

# 改进 YOLOv4 的油田作业现场烟火检测<sup>①</sup>

田 枫, 冯建臣, 刘 芳

(东北石油大学 计算机与信息技术学院, 大庆 163318)

通信作者: 刘 芳, E-mail: lfliufang1983@126.com



**摘 要:** 为解决油田作业现场烟火预警依赖人工巡检、不能及时发现等问题, 本文提出了改进的 YOLOv4 烟火检测算法. 具体针对摄像头距离远导致的烟火目标小、不易被识别的问题, 改进了网络特征融合部分, 并添加金字塔卷积 PyConv, 增强细节提取能力、增大局部感受野. 针对油田作业现场的复杂背景干扰问题, 加入注意力机制, 用于加强网络对重要特征的权重计算, 同时减少非关键数据的计算量. 最后通过聚类算法对目标样本锚定框优化, 利用自建烟火数据集进行实验, 实验结果证明, 改进后的算法模型具有相当高的性能, MAP 达到 90% 以上, 能够在复杂背景下对较小烟火目标保持较高的识别率, 说明改进后的算法在油田作业现场烟火识别中具有较高实用价值.

**关键词:** 特征融合; 金字塔卷积; 注意力机制; 烟火检测; YOLOv4; 目标检测

引用格式: 田枫, 冯建臣, 刘芳. 改进 YOLOv4 的油田作业现场烟火检测. 计算机系统应用, 2022, 31(6): 300-306. <http://www.c-s-a.org.cn/1003-3254/8366.html>

## Fire Detection Based on Improved YOLOv4 in Oil Field

TIAN Feng, FENG Jian-Chen, LIU Fang

(School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China)

**Abstract:** Considering that fire warning in oilfield operation sites depends on manual inspection and cannot be realized in time, this study proposes an improved YOLOv4 fire detection algorithm. Specifically, due to the long distance between the camera and the fire target, fires are too small to be identified. Given this problem, the network feature fusion is improved and the pyramid convolution (PyConv) is added to enhance the detail extraction ability and increase the local receptive field. In response to the complex background interference in the oilfield operation sites, the attention mechanism is adopted to strengthen the network's ability in weight calculation of important features, reducing non-critical data calculation. Finally, the anchor boxes of target samples are optimized through a clustering algorithm, and the self-built fire dataset is used for experiments. The experimental results prove that the improved algorithm model has quite good performance, has a mean average precision (mAP) of more than 90%, and can maintain a high recognition rate for small pyrotechnic targets in complex background. It shows that the improved algorithm has high practical value in pyrotechnic recognition in oil field operation site.

**Key words:** feature fusion; pyramid convolution (PyConv); attention mechanism; fire detection; YOLOv4; object detection

油田作业现场的环境具有易燃易爆的特征, 烟火一旦发生, 企业就会面临巨大的损失, 使企业生产停滞, 甚至对烟火灾害发生地工作人员造成伤亡. 目前油田作

① 基金项目: 黑龙江省自然科学基金 (LH2021F004); 黑龙江省高等学校教改工程 (SJGZ20200037); 东北石油大学研究生教育创新工程 (JYCX\_11\_2020); 黑龙江省省属本科高校基本科研业务费 (KYCXTD201903); 东北石油大学引导性创新基金 (2020YDL-11); 黑龙江省优秀青年科学基金 (YQ2020D001); 黑龙江省教育科学规划重点课题 (GJB1421113)

收稿时间: 2021-05-08; 修改时间: 2021-06-08; 采用时间: 2021-06-24; csa 在线出版时间: 2022-05-26

业现场主要采用人工巡检的方式预防烟火灾害,费时费力且实时性较低.随着油田安防监控的广泛应用,使得视频数据获取变得方便快捷.因此,如何充分利用油田安防监控的优势进行烟火检测,并在保证准确率的前提下,实现实时性检测,一直都是一个热点的研究课题.

