

基于YOLO-v3模型压缩的卫星图像船只实时检测

陈科峻, 张叶

引用本文:

陈科峻, 张叶. 基于YOLO-v3模型压缩的卫星图像船只实时检测[J]. 液晶与显示, 2020, 35(11): 1168–1176.

CHEN Ke-jun, ZHANG Ye. Real-time ship detection in satellite images based on YOLO-v3 model compression[J]. *Chinese Journal of Liquid Crystals and Displays*, 2020, 35(11): 1168–1176.

在线阅读 View online: <https://doi.org/10.37188/YJYXS20203511.1168>

您可能感兴趣的其他文章

Articles you may be interested in

基于深度学习的航空对地小目标检测

Detection of small target in aerial photography based on deep learning

液晶与显示. 2018, 33(9): 793–800 <https://doi.org/10.3788/YJYXS20183309.0793>

基于YOLO模型的宫颈细胞簇团智能识别方法

Intelligent recognition method of cervical cell cluster based on YOLO model

液晶与显示. 2018, 33(11): 965–971 <https://doi.org/10.3788/YJYXS20183311.0965>

基于改进YOLOv3的棉花异性纤维检测

Detection of foreign fiber in cotton based on improved YOLOv3

液晶与显示. 2020, 35(11): 1195–1203 <https://doi.org/10.37188/YJYXS20203511.1195>

应用图像融合与多样性的舰船显著性检测

Ship-target saliency detection via image fusion and graph-based manifold ranking

液晶与显示. 2016, 31(10): 1006–1015 <https://doi.org/10.3788/YJYXS20163110.1006>

基于YOLO和嵌入式系统的车流量检测

Vehicle flow detection based on YOLO and embedded system

液晶与显示. 2019, 34(6): 613–618 <https://doi.org/10.3788/YJYXS20193406.0613>

文章编号:1007-2780(2020)11-1168-09

基于 YOLO-v3 模型压缩的 卫星图像船只实时检测

陈科峻^{1,2}, 张 叶^{1*}

(1. 中国科学院 长春光学精密机械与物理研究所 应用光学国家重点实验室, 吉林 长春 130033;

2. 中国科学院大学, 北京 100049)

摘要:常见的目标检测模型由于模型参数量较大,往往难以部署在无人机、卫星等移动嵌入式设备上。为了对船只进行实时监测,将目标检测模型部署在计算能力较弱的设备上,对基于计算机视觉的卫星图像船只目标检测方法进行研究。针对卫星图像中船舰的形状长宽比例特点,采用 K-means++ 聚类算法选取初始的锚点框;接着对模型进行多尺度训练,将多尺度金字塔图像作为模型训练的输入;将 YOLO-v3 目标检测算法的批归一化层的尺度因子作为通道重要性的度量指标,对 YOLO-v3 模型进行剪枝压缩。实验结果表明,采用的模型剪枝和压缩方法能有效地对模型进行压缩,模型的参数量减少了 91.5%,模型检测时间缩短了 60%,极大地减少了系统计算性能的开销。当采用的初始锚点框个数为 6 个时,平均准确率(mAP)达到 77.31%,满足了卫星图像船舰实时性检测的需求。

关键词:船只检测;YOLO-v3;聚类算法;模型压缩;通道剪枝

中图分类号:TP391.4 **文献标识码:**A **doi:**10.37188/YJYXS20203511.1168

Real-time ship detection in satellite images based on YOLO-v3 model compression

CHEN Ke-jun^{1,2}, ZHANG Ye^{1*}(1. *State Key Laboratory of Applied Optics, Changchun Institute of Optics,
Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;*2. *University of Chinese Academy of Sciences, Beijing 100049, China*)

Abstract: Due to the large number of model parameters, common target detection models were often difficult to be deployed on mobile embedded platforms such as unmanned aerial vehicle and satellite. In order to detect ships in real time, and for the purpose of deploying target detection model in weak computing equipment, the ship detection algorithm based on computer vision was researched. According to the feature of ship shape length ratio and width ratio in satellite images, K-means++ cluster-

收稿日期:2020-03-25; **修订日期:**2020-06-23.

