



(12) 发明专利

(10) 授权公告号 CN 111192240 B

(45) 授权公告日 2023.09.01

(21) 申请号 201911333708.8

(22) 申请日 2019.12.23

(65) 同一申请的已公布的文献号
申请公布号 CN 111192240 A

(43) 申请公布日 2020.05.22

(73) 专利权人 北京航空航天大学
地址 100191 北京市海淀区学院路37号
专利权人 上海航天电子通讯设备研究所

(72) 发明人 史振威 陈科研 邹征夏 马小锋
赵睿 王晓雯 张宁 韩传钊
章泉源 朱新忠 张瑞珏

(74) 专利代理机构 北京慧泉知识产权代理有限公司 11232
专利代理师 李娜 王顺荣

(51) Int.Cl.

G06T 7/00 (2017.01)

G06N 3/0464 (2023.01)

G06N 3/08 (2023.01)

(56) 对比文件

CN 108596101 A, 2018.09.28

CN 108491757 A, 2018.09.04

CN 109325407 A, 2019.02.12

WO 2018214195 A1, 2018.11.29

US 2016104058 A1, 2016.04.14

邓志鹏等. 基于多尺度形变特征卷积网络的高分辨率遥感影像目标检测. 《测绘学报》. 2018, 第47卷(第09期), 第1216-1227页.

审查员 刘梦佳

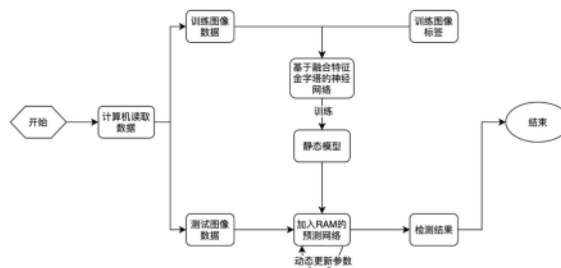
权利要求书3页 说明书8页 附图6页

(54) 发明名称

一种基于随机接入记忆的遥感图像目标检测方法

(57) 摘要

本发明公开一种基于随机接入记忆的遥感图像目标检测方法, 基于SSD网络、最大化后验分布、费希尔信息、拉普拉斯近似等方法, 具体步骤如下: 一、读入图像数据并进行预处理; 二、构造基于融合特征金字塔的卷积神经网络; 三、训练卷积神经网络, 得到静态模型参数; 四、提出随机接入记忆思想; 五、遥感图像目标检测, 动态更新模型参数。本发明通过提出统一框架下的遥感图像多类目标检测算法适应遥感图像目标的检测, 并利用最大后验分布提出随机接入记忆思想来在检测阶段动态调整模型参数, 使模型对新的数据有很好适应能力和泛化能力。输入为遥感图像, 输出为目标的类别和位置信息, 自动化程度高, 大幅度提高效率, 降低成本。



1. 一种基于随机接入记忆的遥感图像目标检测方法,其特征在于:该检测方法通过卷积神经网络实现,分为网络检测部分和预测部分,在预测阶段通过最大后验分布来动态更新模型参数;该方法的具体步骤如下:

步骤一:计算机读取数据;首先读取遥感图像数据;读入数据后,将图片进行数据增广和零均值处理;

步骤二:构造基于融合特征金字塔的卷积神经网络;

基于SSD的网络原型,采用基于特征金字塔的检测方式;移除池化层,使用小卷积和带洞卷积核来提高遥感图像中小目标的召回率;增加可学习的反卷积特征融合层来提高目标检测的精度;设计适应目标尺度的检测先验框来提高先验框的利用率和拟合效果;

具体表达如下:

$$L_{total}(x, c, l, g) = \frac{1}{N} (L_{cls}(x, c) + mL_{loc}(x, l, g))$$

式中 $x_{ij}^p = \{0,1\}$ 表示第i个预设框与第j个参考框匹配上且存在目标的类别为p则为1,否则为0,c为多类别的置信度,l为预测框,g为参考框,N为匹配成功的预设框数目,m用于调整类别预测损失和边框回归损失之间的比例,默认设置为1, $L_{cls}(x, c)$ 是类别预测损失, $L_{loc}(x, l, g)$ 是边框回归损失;

对于类别预测损失 $L_{cls}(x, c)$ 的计算采用Softmax:

$$L_{cls}(x, c) = - \sum_{i \in Pos} x_{ij}^p \log(\hat{c}_i^p) - \sum_{i \in Neg} \log(\hat{c}_i^0)$$

$$\hat{c}_i^p = \frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)}$$

式中 $x_{ij}^p \log(\hat{c}_i^p)$ 表示第i个预测框与第j个参考框关于类别p匹配,则预测为类别p的概率越高,损失越小, $\log(\hat{c}_i^0)$ 表示预测框中没有目标,则预测为背景的概率越高,损失越小, \hat{c}_i^p 通过Softmax产生概率;

对于边框回归损失 $L_{loc}(x, k, g)$ 的计算采用 $smooth_{L1}$:

$$L_{loc}(x, l, g) = \sum_{i \in Pos} \sum_{m \in \{cx, cy, w, h\}} x_{ij}^p smooth_{L1}(l_i^m - \hat{g}_j^m)$$

$$smooth_{L1}(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & |x| \geq 1 \end{cases}$$

步骤三:训练卷积神经网络;如步骤二构造好网络后,在Tensorflow深度学习框架下,利用带标签的训练数据对网络进行训练,直至网络达到最优,并记录此时的网络参数;在训练过程中,利用ImageNet预训练的模型对网络的特征提取器VGG进行参数初始化;使用样本均衡化处理;

步骤四:提出随机接入记忆思想;利用训练数据通过神经网络的训练过程得到一个先验的模型参数分布,然后在检测阶段通过检测数据来修正这个先验分布,得到更适应检测

数据的后验参数分布模型；在模型参数更新时通过最大化后验分布、费希尔信息和拉普拉斯近似来指导参数调整；通过该方式建立检测网络；其具体的公式表达为：

$$\arg \max_{\theta} p(\theta|D_t) = \arg \max_{\theta} \frac{p(D_t|\theta)p(\theta)}{p(D_t)} = \arg \max_{\theta} p(D_t|\theta)p(\theta)$$

$$\min_w M(w|D_t) = -\log p(f(X_t|\theta) = Y_t) + \lambda(w - \hat{w})^T I(\hat{w})(w - \hat{w})$$

