



(12)发明专利申请

(10)申请公布号 CN 108564109 A

(43)申请公布日 2018.09.21

(21)申请号 201810235045.5

(22)申请日 2018.03.21

(71)申请人 天津大学

地址 300072 天津市南开区卫津路92号

(72)发明人 侯春萍 夏晗 杨阳 管岱

莫晓蕾

(74)专利代理机构 天津市北洋有限责任专利代

理事务所 12201

代理人 程毓英

(51) Int. Cl.

G06K 9/62(2006.01)

G06N 3/04(2006.01)

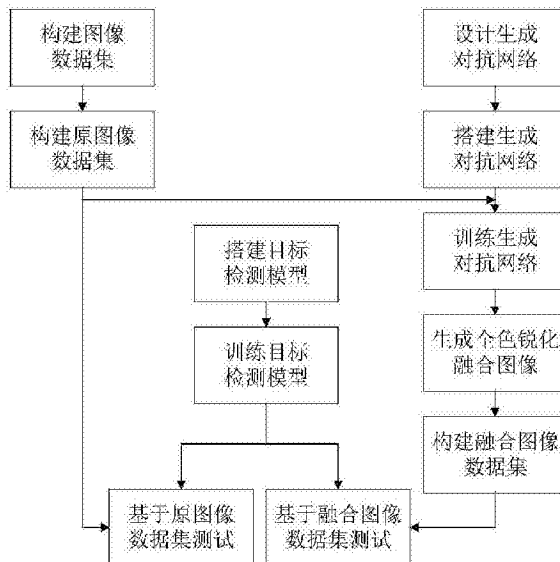
权利要求书1页 说明书6页 附图2页

(54)发明名称

一种基于深度学习的遥感图像目标检测方法

(57)摘要

本发明涉及一种基于深度学习的遥感图像目标检测方法,包括:利用遥感图像构建相关数据集;对遥感图像进行分类和标注后,图像数据集及经过标记工作生成的类别标签;搭建基于生成对抗网络的全色锐化模型;搭建基于深度卷积神经网络的目标检测模型,通过反向传播和随机梯度下降等方法对模型进行端到端训练;对构建好的模型进行端到端测试。本发明具有准确性高的优点。



1. 一种基于深度学习的遥感图像目标检测方法,包括下列步骤:

1) 利用遥感图像构建相关数据集:对遥感图像进行分类和标注后,图像数据集及经过标记工作生成的类别标签,并划分训练集和测试集,用于后续的网络训练和测试;

2) 搭建基于生成对抗网络的全色锐化模型:GAN的生成网络G是可以学习从随机噪声向量 z 和 I 中所述数据集中的图像 x ,到生成的样本图像 y 的映射,即 $G: \{x, z\} \rightarrow y$,生成模型采取增加了跳转连接的U-Net结构,分为编码层和解码层两部分,每编码一层,特征图长和宽减半,特征层数增加一半,每解码一层,特征图的长和宽加倍,特征层数增加加倍,和对应的编码层,通过通道串接,然后进行反卷积处理;基于用于分类的卷积神经网络CNN,设计判别网络模型,该网络被设计为含有一个串接层和四层卷积层;

3) 搭建基于深度卷积神经网络的目标检测模型:按照目标检测算法的候选区域生成,特征提取,分类,位置精修的四个步骤,将上述步骤统一到一个深度网络框架之内,在GPU内并行运算,特征提取以残差网络ResNet作为基础分类网络,其中包含若干卷积层和线性单元ReLU,设计区域生成网络结构,在提取好的特征图上,对所有可能的候选框进行判别,通过共享卷积,减少计算建议框的边际成本,通过反向传播和随机梯度下降方法对模型进行端到端训练;

4) 对构建好的模型进行端到端测试:基于步骤1中构建好的数据集,训练目标检测模型并测试模型。

一种基于深度学习的遥感图像目标检测方法

技术领域

[0001] 本发明涉及遥感图像处理、深度学习、模式识别等领域,尤其是涉及一种基于深度学习,使用生成对抗网络对光谱图像进行全色锐化处理和目标检测的方法。

背景技术

[0002] 由于信号传输波段和成像传感器存储的限制,大多数遥感卫星仅提供具有高光谱分辨率的多光谱(MSI)图像,和高空间分辨率的全色(PAN)图像。利用两种图像的优势互补,融合成具有清晰空间细节和丰富的光谱信息的融合遥感图像,这种融合技术,也被称为全色锐化技术。