基金项目:军委科技委国防科技创新特区基金项目(No.18-H863-00-TS-002-018-01); 中国科学院青年创新促进协会项目(No.2016201)

Supported by Special Zone of National Defense Science and Technology Innovation of the CMC(No.18-H863-00-TS-002-018-01); Youth Innovation Promotion Association, CAS (No.2016201)

* 通信联系人, E-mail: yolanda@spirit.ai

ring algorithm was used to select the initial candidate anchor boxes. Multi-scale pyramid images were used as the input of model training. The scale factor of the batch normalization layer of the YOLO-v3 target detection algorithm was taken as the measure index of channel importance, and the YOLO-v3 model was pruned and compressed. Experimental results show that model pruning and compression method can effectively compress the model. The number of parameters of the model size is reduced by 91.5% and the time of model detection is shortened by 60% compared with the original model, which greatly reduces the overhead of system computing performance. When the initial number of candidate boxes is 6, the mAP reaches at 77.31%, which meets the requirements of real-time detecting ship in satellite images.

Key words: ship detection; YOLO-v3; clustering algorithm; model compression; channel pruning

1 引言

随着海上运输业的快速发展,海上遇难船只搜救、船舶避撞、非法捕鱼、非法货运船只的管控等任务,对海上船只精确定位与识别技术有着广泛的需求和重要的现实意义^[1]。因此,从环境保护机构到国家政府当局,许多组织对海上船只进行更密切的监视。随着卫星技术的日益成熟和卫星航空成像精度的日益提高,卫星图像中包含的信息多样,利用卫星图像进行海上船只的检测受到了广泛的关注。然而,视点、旋转、背景等多种因素变化的影响使得卫星图像船只检测任务依然存在严峻的挑战。

目前,国内外研究人员对图像舰船目标检测做了大量的相关研究。文献[2]通过分析遥感图像中的明暗像素的关联参数对船只进行目标检测,实现图像中明暗场景的快速识别与分离。然而基于传统特征提取的方法进行目标检测往往精度不高。由于海洋表面微波散射情况复杂,传统舰船检测算法能力有限。文献[3]针对以上问题,通过分辨率归一化制作混合的训练样本数据集,基于 Faster-RCNN^[4]的框架下,构建了一个 3 层卷积神经网络的舰船检测算法,取得了较好的检测效果。但基于 Faster-RCNN 的舰船检测算法模型复杂度较高,检测速度难以达到实时性检测的要求。近年来,随着基于深度学习的目标检测技术的发展,YOLO(you only look once)^[5]作为一个兼顾了检测速度和检测精度的算法被广泛应用于日常的目标检测任务中。文献[6]提出基于 YOLO 的车流量检测方法,可以较好地混合车流中的车辆进行跟踪和统计。但基于原始

YOLO 算法的目标检测模型参数量较大,对嵌入式设备计算性能要求较高。文献[7]提出了一种基于改进 YOLO 模型的舰船目标识别方法。通过精简原始模型的卷积层,减少模型的参数量。同时在训练过程中引入迁移学习的概念,防止模型过拟合,达到减少计算资源开销的目的。但该方法主要针对大型舰船的目标检测,对于卫星图像中的小目标舰船检测效果不理想。基于端到端回归的 YOLO-v3^[8]网络在 YOLO 算法的基础上进行改进,在保证快速检测的同时,也达到了较高的检测精度,使得基于深度学习的目标检测算法的落地应用成为可能。

本文将 YOLO-v3 目标检测网络应用于卫星图像舰船目标检测任务中。针对输入图像中舰船的外形特点,构建图像金字塔对训练输入图像进行多尺度变换,丰富原始图像的特征。由于 YOLO-v3 的初始锚点框^[9](anchor)的尺寸是针对 VOC^[10]等数据集计算设计得到的,并不适用于文本的卫星舰船图像数据集。本文采用 K-means(K 均值)聚类算法的改进版 K-means++ 算法进行初始锚点框数量和尺寸大小的聚类,构建出适合舰船数据集的舰船目标检测网络结构。最后,为了能够将模型应用于嵌入式设备中,对训练好的模型进行模型剪枝,将模型中冗余的通道进行裁剪,减少模型的参数量,提高了模型的检测速度。

2 图像舰船检测原理

2.1 多尺度金字塔图像

针对舰船的图像特征,本文采用多尺度金字塔图像作为模型训练的输入。本文受到文献[11]

的启发,采用高斯金字塔与拉普拉斯金字塔相结合的方式,先采用高斯金字塔对图像进行下采样以获得新的图像,用拉普拉斯金字塔对图像进行上采样重建获得新样本。

2.1.1 高斯金字塔

对于大小为 $w \times h$ 的图像 I ,高斯金字塔 G_i 由 I 的几个分辨率减小的高斯图像 I_i 组成。 I_i 的尺度计算公式如下:

$$I_i = \left(\frac{w}{2^i}\right) \times \left(\frac{h}{2^i}\right), \quad (1)$$