采用对现场的视频流实时分析来检测烟火目标的方式,可大大缩短预警响应时间、增大检测范围.近年来,基于图像的烟火检测研究取得了丰硕的研究成果. Machine learning<sup>[1]</sup>使得基于图像识别的烟火检测技术取得较大进步<sup>[2]</sup>,但是该方式的识别性能受到所设计的手工特征质量的影响<sup>[3]</sup>.深度学习作为机器学习的分支,在基于图像的烟火识别中应用广泛<sup>[4]</sup>,它使用具有精准标注的数据集训练深层神经网络,使模型在求解过程中自动学习目标的特征.目前,以先产生候选区再进行分类的 RCNN<sup>[5]</sup>为代表的两阶段算法及直接输出类别和定位的 YOLO 系列<sup>[6]</sup>为代表的一阶段算法是目标检测主流算法.通常,两阶段算法检测精度高检测帧率低.单阶段算法检测帧率高,能够满足实时性的要求.基于监控视频的油田作业现场烟火检测,对实时性要求较高,所以本文选择 YOLOv4 为基础算法.

YOLOv4 作为通用的目标检测算法,直接应用于油田作业现场的烟火检测,其检测效果并不理想.主要由于视野较开阔,距离较远导致的烟火目标在监控视频中占比较小,现有算法漏检误检严重.并且对油田复杂背景的干扰不能很好处理.例如,实际监控视频中,发生火灾的区域范围比较小,很难从复杂背景中识别出来.场地中具有大量类似烟火的区域,比如黑色的油、深色的树、红色的工装等,都可能造成检测失败.

针对油田作业现场烟火特征,本文在 YOLOv4 深度学习算法的基础上,提出了改进的 YOLOv4 烟火识别模型.该算法模型改进了特征融合网络模块,将骨干网提取的浅层特征进一步融合,并使用多尺度的金字塔卷积 PyConv<sup>[7]</sup>替代标准卷积,增强对不同层细节的提取能力,增大局部感受野<sup>[8]</sup>,增强了小目标<sup>[9]</sup>的特征描述能力,提高了小目标的检测精度.同时在骨干网络中加入通道和空间注意力机制 CBAM<sup>[10]</sup>,增强网络对重要特征的提取能力并忽略特征中的噪声信息,以应对复杂背景,同时降低计算量,增强网络对复杂背景下烟火的识别能力,提高分类准确率.

## 1 网络 Neck 部分的改进

### 1.1 改进的特征融合

抽象语义在网络浅层较少,但纹理、坐标等信息丰富;网络深层含有较多的抽象语义信息,但纹理、坐标信息则比较少. YOLOv4 中,FPN 网络<sup>[11]</sup>将较强的语义特征自顶向下传递用于增强浅层类别的判定信息, PAN 模块<sup>[12]</sup>将浅层较强的纹理、坐标信息通过自底向上的方式向上传递,增强深层定位能力.在识别小目标时,由于网络深层小目标特征信息较少,造成对小目标的漏识别和误识别.为了强化网络对较小目标的检测能力,充分利用浅层信息本文重新设计了 FPN+PAN 层,改进了 YOLOv4 算法,如图 1(a) 中的 Neck-2 部分.

