

# 显著性偏振参量深度稀疏特征学习的目标检测方法

王美荣<sup>1</sup>, 徐国明<sup>1,2\*</sup>, 袁宏武<sup>1,2</sup>

<sup>1</sup>安徽新华学院信息工程学院, 安徽 合肥 230088;

<sup>2</sup>中国人民解放军陆军炮兵防空兵学院偏振光成像探测技术安徽省重点实验室, 安徽 合肥 230031

**摘要** 基于偏振成像特点和深层特征分类需求, 提出一种显著性偏振参量深度稀疏特征学习的目标检测方法。首先在偏振解析基础上构造显著性偏振参量图像作为检测源图像; 然后在判别式字典对下对待检测图像进行稀疏特征学习, 并以字典对作为分类器在卷积神经网络(CNN)框架下进行目标分类和定位; 最后结合偏振成像探测实际应用需求, 选择典型目标和应用场景进行数据采集和模型训练, 并进行仿真验证。结果表明该方法在检测得分和平均检测精度上都比直接偏振方向方法有所提高, 验证了其有效性, 该方法对于有效提升偏振成像探测能力具有应用价值。

**关键词** 成像系统; 偏振参量; 稀疏特征; 显著性; 深度学习

**中图分类号** TP391.4; TN911.73

**文献标识码** A

**doi:** 10.3788/LOP56.191101

## Object Detection by Deep Sparse Feature Learning of Salient Polarization Parameters

Wang Meirong<sup>1</sup>, Xu Guoming<sup>1,2\*</sup>, Yuan Hongwu<sup>1,2</sup>

<sup>1</sup>Institute of Information Engineering, Anhui Xinhua University, Hefei, Anhui 230088, China;

<sup>2</sup>Anhui Province Key Laboratory of Polarized Imaging Detecting Technology, Army Artillery and Air Defense Forces Academy, Chinese People's Liberation Army, Hefei, Anhui 230031, China

**Abstract** Based on polarization imaging characteristics and deep feature classification requirements, an object detection method based on deep sparse feature learning of salient polarization parameters is proposed. First, the salient polarization parameter image is constructed as the source image based on polarization analysis. Then the sparse feature of the image to be detected is learned by discriminant dictionary pair, and the object is classified and located by the dictionary pair which is used as the classifier in CNN framework. Finally, the typical object and scene are selected for data acquisition and model training according to the practical application requirements of polarization imaging detection, and some simulation experiments are conducted. The results show that the detecting score and average detection precision of the proposed method are improved at different degrees by comparing to the polarization direction detection methods and the effectiveness of this method is verified. The proposed method has application value for improving the detection ability of polarization imaging effectively.

**Key words** imaging system; polarization parameter; sparse feature; saliency; deep learning

**OCIS codes** 110.5405; 330.1880; 100.4996

## 1 引言

同传统的成像探测技术相比, 偏振成像技术在获取光强、光谱和空间信息的前提下, 能同时得到各

观测目标直观的偏振特性, 从而为实现目标空间结构的反演提供基础<sup>[1]</sup>。随着偏振成像探测技术的快速发展, 其在军事应用的需求也从人工离线判读辅助侦察行动逐步发展到对目标的快速自动检测和识

**收稿日期:** 2019-04-10; **修回日期:** 2019-05-14; **录用日期:** 2019-05-20

**基金项目:** 国家自然科学基金(61379105)、安徽省高校自然科学研究重点项目(KJ2018A0587, KJ2019A0906)、中国博士后科学基金(2016M592961)、安徽省自然科学基金(1608085MF140)

\* **E-mail:** xgm121@163.com

别。在目标检测方法研究领域,利用大规模训练数据集<sup>[2]</sup>,在基于区域的卷积神经网络(CNN)<sup>[3]</sup>框架下,基于候选区域、特征学习和目标分类等基本步骤的目标检测算法取得了显著检测效果,是目前的主流方法之一<sup>[4]</sup>。其中,特征学习是目标检测的必备步骤,选择合适的特征模型将图像区域映射为特征向量,然后利用从训练样本学习到的分类器对该特征向量进行分类,判断其所属类型。在此过程中如何根据不同成像方式和目标特征选择合适的特征模型进行特征学习,对其后的候选框回归计算以及目标分类等有重要影响,同时特征的表达能力也影响分类器精度。

特征学习包括人工设计特征和自动学习特征两个方面。典型的人工特征从早期的边缘检测算子<sup>[5]</sup>和角点检测算子<sup>[6]</sup>,到 SIFT(scale-invariant feature transform)<sup>[7]</sup>和 HOG(histogram of oriented gradients)<sup>[8]</sup>等,具有良好的可扩展性。为弥补利用单一特征进行目标表示的不足,后来有学者提出多种特征组合方法<sup>[9]</sup>,该方法被广泛应用于行人检测<sup>[10]</sup>、目标跟踪<sup>[11]</sup>以及人体姿态估计<sup>[12]</sup>等任务并取得了良好效果。但是依靠人工设计特征,需要丰富的专业知识并且花费大量的时间,特征的好坏在很大程度上还要依靠经验和运气。近年来,一方面,深度学习在图像分类和目标检测等领域取得了突破性进展,成为目前最有效的自动特征学习方法,且深度学习模型具有强大的表征和建模能力,避免了人工设计特征的繁琐低效<sup>[13]</sup>;另一方面,辨别字典学习(DDL)近年来在稀疏表示领域也取得了巨大的成功<sup>[14-17]</sup>,并且在计算机视觉领域如行人重识别<sup>[18]</sup>任务中取得良好效果。

DDL的目的是学习一个字典,在此字典下对图像信号形成稀疏表示,并考虑其表示精度和判别能力,而稀疏特征更适合作为目标类别分类的分类器,通常有两个方法可以提高字典的判别能力。一种方法是提高编码向量的判别分类能力,如:Jiang等<sup>[17]</sup>提出二值分类标签稀疏编码矩阵以提高相同分类中具有相似稀疏编码的样本泛化能力;Mairal等<sup>[14]</sup>提出任务驱动字典学习框架,针对不同任务求解编码系数的不同目标函数;Yang等<sup>[16]</sup>提出 Fisher 判别字典学习方法,利用 Fisher 准则刻画稀疏表示系数。另一种方法是通过学习结构化字典提高判别能力,如:Ramirez等<sup>[19]</sup>提出用结构不相关项来描述字典的独立性;Gao等<sup>[20]</sup>提出学习一个额外字典供所有的分类共享共有特征编码;Gu等<sup>[21]</sup>还提出投影

字典对学习方法,利用分析字典来估计表示系数,大大提高了稀疏特征表示性能。如果将这些方法直接应用于偏振成像,虽然能够进行目标检测,但是由于无法有效利用不同目标的偏振信息,在图像目标较小或者不良成像条件下探测精度就会受到影响。

