

基于 YOLO v3 的红外末制导典型目标检测

陈铁明*, 付光远, 李诗怡, 李源

火箭军工程大学作战保障学院, 陕西 西安 710025

摘要 导弹末制导飞行过程中, 基于传统方法检测红外目标时准确率和实时性不足。针对这一问题, 提出一种基于改进 YOLO v3 的红外末制导目标检测方法。从红外末制导背景出发, 优化损失权重, 提高了网络定位和分类能力。充分利用 Adam 算法自适应和动量法稳定的特点, 运用“预训练”的思想, 提出一种联合训练的方法, 大幅提高模型检测精度。实验表明, 改进算法在设计红外目标数据集上进行训练和测试, 检测效果理想, 平均准确率达到 77.89%, 检测速度达到 25 frame/s, 虚警率和漏检率都得到有效降低。

关键词 机器视觉; 红外图像; 末制导; 目标检测; YOLO v3; 联合训练

中图分类号 TP183

文献标识码 A

doi: 10.3788/LOP56.161502

Typical Target Detection for Infrared Homing Guidance Based on YOLO v3

Chen Tieming*, Fu Guangyuan, Li Shiyi, Li Yuan

Department of Information Engineering, Rocket Force University of Engineering, Xi'an, Shaanxi 710025, China

Abstract The traditional infrared target detection method for missile homing guidance is flawed because of low accuracy and lack of real-time feedback. Therefore, an infrared homing guidance target detection method based on the improved YOLO v3 is proposed, and it involves the optimization of the weight loss by considering the background of infrared homing guidance, improving the accuracy of positioning and classification. Subsequently, the adaptive moment estimation (Adam) and stable stochastic gradient descent (SGD) with momentum are fully exploited. Further, a joint training method predicated on pre-training, which significantly improves the accuracy of detection, is proposed herein. The improved algorithm is ideally trained and tested on the infrared target dataset designed in this work. The best mean average precision is 77.89%, and all the detection rates are greater than 25 frame/s. The false and missing alarm probabilities are effectively reduced.

Key words machine vision; infrared image; homing guidance; target detection; YOLO v3; joint training

OCIS codes 150.1135; 130.3060; 110.2970

1 引言

随着人工智能的迅猛发展, 大量先进技术广泛应用在制导武器系统中。红外成像凭借其抗干扰能力强、空间分辨率高、昼夜工作等特点成为精确制导中重要的研究领域。作为导弹武器的“眼睛”, 准确并及时发现目标对提高导弹的命中率具有重要意义^[1]。

由于目标的快速性和机动性日益增强, “人在回路中”的方式已经无法满足探测高时敏性和高覆盖率的要求^[2], 这就要求制导系统具备自动目标识别的能力。深度学习方法能够自主学习目标特征, 不断完善网络模型, 真正实现自动目标识别定位。

任何温度高于 0 K 的物体都向外界发出热辐射, 通过描述不同物体热辐射的差异可以得到红外图像。凭借不受光照影响、可昼夜工作的优势, 红外

收稿日期: 2019-01-29; 修回日期: 2019-03-12; 录用日期: 2019-03-27

基金项目: 国家自然科学基金(61403397, 61503389, 61202332)、中国博士后科学基金(2012M521905)、陕西省自然科学基金基础研究计划(2015JM6313)

* E-mail: chentieming1995@163.com

图像广泛应用于军事侦察领域。由于舍弃了物体的颜色、局部纹理等特征,红外图像的对比度和信噪比低,较为模糊^[3]。传统方法中,获取红外图像的质量对目标的检测效果有着十分重要的影响,因此易翔等^[4]提出一种结合帧间时空域特征的红外检测算法,实现了红外图像的增强,检测精度较高,但是较为突出的问题是检测时间长;熊斌等^[5]则针对红外场景中的小目标,提出一种自适应参数目标图像恢复算法,具有较低误检率,其本质也是对红外图像进行增强。传统方法虽然能够较好地完成红外图像的目标检测,但大多是在特征提取和图像增强方面进行改进,对检测目标具有一定的局限性。