式中, $i = (0, 1, \dots, j)$ 代表金字塔的级数。

图像 G_i 通过高斯函数生成的卷积核进行卷积。公式如下:

$$G_{L+1}(i, j) = \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) G_L(2i + m, 2j + n), \quad (2)$$

式中, m, n 代表原图像的尺度, L 代表金字塔的层级, w 代表窗口函数。

2.1.2 拉普拉斯金字塔

拉普拉斯金字塔由高斯金字塔计算得来,公式如下:

$$L_i = G_i - \text{up}(G_i) \otimes g_{5 \times 5}, \quad (3)$$

式中, G_i 表示第 i 层的图像,而 up 操作表示上采样。 $\otimes$ 表示卷积运算, $g_{5 \times 5}$ 表示 5×5 的高斯卷积核。最终得到的图像金字塔效果图如图 1 所示。

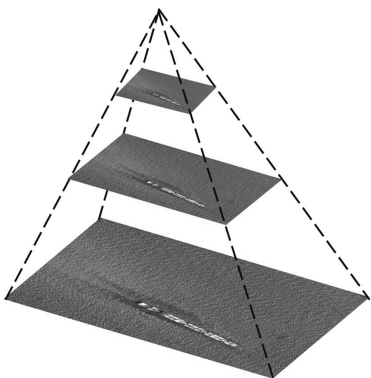


图 1 多尺度图像金字塔

Fig.1 Multi-scale image pyramid

2.2 YOLO-v3 目标检测原理

YOLO-v3 算法将目标区域检测与物体类别预测看作回归问题。YOLO-v3 算法的特征提取网络为 Darknet53。Darknet53 包含 52 个卷积

层,用于对图像进行特征提取,为了防止深层网络训练过程中发生模型退化,YOLO-v3 借鉴了 ResNet^[12] 的做法,在层与层之间进行了链路的连接。最后一层为 1 个全连接层。YOLO-v3 通过单个神经网络,即可实现对图像中目标的定位和分类。Darknet-53 网络结构如图 2 所示。

	类型	卷积信息		特征图大小
1×	卷积	32	3×3	416×416
	卷积	64	3×3/2	208×208
	卷积	32	1×1	208×208
	卷积	64	3×3	
	残差			
	卷积	128	3×3/2	104×104
2×	卷积	64	1×1	104×104
	卷积	128	3×3	
	残差			
	卷积	256	3×3/2	52×52
4×	卷积	128	1×1	52×52
	卷积	256	3×3	
	残差			
	卷积	512	3×3/2	26×26
8×	卷积	256	1×1	26×26
	卷积	512	3×3	
	残差			
	卷积	1 024	3×3/2	13×13
16×	卷积	512	1×1	13×13
	卷积	1 024	3×3	
	残差			
	平均池化			Global
全连接			1 000	
SoftMax				

图 2 Darknet53 网络结构

Fig.2 Structure of Darknet53

Darknet-53 网络对图像进行了 5 次下采样,每次采样的步长为 2。其中卷积模块的结构如图 3。

YOLO-v3 算法假设目标散落在图像中各个区域,因此将图像划分为多个子区域。接着对各个区域预测多个锚点框(Bounding Box),并求

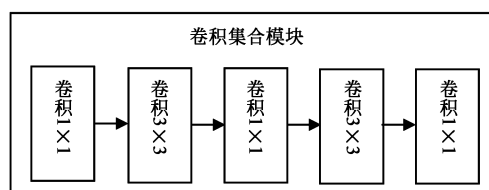


图 3 Darknet53 网络结构

Fig.3 Module of Darknet53 convolution set

出每个边框的置信度(Confidence)。置信度分数的大小,反映了锚点框内存在目标物体的准确程度。置信度定义如下式:

$$\text{Confidence} = \text{Pr}(\text{Object}) \cdot \text{IoU}_{\text{pred}}^{\text{truth}}, \quad (4)$$

式中, $\text{Pr}(\text{Object})$ 表示锚点框中存在目标图像的概率大小。IoU 表示预测的目标框与真实的标注框之间的交并比, IoU 的计算公式如下:

$$\text{IoU} = \frac{\text{area}(C) \cap \text{area}(G)}{\text{area}(C) \cup \text{area}(G)}, \quad (5)$$