其中， D_t 表示测试样本数据空间， θ 为训练完成后得到的模型参数， w 为模型的最后一层参数， \hat{w} 为模型的最后一层的最似然参数，设 $\lambda > 0$ ， λ 是在费希尔信息指导下的正则化项， $f(X_t|\theta)$ 为测试样本网络前馈的输出；

步骤五：遥感图像目标检测；利用步骤三训练好的网络模型和步骤四建立的检测网络对检测数据进行目标检测，输出检测目标的类别信息和位置信息；

步骤四所述的在模型参数更新时通过最大化后验分布，利用训练数据通过神经网络的训练过程得到一个先验的模型参数分布，然后在检测阶段通过检测数据来修正这个先验分布，得到更适应检测数据的后验参数分布模型；在训练阶段假设训练集数据服从独立同分布原理，在训练阶段当成一个极大似然估计的过程；设模型是确定已知的，模型的参数是未知的，由某种方法通过训练数据集 D_s 的数据估计假设模型的参数，选择参数解 $\hat{\theta}$ 来使得在该参数下模型产生 D_s 是最似然的；设训练域数据为 $D_s = \{(x_i, y_i)\}_{i=1}^n$ ，模型为最大似然估计是：

$$\hat{\theta} = \arg \max_{\theta} p(D_s|\theta)$$

如果训练域样本独立同分布，则有：

$$\hat{\theta} = \arg \max_{\theta} \sum_{i=1}^n \log p(y_s^i|x_s^i;\theta)$$

其中 y_s^i 表示源训练数据域的数据标签， x_s^i 表示源训练数据域的样本数据或特征， $\hat{\theta}$ 指深度网络学习到的最优模型参数；引入费希尔信息，费希尔信息的值越大，表示信息量越丰富，越有利于估计；费希尔信息表示为 $I(\theta)$ ：

$$I(\theta) = E(S^2) = -E\left(\frac{\partial^2}{\partial \theta^2} \log p(D_s|\theta)\right)$$

在最大似然解附近将似然函数二阶泰勒展开；设 $X \sim N(\mu, \sigma^2)$ ，则其具体表达为：

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

在最大似然解处把对数似然函数二阶泰勒展开结合上式得出：

$$\theta \sim N\left((D_s|\hat{\theta}), I(\hat{\theta})^{-1}\right)$$

在模型参数更新时通过最大化后验分布的原理，其公式表达如下：

$$\arg \max_{\theta} p(\theta|D_t) = \arg \max_{\theta} \frac{p(D_t|\theta)p(\theta)}{p(D_t)} = \arg \max_{\theta} p(D_t|\theta)p(\theta)$$

只对最后一个卷积层参数进行自适应调整，其他网络层的参数固定不变，在该场景下

转化为数学优化问题：

$$\min_w M(w|D_t) = -\log p(f(X_t|\theta) = Y_t) + \lambda(w - \hat{w})^T I(\hat{w})(w - \hat{w})$$

其中 w 为最后一层参数,设 $\lambda > 0$, λ 是在费希尔信息指导下的正则化项, $f(X_t|\theta)$ 为测试样本网络前馈的输出;结合步骤二中的内容,得到下式对参数进行更新:

$$\min_w M(w|D_t) = E_{(x_t, y_t)} \left(E_{Pos} \left(L_{cls}^{(\pi, w)}(x, c) + mL_{loc}^{(\pi, w)}(x, l, g) \right) \right) + \lambda(w - \hat{w})^T I(\hat{w})(w - \hat{w})$$

上式中 (π, w) 为最后一个卷积层的输入特征和卷积核的参数,该问题是一个凸优化问题,通过Python中的CVXOPT包对其进行求解。

2. 根据权利要求1所述的一种基于随机接入记忆的遥感图像目标检测方法,其特征在于:步骤二所述的增加可学习的反卷积特征融合层,是在网络模型经过连续卷积后,添加多个反卷积层来扩展模型在小尺度上的高层次特征信息;先把得到的特征图做反卷积操作,使得到反卷积后的特征图尺寸与上一层特征图的尺寸大小相同,再进行特征融合,把两个方向上具有相同尺度的特征图叠加后再引出检测头。

3. 根据权利要求1所述的一种基于随机接入记忆的遥感图像目标检测方法,其特征在于:步骤二所述的设计适应目标尺度的检测先验框,具体是先对检测数据域的待检测目标在不同尺度、不同长宽比上的分布进行统计,根据不同检测头对不同尺度目标的信息容量不同、不同尺度目标的检测对于不同的检测头的偏好性,在不同的检测头设置不同宽高比的预设框。

一种基于随机接入记忆的遥感图像目标检测方法

技术领域

[0001] 本发明一种基于随机接入记忆的遥感图像目标检测方法,尤其涉及基于最大后验分布MAP(Maximum A Posteriori)和深度学习中的SSD(Single Shot Multibox Detector)网络的一种高分辨率可见光遥感图像目标检测深度学习的方法,属于高分辨率遥感图像目标检测技术领域。

背景技术

[0002] 遥感(Remote Sensing)科学技术萌芽于17世纪的无记录地面遥感,随着科技的发展和时代的进步,现在已经发展为多种类型传感器、多种搭载平台、多种系统组成、多种应用环境的先进实用的探测技术。遥感图像目标(例如飞机、船只、机场、油罐、港口等)的检测识别技术作为遥感技术重要组成,其发展与遥感技术相辅相成,已经成为了遥感科学领域的研究热点内容。在国家领域,农作物估产和种植区域统计调查、森林火灾的防控、土地管理、环境保护等都离不开遥感图像目标检测技术的发展;在军事领域,检测军用飞机、舰船、机场、港口等重要军事目标对于了解敌方军事情报,监视敌方军港舰船和军机的部署与动态,快速合理的进行信息化作战有着举足轻重的作用和意义;在民用领域,遥感图像目标检测技术在车辆检测和城市交通智能化管理、地面导航系统、资源搜索和勘察等方面有着广阔的应用价值和前景。

