

# 基于半监督学习的遥感飞机图像检测方法

杜泽星\*, 殷进勇, 杨建

江苏自动化研究所计算机事业部, 江苏 连云港 222002

**摘要** 现有基于深度学习的遥感飞机图像检测方法在训练时需要大量的带标记数据集和较长的训练时间, 鉴于此, 提出了一种基于生成式对抗网络的半监督学习方法。采用两种粒度的深度卷积生成式对抗网络, 分别提取待检测目标的边缘特征信息和深层语义特征。通过结合两种粒度的生成式对抗网络的判别器网络模型, 设计了目标检测网络模型。实验结果表明, 实验设计的半监督学习训练方法具有更快的收敛速度, 且在训练时需要的标记样本更少。

**关键词** 图像处理; 半监督学习; 生成式对抗网络; 遥感图像; 目标检测

中图分类号 TP183

文献标志码 A

doi: 10.3788/LOP57.061009

## Remote Sensing Aircraft Image Detection Based on Semi-Supervised Learning

Du Zexing\*, Yin Jinyong, Yang Jian

Computer Division of Jiangsu Automation Research Institution, Lianyungang, Jiangsu 222002, China

**Abstract** Aiming at the existing remote sensing aircraft image detection methods based on deep learning, which require a large number of tagged data sets and a long training time, we propose a semi-supervised learning method based on generative adversarial networks (GANs). Two granularity deep-learning generative adversarial networks are used to get the edge feature and deep semantic feature information. By combining these two discriminator networks of the GANs, we design the object detection model. The experiment shows that the proposed method has a faster training speed and less labeled dataset is needed during the training process.

**Key words** image processing; semi-supervised learning; generative adversarial networks; remote sensing image; object detection

**OCIS codes** 100.3008; 100.4996; 100.2960; 100.5010

## 1 引言

近年来, 遥感图像凭借获取信息速度快、更新周期短、视野范围大、不受地形阻隔等特点, 在军事和民用领域获得了广泛的应用, 对遥感图像的目标检测任务也成为了当前研究的热点。其中, 飞机目标的检测作为战场检测和对地打击的重要环节<sup>[1]</sup>是军事领域研究的重点。因此, 如何快速并准确地在遥感图像中对飞机进行检测具有很高的研究价值。另外, 深度学习已经广泛地应用到目标检测任务中, 与传统的目标检测算法相比拥有更快的检测速度和更高的检测精度, 且不需要过多的人为干预。通过深

度学习实现智能化目标检测, 是未来遥感图像中目标检测的发展趋势。

近年来, 基于深度学习的遥感飞机检测方法成为研究热点。其中朱明明等<sup>[2]</sup>在 Faster-RCNN<sup>[3]</sup>的基础上, 通过特征融合和软判决的非极大值抑制 (NMS) 方法优化了网络对小目标的检测性能, 从而在飞机检测中取得了优异的检测结果。辛鹏等<sup>[1]</sup>将浅层信息和深层信息经过采样后在同一尺寸进行融合, 并且使用卷积网络代替全连接网络设计了一种全卷积网络结构, 提升了神经网络的检测精度和检测速度。郭进祥等<sup>[4]</sup>在 YOLOv3<sup>[5]</sup>网络的基础上, 使用空洞卷积替换卷积层, 并利用线性衰减置信得

收稿日期: 2019-07-23; 修回日期: 2019-08-05; 录用日期: 2019-08-27

\* E-mail: duzexing@outlook.com

分衰减的方式对 NMS 进行优化,从而提升了 YOLOv3 网络对机场场面飞机的检测精度。王俊强等<sup>[6]</sup>使用密集连接网络对 SSD (Single Shot multibox Detector)<sup>[7]</sup>进行了优化,减少了网络参数,提升了检测速度。

深度学习在目标检测领域所取得的进展,一方面是基于深度学习的检测算法的进步,使得卷积神经网络(CNN)可以提取到越来越多的目标特征;另一方面则要归功于出现了大规模带标记信息的数据集,使得训练这些复杂的神经网络成为可能。在数据集较小的情况下,训练过程中很容易出现过拟合的情况,使得神经网络在训练数据集上具有十分优秀的性能,而在测试数据集上却性能很差,即网络的泛化能力很差。对于遥感图像的飞机目标检测而言,受限于遥感图像的高分辨率和飞机目标的密集性,遥感图像的标记工作需要耗费大量的人力和物力,因此难以获得大规模带标记信息的数据集,从而难以训练出泛化能力很强的神经网络。