式中, C 代表预测出物体的边界框, G 代表实际标注的物体边界框(Ground Truth)。

2.3 初始锚点框聚类算法

常见的目标检测算法中,需要预先对初始的锚点框(Anchor)进行设定。锚点框与检测目标的契合度对目标检测算法的精度产生极大影响。针对卫星图像船舰的形状特征,本文采用聚类算法对 YOLO-v3 的锚点框进行初始化设定。

理想的聚类距离度量标准应该是计算误差与锚点框尺寸大小没有太大的关系,因此定义 IoU 距离作为聚类算法距离的度量。IoU 距离如式(6)所示:

$$d_{\text{IoU}}(\text{box}, \text{centroid}) = 1 - \text{IoU}(\text{box}, \text{centroid}), \quad (6)$$

式中, $\text{IoU}(\text{box}, \text{centroid})$ 表示某一锚点框 box 与聚类中心锚点框的 IoU, d_{IoU} 即代表计算得到的 IoU 距离。

本文采用 K-means++ 聚类方法对 YOLO-v3 的初始锚点框进行聚类。传统的 K-means 算法通过随机初始化的方式选定初始聚类中心,往往容易陷入局部最优,得到的聚类结果无法反映真实的数据分布情况。K-means++ 聚类算法在 K-means 算法基础上进行了改进。具体计算步骤如下:

(1)在数据集中随机初始化选择 1 个点作为第一个簇的中心点;

(2)计算数据集中各个样本点到每个已知的簇中心的距离 $D(x)$;

(3)定义概率 P_x 表示每个样本被选为下一个簇的中心的概率:

$$P_x = \frac{D(x)^2}{\sum_{x \in X} D(x)^2}. \quad (7)$$

(4)重复(2),(3)步骤,直至确定出 K 个聚类中心;

(5)将第(4)步得到的 K 个聚类中心作为初始聚类中心点,代入传统的 K-means 算法中,得到最终的聚类中心点。

从 K-means++ 算法的计算过程可以看出,聚类中心的初始化以距离作为依据,和当前聚类中心距离较大的点,属于当前类的概率则比较小,因此有更大的可能性被选为新的聚类中心,聚类中心的初始化更为合理。

3 基于 YOLO-v3 的模型剪枝

图像船舰检测任务的模型往往被部署在计算能力有限的设备上,对模型的检测时间和模型的参数量大小有着严格的要求。因而本文对原始的模型进行模型剪枝压缩。

模型剪枝是模型压缩的常用方法,对深度神经网络的稠密连接引入稀疏性,通过将“不重要”的权值直接置零来减少非零权值数量。典型的模型剪枝方法包含 4 个步骤^[13]: (1)分析预训练模型中每一个神经元的重要程度; (2)剔除模型中对模型推理激活程度较低的神经元; (3)微调模型,提高修剪后模型的精度; (4)对微调后的模型进行测试,确定修剪后的模型是否满足要求。模型剪枝的流程如图 4 所示。

3.1 稀疏训练

对 YOLO-v3 模型进行压缩,需要对训练好的模型进行稀疏化训练。稀疏训练的目的即在模型训练过程鉴别重要性较低的通道,以便于对重要性较低的通道进行裁剪。由于 YOLO-v3 网络的每个卷积层后面都有一个批归一化层来加速网络收敛和改进泛化性能。本文将批归一化层(Batch Normalization)^[14]中的尺度因子作为通道重要性的指标,并对这些尺度因子计算 $L1$ 范数并融入损失函数中训练,得到每个通道的重要性分数^[15]。应用于精简基于卷积神经网络(CNN)

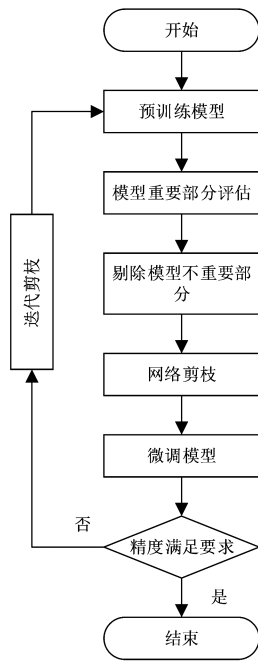


图 4 模型剪枝流程图

Fig.4 Flow chart of model pruning

的图像分类器,可以显著地减小模型参数量。

批归一化层的更新过程如下。求解上一层输出数据的均值和标准差:

$$\mu_{\beta} = \frac{1}{m} \sum_{i=1}^m x_i, \quad (8)$$

$$\sigma_{\beta}^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu_{\beta})^2, \quad (9)$$