为此,本文针对偏振成像特点并结合实际应用需求,提出基于斯托克斯(Stokes)偏振信息解析和深度稀疏特征学习的偏振图像目标检测方法。该方法首先进行偏振参量图像解析,为充分利用目标图像的偏振信息,对获得的不同偏振方向图像利用 Stokes 进行解析,得到偏振参量图像;其次,受显著性检测方法启发<sup>[22]</sup>,在目标检测时不直接对偏振方向图像进行检测,而是通过 Choquet 模糊积分<sup>[23]</sup>构造显著性偏振参量图像,以人眼视觉和信噪比高的偏振参量图像进行目标检测;然后,对得到的待检测偏振图像在判别式字典下进行稀疏特征学习;最后,以偏振参量图像的稀疏特征设计字典对分类器层在 CNN 框架完成目标检测。实验验证结合偏振成像目标探测实际应用需求,在某型偏振成像设备采集缩放比仿真目标和实际空中小目标等不同场景下展开。

## 2 偏振参量图像原理

### 2.1 Stokes 变换

同光强度成像探测相比,偏振成像技术是一种探测物体偏振态的成像技术,在获取光强、光谱和空间信息的前提下,能同时得到更多的易于区分目标的特征信息,如偏振度、偏振角和偏振椭圆率等,可以显著提高目标与背景,及不同目标之间的对比度。Stokes 向量是 Stokes 在 1852 年研究部分偏振光时提出的,它既可以用来表示完全偏振光,也可以表示部分偏振光甚至自然光,同时 Stokes 向量带有丰富的信息。因此,偏振成像探测多用 Stokes 参量  $S=[I Q U V]^T$  来描述,其中, $I$  表示光的总强度, $Q$  表示水平方向上的线偏振光的强度, $U$  表示  $45^\circ$  方向上线偏振光的强度, $V$  表示左、右圆偏振光的强度差。在实际目标探测中,由于光的左、右圆偏振特征难以准确测量,而且在通常情况下目标的反射光或者自身的辐射都是线偏振分量占主要部分,因此,常忽略  $V$  分量。当改变偏振片透光轴与所选参考坐标轴的夹角  $\theta$ ,例如  $\theta$  分别取  $0^\circ$ 、 $60^\circ$  和  $120^\circ$ (也可以取  $0^\circ$ 、 $45^\circ$  和  $90^\circ$ ),可获得三个偏振方向的出射光强  $I(\theta_1)$ 、 $I(\theta_2)$ 、 $I(\theta_3)$ ,Stokes 参量  $I$ 、 $Q$ 、 $U$  的解析式为

$$\begin{cases} I = \frac{2}{3} [I(0^\circ) + I(60^\circ) + I(120^\circ)] \\ Q = \frac{2}{3} [2I(0^\circ) - I(60^\circ) - I(120^\circ)] \\ U = \frac{2}{\sqrt{3}} [I(60^\circ) - I(120^\circ)] \end{cases} \quad (1)$$

偏振度  $P$  和偏振角  $A$  反映目标反射面的导电特性。 $P$  表示反射光中线偏振成分的大小,当  $|P|=0$  时,表示非偏振光, $|P|=1$  表示全偏振光, $0<|P|<1$  时,表示部分偏振光; $A$  表示反射光两分量之间的相位差,也就是入射光的偏振方向相对于  $x$  轴的夹角,对于部分偏振光来说,就是能量最大的偏振方向相对于  $x$  轴的夹角。偏振度  $P$  和偏振角  $A$  分别表示为

$$\begin{cases} P = \sqrt{Q^2 + U^2} / I \\ A = \frac{1}{2} \arctan\left(\frac{U}{Q}\right) \end{cases} \quad (2)$$

另外,还可以得到  $X$  方向振动矢量图  $E_x$ 、 $Y$  方向振动矢量图  $E_y$ 、振动矢量差图  $\Delta E$ 、方位角图  $\beta$ <sup>[24]</sup> 为

$$\begin{cases} E_x = \sqrt{IP + Q} \\ E_y = \sqrt{IP - Q} \\ \Delta E = E_x - E_y \\ \beta = \arctan(E_y/E_x) \end{cases} \quad (3)$$

为描述方便,这里把 Stokes 参量 ( $I$ 、 $Q$ 、 $U$ )、偏振度  $P$ 、偏振角  $A$  以及  $X$  方向振动矢量图  $E_x$ 、 $Y$  方向振动矢量图  $E_y$ 、振动矢量差图  $\Delta E$ 、方位角图  $\beta$  所成的图像统称为偏振参量图像。在偏振成像探测中,物体的偏振特征可以用偏振参量图像来完整描述,各参量图像从不同的角度反映物体的本征偏振信息,从而为目标空间结构反演和目标检测提供基础。

## 2.2 显著性偏振参量图像

由于偏振参量是多通道图像,偏振图像目标检测应根据检测具体任务需求在不同的偏振参量上进行,同时也可以大大提高检测速度。就人眼视觉及识别任务来说,要求偏振参量图像在图像细节、纹理和清晰度上表现较好。标准差是图像对比度测量,反映图像细节信息量和图像中数据的集中与离散程度,信息熵反映图像纹理,纹理越丰富,信息熵越大,清晰度反映图像清晰程度以及纹理变换程度。因此,首先选用标准差、信息熵和清晰度这 3 个属性来衡量偏振参量图像;然后构造这 3 个属性的信任函数;最后采用 Choquet 积分自适应选择显著性偏振

参量作为目标检测的图像。

### 1) 评价指标

采用标准差  $\sigma$ 、信息熵  $E_n$ 、清晰度  $\bar{g}$  对各参量图像进行评价,定义为

$$\sigma = \sqrt{\sum_{i=1}^n (s_i - \eta)^2 / n}, \quad (4)$$

$$E_n = - \sum_{b=0}^{M-1} P(b) \log_2 P(b), \quad (5)$$

$$\bar{g} = \frac{1}{n} \sum \sqrt{[(\Delta I_x)^2 + (\Delta I_y)^2] / 2}, \quad (6)$$

式中: $\eta$  为图像的像素均值; $M$  为图像的灰度等级; $n$  为图像的大小; $s_i$  为图像第  $i$  个像素; $P(b)$  为像素灰度值为  $b$  的概率; $\Delta I_x$  和  $\Delta I_y$  分别为图像  $I$  在  $x$  和  $y$  方向上的偏导数。

### 2) 信任函数

对 3 个属性构建信任函数,表示为

$$f(X_j) = \frac{X_j \sqrt{X}}{\max(X_j)}, \quad (7)$$

式中: $X_j$  为第  $j$  个偏振参量图像的  $\{\sigma, E_n, \bar{g}\}$ ;  $X$  为偏振参量图像均值; $j$  取  $\{U, Q, P, \theta, E_x, E_y, \Delta E, \beta\}$ 。

