

基于仿射不变离散哈希和条件随机场的遥感图像目标检测

孔 颖^{1,2}, 孙权森¹

(1. 南京理工大学计算机科学与工程学院, 南京 210094; 2. 江苏科技大学计算机学院, 镇江 212100)

摘要: 建立了一种结合仿射不变离散哈希 (Affined-invariant discrete hashing, AIDH) 和条件随机场 (Confidential random field, CRF) 的模型, 实现遥感图像的目标检测。对遥感图像进行超像素分割, 构建适用于 CRF 的以超像素块为顶点的无向图结构。以超像素块作为测试样本, 使用 AIDH 学习方法作为 CRF 一元势函数, 生成初始类别标签。采用 Potts 模型构建 CRF 的二元势函数进行标签的再学习, 平滑目标邻域信息, 解决目标检测中的漏判问题。最后, 使用基于凸壳边界的方法生成最小外接目标框作为目标检测结果。实验表明, 本文方法在目标检测的精度和效率上取得了较好的平衡。

关键词: 遥感; 仿射不变离散哈希; 条件随机场; 目标检测

中图分类号: TP391 **文献标志码:** A

Object Detection of Remote Sensing Image Based on AIDH and CRF

KONG Jie^{1,2}, SUN Quansen¹

(1. School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China; 2. School of Computer, Jiangsu University of Science and Technology, Zhenjiang 212100, China)

Abstract: By constructing a model combining affine-invariant discrete hashing (AIDH) and confidential random field (CRF), the object detection of remote sensing image is achieved. Firstly, the remote sensing image is reconstructed by superpixel segmentation, and the undirected graph structure with superpixel block as vertex is constructed for CRF. Then, the superpixel block is used as the test sample for AIDH learning which is used as CRF unary potential function to generate the initial category label. Then, the pairwise potential function of CRF is constructed by using Potts model for label re-learning, while the object neighborhood information is smoothed and the missing area of object detection is resolved. Finally, the convex hull boundary method is used for generating minimum external rectangular frame as object detection result. Experimental results demonstrate that the proposed method achieves the tradeoff of accuracy and efficiency for objection detection tasks.

Key words: remote sensing; affined-invariant discrete hashing; confidential random field; object detection

引 言

遥感图像目标检测是从一幅遥感图像中检测出所有目标物, 并确定其类别信息。遥感图像以其尺

度多样性、视角特殊性、多方向问题、小目标问题以及背景复杂度高等特殊性的,为解译图像处理任务带来了很大的挑战。传统的基于机器学习的遥感图像目标检测方法有K近邻法(K-nearest neighbour, KNN)^[1]、支持向量机(Support vector machine, SVM)^[2]和神经网络(Neural network, NN)^[3]等。然而这些方法会面临以下问题:(1)时间复杂度与样本数相关;(2)随样本集的增大,机器内存空间需求不断增大;(3)高效的特征提取方法与高效分类器的兼容性问题。

为解决以上问题,考虑引入哈希学习方法,提升检测效率。哈希技术以其低存储、高效率的优势已经广泛应用于图像处理领域。经典的哈希方法包括局部敏感哈希(Locality sensitive hashing, LSH)^[4]、谱哈希(Spectral hashing, SH)^[5]、核哈希(Kernel supervised hashing, KSH)^[6]以及监督离散哈希(Supervised discrete hashing, SDH)^[7]方法等。然而,要直接引入到遥感图像领域,需要考虑遥感的特殊性。在先前的工作中,通过将SDH中引入仿射不变因子构造仿射不变离散哈希(Affined invariant discrete hashing, AIDH),可以解决因遥感成像中出现的仿射变换引起的语义理解问题^[8]。然而,遥感数据的复杂性使得AIDH在目标检测的总体精度提升有限,一些漏判和误判的问题难以解决。目前流行的深度学习方法具备超强的特征表达优势并且取得了不错的效果,却是对样本集有很大的依赖性。从解决漏判和误判问题入手,以传统的数据驱动的目标检测思想为指导,通过引入空间信息的学习提升检测精度。于是,引入条件随机场模型^[9],利用其在标签学习方面的优势,通过对AIDH的目标检测结果进行邻域空间信息学习与消息传递复原图像中的目标区域。而为了更好地利用空间邻域信息,通常采用超像素分割方式。本文使用了基于简单线性迭代聚类(Simple linear iterative clustering, SLIC)生成超像素块^[10],既能满足条件随机场构建概率无向图的模型需求,也为AIDH方法提供了机器学习所需的测试样本。此外,在最终结果生成时为了更好地生成目标框,本文引入了一种基于凸壳边界的最小外接矩形框生成方法,生成了一种更有效的、更能代表目标信息的检测框^[11-12]。