深度学习方法和传统方法相比,最大的优势在于其强大的自学习和泛化能力。通过学习,深度模型不仅可以自动提取目标的形状、轮廓等表象特征,还可以挖掘出肉眼不可见的深层特征,从而弱化传统方法中对红外图像的增强操作,将研究的重点转移到网络模型的改进。2012年,Krizhevsky等^[6]提出的 AlexNet 以绝对的优势获得了 ImageNet 图片分类竞赛的冠军,自此引发了将深度学习应用于目标检测的热潮。Sermanet等^[7]提出的 OverFeat 是最早应用于目标检测的卷积神经网络方法之一,检测效果较传统算法有显著改进,但错误率依旧较高。Ren等^[8]提出的 Faster R-CNN 在 VOC2012 数据集上的平均准确率(mAP)达到 75.9%,但检测速度只有 5 frame/s,无法满足导弹高速飞行时实时检测目标的要求。2016年,Redmon等^[9]提出的 YOLO 算法真正实现了对目标的实时检测,检测速度达到 45 frame/s。在 YOLO 的基础之上,又相继改进提出 YOLO v2^[10]和 YOLO v3^[11]算法,大幅提高了检测速度和精度以及对小目标的检测效果。

本文将深度学习方法应用于红外图像目标检测,大大减少人为特征提取和处理的过程,借助于深度学习强大的性能,使红外目标的检测效率和精度得到大幅提高。为了在红外末制导背景下实现对目标又快又准的检测,针对红外图像的特点对 YOLO v3 网络结构进行改进,并从损失函数和训练模式入手,使模型能够较好地实现对红外图像的目标检测。以特种车辆、舰船、飞行器三类典型目标作为末制导研究对象,为解决红外目标数据集稀缺的问题,采取自行拍摄、标注的方式,制备数据集。同时,从边界框损失、置信度损失和分类损失三方面入手,优化损失函数权重系数,实验证明该方法能小幅提高模型的检测精度,并降低检测虚警率。为了深层挖掘红

外图像特征,结合动量法和 Adam 算法的特点设计联合训练方法,利用自适应算法 Adam 使网络快速收敛,再通过动量法对网络进行进一步优化和精修。Adam 算法加快网络收敛速度,有益于动量法初始学习率的设定,改善网络检测效果。

2 YOLO v3 原理及改进

2.1 YOLO v3 原理

YOLO 采用单个神经网络直接将整张图像作为输入进行训练,省去生成候选区域的中间步骤,能够快速区分背景区域和目标,从而实现对目标物体的实时监测。

YOLO v3 的检测过程是:首先将输入图像缩放至 416 pixel×416 pixel,接着将图像划分成 $S \times S$ 个相同大小的单元格,并输入神经网络进行训练。每个单元格只负责对中心位于该单元格中的目标进行预测,一个单元格预测 B 个边界框,每个边界框中包含的信息可以表示为 $(t_x, t_y, t_w, t_h, t_{conf})$,其中, (t_x, t_y, t_w, t_h) 和 t_{conf} 分别为预测的坐标信息和置信度信息。假设以图像左上角顶点为坐标原点,则单元格相对于图像的坐标为 (x, y) ,且锚点框的宽度和高度分别为 p_w 和 p_h ,那么预测出的边界框可以表示为

$$\begin{cases} b_x = \sigma(t_x) + x \\ b_y = \sigma(t_y) + y \\ b_w = p_w \cdot \exp(t_w) \\ b_h = p_h \cdot \exp(t_h) \end{cases}, \quad (1)$$

式中: (b_x, b_y) 表示预测的边界框中心的坐标; b_w 和 b_h 分别表示预测边界框的宽度和高度; $\sigma(t_x)$ 和 $\sigma(t_y)$ 分别为相较于选定单元格在横、纵轴方向上的预测偏移。由于预测的参数存在误差,因此预测的边界框(predicting box)与标定的边界框(bounding box)之间会存在一定偏移,如图 1 所示。