[0003] 几十年来,遥感图像目标检测技术主要分为两个大类,传统检测方法和深度学习的方法,传统检测方法主要是基于手工提取的特征构造分类器,而且同一特征提取方法对不同目标有不同的表达能力,所以经常需要根据目标具有的特点设计不同的特征提取方法。数据和特征限制了机器学习到达的高度和上限,而算法和模型只能是无限的逼近这个理想状态,传统多阶段的目标检测算法依赖于提取大量且有效的特征,不足以满足海量遥感数据信息提取的需求。

[0004] 随着计算机并行计算硬件的支撑和大数据时代的来临,基于深度学习的方法在目标检测等计算机视觉领域得到了广泛的应用,在检测准确率上表现出的效果远优于传统方法。尽管如此,但是由于遥感图像有别于自然图像,遥感图像地物复杂,覆盖区域广阔等使得数据在空域上分布广泛,同时遥感图像的生产受天气状况,季节变化,太阳光照角等影响,使得数据在时域上分布多样,并且目前绝大部分遥感图像目标检测算法将目标检测问题看成一个最大似然估计(Maximum Likelihood Estimate,MLE)的过程,得到一种静态模型,当训练过程结束后,模型只会前馈传播检测数据,并不会将产生的损失值反馈传播以更新模型参数。在实际工程化应用时,训练数据的获取往往是在全数据域空间对其非均匀采样,因为生成遥感图像数据的时域空域更为广阔,收集得到的训练数据非均匀分布性更加明显,训练得到的参数模型并不能完全拟合到整个数据空间,所以会使深度网络模型在实际工程化中,表现出来的检测效果时好时坏,严重时模型需要重新进行训练调整,耗时耗力,难以达到工程化应用的要求。

发明内容

[0005] 本发明的目的在于提供一种基于随机接入记忆的遥感图像目标检测方法,该方法是基于最大后验分布和深度学习中的SSD网络的一种高分辨率可见光遥感图像目标检测深度学习的方法,该方法改进了SSD网络并利用最大后验分布在模型预测阶段动态更新模型参数,可以使网络更好地契合遥感图像目标的特点,同时也可以使网络适应新场景目标的分布,从而提升检测效果。

[0006] 本发明是通过以下技术方案实现的:

[0007] 本发明是一种基于随机接入记忆的遥感图像目标检测方法。该检测方法主要通过卷积神经网络实现,可以分为网络检测部分和预测部分,在预测阶段通过最大后验分布来动态更新模型参数。该方法的具体步骤如下:

[0008] 步骤一:计算机读取数据。首先使用计算机读取遥感图像数据;读入数据后,将图片进行数据增广和零均值处理。

[0009] 步骤二:构造基于融合特征金字塔的卷积神经网络。

[0010] 基于SSD的网络原型,采用基于特征金字塔(Pyramidal Feature Hierarchy)的检测方式;移除池化层,使用小卷积和带洞卷积核来提高遥感图像中小目标的召回率;增加可学习的反卷积特征融合层来提高目标检测的精度;设计适应目标尺度的检测先验框(Prior Box)来提高先验框的利用率和拟合效果。

[0011] 其中,所述的基于特征金字塔的检测方式,具体是指同时利用低层特征图和高层特征图的信息,在不同的特征图上进行目标检测,大的特征图来检测小的目标,小的特征图来检测大的目标,可以提高目标的检测率。

[0012] 其中,所述的移除池化层使用小卷积和带洞卷积核,是在保留主要特征的同时,不丢失小目标的特征信息,提高小目标的检测率。

[0013] 其中,所述的增加可学习的反卷积特征融合层,是在网络模型经过连续卷积后,添加多个反卷积层来扩展模型在小尺度上的高层次特征信息;先把得到的特征图做反卷积操作,使得到反卷积后的特征图尺寸与上一层特征图的尺寸大小相同,再进行特征融合,把两个方向上具有相同尺度的特征图叠加后再引出检测头;

[0014] 其中,设计适应目标尺度的检测先验框,具体是先对检测数据域的待检测目标在不同尺度、不同长宽比上的分布进行统计,根据不同检测头对不同尺度目标的信息容量不同、不同尺度目标的检测对于不同的检测头的偏好性,在不同的检测头设置不同宽高比的预设框。

[0015] 步骤三:训练卷积神经网络。如步骤二构造好网络后,在Tensorflow深度学习框架下,利用带标签的训练数据对网络进行训练,直至网络达到最优,并记录此时的网络参数。在训练过程中,利用ImageNet预训练的模型对网络的特征提取器VGG进行参数初始化;使用样本均衡化处理,来应对遥感图像中目标的数量远小于背景负样本的数量,导致模型训练失败。

[0016] 步骤四:提出随机接入记忆思想。利用训练数据通过神经网络的训练过程得到一个先验的模型参数分布,然后在检测阶段通过检测数据来修正这个先验分布,得到更适应检测数据的后验参数分布模型。在模型参数更新时通过最大化后验分布、费希尔信息和拉普拉斯近似来指导参数调整。通过该方式建立检测网络。其中:在模型参数更新时通过最大

化后验分布,具体为:

$$[0017] \quad \arg \max_{\theta} p(\theta|D_t) = \arg \max_{\theta} \frac{p(D_t|\theta)p(\theta)}{p(D_t)} = \arg \max_{\theta} p(D_t|\theta)p(\theta)$$

[0018] 费希尔信息和拉普拉斯近似来指导参数调整,具体为:

$$[0019] \quad \min_w M(w|D_t) = -\log p(f(X_t|\theta) = Y_t) + \lambda(w - \hat{w})^T I(\hat{w})(w - \hat{w})$$

[0020] 其中, D_t 表示测试样本数据空间, θ 为训练完成后得到的模型参数, w 为模型的最后一层参数, \hat{w} 为模型的最后一层的最似然参数,设 $\lambda > 0$, λ 是在费希尔信息指导下的正则化项, $f(X_t|\theta)$ 为测试样本网络前馈的输出。

[0021] 步骤五:遥感图像目标检测。利用步骤三训练好的网络模型和步骤四建立的检测网络对检测数据进行目标检测,输出检测目标的类别信息和位置信息。

[0022] 本发明是一种基于随机接入记忆的遥感图像目标检测方法。本发明的优点是:本发明通过提出统一框架下的遥感图像多类目标检测算法适应遥感图像目标的检测,并利用最大后验分布提出随机接入记忆思想来在检测阶段动态调整模型参数,使得模型对新的数据有很好的适应能力和泛化能力。输入为遥感图像,输出为目标的类别和位置信息,自动化程度高,能够大幅度提高效率,降低成本。