[0003] 目前,遥感领域主流的全色锐化方法有分量替换法、多尺度分析法等。分量替换法主要通过主成分分析,施密特正交化,和强度、色调、饱和度变换等方法,在颜色空间域上对光谱图像进行变换,用全色图像替换多光谱图像的空间信息通道,由逆变换得到融合图像。

[0004] 多尺度分析法,指的是基于小波变换、拉普拉斯金字塔和多尺度几何分析等途径,将源图像运用多分辨率分析工具分解为一序列分解系数,然后将这些分解系数通过某种融合准则合并成融合图像的分解系数,最后通过多分辨率分析工具逆变换获得融合图像。

[0005] 近年来,基于大规模数据的出现和深度神经网络的发展,深度学习方法成为了机器学习领域的重要研究方向。基于深度学习技术和博弈论的思想,能够由低维特征生成高维样本的生成对抗网络(Generative Adversarial Networks, GAN)可以被引入用于全色锐化过程,GAN由生成网络模型和判别网络模型所构成。生成模型可以帮助生成相关样本数据,而判别模型可以判断样本的真实度,两者同时训练,生成模型则不断加强,通过不断迭代,使生成样本越来越接近真实样本。

[0006] 作为模式识别在遥感领域的一个重要应用,基于遥感图像下的多种类多尺度目标检测与识别是地理勘测、军事侦察和精准打击等领域的一项关键技术,如何提高目标检测的精度,也一直是遥感应用领域的研究热点和难点,有着重要的军事和民用价值。随着高分辨率遥感技术的快速发展,得以搭建大规模高分辨率遥感影像数据集,为开发更加智能的遥感影像目标检测系统提供了可能,从海量数据中提取目标有效特征就成为遥感图像应用的关键技术。

[0007] 传统检测算法几乎均是在某种给定的特征基础上,完成分类和检测工作。提取的特征及检测模型作为决定检测效果的两大重要因素,对模型起到至关重要的作用。这就要求对输入的特征有着严格的要求,并且找到匹配该特征的检测模型。

[0008] 然而上述要求无疑是复杂而耗时的,并且强烈依赖于专业知识和数据本身的特征,此外,很难从大规模数据中学习出一个有效的分类模型,来充分挖掘大规模数据之间的相互关联。

[0009] 随着深度学习方法的发展,使得原始数据作为输入,实现端到端(End-to-End)的学习过程成为可能。深层人工神经网络有很强的特征学习能力,深度学习模型学习得到的特征数据对原数据有更本质的代表性,通过大规模数据训练好的基于深度学习技术的目标

检测模型更能提取其丰富的内在信息,有利于可视化和分类问题处理。因此基于卷积神经网络,可以设计一种能够自动学习特征的方法,通过对大量数据本身的学习,获取其中最有效的深层特征,并通过建立相对复杂的网络结构,充分挖掘数据之间的关联。

发明内容

[0010] 本发明的目的是提供一种基于深度学习的遥感图像目标检测方法,本发明应用于全色锐化的生成对抗网络模型,能够实现遥感图像中信息含量的扩充;应用于目标检测的深度卷积神经网络模型,准确性更高,实时性更好,鲁棒性更强。为实现上述发明目的,本发明采用如下的技术方案:

[0011] 一种基于深度学习的遥感图像目标检测方法,包括下列步骤:

[0012] 1) 利用遥感图像构建相关数据集:对遥感图像进行分类和标注后,图像数据集及经过标记工作生成的类别标签,并划分训练集和测试集,用于后续的网络训练和测试;

[0013] 2) 搭建基于生成对抗网络的全色锐化模型:GAN的生成网络G是可以学习从随机噪声向量 z 和 I 中所述数据集中的图像 x ,到生成的样本图像 y 的映射,即 $G: \{x, z\} \rightarrow y$,生成模型采取增加了跳转连接的U-Net结构,分为编码层和解码层两部分,每编码一层,特征图长和宽减半,特征层数增加一半,每解码一层,特征图的长和宽加倍,特征层数增加加倍,和对应的编码层,通过通道串接,然后进行反卷积处理;基于用于分类的卷积神经网络CNN,设计判别网络模型,该网络被设计为含有一个串接层,和四层卷积层;