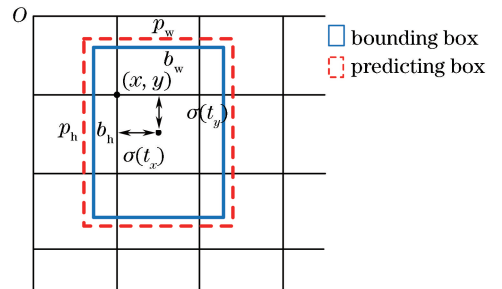


图 1 预测框和标定框的位置信息图

Fig. 1 Illustration of positions of predicting and bounding boxes

通常情况下,红外图像是以单通道的灰度图像形式存储,图2和图3分别是可见光图像及其红外图像示意图。由于红外图像缺失了颜色、轮廓、纹理等局部信息,对同一目标进行检测时,神经网络需要从红外图像中挖掘出更多细节特征才能实现较好的检测效果。



图2 机场可见光图像

Fig. 2 Visible image of airport



图3 机场红外图像

Fig. 3 Infrared image of airport

YOLO v3 原本的网络结构适用于可见光图像,即输入的图像大小为 $416 \text{ pixel} \times 416 \text{ pixel} \times 3$,一种解决方案是将红外单通道图像转换成三通道格式输入。虽然这种方案可以满足网络的输入格式,但只是将单通道进行简单的复制获得三通道图像,并没有增加颜色、纹理等特征,对最终的检测结果没有实质性的提高。

因此,从网络结构入手,将网络的输入调整为 $416 \text{ pixel} \times 416 \text{ pixel} \times 1$,这样做的好处在于输入单通道图像,使网络参数的计算更加集中,减少参数冗余。

2.2 损失函数加权

假设每个单元格的预测值为 p_* , 真实值为 \hat{p}_* , 损失可以表示为真实值和预测值的差,即 $p_* - \hat{p}_*$ 。在 YOLO v3 训练过程中,损失为所有单元格预测值和真实值的均方差。损失函数(l)共由三部分组成,分别为边界框损失($l_{\text{coordinate}}$)、置信度损失($l_{\text{confidence}}$)和分类损失(l_{class}),可以表示为

$$l = l_{\text{coordinate}} + l_{\text{confidence}} + l_{\text{class}} \quad (2)$$

边界框损失由中心坐标的误差和边界框宽与高的误差构成,可以描述为

$$l_{\text{coordinate}} = \sum_{i=0}^{S^2} \sum_{j=0}^B \text{Pr}_{\text{obj}}(i, j) [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] + \sum_{i=0}^{S^2} \sum_{j=0}^B \text{Pr}_{\text{obj}}(i, j) \cdot [(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2], \quad (3)$$

式中: (x_i, y_i) 表示预测框的坐标; w_i 和 h_i 则分别表示该预测框的宽和高; (\hat{x}_i, \hat{y}_i) 表示标定框的坐标, \hat{w}_i 和 \hat{h}_i 则分别表示其宽和高; $\text{Pr}_{\text{obj}}(i, j)$ 表示单元格 i 内存在目标,并由第 j 个锚点框负责预测; $\sum_{i=0}^{S^2} \sum_{j=0}^B \text{Pr}_{\text{obj}}(i, j) [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2]$ 为所有边界框中心坐标的误差项之和; $\sum_{i=0}^{S^2} \sum_{j=0}^B \text{Pr}_{\text{obj}}(i, j) \cdot [(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2]$ 为所有边界框宽与高的误差项之和。

置信度损失和分类损失的计算公式可以分别描述为

$$l_{\text{confidence}} = \sum_{i=0}^{S^2} \sum_{j=0}^B \text{Pr}_{\text{obj}}(i, j) (C_i - \hat{C}_i)^2 + \sum_{i=0}^{S^2} \sum_{j=0}^B \text{Pr}_{\text{noobj}}(i, j) (C_i - \hat{C}_i)^2, \quad (4)$$

$$l_{\text{class}} = \sum_{i=0}^{S^2} \text{Pr}_{\text{obj}}(i, j) \sum_{c \in \text{classes}} [p_i(c) - \hat{p}_i(c)]^2, \quad (5)$$

式中: C_i 为预测框的置信度; \hat{C}_i 为标定框的置信度; 分类标签都经过 one-hot 编码, $p_i(c)$ 为编码后预测类别; $\hat{p}_i(c)$ 为编码后的真实类别; c 表示当前检测的目标种类; classes 表示数据集的所有目标种类; $\text{Pr}_{\text{noobj}}(i, j)$ 表示单元格 i 内不存在目标。