为解决上述问题,本文提出了一种基于生成式对抗网络(GAN)<sup>[8]</sup>的半监督学习方法优化小样本数据集的训练问题。近年来,国内外研究人员针对半监督学习开展了广泛的研究,文献[9]指出,半监督学习是在含有少量被标记数据的情况下,利用大量未标记图像进行无监督学习从而改善监督学习的性能,并且利用半监督学习的方法,通过大量未标记数据集实现了 X 光图像的分类任务。文献[10]在深度卷积的生成式对抗网络的基础上增加了条件扩

展,将原有的网络模型扩展为条件模型,从而生成了特定条件下的样本。

为优化小样本数据集训练时出现的过拟合问题,同时减小神经网络对大量标记数据集的依赖性,本文设计了一种基于半监督学习的训练方法。在标记数据集较小的情况下,通过粗粒度的 GAN,得到了未标记样本的背景和边缘特征。再利用细粒度的 GAN 模型,得到待检测目标的深层语义特征。利用两种粒度的 GAN 的判别器模型,设计了目标检测网络模型,从而实现了对遥感图像中飞机的检测。实验表明,通过半监督学习方法提升了网络的检测效果,并使网络不再依赖于大量的标记数据集。

## 2 网络结构

为了验证每层卷积神经网络提取到的特征信息,首先使用卷积神经网络对图像进行处理,并且得到每层网络提取到的信息,如图 1 所示。在神经网络的初始阶段,提取得到较为简单的边缘特征信息,随着神经网络的加深,提取得到的特征信息也越来越抽象,这意味着深层网络提取的是更为深刻的语义特征。因此,将网络结构分为粗粒度网络结构和细粒度网络结构,分别用来提取简单的边缘特征信息和复杂的深层语义特征。其中粗粒度网络模型采取大量未标记的图像数据信息进行无监督学习,通过深度卷积的 GAN<sup>[11]</sup>模型得到未标记图像的目标和背景的潜在分布规律,提取边缘特征。在此基础上,设计细粒度的 GAN 模型,进一步提取待检测目标的深层语义特征。

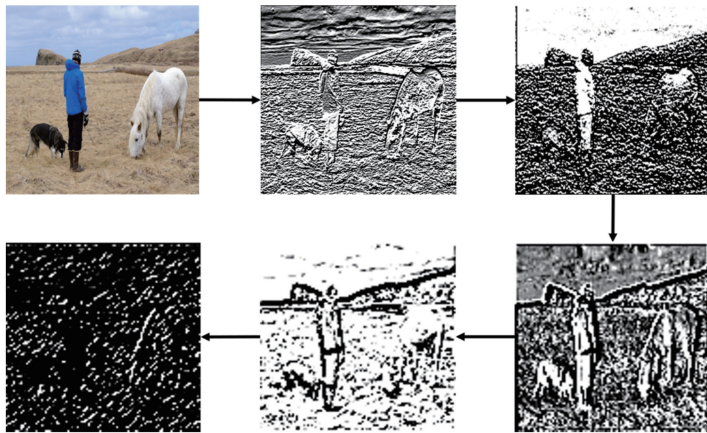


图 1 CNN 结构中每层卷积结构提取到的信息

Fig. 1 Information extracted by the convolutional structure of each layer in CNN structure

### 2.1 粗粒度网络模型

卷积神经网络能够较好地保留图像数据位置信息且具有较少的网络参数,因而在目标检测领域取

得了广泛的应用。本文在深度卷积对抗生成网络(DCGAN)模型的基础上,设计了图 2 所示 G 网络模型和图 3 所示 D 网络模型。

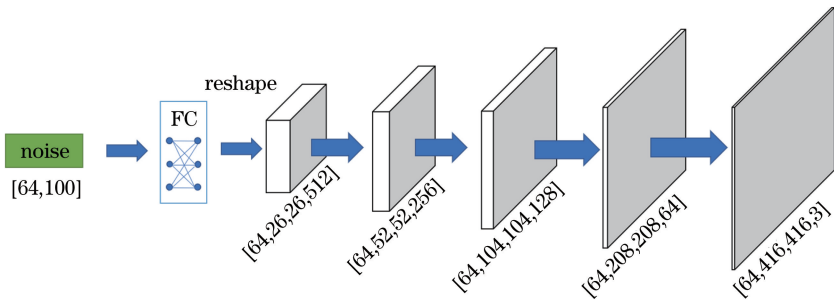


图2 粗粒度网络的生成器网络模型

Fig. 2 Generator network model for coarse-grained network

在G网络中,首先向网络模型中输入100维高斯分布的随机变量 $z$ ,随机变量通过全连接(FC)网络和调整矩阵形状(Reshape)操作后生成4维张量数据,生成的张量数据通过图2所示含有4层反卷积结构的G网络模型结构。其中,反卷积网络采用的卷积核尺寸为 $3 \times 3$ ,步长为2,在每个反卷积层后都带有批量归一化处理(BN)对每一层