1 相关工作

1.1 简单线性迭代聚类算法

基于SLIC的超像素分割方法,在2010年的MICCAI上首先由Lucchi等^[13]提出,被用于医学图像处理中的细胞显影,其后由Achanta等将其应用到自然图像分割处理中,而逐渐为大家熟知^[10]。该算法生成的超像素块紧凑、均匀,并且运算速度、物体轮廓保持方面也都能满足后续算法需求。该方法的主要思想是:首先确立以颜色和坐标为主体的相似性度量方式;再针对用户需求(超像素块个数 k)使用种子生长的方式聚合相似的像素点,通过限定种子搜索范围和超像素块的连通性保证超像素块都不会产生邻域失衡的情况。

SLIC的优势为:(1)生成的超像素做到了紧致性和均匀程度的平衡,有利于完成基于邻域信息的机器学习任务;(2)具有线性计算复杂度低、存储效率高等优点,能够保持图像的全局特性;(3)既可以分割彩色图像,也可以用于分割灰度图像;(4)算法需要设置的参数很少,只需要设置预分割的超像素块数量即可。

1.2 仿射不变离散哈希

监督离散哈希的目标函数可以表示为

$$\min_{B, W, H} \|Y - W^T B\|^2 + \lambda_1 \|W\|^2 + \lambda_2 \|B - H(X)\|^2 \quad (1)$$

s.t. $B \in \{-1, 1\}^{L \times n} \quad i = 1, \dots, n$

式中: X 为样本数据矩阵; W 为哈希码在 C 个分类上的投影矩阵; Y 为样本的标签矩阵; B 为哈希码矩阵; $H(X)$ 为哈希函数; λ_1 为投影矩阵的正则化参数; λ_2 为惩罚因子。

哈希函数 $H(X)$ 采用映射学习算法获取,为保持样本的非线性结构,采用如下形式

$$H(X) = P^T \phi(X) \quad (2)$$

式中 $\phi(X)$ 为径向基核映射产生的矩阵;矩阵 P 为降维矩阵。

依据多尺度自卷积(Multi-scale autoconvolution, MSA)特征提取的方法,图像的仿射变换不会改变图像所代表物体的语义信息^[14]。由于仿射变换前后的图像对应着相同标签模式类的信息,那么可以将仿射后扩大的样本加入训练样本集,并加以仿射约束信息优化原有目标函数,生成更优质的具有仿射不变性的哈希码。AIDH的构造如下。

定义 R 个旋转角度 $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_R\}$ 及其旋转变换 $T_\alpha = \{T_{\alpha_1}, T_{\alpha_2}, \dots, T_{\alpha_R}\}$, 其中 T_{α_i} 表示经 α_i 角度旋转变换 $\begin{bmatrix} \cos\alpha_i & -\sin\alpha_i \\ \sin\alpha_i & \cos\alpha_i \end{bmatrix}$; 定义 S 个缩放因子 $\beta = \{\beta_1, \beta_2, \dots, \beta_S\}$ 及其缩放变换 $T_\beta = \{T_{\beta_1}, T_{\beta_2}, \dots, T_{\beta_S}\}$, 其中 T_{β_i} 表示 β_i 倍缩放变换 $\begin{bmatrix} \beta_i & 0 \\ 0 & \beta_i \end{bmatrix}$ 。那么仿射变换可表达为旋转变换和缩放变换的笛卡尔积: $T_{\alpha,\beta} = T_\alpha \times T_\beta$ 。将 $T_{\alpha,\beta}$ 应用于所有训练样本 X , 结合原始训练集, 即构成了新的用于仿射不变离散哈希模型的样本集 $X_{\alpha,\beta} = \{X, T_{\alpha,\beta}X\}$ 。

由于仿射变换前后不改变物体本身的语义信息,那么所得到的哈希码应该是相似的,生成的哈希码对数据集具有仿射不变性。因此,可以在目标函数中添加仿射不变约束项,即有

$$Q(B, T_{\alpha,\beta}B) = \lambda_3 \sum_{i=1}^n \|b_i - \overline{T_{\alpha,\beta}b_i}\|^2 = \lambda_3 \|B - \overline{T_{\alpha,\beta}B}\|^2 \quad (3)$$

式中: $T_{\alpha,\beta}x_i = \{T_{\alpha_j}T_{\beta_k}x_i\}$ 表示样本 x_i 所有仿射变换张成的样本子空间, 而 $\overline{T_{\alpha,\beta}b_i}$ 为样本空间 $T_{\alpha,\beta}x_i$ 对应哈希码的均值。

AIDH的目标函数为

$$\begin{aligned} \min_{B, W, P} & \|Y - W^T B\|^2 + \lambda_1 \|W\|^2 + \lambda_2 \|B - P^T \phi(X)\|^2 + \lambda_3 \|B - \overline{T_{\alpha,\beta}B}\|^2 \\ \text{s.t. } & B \in \{-1, 1\}^{L \times n} \quad i = 1, \dots, n \end{aligned} \quad (4)$$

新的约束项可以使每个样本生成哈希码更接近于其张成的仿射样本子空间的均值,从而实现整体样本集 $X_{\alpha,\beta}$ 与其二值哈希编码 B 尽可能相似。