模糊测度  $g$  的值反映每个属性的重要程度,需满足  $\sum g(X_j) = 1$ ,根据信任函数值大小来为每个属性分配固定的权重:

$$g(X_j) = \frac{f(X_j)}{\sum f(X_j)}. \quad (8)$$

### 3) Choquet 积分

在得到各偏振参量图像的信任函数和模糊测度后,采用 Choquet 积分自适应选择最佳的偏振参量图像作为显著性偏振参量进行目标检测和显示等。

设非负函数  $f: X \in [0, +\infty)$ ,  $g(\cdot)$  是在  $X$  上的模糊测度,  $f$  关于  $g(\cdot)$  的 Choquet 模糊积分<sup>[23]</sup> 定义为

$$\int_X f(x) g(\cdot) = \int_0^\infty g(F_a) da, \quad (9)$$

式中: $F_a = \{x | f(x) \geq a, x \in X_j\}$ ,  $x$  取(7)式中不同偏振参量的属性。

当  $X$  是一个有限集合,取值同(7)式的  $X_j$ ,且  $f: X \rightarrow [0, 1]$  时,Choquet 模糊积分相应变为

$$\int f(X_j) dg = \sum_{j=1}^n [f(X_j) - f(X_{j-1})] \times g(X_j), \quad (10)$$

式中: $f(X_j)$  为信任函数; $g(X_j)$  为模糊测度。

由(9)式和(10)式可知,模糊测度利用与模糊测度

相关的模糊积分代替了普通的加权求和法,可以被看作是非线性可加函数。根据(10)式求得模糊积分值,选取最大的模糊积分值作为显著性偏振参量图像。

### 3 深度稀疏特征表示

#### 3.1 分析-合成字典对

字典学习是图像和信号稀疏表示的核心问题,DDL 在过去几年获得了极大发展,而且在计算机视觉领域如目标跟踪应用中取得良好效果,因此字典编码所得到的稀疏特征更适合作为目标检测分类器。将判别式特征学习与稀疏特征表示模型相结合,能够充分利用 CNN 的复杂特征结构以提高目标检测效果<sup>[24]</sup>。

在 CNN 框架下进行目标分类需要两个字典,一个是分析字典  $\mathbf{P}' = [\mathbf{P}_0, \dots, \mathbf{P}_k, \dots, \mathbf{P}_K] \in \mathbf{R}^{m(K+1) \times d}$ ,其中  $K$  表示分类数目, $\mathbf{P}_k \in \mathbf{R}^{m \times d}$  表示第  $k$  个分类子字典, $m$  是字典原子的数量,该字典用来对输入数据  $X$  进行分析编码。对于输入图像区域  $I$ ,令  $\mathbf{X} = [\mathbf{X}_0, \dots, \mathbf{X}_k, \dots, \mathbf{X}_K]$ ,其中, $n_k$  为第  $k$  类训练样本数量, $\mathbf{X}_k \in \mathbf{R}^{d \times n_k}$  为第  $k$  类训练样本。另外一个为合成字典  $\mathbf{D} = [\mathbf{D}_0, \dots, \mathbf{D}_k, \dots, \mathbf{D}_K] \in \mathbf{R}^{d \times m(K+1)}$ , $\mathbf{D}_k \in \mathbf{R}^{d \times m}$  表示第  $k$  个分类子字典,用来对  $X$  进行重建。由此,第  $k$  个类别的分析-合成字典对由  $\mathbf{P}_k$  和  $\mathbf{D}_k$  构成。

与用于图像超分辨率的高-低分辨率双字典学习方法<sup>[25]</sup>相似,同时学习分析-合成双字典  $\mathbf{P}'$  和  $\mathbf{D}$  的目标函数描述为

$$\{\mathbf{P}^*, \mathbf{D}^*\} = \arg \min_{\mathbf{P}, \mathbf{D}} \sum_{k=0}^K \left\| (\mathbf{X}_k - \mathbf{D}_k \mathbf{P}_k \mathbf{X}_k) \mathbf{W}_k \right\|_{\text{F}}^2 + \lambda \left\| \mathbf{P}_k \bar{\mathbf{X}}_k \right\|_{\text{F}}^2 + \kappa \left\| \mathbf{D}_K \right\|_{\text{F}}^2, \quad (11)$$

式中:上标 \* 表示目标预测值(下文同); $\lambda > 0, \kappa > 0$ ,都是平衡参数; $\bar{\mathbf{X}}_k$  是  $\mathbf{X}_k$  相对于整个训练集  $X$  的互补数据矩阵; $k$  表示训练样本的第  $k$  个分类;为避免求解过程中出现零解  $\mathbf{P}_k = 0$ ,增加了  $\left\| \mathbf{D}_K \right\|_{\text{F}}^2$  约束项; $\mathbf{W}_k$  是对角权值矩阵,表示在目标分类时不同的训练样本在训练判别式模型中的作用权重不同, $\mathbf{W}_k$  的引入能够有效提高目标检测定位的性能,通过给样本高权重可以得到更好的定位效果。

在得到分析-合成字典对  $\mathbf{P}_k$  和  $\mathbf{D}_k$  后,合成字典  $\mathbf{D}_k$  对应的第  $k$  个训练样本  $\mathbf{X}_k$  的编码系数  $\mathbf{A}_k$  可表示为  $\mathbf{A}_k = \mathbf{P}_k \mathbf{X}_k$ ,在此过程中,给较低的重建残差赋予高权重。

#### 3.2 字典对学习

为求解(11)式,将约束条件进一步放松,引入编

码系数矩阵  $\mathbf{A}$ ,此时的求解模型为

$$\{\mathbf{P}^*, \mathbf{A}^*, \mathbf{D}^*\} = \arg \min_{\mathbf{P}, \mathbf{A}, \mathbf{D}} \sum_{k=0}^K \left\| (\mathbf{X}_k - \mathbf{D}_k \mathbf{A}_k) \mathbf{W}_k \right\|_{\text{F}}^2 + \tau \left\| (\mathbf{P}_k \mathbf{X}_k - \mathbf{A}_k) \mathbf{W}_k \right\|_{\text{F}}^2 + \lambda \left\| \mathbf{P}_k \bar{\mathbf{X}}_k \right\|_{\text{F}}^2 + \kappa \left\| \mathbf{D}_K \right\|_{\text{F}}^2, \quad (12)$$

式中, $\tau$  是标量参数。

由于目标函数的所有项都以相同的 Frobenius 范数来刻画,可以采用交替最小化算法进行求解。在初始化阶段,利用归一化 Frobenius 范数构成随机矩阵,以生成  $\mathbf{P}'$  和  $\mathbf{D}$  的初始化值。交替最小化算法的基本思想是当求解其中一个目标变量时固定其余变量,其求解过程的详细描述见文献<sup>[24]</sup>。

#### 3.3 CNN 参数更新

在 CNN 的端到端工作模式下,当利用基于字典对的稀疏特征进行目标分类时,需要进行字典对反向传播计算以联合完成分类器层设计与 CNN 参数训练。由于字典对  $(\mathbf{P}_k, \mathbf{D}_k)$  能够分开独立求解,(11)式可以分解为  $K+1$  个子问题,表示为

$$\arg \min_{\mathbf{P}_k, \mathbf{D}_k} \mathbf{L}_k(\mathbf{P}_k, \mathbf{D}_k) = \arg \min_{\mathbf{P}_k, \mathbf{D}_k} \left\| (\mathbf{X}_k - \mathbf{D}_k \mathbf{P}_k \mathbf{X}_k) \mathbf{W}_k \right\|_{\text{F}}^2 + \lambda \left\| \mathbf{P}_k \bar{\mathbf{X}}_k \right\|_{\text{F}}^2 + \kappa \left\| \mathbf{D}_K \right\|_{\text{F}}^2, \quad (13)$$