如果直接使用(2)式计算损失,那么三类损失的权重均为1,可以理解为三种损失的重要程度相当,这显然与实际不符,故对损失函数进行改进加权,以提高未制导过程中目标的定位精准度和识别准确度。加权后的损失函数(l_w)为

$$l_w = (\alpha_{xy} l_{xy} + \alpha_{wh} l_{wh}) + (\beta_{\text{obj}} l_{\text{conf_obj}} + \beta_{\text{no_obj}} l_{\text{conf_noobj}}) + \lambda_{\text{class}} l_{\text{class}}, \quad (6)$$

式中: α_{xy} 为边界框中心坐标损失权重; α_{wh} 为边界框宽高损失权重; β_{obj} 为单元格内包含目标的置信度损失权重; $\beta_{\text{no_obj}}$ 为单元格内不包含目标的置信度损失

权重; λ_{class} 为分类损失权重。

为了提高边界框损失的地位,可以适当提高边界框损失的权重。同时,要想在红外末制导中精准瞄准目标,必须首先提高预测目标中心坐标精度,其次才是目标的形状,故需要增大权重 α_{xy} 和 α_{wh} ,且 $1 < \alpha_{wh} < \alpha_{xy}$ 。通常情况下,目标所占单元格数量要小于背景,因此需要赋予背景较小的置信度损失,即 $0 < \beta_{\text{no_obj}} < 1 < \beta_{\text{obj}}$ 。对末制导而言,目标检测也需要具有较高目标分类能力,需要提高分类损失权重,且 $1 < \lambda_{\text{class}}$ 。

综合考虑各类损失因素对红外末制导目标检测的重要程度,可以得到权重之间的关系为

$$0 < \beta_{\text{no_obj}} < 1 < \beta_{\text{obj}} < \lambda_{\text{class}} < \alpha_{wh} < \alpha_{xy}。 (7)$$

2.3 随机梯度下降联合训练

梯度下降是应用最广泛、效果最好的优化算法之一,并且也是目前为止优化神经网络最常用的方法^[12]。按照能否自动调节学习率分类,梯度下降方法可以分为自适应方法和非自适应法。其中,常用的自适应方法有 AdaGrad、RMSProp 和 Adam 等,非自适应方法有随机梯度下降(SGD)、SGD with Momentum 和 Nesterov Momentum 等。

2.3.1 动量法

SGD 算法容易进入陡谷,陡谷指当前维度的曲率远大于邻近维度,故难以出来,从而陷入局部最优。动量法(SGD with Momentum)能帮助 SGD 算法在相关方向上加速并抑制振荡,具有一定摆脱局部最优的能力^[13],如图 4 所示。

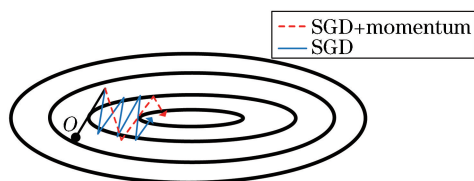


图 4 动量法加速收敛

Fig. 4 Accelerated convergence by SGD with momentum

顾名思义,动量法借鉴物理中的动量概念,将上一时刻的绝大部分更新向量添加到当前更新向量中,可以描述为

$$\begin{cases} v_t = \tau v_{t-1} + \xi \nabla_{\theta} J(\theta) \\ \theta_{t+1} = \theta_t - v_t \end{cases}, (8)$$

式中: v_t 表示从开始时刻到 t 时刻梯度的累加; $\nabla_{\theta} J(\theta)$ 表示在 θ 处函数 $J(\theta)$ 的梯度大小; ξ 表示学习率; τ 表示动量因子,通常取逼近于 1 的实数,这表明每次累加梯度时具有一定衰减,防止无限增大。

这样的好处在于:当前梯度点的方向与前一时刻保持的方向一致时,可以增大动量项并加速,当方向发生改变时,动量项会减小,故而能起到加快收敛和减小振荡的作用。

刻保持的方向一致时,可以增大动量项并加速,当方向发生改变时,动量项会减小,故而能起到加快收敛和减小振荡的作用。

2.3.2 Adam 自适应算法

Adam 算法作为一种学习率自适应调整的一阶梯度优化方法,具有实现简单、计算量小、运算速度快等优点,和其他自适应学习算法相比更具优势^[14]。和动量法类似,Adam 算法不断累积存储历史梯度信息,可以描述为

$$\begin{cases} m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \\ v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \end{cases}, (9)$$