输出的数据分布进行优化处理,提高网络的收敛速度和稳定性。在前3层使用Relu激活函数,最后一层使用Tanh激活函数来获得非线性映射。经过反卷积网络多次上采样之后,得到生成的图像 $G_z$ ,由于实验中采用的数据集的分辨率为 $416 \text{ pixel} \times 416 \text{ pixel}$ ,所以用生成器生成相同尺寸的图像。

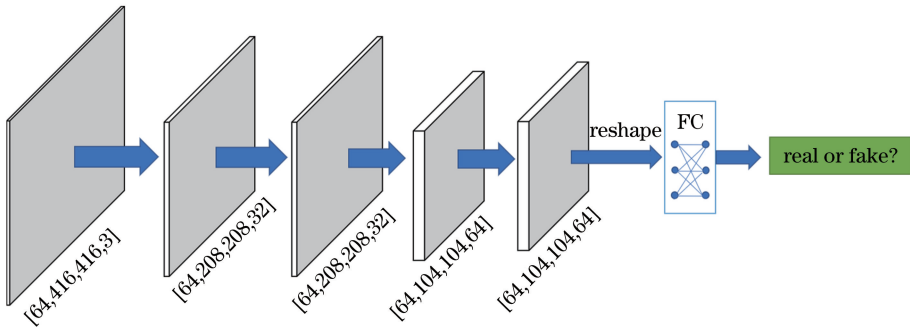


图3 粗粒度网络的判别器网络模型

Fig. 3 Discriminator network model for coarse-grained network

得到G网络生成的图像后,将得到的图像 $G_z$ 和真实的图像输入判别网络模型中,判别网络模型如图3所示。网络将接收的图像输入含有4层卷积结构的D网络中,为了后续和细粒度的网络结构进行级联,将判别器网络的输出尺寸设置为 $104 \times 104$ 。其中,卷积结构采用的卷积核尺寸为 $3 \times 3$ 。整个网络结构中,为了减少池化操作造成的特征损失,实验设计的网络结构均采用步长为2的卷积结构代替池化操作,与G网络相同,卷积结构后采用BN结构加速收敛。此外,使用Leaky Relu激活函数来避免梯度消失情况的出现,将负半轴的斜率设置为0.1。经过多层卷积网络后,将得到的结果输入全连接层和Softmax中得到类别概率,从而判别该图像是否由G网络生成。

## 2.2 细粒度网络模型

为了进一步提取检测目标的深层语义特征,设

计了一个细粒度网络模型,该模型和粗粒度网络模型基本相似,两者之间除了输入输出的尺寸不同外,最大的不同点在于细粒度网络模型不再将整张真实图像输入D网络中,而采用仅包含标记目标的图像,其判别器网络结构模型如图4所示。网络的输入尺寸和粗粒度判别器网络的输出尺寸一样,都是 $104 \times 104$ ,这便于构建之后的目标检测网络。在细粒度网络模型中,卷积核的尺寸为 $3 \times 3$ ,且和粗粒度网络一样使用了BN、激活函数等操作来优化网络结构。

细粒度网络训练过程中的图像来源于带标记信息的图像,通过裁剪获得图5所示的目标图像。经过DCGAN训练后,得到的判别器网络模型可以十分高效地提取待检测目标包含的深层语义信息。