式中, $\mathbf{L}_k$  为第  $k$  个子问题。

将(13)式分别对  $\{\mathbf{P}_k, \mathbf{D}_k\}$  求偏导数,得到

$$\begin{cases} \frac{\partial \mathbf{L}_k}{\partial \mathbf{P}_k} = -2\mathbf{D}_k (\mathbf{I} - \mathbf{D}_k \mathbf{P}_k) \mathbf{X}_k \mathbf{W}_k \mathbf{W}_k^T \mathbf{X}_k^T + 2\lambda \mathbf{P}_k \bar{\mathbf{X}}_k \bar{\mathbf{X}}_k^T \\ \frac{\partial \mathbf{L}_k}{\partial \mathbf{D}_k} = -2\mathbf{X}_k \mathbf{W}_k (\mathbf{I} - \mathbf{D}_k \mathbf{P}_k) \mathbf{W}_k^T \mathbf{X}_k^T \mathbf{P}_k^T + 2\kappa \mathbf{D}_k \end{cases} \quad (14)$$

$$\mathbf{L} = \sum_{k=0}^K \mathbf{L}_k, \text{ 其对 } \mathbf{X}_k \text{ 的偏导数为}$$

$$\frac{\partial \mathbf{L}}{\partial \mathbf{X}_k} = 2(\mathbf{I} - \mathbf{P}_k^T \mathbf{D}_k^T) (\mathbf{X}_k - \mathbf{D}_k \mathbf{P}_k \mathbf{X}_k) \cdot \mathbf{W}_k \mathbf{W}_k^T + \sum_{k' \neq k} 2\lambda \mathbf{P}_{k'}^T \mathbf{P}_{k'} \mathbf{X}_k, \quad (15)$$

式中, $k'$  表示除了第  $k$  个之外的其余所有  $k-1$  个子问题。

求得所有的  $\frac{\partial \mathbf{L}}{\partial \mathbf{X}_k}$  后,直接利用标准反向传播算法<sup>[26]</sup>就可以完成 CNN 的参数更新任务。

## 4 字典对分类器层

在 CNN 框架下利用稀疏特征进行目标检测有两

个主要过程,首先由 CNN 提取偏振参量图像特征,然后由分类器层进行目标检测和定位,而分类器层又进一步分为目标分类器和类别分类器。目标分类器负责将目标从背景中分辨出来,而类别分类器则用于判定目标所属类别,两个分类器都由字典对  $\{\mathbf{D}, \mathbf{P}\}$  来构建。给定偏振参量图像上一个候选区域  $I$ ,从  $I$  提取的 CNN 特征为  $\mathbf{x}_f$ ,第  $k$  类别的重构残差定义为

$$\mathcal{L}(\mathbf{x}_f; \mathbf{D}_k, \mathbf{P}_k) = \|\mathbf{x}_f - \mathbf{D}_k \mathbf{P}_k \mathbf{x}_f\|_{\text{F}}^2. \quad (16)$$

而目标分类规则定义为

$$y = \arg \min_{i'} \mathcal{L}(\mathbf{x}_f; \mathbf{D}_{i'}, \mathbf{P}_{i'}), \quad (17)$$

式中:  $i' \neq k$ ,表示除了第  $k$  类别外的其他分类字典,当

$y \neq 0$  时利用联合任务回归对目标框进行重定位更新。

#### 4.1 目标分类器

在目标分类时,目标定义为涵盖所有类别的目标得分。为了判定输入区域内的目标,目标分类器层分别用字典对  $\{\mathbf{D}_o, \mathbf{P}_o\}$  表示所有类别,字典对  $\{\mathbf{D}_b, \mathbf{P}_b\}$  表示背景。如果待检测区域的特征  $\mathbf{x}_f$  能够更好地被背景字典对  $\{\mathbf{D}_b, \mathbf{P}_b\}$  表示,则在该区域里有目标的概率就非常小。为了进一步地进行目标检测,目标分类器层使用阈值参数  $\gamma$  来判别具有大范围背景的区域。根据 (16) 式重建残差,待检测区域的特征  $\mathbf{x}_f$  的目标分数  $Q(\mathbf{x}_f)$  被定义为

$$Q(\mathbf{x}_f) = \begin{cases} 1 - \frac{\mathcal{L}(\mathbf{x}_f; \mathbf{D}_o, \mathbf{P}_o)}{\sum_{i^* \in \{o, b\}} \mathcal{L}(\mathbf{x}_f; \mathbf{D}_{i^*}, \mathbf{P}_{i^*})}, & \frac{\mathcal{L}(\mathbf{x}_f; \mathbf{D}_o, \mathbf{P}_o)}{\mathcal{L}(\mathbf{x}_f; \mathbf{D}_b, \mathbf{P}_b)} < \gamma \\ 0, & \text{others} \end{cases}, \quad (18)$$

式中,参数  $\gamma$  用来控制检测精度和检测背景召回率, $\gamma$  值越大,准确度越高而召回率越低。因此,目标分类器可根据  $Q(\mathbf{x}_f)$  是否为零值来识别背景。

#### 4.2 类别分类器

为计算目标所属类别概率,类别分类器由  $K$  个字典对组成, $K$  是目标类别数量,对于待检测区域的特征  $\mathbf{x}_f$ ,类别分类器对  $\mathbf{x}_f$  在  $K$  个类别字典对  $\{\mathbf{D}_k, \mathbf{P}_k\}$  上分别进行稀疏编码,得到每个字典对下  $\mathbf{x}_f$  的稀疏重构残差,而类别分数  $S(\mathbf{x}_f, k)$  表示特征  $\mathbf{x}_f$  属于第  $k$  类的概率, $S(\mathbf{x}_f, k)$  可以由稀疏重建残差来定义:

$$S(\mathbf{x}_f, k) = \frac{\mathcal{L}(\mathbf{x}_f; \mathbf{D}_k, \mathbf{P}_k) \exp[\varphi \mathcal{L}(\mathbf{x}_f; \mathbf{D}_k, \mathbf{P}_k)]}{\sum_{i^\# = 1}^K \mathcal{L}(\mathbf{x}_f; \mathbf{D}_{i^\#}, \mathbf{P}_{i^\#}) \exp[\varphi \mathcal{L}(\mathbf{x}_f; \mathbf{D}_{i^\#}, \mathbf{P}_{i^\#})]}, \quad (19)$$

式中, $\varphi$  为调节参数。

将目标分数  $Q(\mathbf{x}_f)$  和类别分数  $S(\mathbf{x}_f, k)$  进行乘法融合,那么  $\mathbf{x}_f$  属于第  $k$  类的分类分数  $\mathbf{F}_k$  定义为

$$\mathcal{F}_k(\mathbf{x}_f) = S(\mathbf{x}_f, k) Q(\mathbf{x}_f). \quad (20)$$

由分类得分  $\mathcal{F}_k$  得最终分类损失为

$$L_{\text{cls}}(I) = \sum_{k=0}^K \mathbf{1}(y=k) \log_2 \mathcal{F}_k[\phi(I, \omega)] + [1 - \mathbf{1}(y=k)] \log_2 \{1 - \mathcal{F}_k[\phi(I, \omega)]\} + R\{\omega, \mathbf{D}, \mathbf{P}'\}, \quad (21)$$

