

doi: 10.7690/bgzdh.2022.07.004

# YOLOv4-tiny 及其改进算法在航空机务维修照相管理中的应用

张锐丽, 张琦, 高万春, 李江龙

(海军航空大学青岛校区五系, 山东 青岛 266000)

**摘要:** 针对飞机机务维修照相管理存在工作量大、不精确等问题, 提出一种利用深度学习 YOLOv4-tiny 算法来执行照片对比检测的方法。利用一个自制的数据集来训练网络模型, 为解决开口销螺母及其他背景干扰, 引入注意力机制模块以改进 YOLOv4-tiny。测试结果表明: 准确率 (precision, P) 相较原 YOLOv4-tiny 提高了 5%, 召回率 (recall, R) 提高约 8%, 平均准确率均值 (mean average precision, mAP) 提高了 4.9%, 照片识别精度和定位精准性方面都有较优表现, 满足照相管理中对目标精准识别与比对的要求。

**关键词:** 照相管理; 目标检测; YOLOv4-tiny; 注意力机制

**中图分类号:** TJ07 **文献标志码:** A

## Application of YOLOv4-tiny and Its Improved Algorithm in Photographic Management of Aviation Maintenance

Zhang Ruili, Zhang Qi, Gao Wanchun, Li Jianglong

(No. 5 Department, Qingdao Campus, Naval Aviation University, Qingdao 266000, China)

**Abstract:** Aiming at the problems of heavy workload and inaccuracy in the photographic management of aircraft maintenance, this paper proposes a method of using deep learning YOLOv4-tiny algorithm to perform photo comparison detection. A self-made data set is used to train the network model. In order to solve the problem of cotter nut and other background interference, the attention mechanism module is introduced to improve YOLOv4-tiny. Test results show that the precision (P) is improved by 5%, the recall (R) is improved by about 8%, and the mean average precision (mAP) is increased by 4.9%. It has excellent performance in photo recognition accuracy and positioning accuracy, and meets the requirements of accurate target recognition and comparison in photo management.

**Keywords:** photographic management; object detection; YOLOv4-tiny; attention mechanism

### 0 引言

飞机机务维修是各飞行单位的日常性工作。为确保飞行安全, 机务工作结束后, 通常需对工作完成情况实施质量检查。照相管理是一项重要的质量检查环节, 通过对机务工作关键部位实施拍照, 记录其完成状态并再次检查, 落实工作责任。传统的照相管理手段中, 照片比对要依靠人工分析判断, 因此基于现有照相管理流程对照片对比方式进行标准化设计, 开发智能化照相对比识别系统<sup>[1]</sup>, 能极大提高维修保障信息化水平, 有极为重要的军事经济意义。

信息化改革中, 装备维修必须要实现从“粗放型保障”到“精确性保障”的转变。近年来, 深度学习在较多学科领域中得到了广泛发展。基于深度学习的目标检测算法可分为双阶段和单阶段<sup>[2]</sup>2类。如 Ross Girshick 等提出的 R-CNN 双阶段算法, 通过 SVM 和回归器确定目标类别并对其进行定位<sup>[3]</sup>; 任少卿

等提出 Faster R-CNN 改进算法, 通过 2 个全连接层进行分类及回归模型检测, 实现神经网络自我学习生成候选区域的策略<sup>[4]</sup>。单阶段算法可以在保证计算精度的前提下尽可能简化网络, 直接将检测任务当作回归问题处理, 检测目标更快, 同时因图片完整输入使信息充分获取, 不容易出现错误的背景信息。鉴于日常维修保障中照相管理均表现为对某零部件拍照, 集中关注单一目标, 而 YOLOv4 网络层数多 (137 层), 计算成本高, 因此检测算法选择简化版的 YOLOv4-tiny<sup>[5]</sup>, 其网络结构更简单, 检测速度快, 性能满足需求。在此算法的基础上, 引入数据增强和注意力模块机制<sup>[6]</sup>并进行了对比分析。

### 1 算法概述及其原理

YOLOv4-tiny 是单阶段目标检测算法的代表之一, 在主干网络中只有 29 层, 有相对较小的模型尺寸 (<25 MB) 和非常高的角度测量速度, 具有直接

收稿日期: 2022-03-03; 修回日期: 2022-04-06

作者简介: 张锐丽 (1974—), 女, 内蒙古人, 硕士, 讲师, 从事装备保障信息管理与质量控制研究。E-mail: 353685047@qq.com。

将检测任务当作回归问题处理的特点<sup>[7]</sup>。

### 1.1 YOLOv4-tiny 算法及模型

YOLOv4-tiny 精简了特征提取网络, 由骨干网络 (CSPDarknet53-Tiny)、特征金字塔 (feature pyramid network, FPN) 和 YOLO 检测头 (YOLO Head) 组成<sup>[8]</sup>。CSPDarknet53-Tiny 主要用于前期的