式中, m 是本次训练样本的 batch 的大小。归一化处理之后,得到 $\hat{x}_i$:

$$\hat{x}_i = \frac{x_i + \mu_{\beta}}{\sqrt{\sigma_{\beta}^2 + \epsilon}}, \quad (10)$$

式中, ϵ 是为了避免分母为 0 而加入的一个趋近于 0 的值。经过对上面归一化处理得到的数据进行重构即可得到:

$$y_i = \gamma \hat{x}_i + \beta, \quad (11)$$

式中, γ 和 β 是可以学习的参数,用于将归一化后的数据分布进行还原。

本文将 γ 值作为通道重要性的衡量标准,对每一个通道的 γ 值计算 L1 范数,具体计算公式如下:

$$\|\gamma\|_1 = \sum_{i=1}^n |\gamma_i|, \quad (12)$$

最终模型训练的损失函数在 YOLO-v3 的损失函

数基础上增加了对 γ 的正则化项:

$$L = \text{loss}_{\text{yolo}} + \alpha \sum_{\gamma \in \Gamma} \|\gamma\|_1, \quad (13)$$

式中, $\text{loss}_{\text{yolo}}$ 表示 YOLO-v3 的损失函数, α 表示惩罚项。

3.2 模型剪枝

在完成稀疏训练之后,引入一个全局阈值 $\hat{\gamma}$ 来决定是否要裁剪一个特征通道。对通道尺度因子 γ 小于全局阈值的通道进行剪枝。全局阈值设置为 $\|\gamma\|_1$ 的第 n 个百分位数,以控制修剪比例。同时引入局部安全阈值 δ ,防止在卷积层上的过度修剪,以保持网络连接的完整性。局部安全阈值 δ 为特定层中所有通道尺度因子的保留比例。YOLO-v3 中的短接层起到了与 ResNet 中的残差学习类似的作用。因此,所有与短接层连接的层都需要具有相同的通道数。在通道修剪之后对模型进行微调,以提升修剪后模型的检测精度。

4 实验与结果分析

4.1 数据集准备

基于深度学习的目标检测往往需要大量的数据用于模型的训练。为了检验本文方法的有效性,本文采集到网络上的船舰图像数据集^[16],该数据集通过卫星进行图像采集。卫星船舰数据集分辨率大小为 768×768 像素,对图像进行了标注。最终得到精确标注的 5 000 张卫星船舰图像用于模型的训练。部分图像样例如图 5 所示。

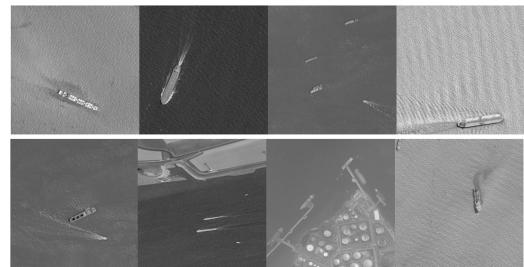


图 5 卫星船舰图像样例

Fig.5 Example of satellite ship images

4.2 锚点框聚类

本文对标注好的数据样本进行聚类,预先选好的锚点框聚类个数为 $k \in (1, \dots, 10)$ 。为了对比不同的聚类算法对模型的影响,在保证各训练

参数固定的情况下,分别采用 K-means 和 K-means++对锚点框进行聚类。表 1 为不同聚类算法下的部分锚点框聚类尺寸。

表 1 初始锚点框聚类结果

Tab.1 Results of initial candidate box clustering

聚类算法	K=5	K=6	K=7
K-means	18,17	15,14	15,14
	29,26	22,20	22,20
	42,44	37,31	35,31
	80,54	80,47	72,44
	158,121	54,91	55,91
		168,109	148,68
K-means++			164,148
		20,19	20,18
	23,22	45,36	38,36
	62,43	53,93	82,47
	60,104	111,55	53,96
	158,72	106,161	166,72
	177,167	215,124	104,157
			225,158

为了评价不同聚类算法的性能,采用平均准确率(mAP)作为评价指标,从整体上反映了目标定位算法的精度。平均准确率综合了检测模型的精确率(Precious)和召回率(Recall):

$$P_{re} = \frac{T_p}{T_p + F_p}, \quad (14)$$

$$R_{ec} = \frac{T_p}{T_p + F_N}, \quad (15)$$

$$mAP = \int_0^1 P_{re}(R_{ec}) d(R_{ec}), \quad (16)$$

式中: T_p 表示真正例, F_p 表示假正例, F_N 表示假负例, P_{re} 表示精确率, R_{ec} 表示召回率。不同个数的锚点框下验证集的平均准确率如图 6 所示。