## 2.3 目标检测网络模型

图像经过粗粒度网络模型和细粒度网络模型训

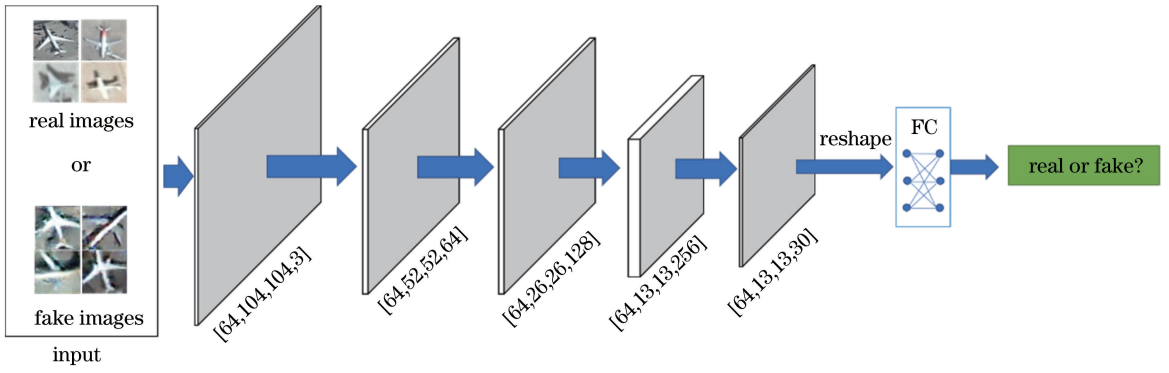


图 4 细粒度网络的判别器网络模型

Fig. 4 Discriminator network model for fine-grained network

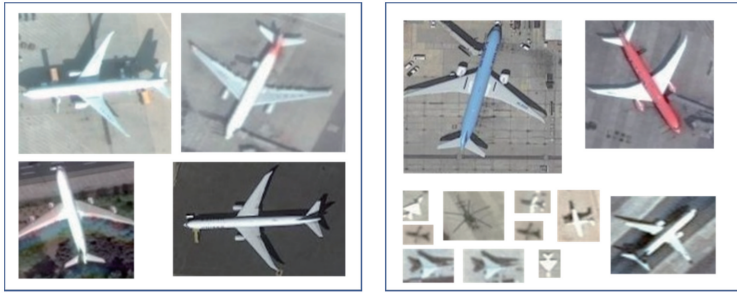


图 5 通过裁剪提取的待检测目标

Fig. 5 Extracted objects to be detected by tailoring

练后,分别提取了未标记数据集中的边缘特征和标记数据集中的目标特征。将这两种粒度的 GAN 判别器网络进行级联,得到图 6 所示目标检测网络。在该网络中,使用 Faster-RCNN 中的 Anchor boxes 设计方法取代原本 D 网络中的 Softmax 层进行目标检测,Anchor boxes 的尺寸由目标尺寸经过聚类

方法获得,分别为  $15 \times 15, 20 \times 10, 10 \times 20, 18 \times 27, 27 \times 18$ 。目标检测网络的训练过程中,采用了和文献[12]相同的坐标处理方法,并没有对目标的边框位置进行直接预测,而是对每个边框中心都预测了其相对于网格(cell)左上角的位置偏移量。训练过程中总的损失函数( $l$ )为

$$l = \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{obj}} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{obj}} [(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2] + \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{obj}} (C_i - \hat{C}_i)^2 + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{noobj}} (C_i - \hat{C}_i)^2 + \sum_{i=0}^{S^2} 1_i^{\text{obj}} \sum_{c \in \text{classes}} [p_i(c) - \hat{p}_i(c)]^2, \quad (1)$$

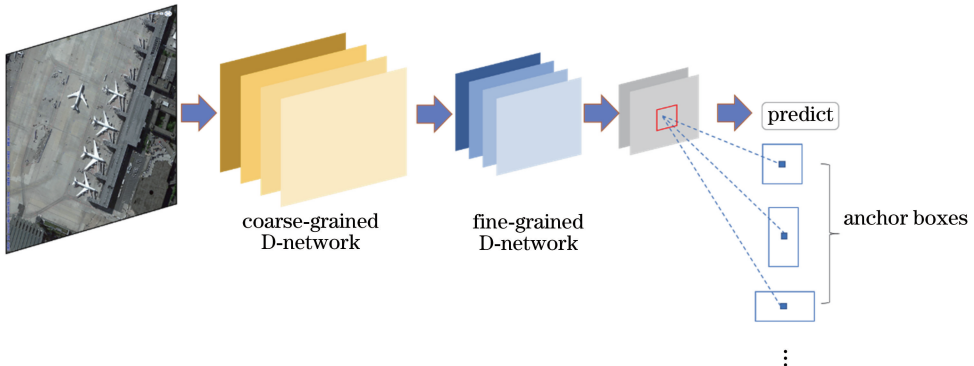


图 6 目标检测网络模型

Fig. 6 Model of object detection network

式中:  $S^2$  和  $B$  分别为网格数量和先验框数量;  $x_i$ ,  $y_i$ ,  $w_i$ ,  $h_i$  和  $\hat{x}_i$ ,  $\hat{y}_i$ ,  $\hat{w}_i$ ,  $\hat{h}_i$  分别为边框位置的预测结果和真实坐标;  $C_i$  和  $\hat{C}_i$  分别为检测结果和真实的置信度;  $p_i$  和  $\hat{p}_i$  分别为预测的和真实的类别;  $\lambda_{\text{coord}}$  为边框的坐标和尺寸损失所占的权重;  $\lambda_{\text{noobj}}$  为不包含目标的置信度损失所占的权重;  $1_{(ij)}^{(\text{obj})}$  用来判断第  $i$  个网格的第  $j$  个先验框是否负责检测目标。