附图说明

[0023] 图1本发明设计的目标检测网络结构图。

[0024] 图2SSD网络原型。

[0025] 图3反卷积特征融合层设计。

[0026] 图4 smooth_{L1} 、L1和L2的函数图像。

[0027] 图5随机接入记忆过程。

[0028] 图6基于随机接入记忆的多类目标检测框架。

[0029] 图7a-f基于随机接入记忆的多类目标检测网络在NWPU-VHR-10数据集上的检测结果。

[0030] 图8本发明方法整体检测流程。

[0031] 表1实验所采用的数据集参数表。

[0032] 表2网络卷积层设置和卷积核的大小。

具体实施方式

[0033] 为了更好地理解本发明的技术方案,以下结合附图对本发明的实施方式作进一步描述:

[0034] 本发明在Tensorflow框架下,使用Python语言编程实现。首先完成网络的搭建,并配置相关的参数;再利用训练数据进行网络的训练,获得调优的网络参数;然后搭建基于最大后验分布的测试模型网络,利用训练好的网络模型和搭建好的检测网络对图像进行目标检测。

[0035] 本发明所基于的SSD网络结构原型如图2所示,其中前面的大立方体代表VGG16的特征提取层,卷积层是对输入数据进行卷积操作,降采样层则是最大池化操作。此外细实线

箭头代表检测头,最后的长方形代表非极大值抑制。计算机配置采用Intel(R)Core(TM) i7-8700K CPU处理器,主频3.70GHz,内存24GB,显卡是NVIDIA GeForce GTX 1080,显存为8GB。如图8所示,本发明的遥感图像目标方法包括如下步骤:

[0036] 步骤一:计算机读取数据。首先使用计算机读取遥感图像数据。本发明使用的数据来自于Levir、NWPU-VHR-10和DOTA三个数据集,来源于Google Earth(谷歌地球)、GF-2(高分二号)和JL-1(吉林一号),分辨率及尺寸等信息见表1。

	数据集	来源	分辨率(m)	图像大小	图像数	类别数	目标数
[0037]	Levir	Google Earth GF-2	0.2-1.0	800×600×3	21952	3	11028
	DOTA	Google Earth	0.11-4.28	387×455×3 6533×3946×3	1591	16	205973
[0038]		GF-2	1	3620×4000×1 4000×5200×1	112	16	3086
		JL-1	0.72	2929×4000×1 4978×4293×1	165	16	1572
	NWPU VHR- 10	Google Earth Vaihingen data set	—	483×776×3 611×944×3	650	10	3896

[0039] 表1

[0040] 上表中,图像大小一栏若为两种大小则说明它们是数据集中图像大小的上下界限定,在数据集中图片大小不是确定的一个值。读入数据后,在训练阶段需要将图片进行数据增广和零均值处理。

[0041] 步骤二:构造基于融合特征金字塔的卷积神经网络。

[0042] 如图1所示,基于SSD网络原型构造了适应遥感小目标检测的多类目标检测网络。SSD算法是Faster RCNN和YOLO的结合,它采用了基于回归的模式达到一步检测的目的,在一个网络中直接回归出物体的类别和位置,因此检测速度很快。同时也利用了区域候选的概念,基于Faster RCNN中的Anchor Box,提出了相似的Prior box,在检测的过程中,使用了许多候选区域作为ROI,在ROI上进行分类和回归。其骨干网络是基于传统的图像分类网络,例如VGG,ResNet等。本发明以VGG16为例进行分析,如图2中所示,经过10个卷积层(conv layer)和3个池化层(max pooling)的处理,可以得到一系列不同尺寸的特征图(feature map)。下一步,在这些特征图中的某些特征图上进行回归,得到物体的位置和类别。

[0043] 图1是本文根据遥感图像特点设计的多类目标检测网络,该网络可以独立完成遥感图像多类目标检测任务,为了在遥感图像上达到优秀的目标检测效果,在检测网络的设计上应用了下述的方法。

[0044] 采用基于特征金字塔(Pyramidal Feature Hierarchy)的检测方式,应对遥感图像中目标尺度变化大的特点。浅层的细节信息对目标检测也有信息补充的作用,同时利用低层特征图和高层特征图的信息,在不同的特征图上进行目标检测,大的特征图来检测小的目标,小的特征图来检测大的目标,可以提高目标的检测率。

[0045] 移除池化层,在池化过程中保留了主要特征,但是会丢失一些小目标的特征信息,

使得小目标的检测率较低;使用小卷积核。用一些尺寸较小的卷积核替代较大的卷积核,不仅可以减少参数数量还可以在保持原感受野范围的同时具备:可以很好的学习到邻域的周边信息,多个小卷积核增加了更多的非线性信息使判决函数效果更好,可以起到隐式正则化的作用;使用带洞卷积核。带洞卷积在不增加网络参数的同时,保留了普通卷积核大感受野的效果,并且还有小卷积核保留特征信息的优点。用上述策略来提高遥感图像中小目标的召回率,其具体的网络卷积层设置和卷积核的大小维数等由表2给出。

Input	512×521 RGB image
Conv1	conv 3×3-64
	conv 3×3-64
Conv2	conv 3×3-128
	conv 3×3-128
Conv3	conv 3×3-256
	conv 3×3-256
	conv 3×3-256
Conv4	conv 3×3-512
	conv 3×3-512
	conv 3×3-512
Conv5	conv 3×3-512
	conv 3×3-512
	conv 3×3-512
Conv6	conv 3×3-1024 rate=6
Conv7	conv 1×1-1024
Conv8	conv 1×1-256
	conv 3×3-512 stride=2
Conv9	conv 1×1-128
	conv 3×3-256 stride=2
Conv10	conv 1×1-128
	conv 3×3-256 stride=2

[0047] 表2

[0048] 增加可学习的反卷积特征融合层来提高目标检测的精度。如图3,低层特征图中包含的细节信息丰富,语义信息较少,目标位置定位精度高;高层特征图中的语义信息丰富,