式中: m_t 表示 t 时刻梯度的一阶矩; v_t 表示 t 时刻梯度的二阶矩; β_1 和 β_2 为指数衰减速率,负责控制梯度的记忆大小; g_t 表示 t 时刻的梯度大小。

Adam 算法不直接使用梯度的累加,而改用一阶矩和二阶矩,为不同参数设计独立的自适应学习率,能够在改变方向时起到平滑的作用。当 m_t 和 v_t 初始化为零向量时,它们都存在向零偏差,尤其是刚开始几次迭代和 β_1 、 β_2 接近于 1 的情况下。通过计算一阶、二阶矩估计值,校正偏差,近似求得一阶矩、二阶矩的无偏估计。并且矩估计可以根据梯度的变化动态调整,每一次迭代都对学习率作动态约束,且有明确范围,使得参数变化较为平稳。

因此,Adam 算法能够实现自适应调整学习率,更新的步长能够限制在一定范围内,能够较好地实现不稳定函数的优化。

2.3.3 基于 Adam 和 SGD with Momentum 的联合训练方法

由于红外图像的特性,神经网络需要花费更多的时间去挖掘图像中的目标特征。对于动量法而言,虽然能够学习红外图像的深层特征,但需要的学习率很小,这给训练过程中学习率的初始值和中间值的设定造成一定的难度。学习率设定过小,网络收敛较慢,设置过大则会阻碍收敛,甚至引发发散^[15]。对于 Adam 自适应算法而言,尽管没有考虑学习率的问题,但训练容易“点到为止”,即收敛能力有限,这会导致对红外图像的深层特征学习不够深入,仍不是最佳的选择。同时,由于 Adam 算法时刻调整学习率,会导致网络权重产生波动,所以训练时误差曲线不太稳定。

针对红外图像的特点,将 Adam 算法的自适应能力和动量法的稳定高效相结合,并借助于“预训练”思想,提出一种基于 Adam 和 SGD with Momentum 的联合训练方法。使用 Adam 优化算法

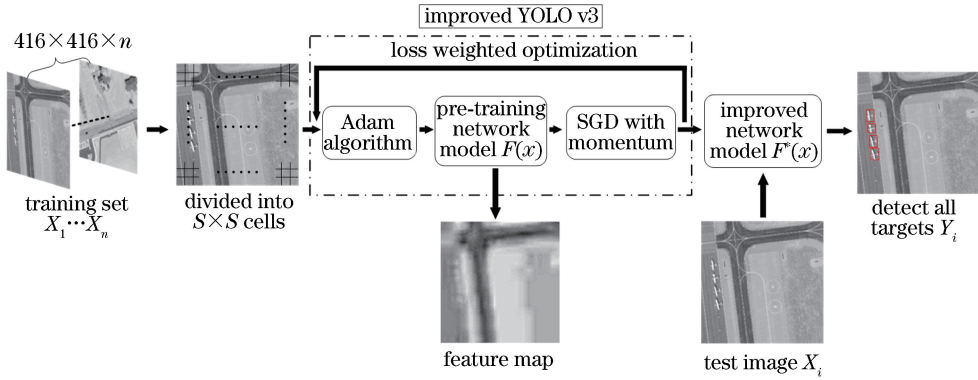


图5 基于改进 YOLO v3 的红外末制导目标检测方法示意图

Fig. 5 Diagram of infrared homing guidance target detection method based on improved YOLO v3

对 YOLO v3 进行预训练,在没有指定学习率的情况下,Adam 算法通过迭代迅速调整神经网络权重,使损失函数迅速收敛,前期 YOLO 网络便能对数据集进行较好的学习。对初步训练的网络权重使用动量法继续训练,由于神经网络已经初步收敛,因此易于设定合适的学习率使已初步收敛的网络逐步趋于稳定,减少损失函数的振荡,最终使网络达到最佳状态。

综上,可以得到基于改进 YOLO v3 的红外末制导目标检测方法示意图,如图 5 所示,图中 $F(x)$ 为使用 Adam 算法后的预训练模型, $F^*(x)$ 为使用联合训练法后得到的网络模型。