### 3 网络训练

#### 3.1 粗粒度网络训练

训练过程中使用了大量的飞机遥感图像数据集, 这些数据集来自于 Google Earth 中的机场遥感图像和部分自己收集的飞机遥感图像, 共 8000 张, 图像的分辨率为  $416 \text{ pixel} \times 416 \text{ pixel}$ , 部分数据集如图 7 所示。这些数据图像均采集于不同的地点, 采集时光照强度、天气状况和拍摄角度等因素也各不相同, 这增加了训练网络的鲁棒性, 可以适应不同的环境状况。

粗粒度网络训练过程中, 随机选用 100 张数据

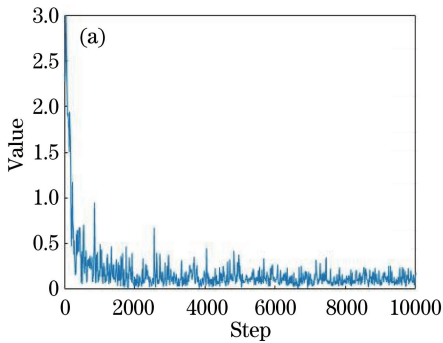


图 7 部分数据集

Fig. 7 Part of dataset

集图像, 采用 Adam 优化方法, 并将学习率设置为 0.0001, 动量设置为 0.9, batch-size 设置为 64。在对抗网络训练过程中, 判别器模型与生成器模型相比更容易取得胜利, 从而导致生成器模型出现梯度消失的情况。为了避免生成器模型出现梯度消失的情况, 设置了 G 网络和 D 网络不同的更新次数, 并经过实验对比, 将判别器网络模型和生成器网络模型的更新次数比例设置为 1:2。经过 10000 轮训练后, 得到 G 网络和 D 网络的损失函数曲线, 如图 8 所示。由图 8 可以看到, 随着训练轮数的增加, 两个网络之间的对抗也呈现出此消彼长的状态。

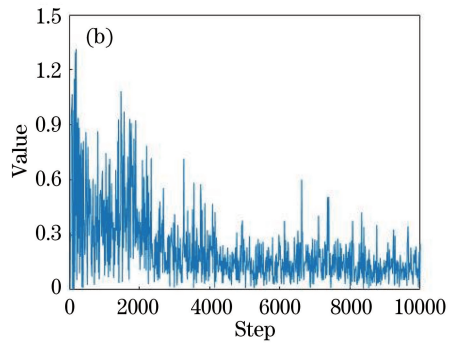


图 8 粗粒度网络中不同模型的损失函数值。(a) 判别器; (b) 生成器

Fig. 8 Loss function values for different models in coarse-grained network. (a) Discriminator network;

(b) generator network

#### 3.2 细粒度网络训练

粗粒度网络模型训练之后, 利用图 5 所示裁剪后的目标图像对细粒度网络模型进行训练。裁剪的图像来源于 100 张带有标记信息的图像, 每张图像大约含有 10 个检测目标, 因此共裁剪了 1000 张左右的检测目标。训练过程中, 该模型和粗粒度网络模型一样, 采用了 Adam 优化方法, batch-size 设置为 64, 学习率设置为 0.0001, 动量设置为 0.9, 并将判别器模型和生成器模型的更新次数比例设置为 1:2。训练过程中, 判别器模型和生成器模型的损失函数曲线如图 9 所示。由图 9 可以看到, 二者的损

失函数值此消彼长, 这是由于两个网络在训练过程中不断对抗所导致, 并且 D 网络的损失值整体出现略微下降的趋势, 可以看到判别器在训练过程中略胜一筹。

经过细粒度的深度卷积 GAN 模型训练后, 由 G 网络生成的部分图像如图 10 所示。由图 10 可以较为明显地看到生成的图像有十分明显的飞机特征, 由此可以得到训练的细粒度网络模型可以非常好地提取待检测目标的特征情况。

#### 3.3 目标检测网络训练

整个数据集中, 除了之前训练 GAN 所用的 100

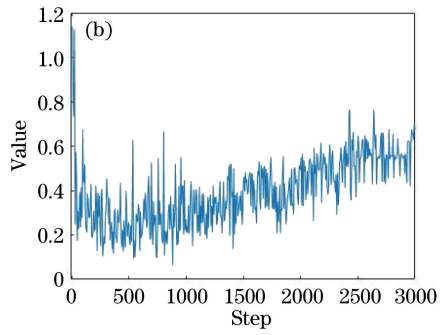
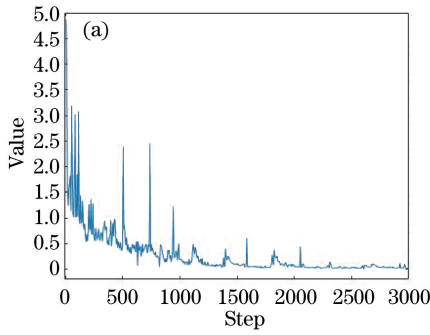