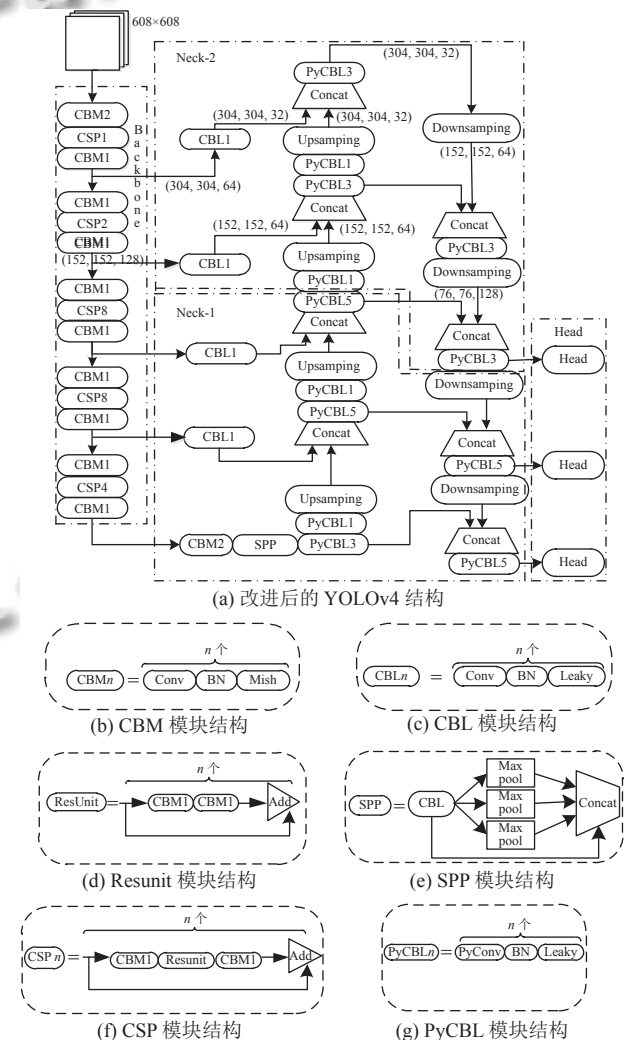


图1 YOLOv4 结构及组成单元

从 Neck 部分可以看出,本文新加入了两个包含更多位置、纹理信息的 Residual Block<sup>[13]</sup> 残差单元模块

提取的浅层特征信息层. 将浅层特征进一步和高层语义特征融合, 将包含更多位置、纹理信息的浅层特征和含有足够语义信息的深层特征进行融合. 加强网络对小目标的判别能力, 在具有较准确语义信息的同时, 为小目标的精准定位提供更多信息依据. 让算法能够准确识别较小的烟火目标.

如图 2 所示, 当输入图片大小为 608×608 时, 本文在原模型融合了 19×19、38×38、76×76 的基础上, 进一步融合了 152×152 和 304×304 大小的特征, 将更多的低层信息融入路径聚合网络.

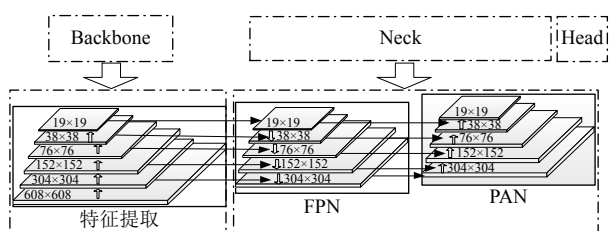


图 2 FPN+PAN 特征融合

最后, 使用整合后的特征图预测目标框、类别概率、置信度.

### 1.2 加入金字塔卷积

卷积层是卷积神经网络的基础构建模块, 卷积构建块的改进往往会增强特征的提取能力. 标准卷积(单一空间大小和深度的单一类型的卷积核)不具备提取多尺度特征的能力<sup>[14]</sup>. 而金字塔卷积(如图 3 所示)包含不同大小和深度的卷积核, 可以在不增加计算量的情况下, 增大局部感受野, 增强不同尺度目标的特征提取能力, 进而提高网络性能. 因此, 为了增强局部特征的提取能力, 解决下采样中细节信息丢失的问题. 本文在网络特征融合模块 Neck-1、Neck-2 部分, 使用金字塔卷积替代标准卷积.

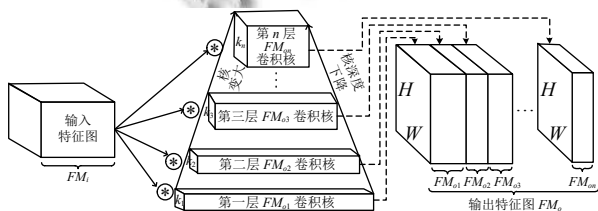


图 3 金字塔卷积 PyConv

金字塔卷积 PyConv(如图 3 所示), 包含一个核(kernel)金字塔, 其中每个层涉及大小(自下而上增大)和深度(随卷积核的增大而减小)不同的 filters, 他

们能够捕获场景中不同层的细节. PyConv 使用多尺度卷积核对输入特征进行提取. PyConv 会将输入特征分成多个组并进行独立的卷积运算.