AIDH求解方案如下:

(1) P -step. 固定 B, W , 优化 P 。

$$P = (\phi(X)\phi(X)^T)^{-1} \phi(X) B^T \quad (5)$$

(2) W -step. 固定 P, B , 优化 W 。

$$W = (BB^T + \lambda_1 I)^{-1} B Y^T \quad (6)$$

(3) B -step. 固定 W, P , 优化 B 。原问题等价于求解

$$\begin{aligned} \min_B & \|W^T B\|^2 - 2\text{Tr}(B^T A) \\ \text{s.t. } & B \in \{-1, 1\}^{L \times n} \end{aligned} \quad (7)$$

式中: $A = WY + \lambda_2 H(X) + \lambda_3 \overline{T_{\alpha,\beta}B}$, $\text{Tr}(\cdot)$ 表示矩阵的迹。

另受SDH优化方式的启发,采用梯度下降法按位求解哈希码,可以获得如下的闭解^[15]

$$b_i = \text{sgn}(a_i - \widetilde{B}^T \widetilde{W} w_i) \quad (8)$$

式中: b_i^T 为矩阵 B 的第 i 行; a_i^T 为矩阵 A 的第 i 行; w_i^T 为矩阵 W 的第 i 行; \widetilde{B} 为矩阵 B 去除 b_i^T 后的矩阵, \widetilde{W} 为矩阵 W 去除 w_i^T 后的矩阵。

1.3 条件随机场

条件随机场是建立在概率无向图模型基础的,在给定随机变量 X 条件下,求随机变量 Y 的马尔可夫随机场;其内涵在于对无向图中任意结点,其条件概率分布只和与之有连接的结点有关,与其他结点无关。

设概率无向图 $G=(V,E)$,当无向图每个顶点对应的的数据 X 与其类别标签 Y 满足相同图结构的线性链条件随机场。那么可以按照决定 Y 的因素,定义条件随机场模型的参数形式为

$$P(Y|X) = \frac{1}{Z} \exp \left[\sum_{i \in V} A_i(y_i, X) + \sum_{i \in V} \sum_{j \in N_i} I_{ij}(y_i, y_j, X) \right] \quad (9)$$

式中 Z 为归一化因子。其表达式为

$$Z = \sum_y \exp \left[\sum_{i \in V} A_i(y_i, X) + \sum_{i \in V} \sum_{j \in N_i} I_{ij}(y_i, y_j, X) \right] \quad (10)$$

在式(9)中, $A(y_i, X)$ 为一阶势能或自相关势能,是指当前顶点的特征信息对条件概率的影响; $I_{ij}(y_i, y_j, X)$ 为二阶势能或交互势能,表示空间相邻顶点间的上下文信息对类别条件概率的影响; N_i 代表与第 i 个顶点空间相邻的超像素块集合; $j \in N_i$ 是与顶点 i 空间相邻的所有顶点组成的顶点集^[16]。

2 基于 AIDH 和 CRF 的遥感图像目标检测

本节重点介绍算法4部分内容:基于 SLIC 的遥感图像超像素分割,基于 AIDH 的遥感图像多目标分类,基于 CRF 的标签学习以及基于凸壳边界的最小目标框生成。算法流程图如图 1 所示。

2.1 基于 SLIC 的遥感图像超像素分割

SLIC 是遥感图像目标检测的第一步,重点是将包含相似信息的区域聚合在同一个超像素块中,这样每个超像素块都代表了目标或负样本的信息,超像素块成为了测试样本的载体。

SLIC 的主要思想是,先确立以颜色和坐标为主体的相似性度量方式;再针对用户需求(超像素块个数 k)使用种子生长的方式聚合相似的像素点在一起,通过限定种子搜索范围和超像素块的连通性保证超像素块均匀分布。

SLIC 的基本步骤如下:

(1)对遥感图像,依照超像素块个数 k 初始化等间距为 S 的聚类中心 $C_i = [l, a_i, b_i, x_i, y_i]^T$ 。

(2)在聚类中心的 8 邻域内搜索梯度最小的位置,将其确定为新的聚类中心。

(3)对每个聚类中心 C_i ,计算其与周围 $2S$ 邻域内所有像素点的相似度距离,将距离最近的像素点与聚类中心标记为同类。

(4)强化聚类像素块间的连通性。

(5)更新聚类中心,并计算更新前后总变

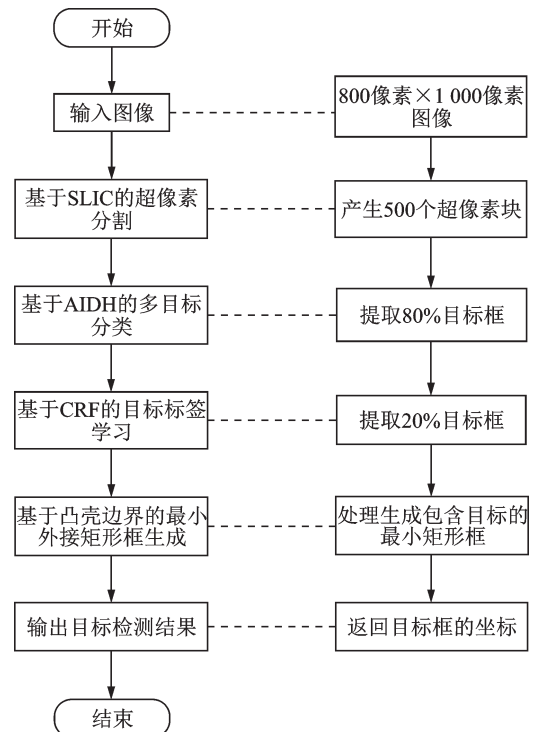


图 1 基于 AIDH-CRF 的目标检测流程图

Fig.1 Flow chart of object detection based on AIDH-CRF

化值 $\sum E(c_i, c'_i) = \sqrt{(\tilde{x}_i - \tilde{x}'_i)^2 + (\tilde{y}_i - \tilde{y}'_i)^2}$ 。

(6)循环执行步骤(3)~(5)直到 $\sum E$ 收敛。

(7)从图像左上方开始,依次给矩阵 Label 赋值,相同像素块的坐标注赋相同的编号。

算法结束后,矩阵 Label 保存的是超像素块的编号,用于计算目标类别标签。

2.2 基于 AIDH 的多目标分类

基于 AIDH 的多目标分类,是指依据数据集中已经训练好的学习机,预测遥感图像上的目标位置及类别的过程^[15]。与传统的机器学习方法中多目标分类的方式不同^[8],本文方法是以超像素块作为目标检测的测试样本。而这些超像素块,除了包含目标信息(正样本集),还包含背景类(负样本集)信息。因此 AIDH 的训练过程,需要由超像素块提炼生成负样本集,与正样本集共同组成样本训练集,这样才能生成优质的哈希码,预测图像中的目标。基于 AIDH 的目标检测框架图如图 2 所示,经过 AIDH 的目标检测后,超像素块变量 Label 存储的是初步预测的类别标签。

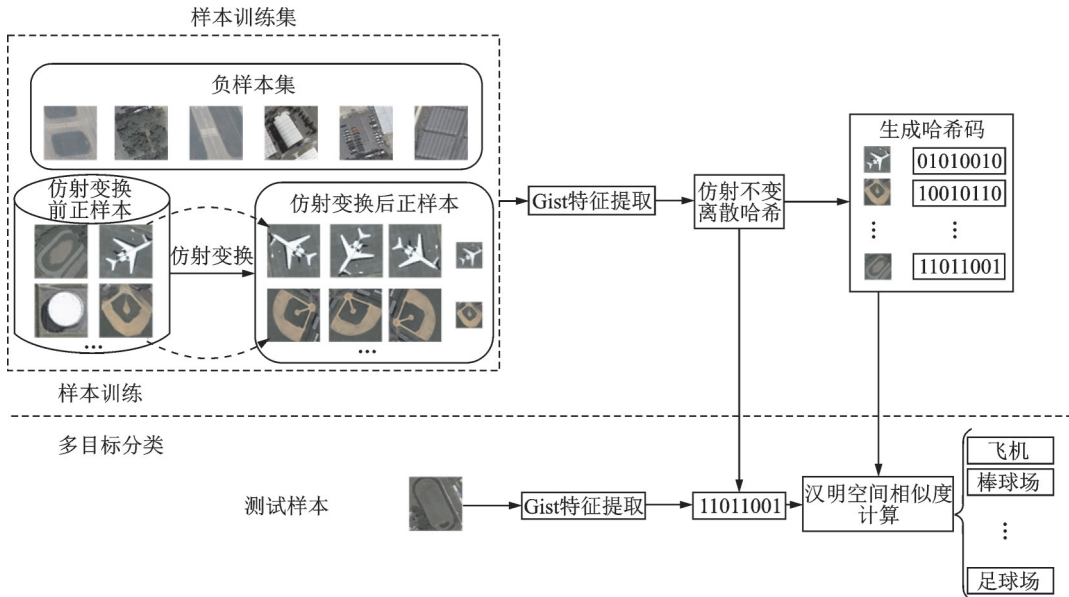


图2 基于 AIDH 的遥感图像目标检测

Fig.2 AIDH-based remote sensing image object detection

2.3 基于 CRF 的标签学习

为了改进 AIDH 方法的分类结果,采用 CRF 方法,进行标签再学习的方式,通过邻域信息传递的方法修正漏判、错判情况。

CRF 模型的一阶势能探讨的是观测序列与状态序列间的决定关系。这里以 AIDH 模型作为一阶势能函数 $A(y_i, X)$, 其形式为

$$A(y_i, X) = \max \left\{ \frac{\sum_k \{ \text{train}X_k^l | \text{Ham}(H(x_i), H(\text{train}X_k^l)) \}}{\text{total}} \right\} \quad (11)$$