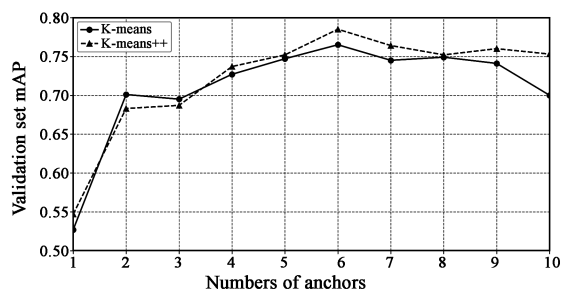


图 6 不同聚类算法结果比较

Fig.6 Results of different clustering algorithms

从图 6 可以看出,采用 K-means++进行锚点框度量可以提高目标检测算法的定位精度。所得到的最高平均准确率高于采用 K-means 进行聚类所得到的最高平均准确率。当锚点框个数为 6 时,所得到的验证集的平均定位精度最高,为 78.54%。相比于原始的 YOLO-v3 采用的 9 个锚点框,更少的锚点框可以提高模型的训练速度,减少模型的推断时间。

4.3 模型训练

对模型进行训练的实验系统环境为 ubuntu16.04,基于 Pytorch 深度学习框架,显卡为 NVIDIA 1660 Ti。本文对模型进行了 2 000 次迭代训练,初始学习率为 0.001,采用动量梯度下降法,设置 momentum 为 0.9。将原始 YOLO-v3 的初始锚点框替换成经过本文聚类之后得到的锚点框。模型训练过程损失曲线如图 7 所示。

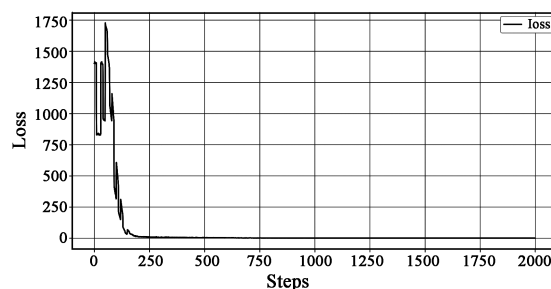


图 7 基础训练损失变化曲线

Fig.7 Loss curve of basic training

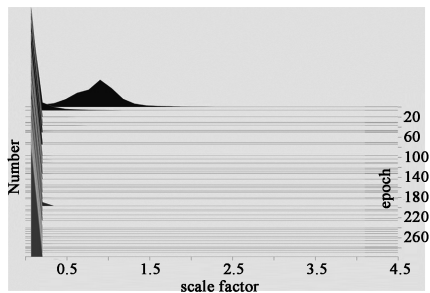
从图 7 可以看出,经过 2 000 步的迭代训练,模型已经趋近于收敛,最终的损失稳定在 0.665。

4.4 模型压缩

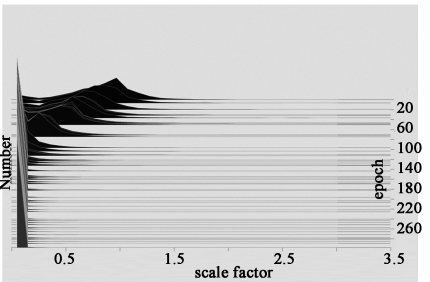
模型剪枝之前需要对模型进行稀疏训练,将批归一化层中的尺度因子作为衡量通道重要性的指标。实验中将稀疏训练的惩罚因子 α 分别设置为 0.01, 0.001, 0.000 1。稀疏训练后的尺度因子分布直方图如图 8 所示。

原始模型的批归一化层因子 γ 近似于均值为 1 的正态分布。从尺度因子分布直方图可以看出,对模型稀疏训练的结果可以使得通道重要性较低的通道 γ 值压缩至接近于 0,因此可以将 γ 值接近于 0 的通道进行裁剪。本文采用不同的惩罚因子对模型进行 300 个 Epoch 的稀疏训练,稀疏训练的损失曲线如图 9 所示。

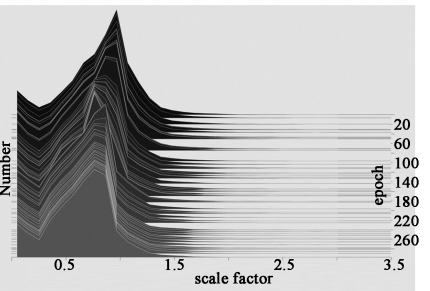
从图 9 可以看到,当稀疏训练的惩罚因子 α 取值为 0.001 时,得到的模型训练损失最小。因