图 9 细粒度网络的损失函数值。(a)判别器;(b)生成器

Fig. 9 Loss function values for different models in fine-grained network. (a) Discriminator network; (b) generator network

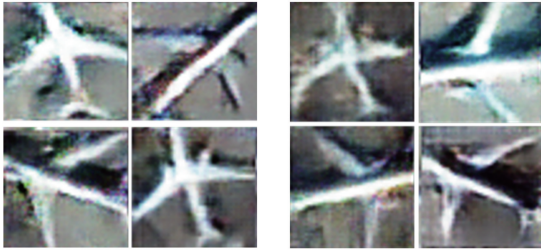


图 10 细粒度网络的生成器产生的飞机图像

Fig. 10 Airplane images produced by fine-grained network

张图像外,再随机挑选出 400 张图像用于监督学习,将挑选出来的图像用开源的数据集标记方法标记目标的边框坐标和类别,再随机选择 100 张图像用于测试,并使用交叉验证方法得到训练结果。当粗粒度网络和细粒度网络经过无监督训练后,得到了两个网络的判别器网络参数。将得到的参数加载到图 6 所示目标检测网络中,用带标记的图像信息输入目标检测网络中进行训练。

训练过程中采用带动量的梯度下降法对(1)式所示损失函数进行优化,学习率采用指数衰减法进行设置,初始值设置为 0.0001,训练过程中 batch-size 设置为 64。训练过程中,通过观察损失函数值来判断网络是否收敛,如图 11 所示。由图 11 可以看到损失函数值经过短短 500 轮训练后便下降到 0.1 左右并基本不再变化,由此可以看到网络的参数

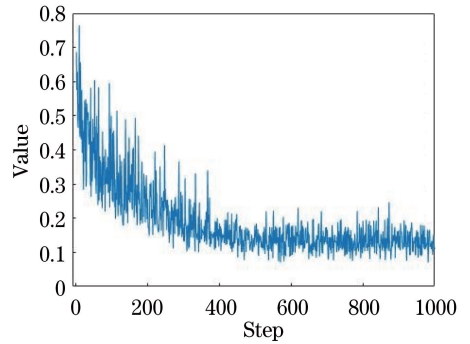


图 11 训练过程中的损失函数值变化曲线

Fig. 11 Change in loss function value during the training process

已经调节到最优。

## 4 结果分析

网络训练时,采用 GPU 为 Tesla P4,操作系统为 Ubuntu16.04 并使用 TensorFlow 深度学习框架。图像经过粗粒度网络模型、细粒度网络模型和目标检测网络模型三个阶段后,利用测试数据集对得到的目标检测模型进行测试,部分检测结果如图 12 所示。为了更好地评价得到的网络模型,采用 mAP 值对模型进行评估,mAP 值是一个综合考虑网络模型查准率( $P$ )和召回率( $R$ )的常用评价指标,代表的是  $P$ - $R$  曲线和坐标轴之间的面积。



图 12 部分检测结果

Fig. 12 Part of detection results

设计了一个实验来评估本文半监督模型的检测精度。如表 1 所示,在相同检测网络模型的前提下,对比了使用 GAN 进行无监督学习和不进行无监督学习两种情况下的神经网络检测精度,两种训练方法采用的都是 500 张带标记的数据集图像。进行无监督学习时,随机选择带标记数据集中的 100 张图像对 GAN 进行训练,除了是否进行无监督学习之外,其余的训练方法都相同。由表 1 和图 13 可以看到,经过粗细两种粒度的 DCGAN 训练后,

表 1 损失函数值随训练轮数的变化

Table 1 Loss function value variation with the number of training steps

Training step	With GAN				Without GAN				
	200	500	1000	2000	500	1000	2000	5000	8000
Loss	0.26	0.14	0.14	0.13	0.50	0.40	0.36	0.19	0.35

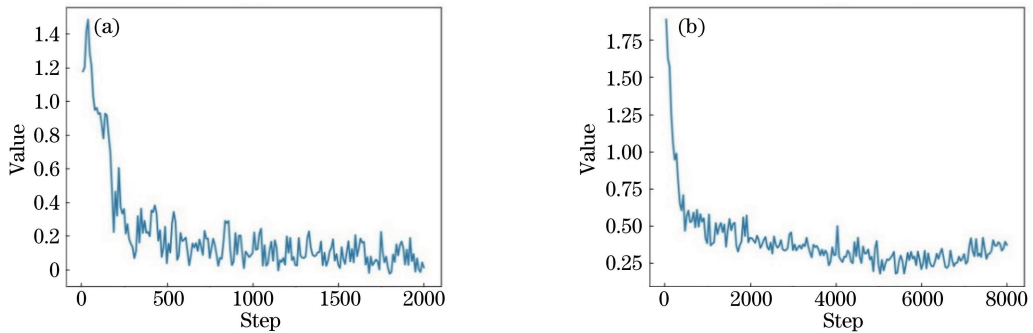


图 13 训练过程中的损失函数值曲线。(a)使用 GAN 进行预训练;(b)未使用 GAN 进行预训练

Fig. 13 Loss function value curves during the training process. (a) With GAN for pretraining;