细节信息较少,目标定位精度低。考虑到上述因素,在网络模型经过连续卷积后,添加了多个反卷积层来扩展模型在小尺度上的高层次特征信息。先把得到的特征图做反卷积操作,使得到反卷积后的特征图尺寸与上一层特征图的尺寸大小相同,再进行特征融合,把两个方向上具有相同尺度的特征图叠加后再引出检测头,这种设计使检测头同时利用不同尺度上的低级特征和高级特征。

[0049] 设计适应目标尺度的检测先验框来提高先验框的利用率和拟合效果。具体来讲,本文对参考窗口的设置利用了一种数据域的先验信息,先对检测数据域的待检测目标在不同尺度,不同长宽比上的分布进行统计,根据不同检测头对不同尺度目标的信息容量不同,不同尺度目标的检测对于不同的检测头也有偏好性,所以在不同的检测头设置不同宽高比的预设框。

[0050] 本检测网络基于VGG16,共设置有10个卷积层,同时利用反向卷积把高层特征与底层特征进行融合,在融合的特征图上引出检测头进行类别预测和边框回归。

[0051] 本文目标检测网络的总体损失函数由类别预测损失和边框回归损失组成,其具体表达如下:

$$[0052] \quad L_{total}(x, c, l, g) = \frac{1}{N} (L_{cls}(x, c) + mL_{loc}(x, l, g))$$

[0053] 式中 $x_{ij}^p = \{0,1\}$ 表示第i个预设框与第j个参考框(Ground Truth)匹配上且存在目标的类别为p则为1,否则为0,c为多类别的置信度,l为预测框,g为参考框,N为匹配成功的预设框数目,m用于调整类别预测损失和边框回归损失之间的比例,默认设置为1, $L_{cls}(x, c)$ 是类别预测损失, $L_{loc}(x, l, g)$ 是边框回归损失。

[0054] 对于类别预测损失 $L_{cls}(x, c)$ 的计算采用Softmax:

$$[0055] \quad L_{cls}(x, c) = - \sum_{i \in Pos} x_{ij}^p \log(\hat{c}_i^p) - \sum_{i \in Neg} \log(\hat{c}_i^0)$$

$$[0056] \quad \hat{c}_i^p = \frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)}$$

[0057] 式中 $x_{ij}^p \log(\hat{c}_i^p)$ 表示第i个预测框与第j个参考框关于类别p匹配,则预测为类p的概率越高,损失越小, $\log(\hat{c}_i^0)$ 表示预测框中没有目标,则预测为背景的概率越高,损失越小, \hat{c}_i^p 通过Softmax产生概率。

[0058] 对于边框回归损失 $L_{loc}(x, l, g)$ 的计算采用 $smooth_{L1}$:

$$[0059] \quad L_{loc}(x, l, g) = \sum_{i \in Pos} \sum_{m \in \{cx, cy, w, h\}} x_{ij}^p smooth_{L1}(l_i^m - \hat{g}_j^m)$$

$$[0060] \quad smooth_{L1}(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & |x| \geq 1 \end{cases}$$

[0061] $smooth_{L1}$ 相比于L1函数而言处处可导,它在0附近收敛缓慢,不至于在最优值附近震荡,图4是 $smooth_{L1}$ 、L1和L2的函数图像。

[0062] 步骤三：训练卷积神经网络。如步骤二构造好网络后，在Tensorflow深度学习框架下，利用带标签的训练数据对网络进行训练，直至网络达到最优，并记录此时的网络参数。在训练过程中，利用ImageNet预训练的模型对网络的特征提取器VGG进行参数初始化；使用样本均衡化处理，来应对遥感图像中目标的数量远小于背景负样本的数量，导致模型训练失败。

[0063] 步骤四：提出随机接入记忆思想。随机接入记忆对思想如图5所示，利用训练数据通过神经网络的训练过程得到一个先验的模型参数分布，然后在检测阶段通过检测数据来修正这个先验分布，得到更适应检测数据的后验参数分布模型。本文中在训练阶段假设训练集数据服从独立同分布原理，在训练阶段当成一个极大似然估计的过程。假设模型是确定已知的，模型的参数是未知的，希望由某种方法通过训练数据集 D_s 的数据估计假设模型的参数，极大似然估计过程可以帮我们选择参数解 $\hat{\theta}$ 来使得在该参数下模型产生 D_s 是最似然的。在本文中设训练域数据为 $D_s = \{(x_i, y_i)\}_{i=1}^n$ ，模型为最大似然估计是：

$$[0064] \quad \hat{\theta} = \underset{\theta}{\operatorname{argmax}} p(D_s|\theta)$$

[0065] 如果训练域样本独立同分布，则有：

$$[0066] \quad \hat{\theta} = \underset{\theta}{\operatorname{argmax}} \sum_{i=1}^n \log p(y_s^i|x_s^i; \theta)$$

[0067] 其中 y_s^i 表示源训练数据域的数据标签， x_s^i 表示源训练数据域的样本数据或特征， $\hat{\theta}$ 指深度网络学习到的最优模型参数。引入费希尔信息，费希尔信息的值越大，表示信息量越丰富，越有利于估计，通俗来讲费希尔信息可以表征观测样本数据能够为深度学习模型参数训练提供的隐含信息容量。费希尔信息表示为 $I(\theta)$ ：

$$[0068] \quad I(\theta) = E(S^2) = -E\left(\frac{\partial^2}{\partial \theta^2} \log p(D_s|\theta)\right)$$

[0069] 在自然界出现的概率问题中，很多分布十分复杂，对其描述较为困难，所以需要将它近似处理为一个已知分布函数的概率分布问题，拉普拉斯近似便是一种方法。在本文中，在最大似然解附近将似然函数二阶泰勒展开。设 $X \sim N(\mu, \sigma^2)$ ，则其具体表达为：

$$[0070] \quad f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

[0071] 在最大似然解处把对数似然函数二阶泰勒展开结合上式可以得出：

$$[0072] \quad \theta \sim N\left((D_s|\hat{\theta}), I(\hat{\theta})^{-1}\right)$$