特征提取, 它由卷积块 (Conv) 和跨阶段连接块 (CSPBlock) 组成; 卷积块中除了包含正常的卷积步骤, 还包含批量标准化与激活函数。特征金字塔通过降低网络的几何细节信息加强对特征的提取能力; YOLO Head 利用得到的特征结果进行最后的预测, 最终形成  $13 \times 13$  和  $26 \times 26$  2 个预测尺度<sup>[9]</sup>。其网络结构如图 1 所示。

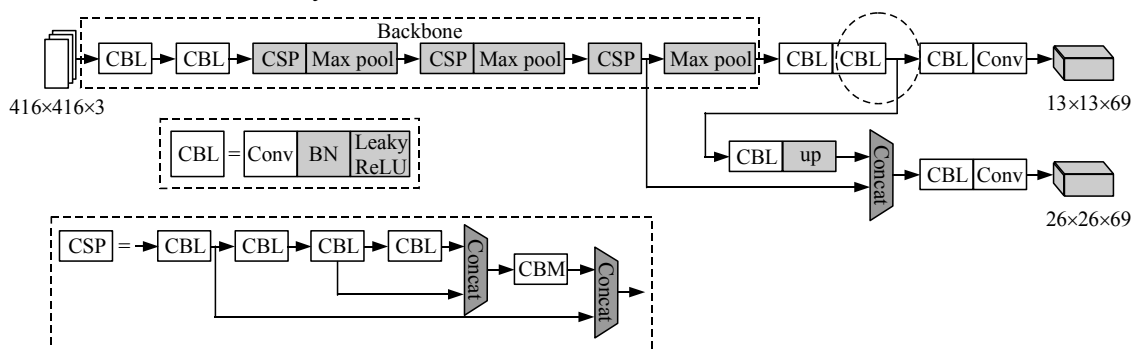


图 1 YOLOv4-tiny 网络结构

### 1.2 注意力机制

在标准 YOLOv4-tiny 程序中, 简单卷积与上采样操作不能有效减少其假阳性预测, 导致整体检测准确率不高<sup>[10]</sup>; 因此在骨干网络 CSPDarknet53-tiny 模块后, FPN 模块前嵌入注意力机制结构 Attention Reseau, 用于融合更多高质量的上下层语义特征信息, 以大幅度提高检测准确率。

## 2 实验与结果分析

为验证 YOLOv4-tiny 算法相较于传统目标检测算法更适用于照相管理系统目标状态检测, 笔者通过创建数据集、设置实验环境、开展与改进实验、结果对比 4 个主要环节进行实施, 选取机务工作日常使用频率较高、具有典型代表性的部位——开口销作为实验的基本对象。开口销呈金属质地, 标准尺寸为  $2.0 \text{ mm} \times 35 \text{ mm}$ 。打开口销所固定的基底为螺帽, 其 6 瓣均匀分布, 相互夹角约为  $60^\circ$ , 相关编号如图 2 所示。



图 2 开口销实物

### 2.1 创建目标数据集

根据标准打开口销方法, 开口销插入孔内后,

需将头部推至紧贴螺帽缺口 (1、2 号位之间), 尾部沿螺帽棱面向两侧分开并紧贴螺帽棱面 (3、5 号位), 最后切除螺帽外侧多余部分, 将开口销两段尾端分别打紧在螺帽间的槽口内; 因此, 在操作过程中容易产生的错误包括如图 3 所示的几种类型: 1) 开口销头部露出过长; 2) 开口销尾端切除过多; 3) 开口销产生水平/垂直歪斜。其中, 问题 2) 由于尾端裁剪长度难把握, 成为出错频率较高的典型问题, 因此实验针对该问题各状态模式进行目标检测。



图 3 开口销各类型问题

确定研究对象后, 通过高性能相机 (型号: EOS 5D Mark IV) 采集实验图像数据, 来源于学生针对某飞机零部件打开口销的练习结果实际拍摄。选取正例目标 6 个, 问题 2) 类型反例目标 4 个。为充分反映维修操作实际情况, 采取均匀多角度拍摄方式, 每个目标获取分辨率为  $1280 \times 960$  像素的 JPG 格式图像 100 余张, 其中约 100 张作为训练集。各目标抽取 10 张图像混合成总数为 100 的测试集, 包含 61 张正确图像, 39 张错误图像, 共获得 1 079 张有效图像。