令  $c_i$  表示金字塔卷积 PyConv 的输入通道数, 令  $k_j$  表示第  $j$  层卷积核的宽度, 则 PyConv 第 1 至  $n$  层的卷积核尺寸为  $k_1^2, k_2^2, \dots, k_n^2$ , 相应深度为  $c_i, \frac{c_i}{\left(\frac{k_n^2}{k_1^2}\right)}, \dots, \frac{c_i}{\left(\frac{k_n^2}{k_1^2}\right)}$ . 令  $c_{ol}$  表示第  $l$  层的输出特征维度, 则第 1 至  $n$  层相应的输出特征维度为  $c_{o1}, c_{o2}, \dots, c_{on}$ , 令  $H, W$  表示输出特征的高度和宽度, 则 PyConv 的参数量  $para$  与计算复杂度  $FLOPs$  分别为式 (1) 和式 (2):

$$para = k_n^2 \times c_{on} \times \frac{c_i}{\left(\frac{k_n^2}{k_1^2}\right)} + \dots + k_3^2 \times c_{o3} \times \frac{c_i}{\left(\frac{k_3^2}{k_1^2}\right)} + \dots + k_1^2 \times c_{o1} \times c_i \quad (1)$$

$$FLOPs = para \times (H \times W) \quad (2)$$

令  $k^2$  ( $k$  为卷积核大小) 表示标准卷积卷积核的空间分辨率, 令  $c_i$  表示输入特征通道数即深度, 使用  $c_o$  个相同分辨率与相同深度的卷积核进行运算. 令  $H, W$  表示输出特征的高度和宽度. 可得出标准卷积的参数量  $para$  与计算量  $FLOPs$  分别为式 (3) 和式 (4):

$$para = k^2 \times c_i \times c_o \quad (3)$$

$$FLOPs = k^2 \times c_i \times c_o \times (H \times W) \quad (4)$$

假设每层输出特征的深度一致, 那么每层的参数量和计算复杂度会呈现出较为均匀的分布. 因此, 可以得出 PyConv 与标准卷积的计算复杂度近似. 如图 1(a) 所示, 在网络的 Neck 部分中的 PyCBLn 模块, 其中 PyC 代表的就是金字塔卷积.

### 1.3 嵌入通道和空间注意力 CBAM

特征网络越深, 背景噪声越明显, 为减少复杂背景噪声的干扰, 并减少计算量, 本文在 backbone 的每个残差单元模块添加通道和空间注意力 CBAM(如图 4 所示). 注意力机制能够对特征图中重要的通道和空间特征增强, 使得网络在训练过程中能抓住目标的重点特征进行学习. 采用 CBAM 模块可以使网络学习到特征的每个通道的重要程度以及特征的每个空间位置的重要程度, 从而实现空间和通道的自适应调整, 优化网

络的权重,提升网络性能.进一步提高识别的准确率.

CBAM 注意力机制模块,同时结合了通道 (channel) 和空间 (spatial) 的注意力.相比于只专注通道和只专注空间的注意力机制可以取得更好的效果.令  $C$ 、 $H$ 、 $W$  分别表示通道、高、宽,则一幅输入的特征图为  $F = R^{C \times H \times W}$  要经过一维的通道注意力  $M_c = R^{C \times 1 \times 1}$  ( $C$  为通道数) 输出特征  $F'$ , 然后经过二维的空间注意力  $M_s = R^{1 \times H \times W}$ , 最终输出特征  $F''$  其过程表示为式 (5) 和式 (6):