(a) $\alpha = 0.01$



(b) $\alpha = 0.001$



(c) $\alpha = 0.0001$

图 8 不同惩罚因子 BN 层尺度因子分布直方图
Fig.8 Histogram statistics of scaling factors in all BN layers with different penalty factor

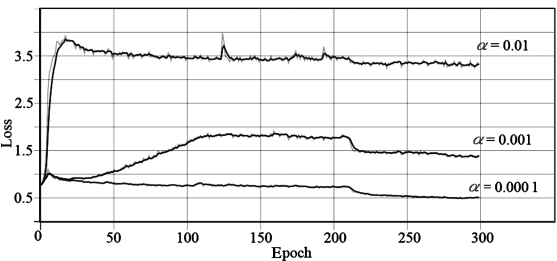


图 9 不同惩罚因子稀疏训练损失曲线
Fig.9 Loss curve of sparsity training with different penalty factor

此,实验对 $\alpha = 0.0001$ 稀疏训练得到的模型进行通道剪枝。设置的全局剪枝阈值 $\hat{\gamma} = 0.7$,每一

层通道的安全阈值 $\delta = 0.01$,即最低保持的通道数比例为 0.01。剪枝前后模型各项参数对比如表 2。

表 2 剪枝前后模型参数对比

Tab. 2 Comparison of model parameters before and after pruning

指标	通道剪枝前	通道剪枝后
mAP/%	78.54	74.41
参数量	62 573 334	5 346 180
单张图像检测时间/ms	30	12

4.5 实验结果分析

通过对批归一化层中尺度因子大小接近于 0 的通道进行裁剪,剪枝后的模型参数量下降为原始模型的 0.15 倍,模型检测的平均准确率下降了 4.13%。为了提高剪枝后模型的精度,采用微调的方式,对剪枝后的模型进行 100 个 Epochs 的微调训练,最终模型的平均检测精度稳定在 77.31%,接近原始模型的精度。

为了验证本文卫星图像船舰检测方法的有效性,本文设置了对比实验,分别采用 YOLO-v3, Faster R-CNN, SSD, R-FCN 等常见目标检测算法对卫星图像船舰数据集进行检测。在同一台服务器上对以上各算法进行测试,测试输入图片大小为 768×768 像素。表 3 为各个算法的检测耗时与平均检测精度对比。

表 3 各算法检测结果对比

Tab.3 Comparison of different algorithm detection result

算法	检测耗时/ (ms/张)	mAP/%
Faster R-CNN(VGG16 ^[17])	167	79.76
SSD ^[18] 300(VGG16)	26	76.93
SSD 512(VGG16)	67	78.97
R-FCN ^[19] (ResNet-101 ^[12])	150	80.1
YOLO-v3	30	78.54
本文方法	12	77.31

从实验结果可以看出,本文模型的 mAP 达到了 77.31%,在与其他检测模型检测精度相当的情况下,检测速度缩短到 12 ms/张,模型参数量

大小仅为原始 YOLO-v3 模型参数量的 8.5%,对计算资源消耗更少,达到了实时性的目标检测要求。

5 结 论

本文针对卫星图像舰船检测问题,通过对训练图像进行多尺度金字塔变换,达到对模型进行多尺度训练的目的,提高了模型的特征提取能力。

同时,采用 K-means++ 聚类算法选定初始锚点框的个数和尺寸,提高了 YOLO-v3 模型的定位精度。接着对 YOLO-v3 目标检测模型进行模型剪枝。实验结果表明,本文的方法在保证检测精度的同时,减小了模型的参数量,检测时间由原来的 30 ms/张缩短至 12 ms/张,模型大小由原来的 238.9 M 缩小至 21.5 M。满足了卫星、无人机等移动小算力设备的实时目标检测需求,为嵌入式目标检测算法的应用提供了参考依据。

参 考 文 献:

- [1] 马啸,邵利民,金鑫,等.舰船目标识别技术研究进展[J].科技导报,2019,37(24):65-78.
MA X, SHAO L M, JIN X, et al. Advances in ship target recognition technology [J]. *Science & Technology Review*, 2019, 37(24): 65-78. (in Chinese)
- [2] 安洁玉,丁斌芬.无人机海监测绘技术应用下舰船遥感图像目标检测[J].舰船科学技术,2019,41(24):187-189.
AN J Y, DING B F. The target detection of the ship's remote sensing image under the UAVs [J]. *Ship Science and Technology*, 2019, 41(24): 187-189. (in Chinese)
- [3] 胡炎,单子力,高峰.基于 Faster-RCNN 和多分辨率 SAR 的海上舰船目标检测[J].无线电工程,2018,48(2):96-100.
HU Y, SHAN Z L, GAO F. Ship detection based on Faster-RCNN and multiresolution SAR [J]. *Radio Engineering*, 2018, 48(2): 96-100. (in Chinese)
- [4] REN S Q, HE K M, GIRSHICK R B, et al. Faster R-CNN: towards real-time object detection with region proposal networks [C]//*Proceedings of Advances in Neural Information Processing Systems* 28. Montreal, Quebec, Canada; NIPS, 2015: 91-99.
- [5] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA: IEEE, 2016: 779-788.
- [6] 马永杰,宋晓凤.基于 YOLO 和嵌入式系统的车流量检测[J].液晶与显示,2019,34(6):613-618.
MA Y J, SONG X F. Vehicle flow detection based on YOLO and embedded system [J]. *Chinese Journal of Liquid Crystals and Displays*, 2019, 34(6): 613-618. (in Chinese)
- [7] 马啸,邵利民,金鑫,等.改进的 YOLO 模型及其在舰船目标识别中的应用[J].电讯技术,2019,59(8):869-874.
MA X, SHAO L M, JIN X, et al. Improved YOLO model and its application in ship target recognition [J]. *Telecommunication Engineering*, 2019, 59(8): 869-874. (in Chinese)
- [8] REDMON J, FARHADI A. Yolov3: an incremental improvement [J]. arXiv: 1804.02767, 2018.
- [9] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]//*Proceedings of 2017 IEEE Computer Vision and Pattern Recognition*. Honolulu, USA: IEEE, 2017: 6517-6525.
- [10] EVERINGHAM M, VAN GOOL L, WILLIAMS C K I, et al. The PASCAL visual object classes (VOC) challenge [J]. *International Journal of Computer Vision*, 2010, 88(2): 303-338.
- [11] 刘梦伦,赵希梅,魏宾.基于多尺度多特征卷积神经网络的肝硬化识别[J/OL].计算机仿真[2020-03-08].<http://kns.cnki.net/kcms/detail/11.3724.TP.20191114.1056.062.html>.
LIU M L, ZHAO X M, WEI B. Cirrhosis recognition based on multi-scale multi-feature convolutional neural network [J/OL]. *Computer Simulation* [2020-03-08]. <http://kns.cnki.net/kcms/detail/11.3724.TP.20191114.1056.062.html>. (in Chinese)
- [12] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]//*Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA: IEEE, 2016: 770-778.

- [13] ZHANG P Y, ZHONG Y X, LI X Q. SlimYOLOv3: narrower, faster and better for real-time UAV applications [C]//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision Workshop*. Seoul, Korea (South): IEEE, 2019: 37-45.
- [14] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]//*Proceedings of the 32nd International Conference on International Conference on Machine Learning*. Lille, France: JMLR, 2015: 448-456.
- [15] LIU Z, LI J G, SHEN Z Q, *et al.* Learning efficient convolutional networks through network slimming [C]//*Proceedings of 2017 IEEE International Conference on Computer Vision*. Venice, Italy: IEEE, 2017: 2755-2763.
- [16] Airbus ship detection challenge [DB/OL]. <https://www.kaggle.com/c/airbus-ship-detection>.
- [17] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C]//*Proceedings of the 3rd International Conference on Learning Representations*. San Diego, USA: ICLR, 2015: 1-14.
- [18] LIU W, ANGUELOV D, ERHAN D, *et al.* SSD: single shot multibox detector [C]//*Proceedings of the 14th European Conference on Computer Vision*. Amsterdam, The Netherlands: Springer, 2016: 21-37.
- [19] DAI J, LI Y, HE K, *et al.* R-FCN: object detection via region-based fully convolutional networks [C]//*Proceedings of the 30th Conference on Neural Information Processing Systems*. Barcelona, Spain: NIPS, 2016: 379-387.

作者简介:



陈科峻(1993—),男,广西南宁人,硕士研究生,2016 年于哈尔滨工业大学获得学士学位,主要从事计算机视觉、模式识别方面的研究。E-mail: ckj409399@sina.com



张 叶(1982—),女,吉林长春人,博士,研究员,主要从事计算机视觉、模式识别算法方面的研究。E-mail: yolanda@spirit.ai