式中:  $\phi$  表示 CNN 层函数; $I$  表示带有类别标签  $y$  的待检测区域,则  $\mathbf{x}_f = \phi(I, \omega)$ , $\omega$  表示 CNN 的微调

参数; $\mathbf{1} \in \{0, 1\}$  表示指示函数; $R\{\omega, \mathbf{D}, \mathbf{P}'\}$  表示 CNN 参数和两个分类器的正则化项。

#### 4.3 联合任务回归

在利用字典对分类器进行目标检测时,由于目标、分类和定位是密不可分的三个任务。综合考虑这三个方面,定义联合多任务损失函数以提高目标检测算法性能。令  $t^k(I) = (t_x^k, t_y^k, t_w^k, t_h^k)$  和  $t^*(I) = (t_x^*, t_y^*, t_w^*, t_h^*)$  分别表示待检测区域  $I$  的估计和理想目标框,其中, $k$  表示  $I$  属于第  $k$  个目标类别。目标框的回归损失函数定义为

$$L_{\text{loc}}[t^k(I), t^*(I)] = \sum_{j^* \in x, y, w, h} H_1(t_j^{k*} - t_j^{**}), \quad (22)$$

式中: $H_1(z)$  是对异常值具有稳健性的 Huber 损失项; $z = t_j^{k*} - t_j^{**}$ ,表示估计值与理想值误差,该损失项定义为

$$H_1(z) = \begin{cases} 0.5z^2, & |z| < 1 \\ |z| - 0.5, & \text{others} \end{cases}. \quad (23)$$

综合  $L_{\text{cls}}$  和  $L_{\text{loc}}$ , $p_l^*$  表示第  $l$  个检测区域  $I_l$  是否为目标的指示标签,则联合多任务损失函数定义为

$$L_{\text{mt}} = -\frac{1}{N} \left\{ \sum_{l=1}^N L_{\text{cls}}(I_l) + p_l^* L_{\text{loc}}[t^k(I_l), t^*(I_l)] \right\}. \quad (24)$$

在字典  $\mathbf{D}_b, \mathbf{P}_b, \mathbf{D}_o, \mathbf{P}_o, \mathbf{D}_k, \mathbf{P}_k$  以及数据集  $X_k$  域内,能够获得  $L_{\text{mt}}$  的偏微分方程数值解,从而可以反向回归到 CNN 和字典分类器,以更新字典对、

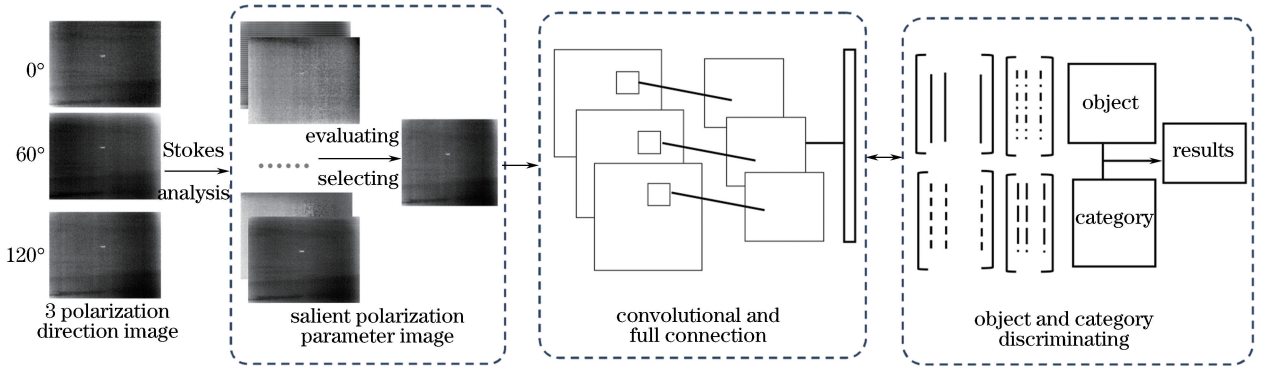


图1 目标检测框架

Fig. 1 Object detection framework

CNN 参数以及目标框回归。

## 5 目标检测算法

### 5.1 目标检测框架

整个目标检测框架由显著性偏振参量图像构造、CNN 特征提取及目标与类别判定 3 个模块组成,如图 1 所示。在显著性偏振参量图像构造阶段,由 Stokes 公式解析获得各偏振参量  $I$ 、 $Q$ 、 $U$ 、 $P$ 、 $A$ 、 $E_x$ 、 $E_y$ 、 $\Delta E$ 、 $\beta$  图。用标准差、信息熵和清晰度 3 个属性对偏振参量图像进行评价,通过信任函数和 Choquet 模糊积分自适应选择显著性偏振参量图像。

CNN 模块由卷积层和全连接层构成,用于提取图像特征。目标与类别判定模块基于 CNN 提取的图像特征,分为目标分类器字典对和类别分类器字典对。利用分类器字典对得到评估分数,判断是否

为目标;利用类别分类器字典对计算特定目标类别的得分。每个图像区域的最后得分是目标和类别的组合。在目标与类别判定过程中,需要进行字典对反向传播计算以联合完成分类器层的字典对更新、边界框回归以及 CNN 参数优化微调。

### 5.2 目标检测算法

在 CNN 架构下,目标检测由卷积网络训练和目标检测两部分组成。CNN 的初始参数通过 ImageNet 模型<sup>[2]</sup>预先训练,字典  $\mathbf{P}$  和  $\mathbf{D}$  的初始值由归一化 Frobenius 范数生成随机矩阵。在使用预先训练的网络初始化 CNN 参数,获得  $L_{mi}$  分别对  $\mathbf{D}_b$ 、 $\mathbf{P}_b$ 、 $\mathbf{D}_o$ 、 $\mathbf{P}_o$ 、 $\mathbf{D}_k$ 、 $\mathbf{P}_k$ 、 $\mathbf{X}_k$  的偏导数后,利用 3.3 节的字典对反向传播来进行端到端的字典对更新、CNN 参数微调 and 边界框回归。结合分析,基于显著性偏振参量图像深度稀疏特征学习的偏振图像目标检测流程描述如图 2 所示。