[0014] 3) 搭建基于深度卷积神经网络的目标检测模型:按照目标检测算法的候选区域生成,特征提取,分类,位置精修的四个步骤,将上述步骤统一到一个深度网络框架之内,在GPU内并行运算,特征提取以残差网络ResNet作为基础分类网络,其中包含若干卷积层和线性单元ReLU,设计区域生成网络结构,在提取好的特征图上,对所有可能的候选框进行判别,通过共享卷积,减少计算建议框的边际成本,通过反向传播和随机梯度下降方法对模型进行端到端训练;

[0015] 4) 对构建好的模型进行端到端测试:基于步骤1中构建好的数据集,训练目标检测模型并测试模型。

[0016] 与现有的技术相比,本发明的提升和优势在于:

[0017] 一、与所有现有的遥感图像目标检测方法的思路不同,本发明创新性地提出基于深度学习方法的级联的先进进行全色锐化处理再进行目标检测的方法。进行全色锐化的过程中,GAN能够利用深度卷积神经网络提取大规模数据中隐含的高维深层特征,其结构也能够最大程度地减少卷积过程的信息损失。进行全色锐化后的遥感图像,具有清晰的空间细节和丰富的的光谱信息,具有相对于全色图像和光谱图像的高空间分辨率和高光谱分辨率的特性,能够提高数据集中遥感数据的基础利用信息的丰度,空间分辨率的提升对于小目标的检测更具有实际应用意义。通过训练好的模型直接对图像进行端到端的检测工作,其更为高效,时间和计算冗余度更低。

[0018] 二、与现有的遥感图像目标检测的传统方法不同,本发明创新性地提出基于深度卷积神经网络的目标检测网络模型。相对于HOG特征等手工特征提取,特征可以直接用数据中经过卷积神经网络后得到的特征图表示,由于卷积操作具有平移不变形,特征图中不仅包含了物体的类别信息,还包含着物体的位置信息,所以特征的分类结果和位置回归具有

更好的准确性和更强的普适性。本发明采用区域生成网络,区域推荐也放在网络里完成,从特征提取到最后检测的全过程都在一个网络中完成,速度提升更高,同时解决了拟合相关问题。

附图说明

[0019] 图1为本发明所需实验的流程图。

[0020] 图2为全色锐化的生成网络结构示意图。

[0021] 图3为遥感图像的全色锐化效果图,(a)全色遥感图像数据;(b)光谱遥感图像数据;(c)融合遥感图像数据;(d)本发明融合遥感图像数据。

[0022] 图4为区域生成网络结构图。

具体实施方式

[0023] 为使本发明的技术方案更加清楚,下面对本发明具体实施方式做进一步地描述。如图1所示,本发明按以下步骤具体实现:

[0024] 1. 构建大规模遥感图像数据集

[0025] 本发明选用网络公开的SpaceNet on AWS、NWPU VHR-10、美国地质勘探局USGS等遥感图像集进行检测任务的数据集构建。

[0026] NWPU VHR-10数据集是一个公开可用的十种类地理空间物体检测数据集。这十类物品是飞机、船舶、储油罐、港口和桥梁等,包含高分辨率图像和图中目标及其标注的标签文件。

[0027] SpaceNet是托管于Amazon公司AWS云服务平台的大规模遥感图像数据集,为DigitalGlobe、CosmiQ Works以及NVIDIA共同完成,其包含卫星图像的在线存储库和已经标注好的训练数据,是公开发布的高分辨率、专用于训练机器学习算法的卫星图像数据平台。除此之外,本发明也结合了中科院地理空间数据云平台、美国地质勘探局(USGS)和谷歌公司的相关遥感数据来搭建训练和测试所需的数据集。

[0028] 将上述数据集中的图像数据按4:1的比例分为了训练集和测试集。本发明对其进行分类和标注后,本发明按照PASCAL VOC挑战赛的格式制作图像数据集及标签,用于后续的网络训练和测试。

[0029] 2. 搭建基于生成对抗网络的全色锐化模型

[0030] 基于1中构建的相关遥感数据库,搭建和训练用于遥感图像全色锐化的生成对抗神经网络,该步骤是为后续的检测提供具有高的空间分辨率和光谱分辨率的遥感图像数据。用于全色锐化的生成对抗网络,由生成网络和判别网络两个网络构成,生成网络和判别网络通常由包含卷积和(或)全连接层的多层网络构成。通过对效果优秀的网络结构进行多次实验测试,本发明构建以U-Net网络为基础的卷积神经网络作为生成网络。