(b) without GAN for pretraining

除了损失函数值随着训练轮数的增加出现下降的情况外,从图 13 还可以看到经过 GAN 训练后的网络具有更小的损失函数值,未经过 GAN 训练的网络即使训练了较多轮次也难以达到经过 GAN 训练情况下的训练精度。此外,未经过无监督学习的网络在经过较多轮次训练后,测试数据集的损失函数值出现了上升状态,这是由于随着训练轮数的增加,所用训练数据集较小,因此出现了过拟合的情况,而实验设计的半监督学习的训练方则很好地避免了这种情况。

如表 2 所示,还设计了一个实验来验证训练样本的大小对训练结果的影响。在实验中,除了对比是否进行无监督学习两种情况外,还对比了几种主流的目标检测网络的性能。为了保证对比实验的公平性,采用网上开源的模型代码和预训练好的网络模型,在不同样本大小的数据集上进行微调,观察训练过程中的损失函数曲线,当损失函数值下降到一定程度不再下降后,说明训练的网络模型达到了最优。训练过程中,将 batch-size 设置为 64,采用带动

得到的目标检测网络在经过很少的轮次训练后,损失函数值就可下降到较低的范围。这是由于数据集经过粗粒度的 GAN 训练后,可以较为准确和清晰地提取到检测图像的边缘信息,再经过细粒度 GAN 训练后,提取到了检测目标包含的深层语义信息。因此目标检测网络在加载完这两部分参数后,只需经过简单的训练即可达到较低损失函数值,未经过无监督学习的网络则需要训练较长的时间。

量的梯度下降法,学习率采用初始值为 0.0001 的指数衰减法进行设置,多种训练方法采用的实验环境完全相同。

为了对表 2 中的数据有更加直观的认识,将表 2 中的数据绘制成图 14 所示曲线。由图 14 可以看到,实验采用的训练方法仅需要很少的训练数据集便可以达到较高的检测精度,这是由于经过两种粒度的 GAN 无监督训练后,检测网络在一定程度上可以较好地提取检测图像的部分特征,因此只需再经过少量训练数据的微调后便可以达到较高的检测精度。未经过无监督学习的网络,仍然需要大量的数据进行训练。

通过表 2 和图 14 中的对比,可以看到粗粒度网络结构和细粒度网络结构对实验结果的影响。实验采取了控制变量的方法分别去掉粗粒度网络和细粒度网络的预训练部分,并且之后的训练和检测方法保持一致。通过实验结果可以看到,细粒度网络对网络的检测性能有非常大的影响,相比而言粗粒度网络对检测结果的影响相对较小,不过粗粒度网络

表 2 样本大小对检测精度的影响

Table 2 Effect of sample size on detection accuracy

Labeled data	mAP / %					
	SSD	Faster-RCNN	YOLOv3	Without		With GAN
				coarse-grained network	fine-grained network	
100	38.40	35.26	37.47	56.82	35.73	60.80
300	43.85	38.75	42.06	67.10	40.21	70.93
500	48.71	42.92	47.83	75.31	42.80	76.19
1000	57.69	50.38	58.04	75.49	51.82	77.27
2000	68.06	65.73	69.41	75.93	64.73	77.49
3000	76.50	74.51	76.62	76.45	74.06	77.93
5000	78.04	77.28	78.12	77.52	76.71	78.17

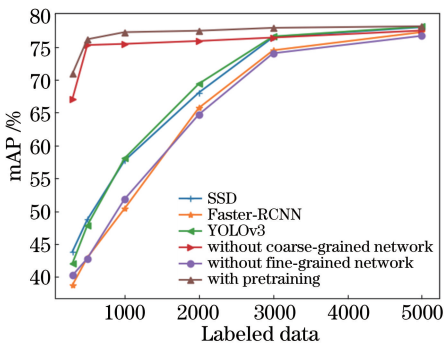


图 14 不同网络模型的检测精度对比

Fig. 14 Comparison of the mAP of different network models

的存在也进一步提升了网络的检测效果。

表 2 中 SSD/YOLOv3 等主流的神经网络模型,虽然也通过 coco 和 VOC 等公开数据集进行了预训练,但是其训练过程中仍然需要大量带标记的数据。这是由于虽然经过了预训练,但其采用的预训练数据集和实验采用的数据集之间相差甚远,预训练的网络模型难以用于实验中数据集的特征提取。当训练中使用的数据集较小时,就会出现过拟合的情况而造成检测精度很低。