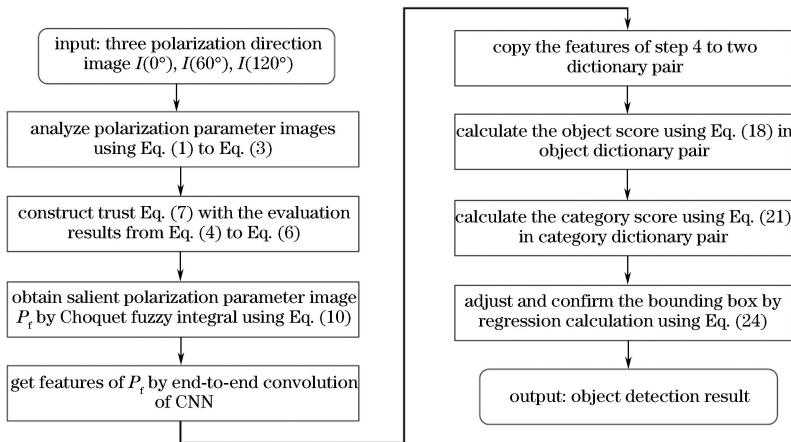


图2 显著性偏振参量图像目标检测算法

Fig. 2 Object detection algorithm for salient polarization parameter image

算法运行效率主要受 3 个方面的影响。首先是显著性选择,由于偏振解析及模糊积分是数值计算,无需循环迭代,具有实时性;其次是特征计算及目标

和类别判定,该过程主要取决于 CNN 卷积层和全链接层的时间复杂度:  $O(2 \times C_i \times K_i^2 - 1) \times H \times W \times C_o + O(2 \times I_{in} - 1) \times I_o$ , 其中,  $C_i$  表示输入通

道,  $K_j$  表示卷积核大小,  $H \times W$  表示输出特征大小,  $C_c$  表示输出通道, 而  $I_{in}$  和  $I_o$  分别表示输入输出神经元数; 第三个方面是最后的回归计算, 可以略去。故该算法的时间复杂度与 CNN 相当, 但能够提高针对偏振图像目标的检测效果。

## 6 实验与分析

### 6.1 实验说明

实验验证主要从两个方面进行, 一是在全场景仿真实验室内模拟无人机平台对地面目标低空侦察, 目标对象为缩比仿真模型; 二是实测对空目标,



图3 测试数据( $0^\circ$ 方向)。(a)民航飞机;(b)草地坦克;(c)沙地卡车

Fig. 3 Test data ( $0^\circ$  polarization direction). (a) Airplane; (b) tank; (c) truck

实验环境为戴尔 PRECISION TOWER 5810 工作站, 具体参数为: Intel(R) Xeon(R) E5-1660 v4 3.2 GHz, 32.0 GB RAM, 8 GB SGRAM, Windows 8(64 位)专业版+Matlab(R2016a)。由于偏振目标检测任务明确, 故将分类器字典的类别参数  $K$  设为 20; 字典更新停止条件是两次相邻迭代的误差值小于设定阈值, 实验中设置阈值为 0.01 能够在算法效率和字典性能之间获得比较平衡的效果; 将(18)式中用来控制检测精度和检测背景召回率的  $\gamma$  值, 以及(19)式中类别分数计算的调节参数  $\varphi$ , 参考原始方法<sup>[24]</sup>分别设置为 0.5 和 0.003。

以 Faster R-CNN<sup>[3]</sup> 为基本模型框架, CNN 参数在 ImageNet<sup>[2]</sup> 上进行预训练, 然后利用采集的数

以民航飞机为对象进行数据获取并进行实验验证。其中, 室内仿真数据采集主要是样本采集和训练, 目标有草地坦克和沙地卡车等, 这些目标在低空场景下能够被识别, 在良好气象条件下对民航飞机小目标进行跟踪采集, 以进行训练样本准备。实验设备为自行研制的 3 方向及红外同时成像偏振相机, 采集的部分数据如图 3 所示, 图中数据都是  $0^\circ$  方向的偏振图像, 为方便显示, 对图像尺寸进行裁剪, 原图像分辨率分别为: 图 3(a)为 576 pixel  $\times$  447 pixel, 图 3(b)为 808 pixel  $\times$  608 pixel, 图 3(c)为 780 pixel  $\times$  580 pixel。

据进行验证微调, 其他参数设置采用文献<sup>[24]</sup>的默认值。在微调过程中的背景与目标的判定阈值设定上, 以候选框与原标记框的交叠率, 即交并比(IoU)  $\alpha_{IoU}$  为调节阈值, 所有  $\alpha_{IoU} < 0.5$  判定为背景, 而  $\alpha_{IoU} \geq 0.5$  判定为目标。

### 6.2 实验结果比较和分析

#### 1) 显著性选择实验

实验首先验证显著性偏振参量图像选择情况。图 4 和表 1 是部分结果展示, 分别称为飞机 1 (airplane 1)、飞机 2 (airplane 2) 和卡车 (truck) 等, 图 4 中第一行是各目标的  $0^\circ$  偏振方向图像, 第二行是对应的显著性选择结果。可以看出, 不同的背景和目标下, 显著性偏振参量图像选择结果不同。

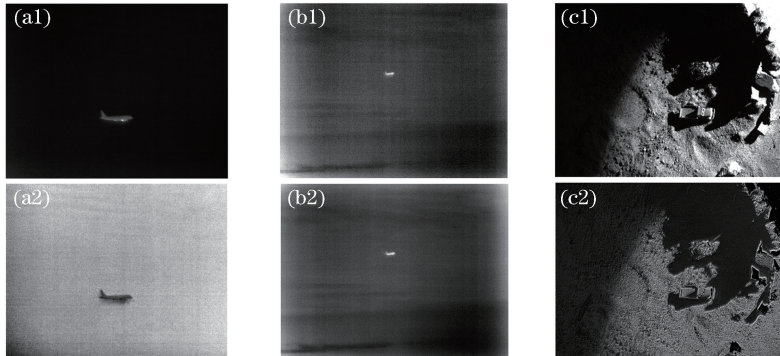


图4 显著性参量图像选择结果。(a1)~(c1)飞机 1, 飞机 2, 卡车图像;(a2)~(c2)显著性选择结果

Fig. 4 Results of salient parameter image selection. (a1)-(c1) Images of airplane 1, airplane 2, and truck; (a2)-(c2) salient selected results

表 1 为偏振解析显著性参量图像选择前后的标准差  $\sigma$ 、信息熵  $E_n$  和清晰度  $\bar{g}$  结果。由实验数据可知,如果仅仅根据某一个指标值,无法判定图像质量,如根据信息熵可以看出三个图像的显著性选择后的值都增加,但是其清晰度并不是都提升,卡车图像的清晰度反而下降,标准差亦是如此。

表 1 偏振解析显著性参量图像选择前后评价结果

Table 1 Results before and after polarization analysis and salient parameter image selection

| Criteria  | Airplane 1 |         | Airplane 2 |         | Truck |         |
|-----------|------------|---------|------------|---------|-------|---------|
|           | 0°         | Salient | 0°         | Salient | 0°    | Salient |
| $E_n$     | 3.75       | 6.84    | 5.29       | 6.54    | 6.22  | 6.77    |
| $\bar{g}$ | 0.31       | 0.29    | 0.36       | 0.39    | 0.87  | 0.61    |
| $\sigma$  | 5.55       | 29.94   | 30.22      | 25.65   | 75.69 | 44.19   |

根据上述三个指标综合评价来选择待检测图像时,在 3.2 节的模糊积分中,不同的目标或同一目标在不同的背景条件下,选择的偏振参量图像应该不同。由实验结果可以看出,三组结果的显著性目标选择分别为偏振度图像、 $I$  图像以及  $E_y$  图像,该结果在视觉效果上能够更好的凸显目标,利于下一步的目标检测和判定。