[0031] 使用全卷积结构的U-Net架构搭建生成网络,并搭建不同大小感受野的判别网络架构。通过使用U-Net网络中的卷积核来实现下采样,不仅能够减少操作的冗余度,并且在一定程度上还能提取目标的抽象特征;用多种不同的卷积核对图像进行卷积操作,可以得到不同核上的响应,作为图像的特征。输入的图像矩阵经过卷积核(kernel)卷积运算之后得到的一个新的图像矩阵,即特征图(feature map)。

[0032] 在此后连接的单元能够保持特征图尺度不变,此外,将网络中的池化层代替为特征图尺度不变的卷积层;删除网络中的全连接层,用反卷积层来实现图像的上采样,这里能够将浅层卷积层与深层卷积层输出的特征进行处理,提高特征提取的准确度。

[0033] 在如上所述生成网络中输入随机的噪声信号 z 向量和数据库中的全色图像 x ,将生成网络生成的图像数据 y 作为判别网络的输入,即 $G: \{x, z\} \rightarrow y$,经过训练,生成样本不能被判别网络模型判别为假。而判别网络模型 D ,经过训练,能够尽可能好地完成判别生成样本的分类问题。

[0034] GAN的训练目标可用如下损失函数公式表示,其中, x 为输入的现有图像, y 为输出的样本图像, z 为随机噪声向量:

$$[0035] \quad L_{cGAN}(G, D) = E_{x,y} [\log D(x, y)] + E_{x,z} [\log (1 - D(x, G(x, z)))]$$

[0036] 生成模型和判别模型都是采用卷积层-批规范化-线性单元的结构,GAN来处理图像中高频的结构信息等细节部分,在训练的过程中,生成模型来使该目标最小化,而判别模型使其最大化,即

$$[0037] \quad G^* = \operatorname{argmin}_G \max_D L_{cGAN}(G, D)$$

[0038] 在最终样本图像的生成过程,输入的全色图像和输出的融合图像具有相同的底层结构,共享突出边缘的位置。为了使生成模型获取该信息,生成模型增加了跳转连接,采用U-Net的整体结构。

[0039] U-Net是一种全卷积结构,它在传统的编码器-解码器架构的基础上,于编码模块与解码模块的对应层(具有同样大小的特征图的层)之间加入了跳跃链接。

[0040] U-Net网络分为编码层(共八层),解码层(共八层)两部分,每编码一层,特征图(feature map)长和宽减半,特征层数增加一半,每解码一层,特征图长和宽加倍,特征层数增加加倍,即还和对应的编码层,通过通道串接,然后进行反卷积处理。

[0041] 对输入图像的四周做了镜像操作,卷积层的数量设计在20个,4次下采样,4次上采样。具体地,对于 n 层网络,本发明在每一个第 i 层和第 $n-i$ 层之间添加跳转连接,把第 i 层和第 $n-i$ 层中的所有通道相连接。

[0042] 基于用于分类的卷积神经网络(CNN),设计判别网络模型。判别网络的训练次序在生成网络之前,判别模型实际上是可为生成模型充当其损失函数,因此判别器要比生成器训练地更加充分从而为生成器的收敛提供正确的目标。该网络被设计为含有一个串接层,和四层卷积层。减少参数设计的CNN,只是对生成的融合图像中的每个区块的真假性做分类,在图像上卷积运行该网络,对所有响应做均值,来提供 D 的最终输出。

[0043] 3. 搭建基于深度卷积神经网络的目标检测模型。基于深度学习,本发明通过构建具有多隐层的神经网络模型,能实现从大规模训练数据中学习更有用的特征,从而最终提升分类或预测的准确性。为实现遥感目标检测的高精度和高适应性,在这里,本发明采用基础特征提取网络+区域生成网络+分类网络的目标检测网络结构来进行检测网络模型的构建。

[0044] 本发明设计的基于深度网络的检测网络算法的实现步骤如下:

[0045] (1) 输入经全色锐化后的遥感图像;

[0046] (2) 将整张图片输入卷积神经网络,进行特征提取;

[0047] (3) 用RPN生成建议窗口,每张图片生成300个建议窗口;

[0048] (4) 把建议窗口映射到卷积神经网络的最后一层卷积特征图上;

[0049] (5) 通过池化层使每个感兴趣区域生成固定尺寸的特征图;