下一步, 为训练集添加图像标注。借助数据集标注工具 Labeling 软件, 手动为目标添加标签, 生

成如图 4 所示 xml 文件，并提取转化为 txt 文件。本实验中设置正确标签为 1，3、5 号位缺失的标签设为 2，3 号位单独缺失设为 3，5 号位单独缺失设为 4。目标检测数据集如表 1 所示。

```
- <bndbox>
  <xmin>2180</xmin>
  <ymin>1030</ymin>
  <xmax>3698</xmax>
  <ymax>2591</ymax>
</bndbox>
```

图 4 图像标签信息 xml 文件(部分)

表 1 目标检测数据集

目标类型	训练集	测试集	总计
正确开口销成品(1)	578	60	638
3、5号位缺失成品(2)	96	8	104
3号位缺失成品(3)	156	15	171
5号位缺失成品(4)	149	17	166

### 2.2 实验环境配置

实验网络训练环境如下：迭代次数 50，初始学习率 0.000 1，结束学习率 0.000 001，共计 4 种状态类别的检测目标，神经网络模型基于高性能计算机开展训练，具体参数如表 2 所示。

表 2 实验环境及参数

参数	配置信息
CPU	Intel Xeon Silver 4110
内存	16G
GPU	NVIDIA RTX 2080Ti
显存	8G
操作系统	Ubuntu16.03
深度学习框架	TensorFlow1.14

### 2.3 实验操作与改进

本实验共采取 3 种不同方案训练网络模型，即原 YOLOv4-tiny 算法、YOLOv4-tiny+数据增强、YOLOv4-tiny+Attention(注意力机制)。各方案简要操作如下：

1) YOLOv4-tiny: 将图像文件与对应的 txt 文件输入程序，经过多轮迭代学习，损失函数趋近收敛，得到神经网络模型文件——1.h5(17.51 MB)，将测试集数据按照随机顺序经由模型判定，最终输出附带判定标签的图像，如图 5 所示，测试集每张图片输入模型实施判定后，会附带一个判定标签。

2) YOLOv4-tiny+数据增强: 首先对 1 079 张数据集图片依次进行了翻转、色彩变换和加噪声等操作，进一步将数据集扩充为 4 416 张图像，训练集与测试集数量比例不变。最后按 YOLOv4-tiny 方法输入数据并迭代，得到模型文件——2.h5(18.07 MB)，将测试集数据按照随机顺序经由模型判定，

最终输出附带判定标签的图像。

3) YOLOv4-tiny+Attention: 在 YOLOv4-tiny CSPDarknet53-tiny 模块后，FPN 模块前嵌入注意力机制模块，按 YOLOv4-tiny 方法输入数据并迭代得到模型文件——3.h5(17.66 MB)，将测试集数据按照随机顺序经由模型判定，最终输出附带判定标签的图像。



图 5 附带判定标签的结果

### 2.4 实验结果与对比

获得模型后，按照随机顺序将测试集数据输入模型进行状态判断，判断结果可分为 4 种类型，如表 3 所示。

表 3 判断结果类型

实际	判断为正(Positive)	判断为负(Negative)
实际为正(True)	TP	TN
实际为负(False)	FP	FN

实验通过以下模型指标进行实验效果评价。

经统计与对比分析，3 种学习方案所获得的模型指标如表 4 所示，算法性能如表 5 所示。

表 4 YOLOv4-tiny 系列算法实验对比

算法	P	R
YOLOv4-tiny	89.00	81.97
YOLOv4-tiny+数据增强	91.00	86.44
YOLOv4-tiny+Attention	94.00	90.08

表 5 YOLOv4-tiny 系列算法性能对比

算法	帧率	模型体积/MB	mAP/%
YOLOv4-tiny	54.9	17.51	88.05
YOLOv4-tiny+数据增强	51.7	18.07	91.39
YOLOv4-tiny+Attention	49.6	17.66	92.95

其中 YOLOv4-tiny+Attention 算法的某目标类型 P 值、R 值效果如图 6、7 所示。

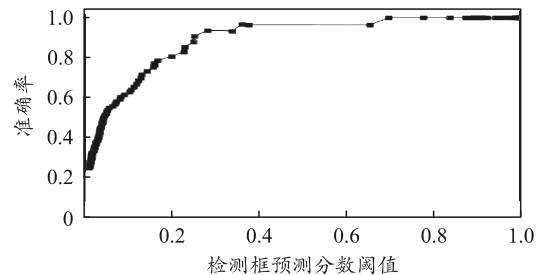


图 6 YOLOv4-tiny+Attention 判定为正确类型的准确率走势