish 激活函数的有效性, 该实验除了将 YOLOv5s 的激活函数由 SiLU 更换为 Mish 函数, 还分别使用 Hardswish、FReLU、MetaAconC 三种激活函数进行对比, 根据最终得到的实验数据进行评价, 改进激活函数实验结果如表 2 所示。使用 Hardswish 函数会得到相对最快的检测速度, 但是 mAP 要比使用 Mish 函数时低 2%。对比 FReLU、MetaAconC 来说, 虽然加入 Mish 函数的网络在检测速度上稍慢, 但是在 mAP 方面提高了 5.7%、4%, 查准率、查全率方面也都得到了提升。所以, 将激活函数更换为 Mish 对网络性能提升最为明显。

表 2 激活函数对比实验结果

算法	权重/MB	查准率/%	查全率/%	mAP@0.5/%	FPS	训练集/验证集图片数量
YOLOv5s	14.5	85.4	76.3	80.8	144.93	670/245
YOLOv5s+Mish	14.5	86.7	76.4	82.2	196.08	670/245
YOLOv5s+Hardswish	14.5	75.5	83	80.2	312.5	670/245
YOLOv5s+FReLU	14.8	80.7	71.5	76.5	243.9	670/245
YOLOv5s+MetaAconC	15.4	85.7	74.7	78.2	112.36	670/245

2.5.3 消融实验

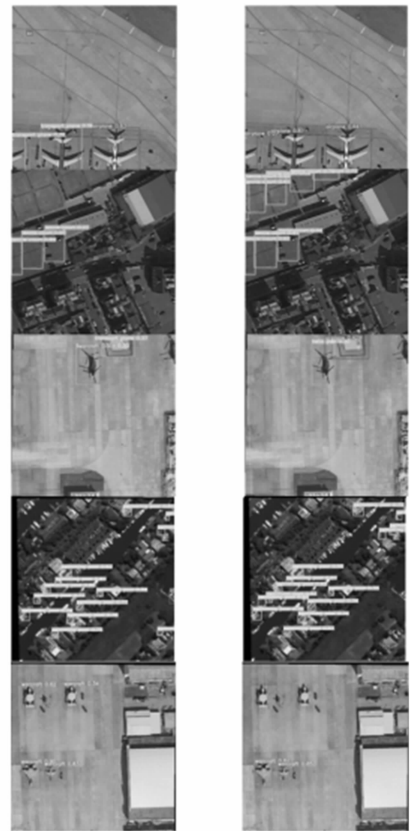
相比于 YOLOv5s, YOLOv5s-MSY 共有三点改进: 由激活函数 SiLU 更换为 Mish 激活函数、在网络模型中使用 SPD-Conv 模块、将 YOLOX 的解耦头应用在 YOLOv5s 模型当中。下面分别对单独加入特定模块、组合加入特定模块做消融实验, 结果如表 3 所示。由表可知, 单独加入 Mish 激活函数时, 由于该函数梯度图像更加平滑, 带来了网络在 mAP 上 1.4% 的提升; 单独加入 SPD-Conv 模块时, 由于该模块提升了对小目标、低分辨率目标的检测能力, 带来了 mAP 上 2.3% 的提升; 单独将 YOLOX 解耦头加入到 YOLOv5 网络中时, 由于 YOLOX 解耦头减少了检测头耦合时造成的任务冲突, 在 mAP 变化不大情况下, 带来查准率 4.8% 的提升。网络在组合加入特定模块时也都有着不同程度的提升。3 个

模块均投入使用时, 将网络命名为 YOLOv5s-MSY, mAP 提升了 7%, 检测速度也提升了 16.95%, 相比于原网络对于遥感目标的检测有较为明显的优势。

表 3 消融实验结果

算法	权重/MB	查准率/%	查全率/%	mAP@0.5/%	FPS	训练集/验证集图片数量
YOLOv5s	14.5	85.4	76.3	80.8	144.93	670/245
YOLOv5s+Mish	14.5	86.7	76.4	82.2	196.08	670/245
YOLOv5s+YOLOX 解耦头	14.4	90.2	70.5	80.4	101.01	670/245
YOLOv5s+SPD-Conv	14.4	91.1	74.9	83.1	232.56	670/245
YOLOv5s+Mish+SPD-Conv	17.6	77.9	81.7	83.2	196.08	670/245
YOLOv5s+SPD-Conv+YOLOX 解耦头	32.2	84.1	78.8	84.1	192.31	670/245
YOLOv5s-MSY	15.3	84.6	87.2	87.8	169.49	670/245

为了能够更加直观地观察 YOLOv5s-MSY 网络的精度提升, 本文利用原网络以及 YOLOv5s-MSY 网络对于测试集进行目标检测。如图 9 所示, 图 9 (a) 和图 9 (b) 分别对应不同网络。



(a) 原 YOLOv5s 检测效果

(b) YOLOv5s-MSY 检

图 9 YOLOv5s 网络与 YOLOv5s-MSY 网络效果对比图

由图可看出,改进之前网络对于运输机、客机、战斗机常会存在误检行为,对于直升机、网球场、游泳池会存在漏检行为。在使用改进后的网络进行检测时,上述情况都有了改善,减少了误检、漏检的行为,检测效果较好。

2.5.4 较大目标的对比实验

YOLOv5s-MSY 网络相比于原始 YOLOv5s 网络更加注重小目标的学习。为了全面体现 YOLOv5s-MSY 的性能变化,再选出共 200 张目标较大的遥感图像进行学习,数据集按照 6:2:2 进行分配,共有 3 个种类:田径场 (Ground-track field)、公路环岛 (Roundabout)、棒球场 (Baseball court)。实验结果如表 4 所示,改进后 mAP 上升了 19.31%、FPS 下降了 27.89%、查准率下降了 9%。由此看出,YOLOv5s-MSY 对于较大的遥感目标的检测速度有所下降。