[0050] (6) 利用探测分类概率和探测边框回归对分类概率和位置回归进行联合训练。

[0051] 其中,特征提取网络采用残差网络结构(ResNet),通过残差网络,实现网络结构的深化和分类效果的显著提升。残差网络相比传统的卷积神经网络如VGG复杂度降低,需要的参数下降可以做到更深,不会出现梯度弥散的问题。深度卷积残差网络是去学习输入到(输出一输入)的映射,由此获得输出由输入部分组成的先验信息。

[0052] 首先构建一个18层和一个34层的残差网络,在简易网络上插入捷径,能够大大地减轻计算量。通过在输出个输入之间引入一个捷径连接,而不是传统方法上简单的堆叠网络,来解决深层网络出现梯度消失的问题。

[0053] 区域生成网络(Region Proposal Network,RPN),能够在ResNet提取好的特征图上,对当前相对稀疏的所有可能的候选框进行判别。利用SPP-Net的映射机制,区域生成网络根据一一对应的点从卷积层映射回原图,根据设计不同的固定初始尺度,来训练网络,根据与参考标准的准确覆盖程度,给其正负标签,令其学习里面是否有目标物体。

[0054] 为了降低区域生成网络的计算复杂度,基于深层网络,可以实现共享卷积计算结果,固定尺度变化、比例尺变化和采样方式,而后得到目标候选区域,即特征的候选窗口。首先按照尺度和长宽比生成9种候选窗口,在卷积的最后一层特征图上使用固定大小的窗口滑动,每个窗口会输出固定大小维度的特征,每一个窗口对候选的9个目标进行回归坐标和分类。

[0055] 区域生成网络的目标函数是分类和回归损失的和。分类采用交叉熵,回归采用稳定的Smooth L1,其公式可表示为:

$$[0056] \quad \text{Smooth}_{L1}(x) = \begin{cases} 0.5x^2 & |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases}$$

[0057] 整体损失函数具体为:

$$[0058] \quad L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*)$$

[0059] 损失函数分为两部分,对应着区域生成网络的两条支路,即目标与否的分类误差和检测框的回归误差,其中 $L_{reg}(t_i, t_i^*)$ 采用平滑L1函数,其比L2形式的误差更容易调节学习率。对于检测框的致信,只考虑判定为有目标的候选窗口,并将其标注的坐标作为致信的目标。此外,计算检测框误差时,不是比较四个角的坐标,而是 t_x, t_y, t_w, t_h ,如下所述,具体四个维度的计算方式:

$$[0060] \quad t_x = (x - x_a) / w_a, t_y = (y - y_a) / h_a,$$

$$[0061] \quad t_w = \log(w / w_a), t_h = \log(h / h_a),$$

$$[0062] \quad t_x^* = (x^* - x_a) / w_a, t_y^* = (y^* - y_a) / h_a,$$

$$[0063] \quad t_w^* = \log(w^* / w_a), t_h^* = \log(h^* / h_a),$$

[0064] 在测试时,感兴趣区域(ROI)池化层从区域生成网络得到候选的ROI列表,通过卷积层拿到所有的特征,进行后面的分类和回归。通过区域生成网络和检测网络公用产生建议窗口的卷积层,能实现生成候选和检测之间的共享。

[0065] 上述网络的训练过程采用四步训练法, 第一步, 单独训练区域生成网络, 网络参数由预训练模型载入; 第二步, 单独训练检测网络, 将第一步区域生成网络的输出候选区域作为检测网络的输入。区域生成网络输出一个候选框, 通过候选框截取原图像, 并将截取后的图像通过几次卷积-池化操作, 再通过ROI池化输出两条支路, 分别是目标分类的探测分类概率 (Softmax Loss) 和探测边框回归 (Smooth L1 Loss)。第三步, 再次训练区域生成网络, 此时固定网络公共部分的参数, 只更新区域生成网络独有部分的参数; 最后, 依据区域生成网络的结果再次微调检测网络结构, 固定公共部分的参数, 只更新检测网络框架独有部分的参数。