[0073] 在贝叶斯学习理论中，利用训练数据通过神经网络的训练过程得到一个先验的模型参数分布，然后通过检测数据来修正这个先验分布，得到更适应检测数据的后验参数分布模型。在模型参数更新时通过最大化后验分布的原理，其公式表达如下：

$$[0074] \quad \underset{\theta}{\operatorname{argmax}} p(\theta|D_t) = \underset{\theta}{\operatorname{argmax}} \frac{p(D_t|\theta)p(\theta)}{p(D_t)} = \underset{\theta}{\operatorname{argmax}} p(D_t|\theta)p(\theta)$$

[0075] 在本发明中只对最后一个卷积层参数进行自适应调整，其他网络层的参数固定不

变,在该场景下转化为数学优化问题:

$$[0076] \quad \min_w M(w|D_t) = -\log p(f(X_t|\theta) = Y_t) + \lambda(w - \hat{w})^T I(\hat{w})(w - \hat{w})$$

[0077] 其中 w 为最后一层参数,设 $\lambda > 0$, λ 是在费希尔信息指导下的正则化项, $f(X_t|\theta)$ 为测试样本网络前馈的输出。结合步骤二中的内容,可以得到下式对参数进行更新:

$$[0078] \quad \min_w M(w|D_t) = E_{(x_t, y_t)} \left(E_{Pos} \left(L_{cls}^{(\pi, w)}(x, c) + mL_{loc}^{(\pi, w)}(x, l, g) \right) \right) \\ + \lambda(w - \hat{w})^T I(\hat{w})(w - \hat{w})$$

[0079] 上式中 (π, w) 为最后一个卷积层的输入特征和卷积核的参数,该问题是一个凸优化问题,可以通过Python中的CVXOPT包对其进行求解,引入随机接入记忆后的检测网络模型如图6所示。

[0080] 步骤五:遥感图像目标检测。利用步骤三训练好的网络模型和步骤四建立的检测网络对检测数据进行目标检测,输出检测目标的类别信息和位置信息。为海量遥感数据自动化数据分析和信息提取提供了便利。

[0081] 实验结果:本发明使用的数据集分别为Levir、NWPU-VHR-10和DOTA。Levir数据集由北京航空航天大学图像处理中心的Levir实验室制作,NWPU-VHR-10数据集由西北工业大学标注,DOTA数据集由武汉大学遥感国家重点实验室的夏桂松和华中科技大学电信学院的白翔联合制作。其中Levir数据集中的70%图像用于训练,其余的用于测试,对于NWPU-VHR-10和DOTA,训练集与测试集的划分方式与其发表论文中的算法保持一致。分别在同种数据集、不同种数据集、不同源数据集上测试算法的适应能力,均取得了一定的性能提升,同时还与其他主流算法对比检测精度和时间效率,本算法表现突出。图7a-f为在NWPU-VHR-10数据集上的测试效果。