表 4 较大目标对比实验

算法	权重 /MB	查准率/%	查全率/%	mAP@0.5/%	FPS	训练集/验证集图片数量
YOLOv5s	14.5	87.5	51.8	59.5	133.33	120/40
YOLOv5s-MSY	32.2	79.9	64.2	67.5	96.15	120/40

3 结束语

由于遥感目标的距离通常较远,造成待检测目标较小、分辨率较低,导致 YOLOv5s 原网络模型的检测精度以及检测速度的效果不好,且存在误检漏检的问题,本文在 YOLOv5s 原网络模型基础上将激活函数更换为 Mish,提升了模型的 mAP ;将 YOLOX 的解耦头使用在 YOLOv5s 模型当中,减少了耦合检测头的任务冲突,提高了模型的查准率;在原模型当中使用 SPD-Conv 模块,很好地提升了网络对小目标、低分辨率图像的检测精度。同时本文将这三方面改进整合在一起,对模型进行了综合优化。实验结果表明,YOLOv5s-MSY 相比于 YOLOv5s: mAP 提升了 7%,达到了 87.8%;检测速度提升了 16.95%,FPS 达到了 169.49。但是 YOLOv5s-MSY 在对小而密集的游泳池、直升机、运输机的检测还存在漏检误检的问题,影响了检测效果;对较大目标的检测存在检测性能下降的问题。由此看出,YOLOv5s-MSY 算法还存在较大的上升空间,在后续工作当中,主要需要提高在检测目标尺寸较小、数据集较少的情况下网络的学习能力以及模型对于较大目标的检测速度。

参考文献:

- [1] 赵文清,康悻瑾,赵振兵,等.改进 YOLOv5s 的遥感图像目标检测 [J].智能系统学报,2023,18(1):86-95.
- [2] 裴宇婵,李琳,梁建虎,等.不均衡小样本遥感图像飞机目标检测方法 [C]//第十六届全国信号和智能信息处理与应用学术会议论文集,2022:150-153.
- [3] 李惠惠,范军芳,陈启丽.改进 YOLOv5 的遥感图像目标检测 [J].弹箭与制导学报,2022,42(4):17-23.
- [4] 龙怡灿,雷蓉,董杨,等.基于 YOLOv5 算法的飞机类型

光学遥感识别 [J].地球信息科学学报,2022,24(3):572-582.

- [5] 张云佐,郭威,蔡昭权,等.联合多尺度与注意力机制的遥感图像目标检测 [J].浙江大学学报(工学版),2022,56(11):2215-2223.
- [6] 成倩,李佳,杜娟.基于 YOLOv5 的光学遥感图像舰船目标检测算法 [J].系统工程与电子技术,2023,45(5):1270-1276.
- [7] 张宏群,班勇苗,郭玲玲,等.基于 YOLOv5 的遥感图像舰船的检测方法 [J].电子测量技术,2021,44(8):87-92.
- [8] 张鹏程,武文波,李强,等.面向星载边缘计算的遥感目标检测算法轻量化优化研究 [J].空间控制技术与应用,2022,48(5):86-94.
- [9] 夏鹏久,赵恒谦,杨姿涵,等.基于 YOLOv5s 的高分辨率遥感影像尾矿库检测方法 [J].信息与电脑(理论版),2022,34(12):172-174.
- [10] 谢政峰.基于深度学习的新材料地板缺陷检测技术研究 [D].成都:四川大学,2021.
- [11] 董丽君,曾志高,易胜秋,等.基于 YOLOv5 的遥感图像目标检测 [J].湖南工业大学学报,2022,36(3):44-50.
- [12] 邢宇驰,李大军,叶发茂.基于 YOLOv5 的遥感图像目标检测 [J].江西科学,2021,39(4):725-732.
- [13] 徐静萍,王芳.基于改进的 S-ReLU 激活函数的图像分类方法 [J].科学技术与工程,2022,22(29):12963-12968.
- [14] 洪惠群,黄风华.基于轻量级神经网络的农作物病害识别算法 [J].沈阳农业大学学报,2021,52(2):239-245.
- [15] GUO X Y, LI Z, XING Y X. Study on analytical noise propagation in convolutional neural network methods used in computed tomography imaging [J]. Nuclear Science and Techniques, 2022, 33(6): 116-129.
- [16] 徐其森.基于深度学习的野生动物目标检测与跟踪 [D].哈尔滨:东北林业大学,2022.
- [17] 顾勇翔,蓝鑫,伏博毅,等.基于几何适应与全局感知的遥感图像目标检测算法 [J].计算机应用,2023,43(3):916-922.
- [18] SUNKARA R, LUO T. No more strided convolutions or pooling: a new CNN building block for low-resolution images and small objects [J]. arXiv Preprint arXiv: 2208.03641, 2022.
- [19] 罗文字,傅明月.基于 YoloX-ECA 模型的非法定野钓现场监测技术 [J].图学学报,2023,44(3):465-472.
- [20] GE Z, LIU S, WANG F, et al. Yolox: Exceeding yolo series in 2021 [J]. arXiv Preprint arXiv: 2107.08430, 2021.
- [21] XIA G S, BAI X, DING J, et al. DOTA: a large-scale dataset for object detection in aerial images [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 3974-3983.
- [22] 张朝阳,张上,王恒涛,等.多尺度下遥感小目标多头注意力检测 [J].计算机工程与应用,2023,59(8):227-238.
- [23] 李嘉新,侯进,盛博莹,等.基于改进 YOLOv5 的遥感小目标检测网络 [J/OL].计算机工程:1-11[2023-01-31]. DOI:10.19678/j.issn.1000-3428.0065935.