[0066] 4. 级联网络, 进行端到端测试。输入数据库测试集中的全色遥感图像, 根据训练好的生成对抗网络对其进行全色锐化处理。然后输入到构建好的检测网络模型中, 对其检测结果进行评价。对于数据库中遥感图像, 在发明内容1构建的测试数据集上, 分别根据融合前后的光谱图像构建基于PASCAL VOC竞赛格式的数据集, 验证了进行全色锐化后的遥感图像具有更好的检测结果; 使用传统图像处理分类方法和本文构建的深度检测网络进行检测实验, 对图中的飞机、轮船、储油罐、桥梁、港口五类军民目标进行检测, 相比于传统算法检测效果有显著提升。

[0067] 检测评价方法如下:

[0068] 将系统测试的所有图片数量即为ALL, 系统识别出有五类待检测目标存在的图片集合1中的图像数量记为F, 其中包括本来无目标而识别出有目标的、以及本来有目标即识别正确的图片数量, 分别记为FP 和FN, 则 $F = FP + FN$; 将系统识别出无目标存在的图片集合2中的图像数量记为T, 其中包括本来无目标即识别正确的、以及本来有目标而没有识别出目标的图片数量, 分别记为TP和TN, 则 $T = TP + TN$ 。本系统根据实际的识别需要, 定义了如下指标:

$$[0069] \quad \text{准确率} = \frac{TP}{TP + FP} \times 100\%$$

$$[0070] \quad \text{召回率} = \frac{TP}{TP + FN} \times 100\%。$$

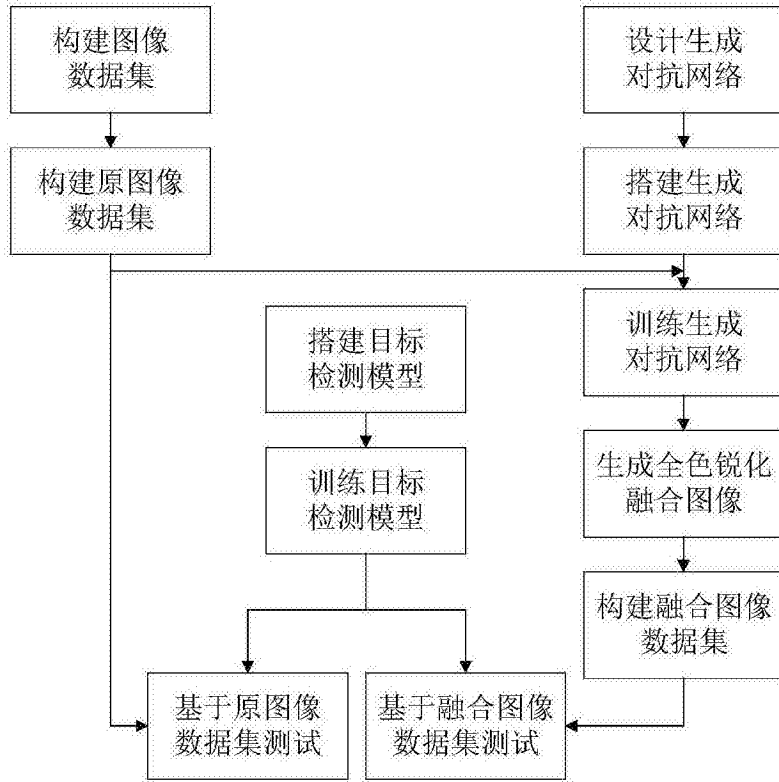


图1

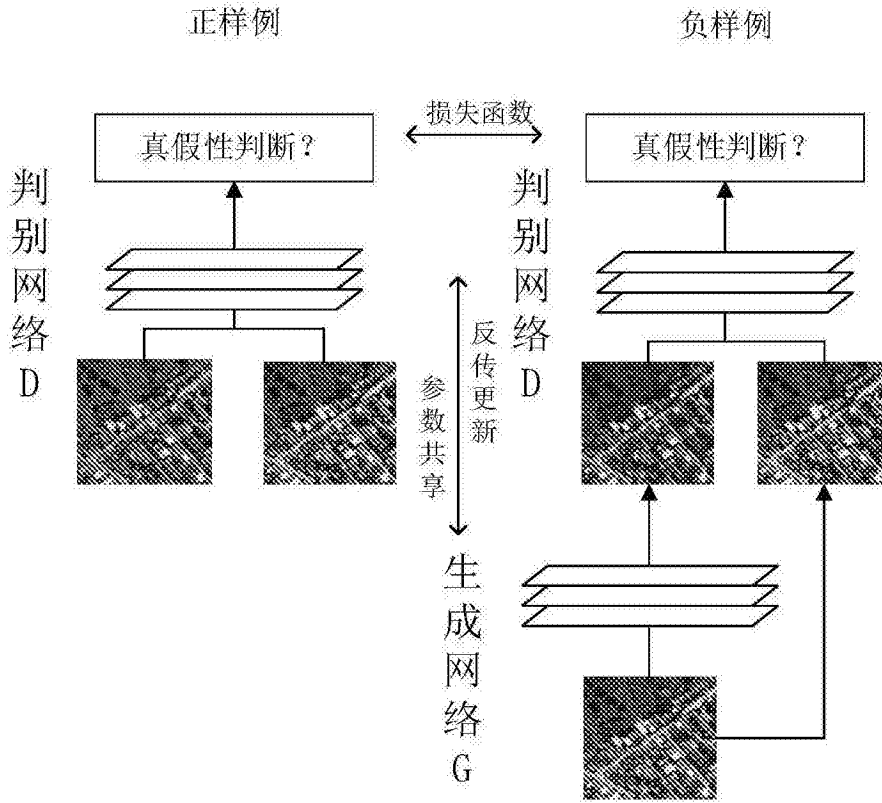
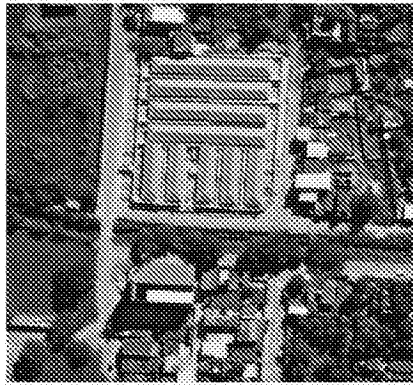
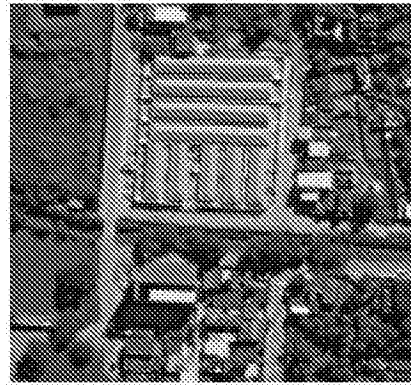


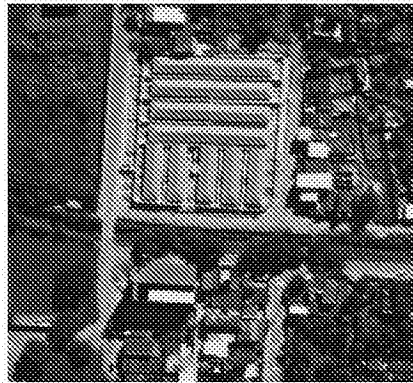
图2



(a)



(b)



(c)



(d)

图3

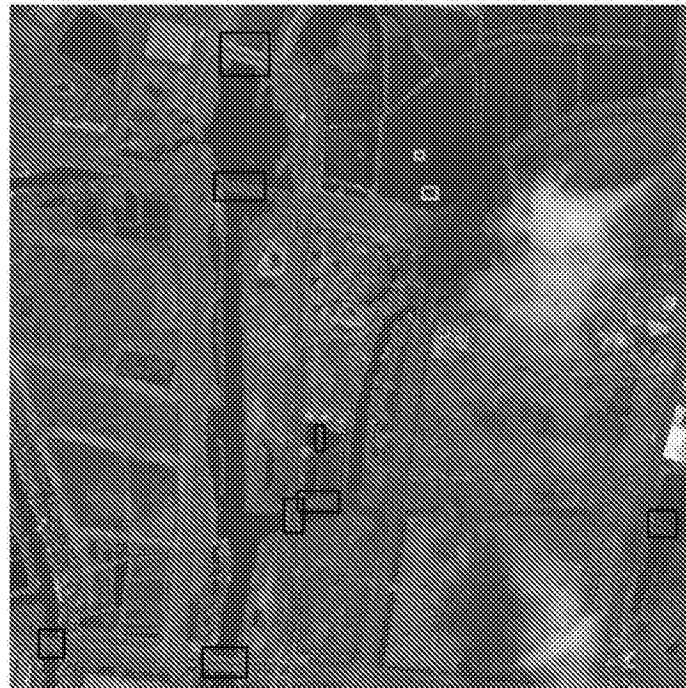


图4