$$F' = M_c(F) \otimes F \tag{5}$$

$$F'' = M_s(F') \otimes F' \tag{6}$$

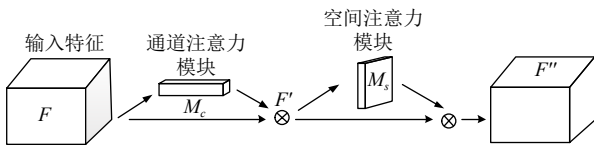


图4 CBAM 注意力结构图

通道注意力模块,将输入的特征进行空间维度压缩时,首先分别经过基于宽和高的全局平均池化和全局最大池化,得到两个不同的一维矢量特征  $F_{avg}^c \in R^{c \times 1 \times 1}$  和  $F_{max}^c \in R^{c \times 1 \times 1}$ , 然后分别经过一个共享的多层感知机 (MLP). 最后将多层感知机输出的特征进行基于对应元素 (element-wise) 的加和操作, 再经过 Sigmoid 激活操作, 生成最终的通道注意特征图. 通道注意力模块可以表示为式 (7):

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \\ = \sigma(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c))) \tag{7}$$

其中,  $F$  表示输入特征,  $\sigma$  代表 Sigmoid 函数,  $W_0$  和  $W_1$  表示多层感知机中模型的参数张量

空间注意力模块使用通道注意力加权后的输出特征作为输入特征. 首先, 基于通道层生成全局最大值和平均值池化, 以获得两个二维特征  $F_{avg}^s \in R^{1 \times C \times H}$  和  $F_{max}^s \in R^{1 \times C \times H}$ . 接下来, 根据通道维度拼接二维特征. 然后在卷积 (卷积核尺寸为  $7 \times 7$ ) 操作之后, 将维度降为单个通道. 最后使用 Sigmoid 激活函数获得空间注意力特征. 空间注意力模块可以表示为式 (8):

$$M_s(F) = \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) \\ = \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])) \tag{8}$$

其中,  $F$  表示输入特征,  $f^{7 \times 7}$  表示卷积核的大小是  $7 \times 7$ . 对于输入网络中的单张图像, 通道注意力关注特征内容的重要程度, 空间注意力关注特征位置的重要程度,

在一定程度上相互补充,提升了网络的特征提取效果.

将 CBAM 模块嵌入至 YOLOv4 网络的骨干网每一个 CSP 模块的每个残差单元的 shortcut linear 融合之前 (如图 5 所示). 由于 CBAM 注意力模块本身是对重要通道和空间特征进行增强, 因此将注意力嵌入到网络的 backbone 中一般可以取得很好的效果.

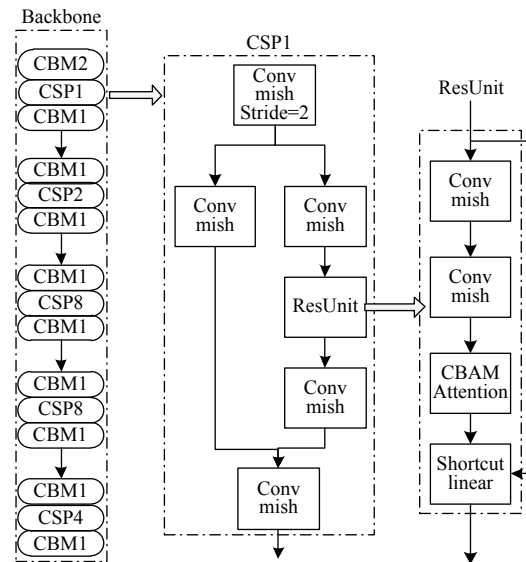


图5 嵌入 CBAM 注意力

### 1.4 先验框优化

为了使先验框的大小更适合烟火数据集,提升检测的精度. 本文采用 K-means++ 聚类方法<sup>[15]</sup>, 根据自建烟火样本数据集, 计算 anchors 值. 步骤如下:

1) 数据预处理: 异常值或噪声数据会对平均值产生很大影响, 并导致中心偏移. 因此, 我们对数据的异常点进行处理, 剔除离群的点.