实验设计的网络结构除了在训练过程中需要更少的数据集且有更快的收敛速度之外,还在检测阶段具有更短的检测时间。如表 3 所示,对比了几种不同的目标检测算法的检测时间。实验中不同检测算法的检测时间都在相同的实验环境下得到。由表 3 可以得到,实验设计的检测方法具有更快的检测速度,这得益于实验设计的训练方法,使得较少的网络层数即可得到较高的检测准确率,从而缩短了检测时间。

表 3 不同检测方法的检测速度

Table 3 Detection speed of different detection methods

Method	SSD	Faster-RCNN	YOLOv3	Ours
FPS / (frame · s <sup>-1</sup> )	46	9	35	49

## 5 结 论

设计了一种基于半监督学习的神经网络训练方法,在无监督学习过程中,采用粗细两种粒度的 GAN 结构提取检测目标的特征。使用了大量未标记的数据集对粗粒度的 GAN 进行训练,获得了未标记图像的边缘特征。在带标记的图像中截取待检测目标,通过截取的目标对细粒度的 GAN 进行训练,获得了待检测目标的深层语义信息。经过无监督学习之后,结合两种粒度的 GAN 判别器网络模型,设计了目标检测网络模型。通过带标记数据集对网络参数进行微调,得到了最优的网络参数。由实验结果对比可以看到,实验设计的训练方法相比于主流神经网络具有更快的收敛速度和更高的检测精度,且在训练时仅需要少量的数据集图像,不易产生过拟合。

## 参 考 文 献

- [1] Xin P, Xu Y L, Tang H, et al. Fast airplane detection based on multi-layer feature fusion of fully convolutional networks [J]. Acta Optica Sinica, 2018, 38(3): 0315003.  
辛鹏, 许悦雷, 唐红, 等. 全卷积网络多层特征融合的飞机快速检测 [J]. 光学学报, 2018, 38(3): 0315003.
- [2] Zhu M M, Xu Y L, Ma S P, et al. Airplane detection based on feature fusion and soft decision in remote sensing images [J]. Acta Optica Sinica, 2019, 39(2): 0210001.  
朱明明, 许悦雷, 马时平, 等. 基于特征融合与软判决的遥感图像飞机检测 [J]. 光学学报, 2019, 39(2): 0210001.
- [3] 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.
- [4] Guo J X, Liu L B, Xu F, et al. Airport scene aircraft detection method based on YOLO v3[J]. Laser & Optoelectronics Progress, 2019, 56(19): 191003.  
郭进祥, 刘立波, 徐峰, 等. 基于 YOLO v3 的机场场面飞机检测方法[J]. 激光与光电子学进展, 2019, 56(19): 191003.
- [5] Redmon J, Farhadi A. Yolov3: an incremental improvement[EB/OL]. (2018-04-08)[2019-07-22]. <https://arxiv.org/abs/1804.02767>.
- [6] Wang J Q, Li J S, Zhou X W, et al. Improved SSD algorithm and its performance analysis of small target detection in remote sensing images[J]. Acta Optica Sinica, 2019, 39(6): 0628005.  
王俊强, 李建胜, 周学文, 等. 改进的 SSD 算法及其对遥感影像小目标检测性能的分析[J]. 光学学报, 2019, 39(6): 0628005.
- [7] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9905: 21-37.
- [8] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]//Advances in Neural Information Processing Systems, December 8-13, 2014, Montreal, Quebec, Canada. New York: Curran Associates., 2014: 2672-2680.
- [9] Liu K, Wang D, Rong M X. X-ray image classification algorithm based on semi-supervised generative adversarial networks [J]. Acta Optica Sinica, 2019, 39(8): 0810003.  
刘坤, 王典, 荣梦学. 基于半监督生成对抗网络 X 光图像分类算法 [J]. 光学学报, 2019, 39(8): 0810003.
- [10] Tang X L, Du Y M, Liu Y W, et al. Image recognition with conditional deep convolutional generative adversarial networks[J]. Acta Automatica Sinica, 2018, 44(5): 855-864.  
唐贤伦, 杜一铭, 刘雨微, 等. 基于条件深度卷积生成对抗网络的图像识别方法[J]. 自动化学报, 2018, 44(5): 855-864.
- [11] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[EB/OL]. (2016-01-07)[2019-07-22]. <https://arxiv.org/abs/1511.06434>.
- [12] 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.