2) 目标检测实验

在目标检测实验中,分别对同一目标及不同目标在不同场景下进行检测,并与直接利用偏振方向图像检测结果进行比较,实验中每种目标分 8 组,每组 10 个样本,室内样本采集时主要模拟不同光照、雾气等变化条件下的图像数据,室外实际样本采集时主要采集不同距离和天气状况下的民航飞机图像数据,检测结果如图 5 所示,图中同时给出目标检测得分情况,为方便显示,各图是在保留目标的情况下将背景图像裁剪,图中显示的是裁剪后的图像尺寸。可以看出,显著性偏振参量图像在一定程度上降低了背景的复杂程度,使得目标更为突出,这种情况下检测得分总体高于直接偏振方向图像。就人眼视觉效果判断而言,目标的整体轮廓如坦克炮管,能够比较清晰的凸显。但是在局部细节上,如图 5(a2) 中的发动机尾焰部分在该显著性参量(偏振度)图像中丢失,而该飞机目标较为清晰的发动机尾焰参量图像是  $Q$  参量图像,如图 5(a2) 中小图所示,在模糊积分评价下并未选择该偏振参量图像作为显著性目标检测图像。

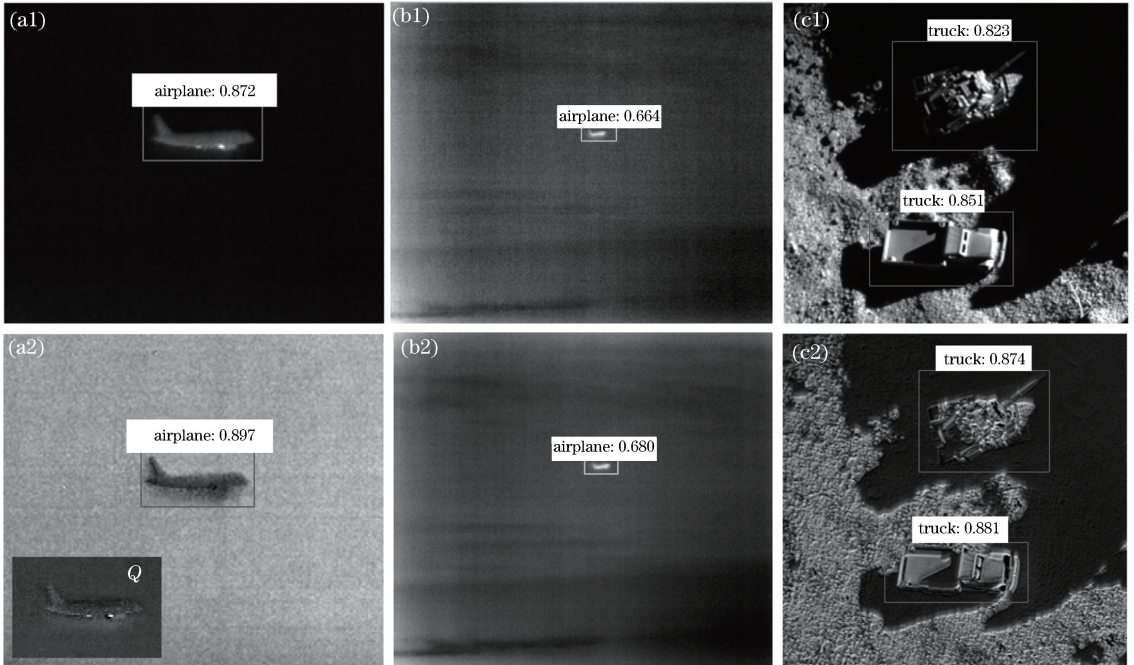


图 5 目标检测结果。(a1)~(c1)飞机 1,飞机 2,卡车图像;(a2)~(c2)显著性图像检测结果  
Fig. 5 Results of object detection. (a1)-(c1) Images of airplane 1, airplane 2, and truck;  
(a2)-(c2) salient image detection results

为进一步评价检测效果,将解析前的各偏振方向图像以及各偏振参量图像的各目标检测精度以均值平均精度(mAP)进行客观评价,表 2 中展示了总体 mAP 以及各类别的平均精度(AP)值,其中黑体

表示同组的优势数据,N 表示该次检测失败,表中 tank 1 表示草地坦克,tank 2 表示沙地坦克。结果表明,各目标的直接偏振方向检测结果总体高于偏振参量检测结果,但是就某一目标来说,显著性偏振



表 2 不同图像目标检测结果比较

Table 2 Comparison of detection results of different image objects

| Object                     | mAP   | AP          |             |             |             |             |
|----------------------------|-------|-------------|-------------|-------------|-------------|-------------|
|                            |       | Airplane 1  | Airplane 2  | Tank 1      | Truck       | Tank 2      |
| Polarization angle of 0°   | 63.88 | 69.3        | <b>61.2</b> | 71.4        | 66.4        | 51.1        |
| Polarization angle of 60°  | 63.08 | 68.5        | 59.8        | 70.6        | 65.9        | 50.6        |
| Polarization angle of 120° | 63.42 | 69.0        | 60.3        | 70.8        | 66.3        | 50.7        |
| <i>I</i>                   | 63.72 | 70.2        | 60.1        | <b>72.3</b> | 65.8        | 50.2        |
| <i>Q</i>                   | 60.64 | 64.4        | 57.4        | 68.8        | 61.3        | 51.3        |
| <i>U</i>                   | 53.7  | 64.9        | 37.6        | 62.9        | 54.4        | 48.7        |
| <i>P</i>                   | 60.46 | <b>70.3</b> | 45.3        | 70.1        | 67.1        | 49.5        |
| <i>A</i>                   | 52.24 | 59.4        | 37.7        | 58.8        | 60.7        | 44.6        |
| <i>E<sub>x</sub></i>       | 61.18 | 70.0        | 57.6        | 67.0        | 63.5        | 47.8        |
| <i>E<sub>y</sub></i>       | 59.56 | 64.7        | 56.8        | 56.2        | <b>67.9</b> | <b>52.2</b> |
| $\Delta E$                 | 57.74 | 64.6        | 53.6        | 55.3        | 65.4        | 49.8        |
| $\beta$                    | 31.85 | N           | N           | N           | 23.4        | 40.3        |

参量检测结果普遍高于偏振方向图像检测结果,如表 2 中 5 组目标,其中 4 组显著性偏振参量检测结果高于偏振方向图像检测结果。同时,表 2 中数据也表明各偏振参量检测精度差异较大,这是因为各偏振参量本身就是从不同维度描述图像目标的物理特征,形成对目标信息刻画的优势互补。

另外,为验证字典对分类器效果,将本方法与基本 CNN 框架 Faster R-CNN<sup>[9]</sup> 进行比较,检测结果如表 3 所示。以显著性偏振参量图像为实验对象,score 表示