3 实验验证

3.1 实验数据及配置环境

将特种车辆、舰船和飞行器作为典型目标进行研究,由于真实红外数据采集较为困难,故拍摄 N 比例模型作为数据来源。为了增强数据可靠性,使用实验室弹上红外系统进行数据采集,并在拍摄过程中不断更换背景,调整目标方向及拍摄角度。接着,对获取的红外图像使用开源工具 LabelImg 进行人工标注,图片标注后,都会对应生成相同文件名且后缀为.xml 的文件,该文件记录了标注框的位置和目标类别等信息。经过筛选和剔除,最终得到 7525 张图片,分辨率为 581 pixel \times 426 pixel。

实验配置环境为:Windows 10 系统,CPU 为 Intel Core i7-7820X,主频 3.6 GHz,GPU 为 NVIDIA GeForce GTX 1080 Ti,内存 32 GB,在 Python 3.5 环境下使用 Keras 框架,并安装 cuda 8.0 加速计算,YOLO v3 网络框架为 DarkNet 53。

3.2 实验测试及结果分析

随机选取 6773 张图片进行训练,676 张图片进

行测试,其余作为交叉验证图片。

1) 聚类优化候选框。使用 K -means 对标注的候选框进行聚类,用交并比 (IOU) 代替欧氏距离作为衡量标准,即

$$d(b_n, b_{center}) = 1 - \text{IOU}(b_n, b_{center}), \quad (10)$$

式中: b_n 表示簇内第 n 个待比较的候选框; b_{center} 表示簇内中心候选框。

画出平均交并比与锚点框的关系,如图 6 所示,通过权衡交并比和模型复杂度,选取 $K=9$,即将候选框聚类分成 9 个簇,分别为 (132, 106), (148, 161), (154, 40), (184, 69), (200, 111), (281, 204), (320, 85), (383, 133), (387, 286)。配合 YOLO v3 多尺度训练,每个尺度分别分配三个锚点框进行训练。

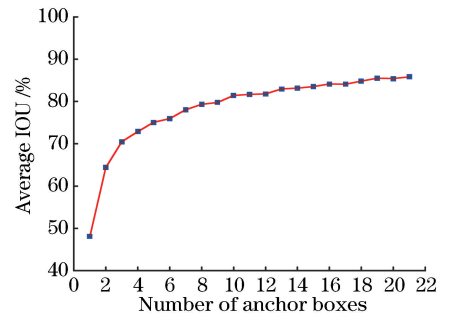


图6 平均交并比与锚点框的关系示意图

Fig. 6 Relationship between IOU and number of anchor box

2) 损失函数加权和联合训练法对比实验。控制 6 组对比实验训练总次数为 200 epoch,选取最优实验结果,实验参数如表 1 所示,图 7、图 8 分别为训练过程的损失曲线和交叉验证损失曲线。

文献[12]中指出,作为自适应算法,在训练时使用 Adam 优化器,选用文献默认参数即可达到较好效果。在固定学习率情况下,过大或过小的学习率都不利于训练,故本文中动量法学习率的设置由离

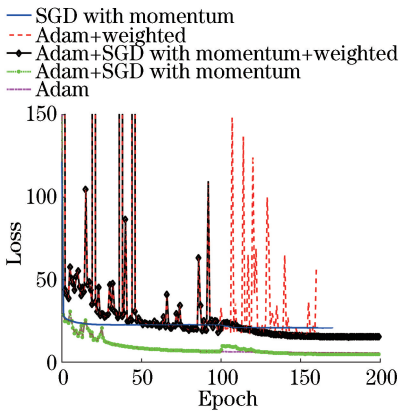


图 7 损失曲线

Fig. 7 Curve of loss

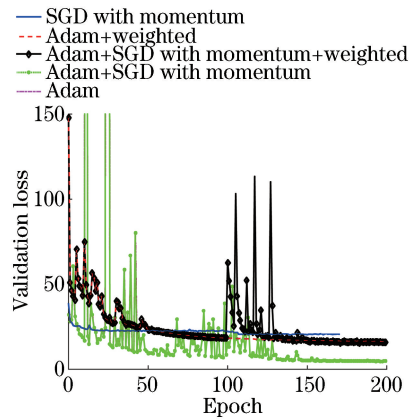


图 8 交叉验证损失曲线

Fig. 8 Curve of cross validation loss

散指数衰减方法得到,使学习率在一定步长范围内保持恒定,且总体呈指数减小。动量法学习率初始值的选取需要通过观察训练损失曲线进行人工调节,设定原则是损失函数能在规定次数内收敛的情况下尽可能大,动量法学习率(L_r)计算公式为