2) 根据数据集中目标大小进行边界框聚类分析. 使用 K-means++ 聚类算法对烟火数据集中目标的 anchor box 进行聚类分析, 首先选取任意一个目标框的值作为初始聚类中心, 其次, 计算所有样本与现有聚类中心之间的距离  $D(x)$ , 并算出任意样本成为下个聚类中心点的概率  $P(x)$  如式 (9):

$$P(x) = \frac{D(x)}{\sum D(x)^2} \tag{9}$$

然后再使用轮盘赌选择法找出下个聚类中心点. 重复上述计算  $P(x)$  和  $D(x)$  的步骤直至得到  $K$  个锚框大小. 最后循环重复计算每个样本点到  $K$  个聚类中心点的距离, 将该样本点划分给距离最小的聚类中心的

类别并更新聚类中心点,直至获得的 Anchorbox 大小不再变化。

### 1.5 损失函数

损失函数用来评估模型的预测值和真实值的距离程度,可通过最小化损失函数求解和评估模型。本文模型中用到的损失函数由3部分组成:类别损失、置信度损失、回归损失。回归损失 CIOU loss 如式(10):

$$LOSS_{CIOU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (10)$$

其中,  $\rho^2(b, b^{gt})$ 表示预测框的中心点和真实值之间的欧式距离  $c$  表示同时包含预测框和真实框的最小闭合区域的对角距离。  $\alpha$ 和 $v$ 分别为式(11)和式(12):

$$\alpha = \frac{v}{1 - IoU + v'} \quad (11)$$

$$v = \frac{4}{\pi^2} \left( \arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (12)$$

YOLOv4 总的损失函数公式为式(13):

$$\begin{aligned} LOSS = & 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \\ & - \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} \left[ \hat{C}_i \log(C_i) + (1 - \hat{C}_i) \log(1 - C_i) \right] \\ & - \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{noobj} \left[ \hat{C}_i \log(C_i) + (1 - \hat{C}_i) \log(1 - C_i) \right] \\ & - \sum_{i=0}^{S^2} I_{ij}^{obj} \sum_{c \in classes} \left[ \hat{p}_i(c) \log(p_i(c)) + (1 - \hat{p}_i(c)) \log(1 - p_i(c)) \right] \end{aligned} \quad (13)$$

其中,  $S^2$ 表示预测特征为  $S \times S$  大小的网格,每个网格单元会对应  $B$  个 Anchor,每个 Anchor 通过模型推理计算得到相应边界框,最后会得到  $S \times S \times B$  个边界框。若预测框中无对象 (noobj),那么只需计算预测框的置信度损失。其置信损失使用交叉熵误差,并分为两部分:有对象 (obj) 和无对象 (noobj)。无对象的损失添加了权重系数  $\lambda$ ,相当于降低了该部分的贡献权值。分类损失也使用了交叉熵误差,当第  $i$  个特征网格的第  $j$  个 Anchor 负责预测某个实际目标时,则由该锚框生成的预测框计算分类损失值。

## 2 实验结果

实验设备的操作系统为 Ubuntu 18.04, i7-6700 3.4 GHz

CPU、NVIDIA RTX 3080Ti GPU,使用 PyTorch 框架。

### 2.1 烟火数据集

烟火数据集通过网络爬虫工具从网站获取,如表1所示,根据“火”“浓烟”“油田”等关键字从百度图片爬取的5000张图片,构成本次实验的训练集。训练集数据分为烟火图片和非烟火图片(如图6所示)。由根据相应关键字爬取的500张烟火数据和250张非烟火数据构成验证集。从网络视频库获取烟火视频和非烟火视频,进行视频帧提取获得5000张图像,构成本次实验的测试集。

表1 数据集描述(张)

数据集名称	烟火图	非烟火图	总数
训练集	3 350	1 650	5 000
验证集	500	250	750
测试集	3 500	1 500	5 000
总数	7 350	3 400	10 750