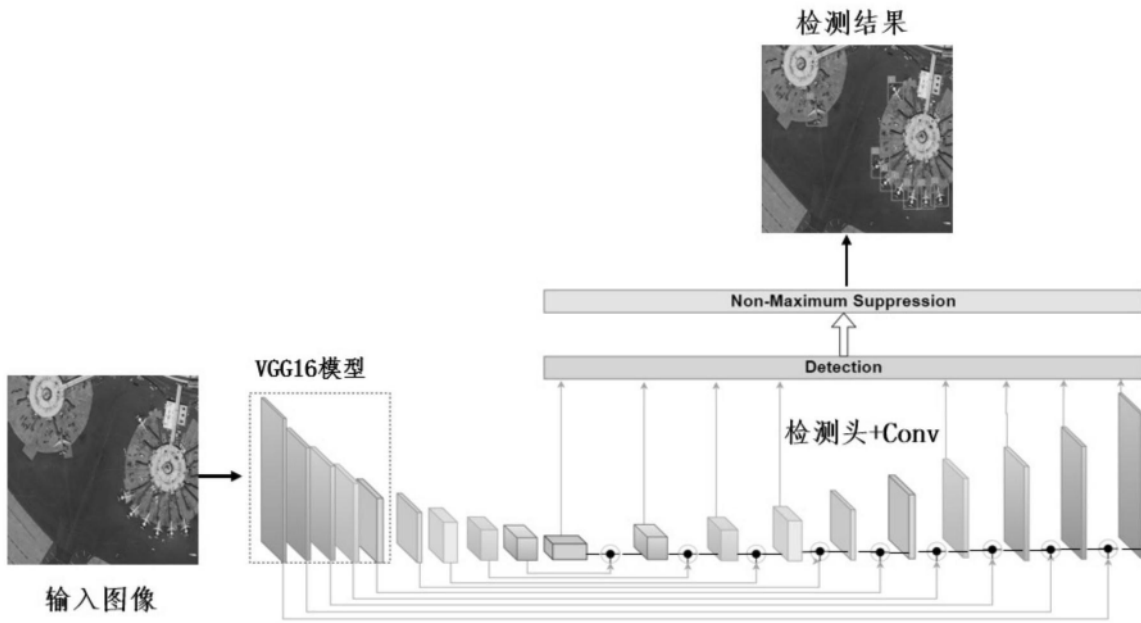


图1

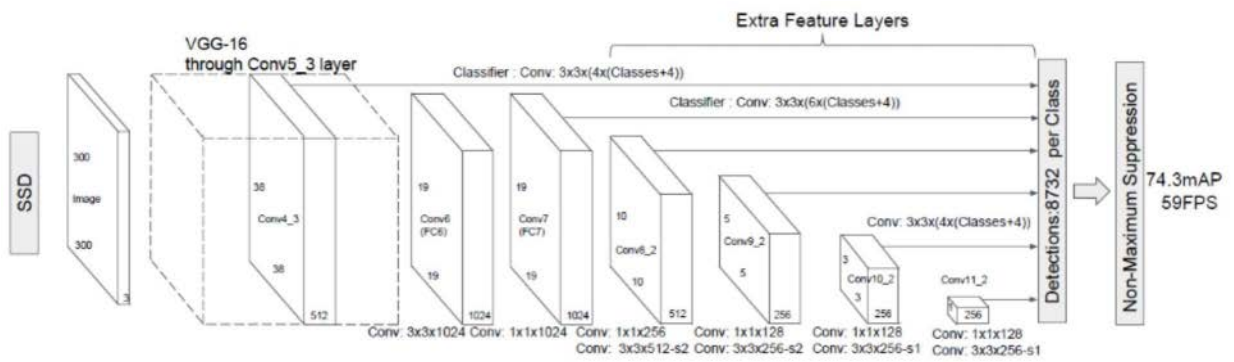


图2

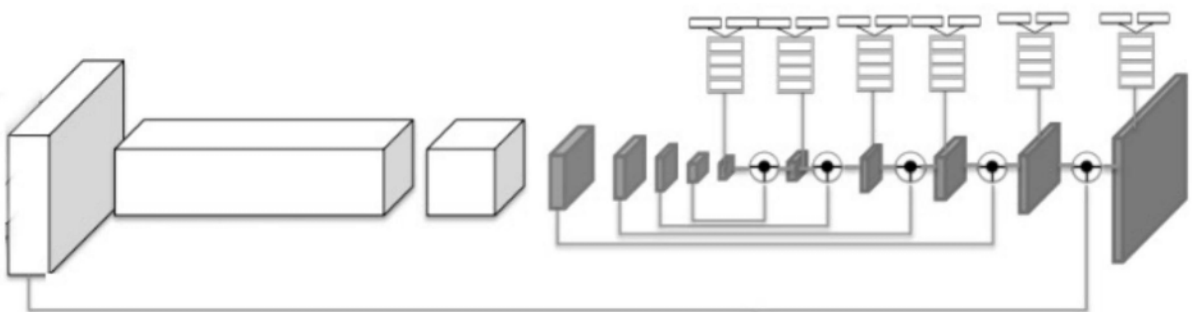


图3

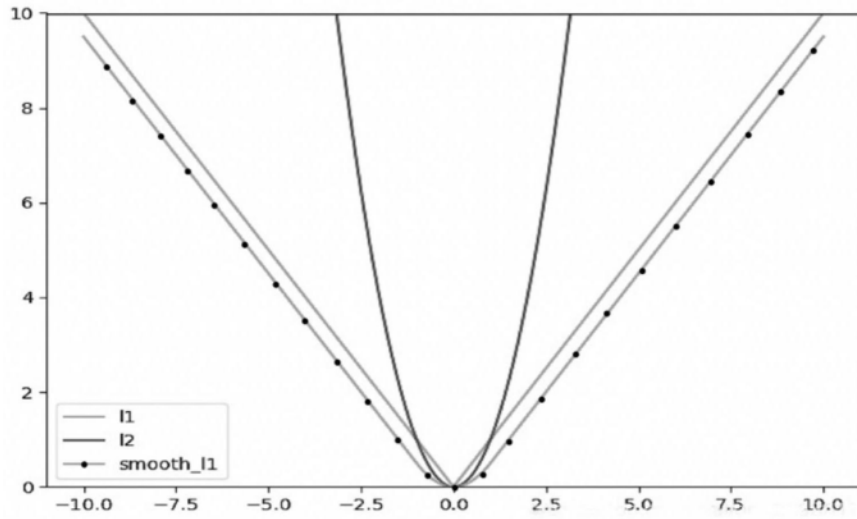


图4

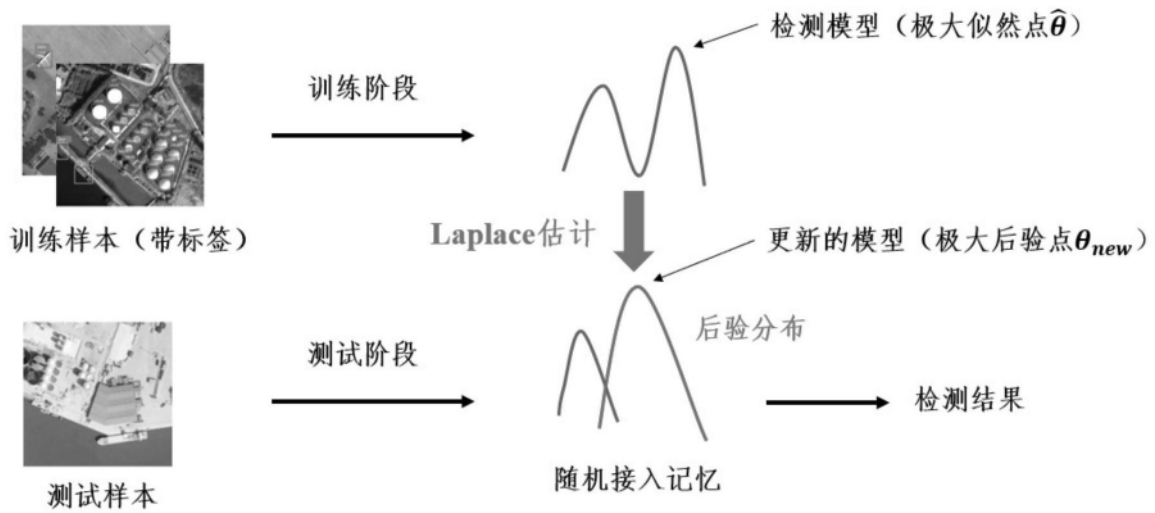


图5