$$L_r = L_{r_0} \times d^{e_{poch}/d - e_{poch}}, \quad (11)$$

式中: L_{r_0} 为学习率初始值; e_{poch} 为当前训练轮数; $d_{e_{poch}}$ 为衰减间隔; d 为衰减系数。

由表1可知,训练相同的次数,动量法的训练时

间普遍略短于 Adam 算法,并且在联合训练法中,因为 Adam 的预优化使得动量法可以使用相对较大的初始学习率,提高了网络训练的潜力。表中:decay 表示衰减系数;decay period 代表衰减间隔。由图 7 和图 8 可看出,当寻找到合适的初始学习率后,动量法较 Adam 算法更为稳定,且能较快收敛;从整体来看,由于学习率不是实时最佳,动量法会有过早收敛的问题,Adam 能够较好地收敛,但收敛程度有限。

表 1 对比实验参数表

Table 1 Parameters for contrastive experiments

Algorithm	Training epoch	Total training time	Parameter
SGD with momentum	1-200	22 h 24 min	$L_{r_0} = 10^{-6}$, decay is 0.5, decay period is 5, momentum is 0.999
Weighted SGD with momentum	1-200	22 h 32 min	$L_{r_0} = 10^{-7}$, decay is 0.5, decay period is 5, momentum is 0.999
Adam	1-200	23 h 20 min	$L_{r_0} = 10^{-3}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$
Weighted Adam	1-200	23 h 28 min	$L_{r_0} = 10^{-3}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$
Adam in united training method	1-100	11 h 40 min	$L_{r_0} = 10^{-3}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$
SGD with momentum in united training method	101-200	11 h 10 min	$L_{r_0} = 10^{-4}$, decay is 0.5, decay period is 10, momentum is 0.999
Adam in weighted united training method	1-100	11 h 44 min	$L_{r_0} = 10^{-3}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$
SGD with momentum in weighted united training method	101-200	11 h 16 min	$L_{r_0} = 10^{-4}$, decay is 0.5, decay period is 10, momentum is 0.999

3)测试及分析。通过设置不同条件进行对比实验,测试结果如表 2 所示。发现所有方法检测速度都能达到 25 frame/s 以上,具有较好的实时性。所有算法都能达到收敛状态,对自制数据集而言,动量法收敛较慢,且检测精度低于自适应算法 Adam,Adam 算法的 mAP 值能够达到 67.55%,加权损失函数能够将其提高至 67.73%,而联合训练法可将其提高至

77.73%。加权损失函数虽然会略微降低对舰船和飞行器的检测精度,但整体 mAP 值可达到 77.89%,改进效果明显。增加虚警率(FA)和漏检率(MA)两个指标,来客观地评价提出的算法性能,两者越小,算法的性能越好。从表中数据看,经过目标函数加权,可以降低虚警率,而联合训练法对漏检率有降低效果,综合所有实验数据,提出的算法达到了预期效果。

表 2 不同方法测试结果对比

Table 2 Comparison of different testing methods

Algorithm	SGD with momentum		Adam		Weighted Adam		United training		Weighted united training	
	100	200	100	200	100	200	100	200	100	200
Epoch	100	200	100	200	100	200	100	200	100	200
mAP / %	—	16.24	32.56	67.55	58.06	67.73	32.56	77.73	58.06	77.89
Special vehicle / %	—	13.05	47.49	63.16	53.32	61.79	47.49	67.36	53.32	69.56
Ship / %	5.30	18.23	21.97	93.54	87.09	86.85	21.97	98.48	87.09	97.97
Aircraft / %	—	17.45	28.21	45.95	33.77	54.54	28.21	67.34	33.77	66.14
MA / %	91.24		62.38		64.20		57.78		57.72	
FA / %	7.27		15.45		8.37		14.29		13.58	
Rate / (frame · s ⁻¹)	26.79	26.88	25.83	25.44	25.62	25.53	25.38	25.27	25.62	25.33