图6 训练集部分数据

### 2.2 评估指标

实验使用准确率 (accuracy) 即正确识别烟火或非烟火的概率、误检率  $P_f$  即在不烟火发生时却判定为有烟火的概率、召回率 (recall)、平均精度均值 (mAP) 和检测速度 FPS 作为实验结果的评价指标。召回率表示样本中有多少正样本被正确预测。MAP 可以全面评价实验结果精度和查全率。此外 FPS 也是本实验的重要评价指标之一。准确率、误检率及召回率公式如下:

$$accuracy = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \quad (14)$$

$$P_f = \frac{F_p}{T_n + F_n} \quad (15)$$

$$recall = \frac{T_p}{T_p + F_n} \quad (16)$$

其中,  $T_p$ 表示被模型预测为正类的正样本数量,  $T_n$ 表示被模型预测为负类的负样本数量,  $F_p$ 表示被模型预测为正类的负样本数量,  $F_n$ 表示被模型预测为负类的正样本数量。

### 2.3 训练过程

训练中将模型初始学习率置为 0.001, 将影响梯度下降到最优的动量因子置为 0.935, 将防止过拟合的权重衰减正则系数置为 0.0005, 共训练 6000 个批次。图 7 为训练过程中的 Loss 和 MAP 变化曲线。

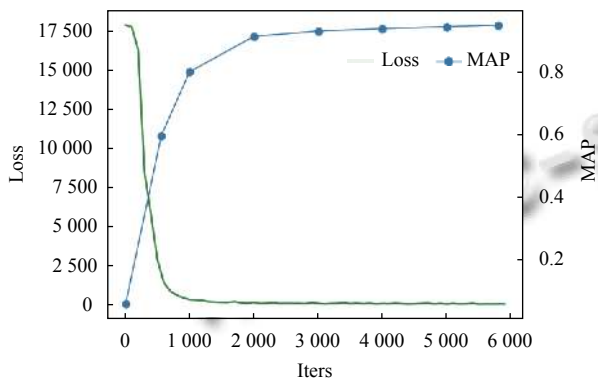


图 7 损失和 MAP

### 2.4 改进 YOLOv4 的烟火识别与分析

为了综合评估基于改进 YOLOv4 算法的烟火识别效果, 本文利用验证集图像数据进行烟火识别测试, 以验证本文模型的可靠性, 实验结果如图 8 所示。

图 8(a) 中, 在烟火目标距离较远且目标较小时, 能够准确检测到目标。在图 8(b) 中, 图片包含大片暗黑色树木等类烟区域及红色灭火器、黄色抽油机等类火区域仍能够准确识别烟火目标。图 8(c) 为其他情况监测结果。

### 2.5 算法比较

为了证明改进算法的有效性, 在自建烟火数据集上, 使用改进后的算法和原算法分别进行实验, 并使用准确率、误检率、MAP、FPS 等指标进行比较。其实验结果如表 2 所示。

从表 2 中可以看出, 本文算法和原算法相比, Fire 分类和 Smoke 分类的准确率分别提高了 (5.3%, 6.6%), 误检率分别降低了 (1.1%, 1.6%), MAP 值均优于原算法。本文提出的算法模型显著提高了对于较小烟火目标的识别能力以及对于类似烟火目标的甄别能力。能够在复杂的油田作业现场背景下过滤类似烟火目标的

干扰并准确地识别出较小的烟火目标。本文算法以 43 fps 的检测速度, 在满足烟火检测的实时性的要求下, 获得了平均 97% 的准确率和平均 1.2% 的误检率, 实验证明了本文算法在烟火实时检测中的可行性。



(a) 较小烟火目标检测



(b) 类烟火干扰区检测



(c) 其他情况检测

图 8 模型验证结果

表 2 模型性能对比

网络模型	类别	准确率 (%)	误检率 (%)	AP@0.5 (%)	MAP@0.5 (%)	FPS
YOLOv4	Fire	93.6	2.0	88.5	88.2	44.3
	Smoke	88.5	3.1	85.3		
Ours	Fire	98.9	0.9	92.5	91.8	43
	Smoke	95.1	1.5	90.3		