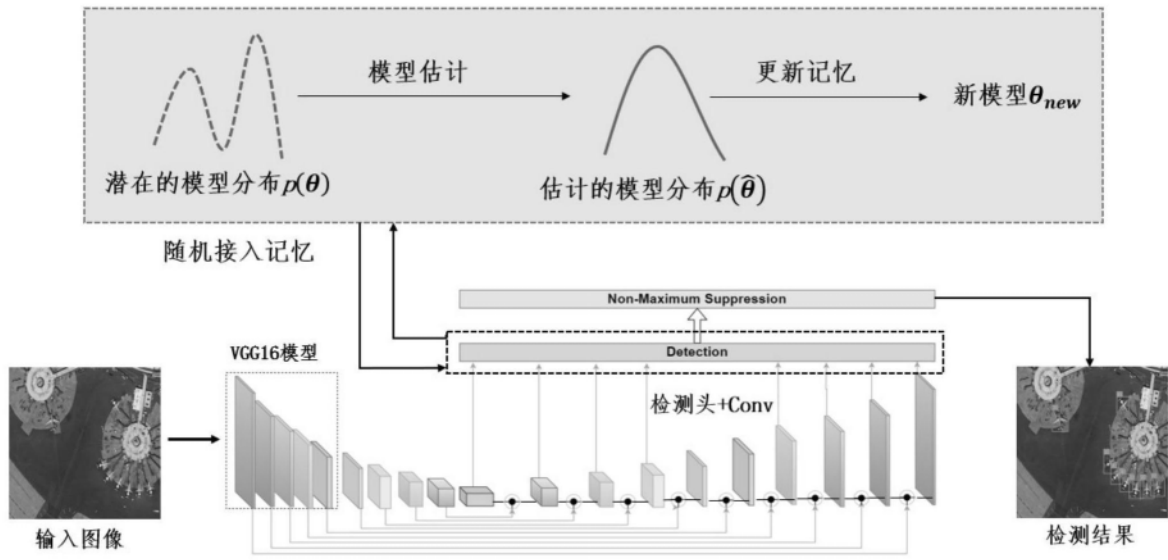


图6



图7a

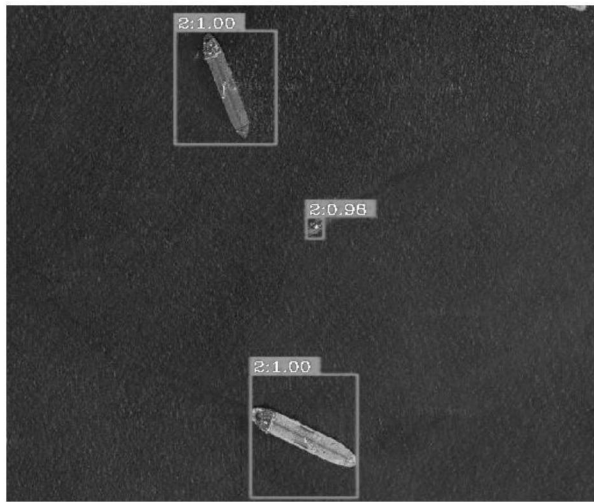


图7b

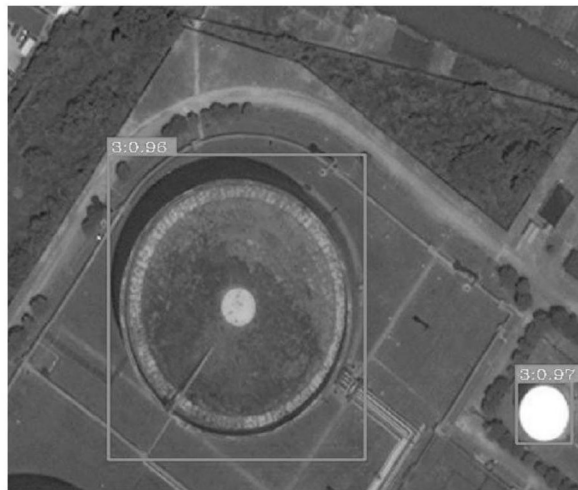


图7c



图7d



图7e

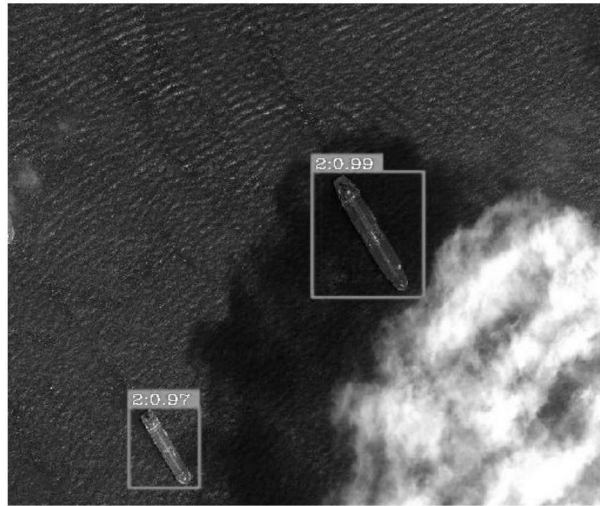


图7f

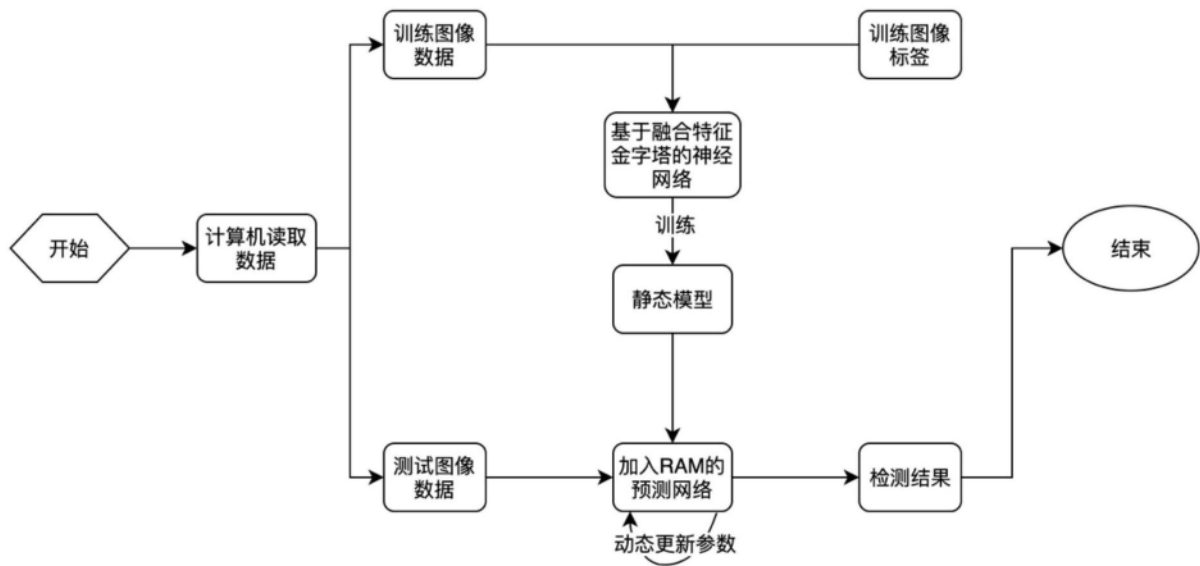


图8