4 结 论

针对红外末制导背景中红外图像局部信息缺失,给传统方法带来极大困难的问题,提出基于改进 YOLO v3 的红外末制导目标检测方法,结合 Adam 算法和动量法的优势,设计联合训练方法,加快网络收敛速度,充分挖掘图像深层信息,同时对 YOLO v3 损失函数模型进行优化加权。通过设置不同条件对比实验进行验证,改进后三类典型目标检测的 mAP 值能够达到 77.89%,虚警率和漏检率都得到降低,改进效果明显。检测速度能够达到 25 frame/s,能较好地满足实时性要求。未来,将从网络模型入手,重点提高检测速度。

参 考 文 献

- [1] Jiang J F. Research on target detection, recognition and tracking technology of infrared image [D]. Xi'an: Northwestern Polytechnical University, 2004. 姜锦锋. 红外图像的目标检测、识别与跟踪技术研究 [D]. 西安: 西北工业大学, 2004.
- [2] Zhang T X, Wang Y H, Zhong S. Guidance information processing methods in aircraft optical imaging seeker [M]. Beijing: National Defense Industry Press, 2014. 张天序, 王岳环, 钟胜. 飞行器光学寻的制导信息处理技术 [M]. 北京: 国防工业出版社, 2014.
- [3] Yao G S, Sun S Y, Fang J A, *et al.* Depth estimation of night driverless vehicle scene based on infrared and radar [J]. Laser & Optoelectronics Progress, 2017, 54(12): 121003. 姚广顺, 孙韶媛, 方建安, 等. 基于红外与雷达的夜间无人车场景深度估计 [J]. 激光与光电子学进展, 2017, 54(12): 121003.
- [4] Yi X, Wang B J. Fast infrared and dim target detection algorithm based on multi-feature [J]. Acta

Photonica Sinica, 2017, 46(6): 0610002.

易翔, 王炳健. 基于多特征的快速红外弱小目标检测算法 [J]. 光子学报, 2017, 46(6): 0610002.

- [5] Xiong B, Huang X H, Wang M. Infrared dim small target detection based on adaptive target image recovery [J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2017, 45(10): 25-30. 熊斌, 黄心汉, 王敏. 基于自适应目标图像恢复的红外弱小目标检测 [J]. 华中科技大学学报(自然科学版), 2017, 45(10): 25-30.
- [6] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
- [7] Sermanet P, Eigen D, Zhang X, *et al.* OverFeat: integrated recognition, localization and detection using convolutional networks [J/OL]. (2014-02-24) [2019-01-06]. <https://arxiv.org/abs/1312.6229>.
- [8] Ren S Q, He K M, Girshick R, *et al.* Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [9] Redmon J, Divvala S, Girshick R, *et al.* You only look once: unified, real-time object detection [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 779-788.
- [10] Redmon J, Farhadi A. YOLO9000: better, faster, stronger [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 6517-6525.
- [11] Redmon J, Farhadi A. YOLOv3: an incremental

- improvement [J/OL]. (2018-04-08) [2019-01-09].
<https://arxiv.org/abs/1804.02767>.
- [12] Wei Y M, Quan J C, Houyu Q Y. Aerial image location of unmanned aerial vehicle based on YOLO v2[J]. Laser & Optoelectronics Progress, 2017, 54(11): 111002.
魏湧明, 全吉成, 侯宇青阳. 基于 YOLO v2 的无人机航拍图像定位研究[J]. 激光与光电子学进展, 2017, 54(11): 111002.
- [13] Ruder S. An overview of gradient descent optimization algorithms[J/OL]. (2017-06-15) [2019-01-05]. <https://arxiv.org/abs/1609.04747>.
- [14] Kingma D, Ba J J C S. Adam: a method for stochastic optimization [J/OL]. (2017-01-30) [2019-01-06]. <https://arxiv.org/abs/1412.6980>.
- [15] Zhang J J. Optimization algorithms of neural networks weights based on stochastic gradient descent [D]. Chongqing: Southwest University, 2018.
张晋晶. 基于随机梯度下降的神经网络权重优化算法[D]. 重庆: 西南大学, 2018.