## 3 消融实验

为了验证改进算法的可行性, 对本文改进的部分进行了多组消融实验。消融实验均使用相同的训练数据集和测试集, 相同的超参数设置。消融实验结果如表 3 所示。

通过消融实验, 可得出以下几点结论, 首先, 从表 3

中第2组实验可看出在重新设计特征融合部分后,在模型特征融合层引入前两个残差块提取的细节信息,提高了浅层信息的利用率,使其平均精度均值有了3个百分点的提升,其次从第3组实验中可得出,在添加CBAM后,评价指标MAP同样得到了改善,其值增加了1.72%。说明CBAM可以有效地抑制噪声,强化通道和空间中目标的信息。最后一组实验中,在FPN和PAN网络结构中,使用PyConv替代标准卷积,减少细节特征的丢失,增大局部感受野,丰富了特征图的信息,其MAP值较上组实验提升了1.5个百分点。

表3 消融实验结果

NET	FPN+PAN	CBAM	PyConv	MAP (%)
1	×	×	×	84.35
2	√	×	×	87.80
3	√	√	×	89.52
4	√	√	√	91.08

#### 4 结论

针对油田作业现场复杂背景下的烟火小目标检测难的问题。本文基于YOLOv4网络模型,并结合CBAM注意力机制、特征融合、金字塔卷积等方式进行优化和改进。使用自建烟火数据集,并使用K-means++聚类算法对烟火数据集进行锚框调整。实验结果表明,改进后的YOLOv4检测模型能够满足实时检测的需要,与原模型相比,提高了准确率、MAP,有效地改善了油田作业现场场景中通用检测器性能不足的问题。本文针对油田有红色工装,黑色油,易与火和烟混淆等问题制作了针对该场景的数据集,如果应用于其他场景,例如学校、工厂等,则需要对数据集做改变,并重新训练模型。

#### 参考文献

- 何英哲, 胡兴波, 何锦雯, 等. 机器学习系统的隐私和安全问题综述. 计算机研究与发展, 2019, 56(10): 2049–2070.
- Tian L, Wang JD, Zhou HM, *et al.* Automatic detection of forest fire disturbance based on dynamic modelling from MODIS time-series observations. International Journal of

- Remote Sensing, 2018, 39(12): 3801–3815. [doi: 10.1080/01431161.2018.1437294]
- 傅天驹. 基于深度学习的林火图像识别算法及实现 [硕士学位论文]. 北京: 北京林业大学, 2016.
- Gaur A, Singh A, Kumar A, *et al.* Video flame and smoke based fire detection algorithms: A literature review. Fire Technology, 2020, 56(5): 1943–1980. [doi: 10.1007/s10694-020-00986-y]
- Girshick R, Donahue J, Darrell T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation. 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014. 580–587.
- Bochkovskiy A, Wang CY, Liao HY. YOLOv4: Optimal speed and accuracy of object detection. arXiv: 2004.10934, 2020.
- Duta IC, Liu L, Zhu F, *et al.* Pyramidal convolution: Rethinking convolutional neural networks for visual recognition. arXiv: 2006.11538, 2020.
- 哈尔滨工业大学深圳研究生院. 一种卷积神经网络感受野量化的方法及其应用. 中国: 108805284A, 2018-11-13.
- 王强. 基于深度学习的小目标检测算法研究 [硕士学位论文]. 成都: 电子科技大学, 2020.
- Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. 15th European Conference on Computer Vision. Munich: Springer, 2018. 3–19.
- Lin TY, Dollár P, Girshick R, *et al.* Feature pyramid networks for object detection. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017. 936–944.
- Yang JF, Fu XY, Hu YW, *et al.* PanNet: A deep network architecture for pan-sharpening. 2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE, 2017. 1753–1761.
- Hutchinson MR, Goldys EM, Ishikawa Y, *et al.* Deep residual learning. 2016, 10013: 100132Y.
- Chua LO, Roska T. The CNN paradigm. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 1993, 40(3): 147–156. [doi: 10.1109/81.222795]
- Arthur D, Vassilvitskii S. K-means++: The advantages of careful seeding. Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms. New Orleans: ACM, 2007. 1027–1035.