表 3 不同模型的检测结果比较

Table 3 Comparison of detection results of different models

| Model        | Time /s | mAP/average | Airplane 1 | Airplane 2 | Tank       |
|--------------|---------|-------------|------------|------------|------------|
|              |         |             | AP/score   | AP/score   | AP/score   |
| Faster R-CNN | 0.7     | 66.1/0.741  | 68.1/0.842 | 59.5/0.619 | 70.6/0.762 |
| Proposed     | 24      | 67.9/0.819  | 70.3/0.897 | 61.2/0.680 | 72.3/0.881 |

## 7 结 论

针对偏振成像目标探测中的偏振参量图像丰富、目标显示及识别任务在不同的目标及背景下用某一种固定偏振参量图像无法达到实际需求的问题,提出基于显著性偏振参量图像深度稀疏特征学习的偏振图像目标检测方法。通过标准差、信息熵和清晰度构造模糊积分,得到显著性偏振参量,并以稀疏表示的字典对改进卷积神经网络框架中的分类器以用于目标检测,利用反向传播对分类器进行端到端学习和 CNN 参数更新。对实际目标探测中的典型目标和背景数据进行测试,验证了其有效性。但是在实际的目标探测任务中,目标及背景会更加复杂,如何充分利用探测目标的物理特性及偏振成像数据特点、结合快速 CNN 研究成果并利用不同偏振参量图像的冗余互

检测得分。从实验数据可以看出,利用字典对进行稀疏特征学习然后进行目标分类和定位的方法,对边界框的标定更为准确,目标的得分普遍提高。就检测精度来说,和采用 softmax 分类器的 Faster R-CNN 方法相比,本文方法在本次实验的检测效果上有所提高。在运行效率上,由于本方法运行时间主要取决于 CNN 时间复杂度,表 3 中的运行时间是以 576 pixel × 447 pixel 的飞机图像检测为例,可以看出本文方法与 Faster R-CNN 相比在运行效率上不占优势。

补来提高预训练效果和设计特征学习模型,从而提高检测性能和效率将是下一步的研究重点。

## 参 考 文 献

- [1] Wang X L, Wang F, Liu X, *et al.* Hyperspectral polarization characteristics of typical camouflage target under desert background [J]. *Laser & Optoelectronics Progress*, 2018, 55(5): 051101.  
王小龙, 王峰, 刘晓, 等. 荒漠背景下典型伪装目标的高光谱偏振特性[J]. *激光与光电子学进展*, 2018, 55(5): 051101.
- [2] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C] // 25th International Conference on Neural Information Processing Systems, December 3-6, 2012, Lake Tahoe, Nevada. New York: ACM,

- 2012: 1097-1105.
- [3] Girshick R. Fast R-CNN [C] // 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 1440-1448.
- [4] Girshick R, Donahue J, Darrell T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation [C] // 2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE, 2014: 580-587.
- [5] Canny J. A computational approach to edge detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 8(6): 679-698.
- [6] Rosten E, Porter R, Drummond T. Faster and better: a machine learning approach to corner detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(1): 105-119.
- [7] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [8] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C] // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2005, San Diego, CA, USA. New York: IEEE, 2005: 886-893.
- [9] Felzenszwalb P F, Girshick R B, McAllester D, *et al.* Object detection with discriminatively trained part-based models [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9): 1627-1645.
- [10] Yan J J, Zhang X C, Lei Z, *et al.* Robust multi-resolution pedestrian detection in traffic scenes [C] // 2013 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2013, Portland, OR, USA. New York: IEEE, 2013: 3033-3040.
- [11] Li J X, Zong Q. Object tracking based on multi-feature and local joint sparse representation [J]. Laser & Optoelectronics Progress, 2017, 54(10): 101502. 李敬轩, 宗群. 基于多特征和局部联合稀疏表示的目标跟踪 [J]. 激光与光电子学进展, 2017, 54(10): 101502.
- [12] Andriluka M, Iqbal U, Insafutdinov E, *et al.* Pose Track: a benchmark for human pose estimation and tracking [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 5167-5176.
- [13] Huang K Q, Ren W Q, Tan T N. A review on image object classification and detection [J]. Chinese Journal of Computers, 2014, 37(6): 1225-1240. 黄凯奇, 任伟强, 谭铁牛. 图像物体分类与检测算法综述 [J]. 计算机学报, 2014, 37(6): 1225-1240.
- [14] Mairal J, Bach F, Ponce J. Task-driven dictionary learning [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(4): 791-804.
- [15] Feng Z Z, Yang M, Zhang L, *et al.* Joint discriminative dimensionality reduction and dictionary learning for face recognition [J]. Pattern Recognition, 2013, 46(8): 2134-2143.
- [16] Yang M, Zhang L, Feng X C, *et al.* Sparse representation based fisher discrimination dictionary learning for image classification [J]. International Journal of Computer Vision, 2014, 109(3): 209-232.
- [17] Jiang Z L, Lin Z, Davis L S. Label consistent K-SVD: learning a discriminative dictionary for recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(11): 2651-2664.
- [18] Chen B, Zha Y F, Li Y Q, *et al.* Person re-identification based on convolutional neural network discriminative feature learning [J]. Acta Optica Sinica, 2018, 38(7): 0720001. 陈兵, 查宇飞, 李运强, 等. 基于卷积神经网络判别特征学习的行人重识别 [J]. 光学学报, 2018, 38(7): 0720001.
- [19] Ramirez I, Sprechmann P, Sapiro G. Classification and clustering via dictionary learning with structured incoherence and shared features [C] // 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 13-18, 2010, San Francisco, CA, USA. New York: IEEE, 2010: 3501-3508.
- [20] Gao S H, Tsang I W H, Ma Y. Learning category-specific dictionary and shared dictionary for fine-grained image categorization [J]. IEEE Transactions on Image Processing, 2014, 23(2): 623-634.
- [21] Gu S H, Zhang L, Zuo W M, *et al.* Projective dictionary pair learning for pattern classification [C] // 27th International Conference on Neural Information Processing Systems, December 8-13, 2014, Montreal, Canada. Canada: NIPS, 2014: 793-801.
- [22] Liu F, Shen T S, Lou S L, *et al.* Deep network saliency detection based on global model and local optimization [J]. Acta Optica Sinica, 2017, 37(12): 1215005.

- 刘峰, 沈同圣, 娄树理, 等. 全局模型和局部优化的深度网络显著性检测[J]. 光学学报, 2017, 37(12): 1215005.
- [23] Mesiar R. Fuzzy measures and integrals[J]. Fuzzy Sets and Systems, 2005, 156(3): 365-370.
- [24] Wang K Z, Lin L, Zuo W M, *et al.* Dictionary pair classifier driven convolutional neural networks for object detection [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 2138-2146.
- [25] Yang J C, Wang Z W, Lin Z, *et al.* Coupled dictionary training for image super-resolution [J]. IEEE Transactions on Image Processing, 2012, 21(8): 3467-3478.
- [26] Cun Y L, Boser B, Denker J S, *et al.* Handwritten digit recognition with a back-propagation network [M] // Touretzky D S. Advances in neural information processing systems 2. San Francisco: Morgan Kaufmann Publishers Inc., 1989: 396-404.