式中: $H(\cdot)$ 表示 AIDH 方法的哈希函数; $\text{train}X_k^l$ 表示训练集中类别标签为 l 样本; $\text{Ham}(\cdot)$ 是以汉明距离度量某类别样本与样本点 i 之间相似性的函数。

CRF 模型的二阶势能探讨的是空间相邻超像素块间的状态转移关系,本文采用的是对差异敏感的 Potts 模型^[17]。该模型通过惩罚相邻超像素块间的不一致性,修正模型标签的分配情况。具体公式为

$$I_{ij}(y_i, y_j, X) = \begin{cases} \beta_s \exp\left(-\frac{\theta_s \|H(x_i) - H(x_j)\|}{R}\right) & y_i = y_j \\ \beta_s \left[1 - \exp\left(-\frac{\theta_s \|H(x_i) - H(x_j)\|}{R}\right)\right] & y_i \neq y_j \end{cases} \quad (12)$$

式中: β_s 、 θ_s 为模型参数,其中 β_s 为权重参数,用于调节空间势能在总势能中的权重; θ_s 为高斯核的尺度参数,用于调节空间势能随特征向量变化的平滑程度。 R 为哈希函数 $H(x_i)$ 产生哈希向量的维度,也即哈希码的比特数。

在 Potts 模型的应用中,当邻域样本同类时,两点间距离越小,空间势能越大,则两点属于同类的可信度越大,此时可以解决目标检测中的漏判情况。此外,当邻域样本不同类时,对两点间特征距离采取惩罚机制。若两点间距离越小,则两点间属于不同类的可信度越小,那么两个点也有可能归于同类,此时可以解决目标检测中的误判情况^[18]。

对于算法的训练过程,由于二阶势能参数 θ_s 采用的是高斯核的尺度函数,无法计算正确的目标函数。于是,本文采用网格搜索的方式估计二阶势能参数。而在算法的预测过程,采用的是环状置信传播算法(Loopy belief propagation, LBP)^[19]。LBP 算法是一种消息传递算法,主要思想是通过将当前节点下接收到的所有来自邻域的消息传递给目标节点,完成节点间信息传递。

LBP 算法消息传递的方式是最大积方式,通过更新节点间(本章为超像素块)消息的传递和计算置信度更新节点势能直至收敛或达到最大迭代次数。使用最大积方式的消息传递公式 $\text{msg}_{i \rightarrow j}(c_m)$ 如式(13)所示,表示的是从超像素块 i 到 j 传递的信息,即超像素块 i “认为”超像素块 j 属于类别 c_m 。

$$\text{msg}_{i \rightarrow j}(c_m) = \max_{c_n \in \Omega} \left[\exp(\phi_i(c_m)) \cdot \exp(\phi_{ij}(c_m, c_n)) \cdot \text{msg}_{l \rightarrow i}(c_m) \right] \quad (13)$$

式中

$$\begin{aligned} \phi_i(c_m) &= A_i(y_i = c_m, x) \\ \phi_{ij}(c_m, c_n) &= I_{ij}(y_i = c_m, y_j = c_n, x) \end{aligned} \quad (14)$$

式(14)是简化表示的一阶、二阶势能表达式,即超像素块 i 属于第 m 类目标的概率、超像素块 i 属于第 m 类且超像素块 j 属于第 n 类目标的概率; $l \in N_i \setminus j$ 表示的是与超像素块 i 对应的除 j 以外的邻域超像素块集合; Ω 表示目标检测的类别种类。

而置信度 $b_i^{(p)}(y_i = c_m)$ 表示的是超像素块 i 取标签 c_m 的置信度,即有

$$b_i^{(p)}(y_i = c_m) = \exp(\phi_i(c_m)) \cdot \text{msg}_{l \rightarrow i}(c_m) \quad (15)$$

式中 p 为迭代次数

通过更新超像素块间消息传递,计算置信度至循环结束,得出超像素块属于某类别标签的最大后验概率(Maximum a posterior estimation, MAP),认为超像素块 i 的置信度最大的类别即为目标所属的类别标签,有

$$y_i^{\text{MAP}} = \operatorname{argmax}_{c_m \in \Omega} b_i(x_i = c_m) \quad (16)$$

2.4 基于凸壳边界的最小目标框生成

为了规范化目标检测的结果,对 CRF 生成的由多个超像素块组成的多边形检测区域,本文采用基于凸壳边界(Convex hull boundary, CHB)的方法,求取最小外接矩形框,保证矩形框的面积尽量小。具

体算法如下:

- (1)使用格雷厄姆法获取CRF检测结果的多边形的顶点和边。
- (2)计算多边形每条边的斜率,确定其反正切角为旋转方向角 angle_i 。
- (3)从 $i=1$ 开始计算每次旋转 angle_i 后,多边形每个顶点的坐标 x_{yr} 。
- (4)计算 x_{yr} 中 x_{\min} 、 x_{\max} 、 y_{\min} 、 y_{\max} , 并以 $M_i=(x_{\max}-x_{\min}, y_{\max}-y_{\min})$ 的 2-范数作为矩形框是否更新的评判标准。当新向量 M_i 的 2-范数更小时,则进入下一步;否则跳过本次循环,计算下一次旋转变换的结果。
- (5)更新最小外接矩形框的坐标为 $\text{rect}=[[x_{\min}, y_{\min}]; [x_{\max}, y_{\min}]; [x_{\max}, y_{\max}]; [x_{\min}, y_{\max}]]$ 。
- (6)当完成所有旋转变换后,以 rect 的最终结果作为最小外接矩形框的顶点。

图3给出了基于凸壳边界方法生成最小外接矩形框的效果。从图3(a)可以看到,经过CRF处理后的目标检测结果虽然能勾勒出飞机的大体轮廓,但是受超像素块质量和飞机阴影的影响,目标检测结果的展示效果并不理想,过多的轮廓线使得飞机目标并不突出。而在图3(b)经过CHB方法处理后,飞机在矩形框中完整地展现出来,在保证飞机两翼的最大臂展边界同时,使得矩形框的面积更小。背景信息越小,目标更突出。这样的矩形框,更能代表目标信息。

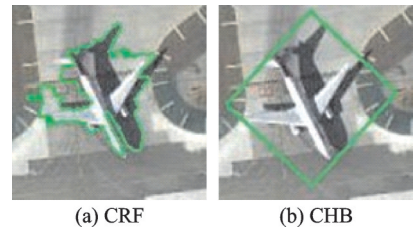


图3 凸壳边界法处理前后目标检测结果对比
Fig.3 Comparison of target detection results before and after convex hull boundary method

3 实验部分

3.1 数据集及算法运行环境

本文以NWPU VHR-10作为检测数据集^[20],以Dense-SIFT作为特征^[21]。AIDH旋转系数为 $R=11$, $\alpha=\{30^\circ, 60^\circ, \dots, 330^\circ\}$, 缩放系数为 $S=2$, $\beta=\{0.5, 0.75\}$; CRF方法的最大迭代次数 $P=4$, CRF的两个参数分别为二阶势能权重参数 $\beta_s=6$, 高斯核尺度参数 $\theta_s=2$ 。

3.2 本文方法目标检测结果

(1)本文方法目标检测各阶段结果如图4所示。图4是本文方法各阶段结果示意图。其中图4(a)是原图,含有12个网球场目标。图4(b)是SLIC超像素分割结果,完成目标物和非目标物的初步分离。图4(c)是AIDH处理结果,可以看到该阶段能够检测出一部分网球场目标,但是存在漏判的情况。图4(d)是CRF处理结果,该阶段恢复了漏判的网球场目标,复原了整幅图像网球场的结果。最后图4(e)是最终目标检测结果,在最小外接矩形框提取后,检测结果中目标信息更加鲜明,背景信息更少,更能代表图像上的目标信息。

(2)本文方法目标检测结果,如图5所示。图5给出了本文方法目标检测的结果图,图中绿色标识为飞机,红色为舰船,粉红色为棒球场,紫色为桥梁。可以看出,本文方法既能检测出单个目标,也能检测出多个目标,还能检测出多个不同类别的目标,并且能很好地勾勒出目标框。而对复杂背景下的桥梁信息,本文方法也能检测出来,这表明了算法的有效性。

3.3 各种目标检测方法的定量对比

为定量分析本文方法有效性,选取了两种经典的基于机器学习的目标检测方法KNN^[1]和SVM^[2]以及两种新的哈希方法KSH^[6]和SDH^[7]方法作对比,此外还列举了先前的工作AIDH-SVM^[15]作为对

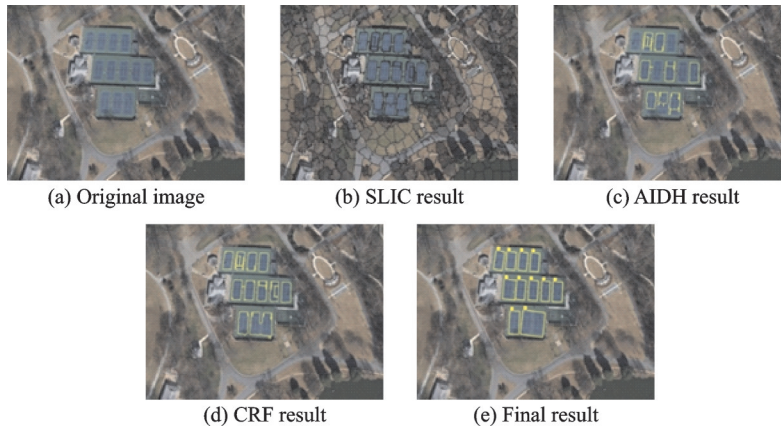


图4 本文方法各阶段结果示意图

Fig.4 Results of each stage of the proposed method



图5 本文方法目标检测结果图

Fig.5 Object detection results of the proposed method

比。在定量数据指标上,采用单类别平均精度 AP 和综合平均精度 mAP 作为评判标准。

表1给出了不同算法下各类别目标的检测精度以及综合平均精度,可以看出:

(1)从横向角度看,几种哈希方法在 AP、mAP 上都优于 KNN、SVM 两种经典方法,体现出哈希方法相对于经典方法的优越性。此外,几种哈希方法间对比发现, AIDH-CRF 的检测精度要优于 KSH-CRF、SDH-CRF,这体现了 AIDH 在哈希系方法中的突出性;而与 AIDH-SVM 对比中,

表1 各算法在 NWPU VHR-10 数据集下的 AP 和 mAP 对比

Table 1 AP and mAP comparison of each algorithm in the NWPU VHR-10 dataset

目标类	KNN	SVM	KSH-CRF	SDH-CRF	AIDH-SVM	AIDH-CRF
飞机	0.222 2	0.367 2	0.542 3	0.678 6	0.684 5	0.788 6
舰船	0.136 5	0.257 3	0.472 7	0.628 8	0.652 9	0.742 6
油桶	0.368 0	0.543 1	0.788 9	0.897 9	0.926 6	0.965 2
棒球场	0.274 5	0.403 1	0.712 6	0.793 8	0.795 7	0.905 3
网球场	0.233 9	0.292 6	0.406 5	0.582 1	0.694 5	0.752 8
篮球场	0.117 5	0.254 8	0.401 8	0.466 5	0.352 5	0.488 4
操场	0.127 9	0.315 6	0.507 1	0.669 0	0.654 0	0.753 9
港口	0.172 2	0.287 9	0.334 3	0.367 8	0.554 6	0.630 8
桥梁	0.086 1	0.178 3	0.308 6	0.344 2	0.439 6	0.502 0
车辆	0.197 1	0.293 4	0.329 0	0.355 9	0.384 6	0.438 2
mAP	0.196 7	0.323 2	0.493 2	0.617 8	0.631 7	0.718 5

AIDH-CRF 在精度上有了明显的提升,说明本文方法使用 CRF 改进的有效性。

(2)从纵向角度看目标类检测的情况,本文方法保持了先前在小目标检测上的精度,如飞机、油桶和棒球场;同时提升了网球场、篮球场等大型运动场的检测率。正因为在大目标检测的成功,才实现了综合平均精度的提升。

表 2 给出了不同目标检测算法的算法用时。首先,可以看到 KNN、SVM 的算法用时明显高于几种哈希方法,而比较几种哈希方法发现,3 种结合 CRF 的哈希方法(KSH-CRF、SDH-CRF、AIDH-CRF)用时高于 AIDH-SVM 方法,表明了 CRF 的引入会增加算法用时。结合表 1 的 mAP 数据来看,AIDH-CRF 以牺牲少许时间为代价,换取了检测精度的大幅提升,做到了精度和速度上的权衡。

4 结束语

本文通过分析传统的遥感图像目标检测方法的不足,引入条件随机场方法,利用其在标签学习中的优势提高目标检测的精度。利用 SLIC 超像素分割方法在空间信息获取上的优势,构建高质量的超像素块,完成不同目标物的初步分离。而基于条件随机场的标签学习方法,利用 SLIC 生成的超像素块构建其模型所需的概率无向图模型,继承了 AIDH 良好的机器学习性能,又通过二阶势能的概率计算改进了检测结果,弥补了大量的漏判和误判情况。此外,采用基于凸壳边界的最小外接矩形框生成方法获得了更能代表目标信息的矩形框。通过在 NW-PU VHR-10 数据集进行实验,取得了不错的目标检测结果,定量分析发现,本文方法以牺牲部分时间复杂度为代价换取了 AP 和 mAP 的大幅提升,实现了检测算法速度和精度的有效权衡。

参考文献:

- [1] MURALIDHARAN R, CHANDRASEKAR C. Combining local and global feature for object recognition using SVM-KNN [C]//Proceedings of International Conference on Pattern Recognition. Erlangen, Germany: Elsevier, 2012.
- [2] MAULIK U, CHAKRABORTY D. Remote sensing image classification: A survey of support-vector-machine-based advanced techniques[J]. IEEE Geoscience & Remote Sensing Magazine, 2017, 5(1): 33-52.
- [3] WU S L, CHEN H D, BAI Y, et al. A remote sensing image classification method based on sparse representation[J]. Multimedia Tools & Applications, 2016, 75(19): 12137-12154.
- [4] GIONIS A, INDYK P, MOTWANI R. Similarity search in high dimensions via hashing[C]//Proceedings of the 25th VLDB Conference. Edinburgh, Scotland, UK: Morgan Kaufmann, 1999, 8(2): 518-529.
- [5] WEISS Y, TORRALBA A, FERGUS R. Spectral hashing[J]. Advances in Neural Information Processing Systems, 2009, 282(3): 1753-1760.
- [6] LIU W, WANG J, JI R R, et al. Supervised hashing with kernels[C]//Proceedings of IEEE Conference on Computer Vision & Pattern Recognition. Rhode Island, USA: IEEE, 2012: 2074-2081.
- [7] SHEN F M, SHEN C H, LIU W, et al. Supervised discrete hashing[C]//Proceedings of IEEE Conference on Computer Vision & Pattern Recognition. Boston, USA: IEEE, 2015: 37-45.
- [8] 孔颖, 孙权森, 徐晖, 等. 基于仿射不变离散哈希的遥感图像多目标分类[J]. 软件学报, 2019, 4: 914-926.
KONG Jie, SUN Quansen, XU Hui, et al. Multi-object classification of remote sensing image based on affine-invariant supervised discrete hashing[J]. Journal of Software, 2019, 4: 914-926.
- [9] BEHERA S, NANDA P K. LBP and weber law descriptor feature based CRF model for detection for man-made structures [C]//Proceedings of International Conference on Man & Machine Interfacing. Bhubaneswar, India: IEEE, 2015.

表 2 各算法的算法用时对比

Table 2 Comparison of time consuming between each algorithm

算法名称	算法用时/s
KNN(128-DenseSIFT)	1 755
SVM(128-DenseSIFT)	441
KSH-CRF(64 bits)	184
SDH-CRF(64 bits)	162
AIDH-SVM(128 bits)	104
AIDH-CRF(64 bits)	176

- [10] ACHANTA R, SHAJI A, SMITH K, et al. SLIC superpixels compared to state-of-the-art superpixel methods [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2012, 34(11): 2274-2282.
- [11] GRAHAM R L. An efficient algorithm for determining the convex hull of a finite planar set [J]. *Information Processing Letters*, 1972, 1(4): 132-133.
- [12] MONFARED M D, MOHADES A, REZAEI J. Convex hull ranking algorithm for multi-objective evolutionary algorithms [J]. *Scientia Iranica*, 2011, 18(6): 1435-1442.
- [13] LUCCHI A, SMITH K, ACHANTA R, et al. A fully automated approach to segmentation of irregularly shaped cellular structures in EM images [C]//*Proceedings of International Conference on Medical Image Computing & Computer Assisted Intervention*. Beijing, China: Elsevier, 2010.
- [14] 张洁玉, 陈强, 白小晶, 等. 基于广义典型相关分析的仿射不变特征提取方法[J]. *电子与信息学报*, 2009, 31(10): 2465-2469.
ZHANG Jieyu, CHEN Qiang, BAI Xiaojing, et al. Affine invariant feature extraction algorithm based on generalized canonical correlation analysis[J]. *Journal of Electronics & Information Technology*, 2009, 31(10): 2465-2469.
- [15] 孔颖, 孙权森, 纪则轩, 等. 基于仿射不变离散哈希的遥感图像快速目标检测新方法[J]. *南京大学学报(自然科学版)*, 2019, 55(1): 49-60.
KONG Jie, SUN Quansen, JI Zexuan, et al. A novel fast object detection method in remote sensing image based on affine-invariant discrete hashing[J]. *Journal of Nanjing University (Natural Science Edition)*, 2019, 55(1): 49-60.
- [16] QI L, LI C. A linear-chain CRF-based learning approach for web opinion mining [C]//*Proceedings of International Conference on Web Information Systems Engineering*. Hong Kong, China: Springer, 2010: 128-141.
- [17] HOBERG T, ROTTENSTEINER F. Classification of settlement areas in remote sensing imagery using conditional random fields [C]//*Proceedings of ISPRS TC VII Symposium 100 Years ISPRS*. Vienna, Austria: Elsevier, 2010: 53-58.
- [18] 汪光亚. 基于条件随机场的多时相遥感影像分类[D]. 南京: 南京理工大学, 2018.
WANG Guangya. New multitemporal remote sensing image classification method based on conditional random fields[D]. Nanjing: Nanjing University of Science & Technology, 2018.
- [19] WEBB G I, SAMMUT C, PERLICH C, et al. *Loopy belief propagation* [M]. London: Springer, 2010.
- [20] CHENG G, ZHOU P C, HAN J W. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images [J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2016, 54(12): 7405-7415.
- [21] LAZEBNIK S, SCHMID C, PONCE J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories[C]//*Proceedings of IEEE Conference of Computer Vision & Pattern Recognition*. Las Vegas, USA: IEEE, 2016: 2169-2178.

作者简介:



孔颖(1985-),男,博士,研究方向:遥感图像处理、机器学习与应用, E-mail: 282754414@qq.com。



孙权森(1963-),通信作者,男,博士,教授,研究方向:模式识别与机器学习、遥感信息处理, E-mail: qs-sun@126.com。

(编辑:刘彦东)