第 39 卷 第 3 期
2020 年 5 月

应用声学
Journal of Applied Acoustics

Vol. 39, No. 3
May, 2020

◇ 研究报告 ◇

一种基于弱监督学习的声图小目标快速检测方法*

徐利刚[1] 朱可卿[2,3] 韦琳哲[2†] 王 朋[2]

(1 海装驻无锡地区军代室 无锡 214061)

(2 中国科学院声学研究所 北京 100190)

(3 中国科学院大学 北京 100049)

摘要： 小目标检测是声呐图像理解中最引人瞩目，同时又极具挑战性的任务之一。该文基于离散余弦变换和K-近邻聚类，提出了一种快速检测方法。离散余弦变换用于生成图像的指纹，是原始图像在二维频域的一种稀疏表达；改进的K-近邻模型对于带有标签数据的需求量相对较低，提升了算法的处理效率和对弱监督场景的适应性。经试验验证，该方法可在准确率和召回率之间达到一个恰当的平衡点，同时在实时成像的合成孔径声呐图像小目标检测中，获得了较为可靠的结果。

关键词： 小目标检测；合成孔径声呐成像；弱监督学习；离散余弦变换；K-近邻

中图法分类号： TP391.41 **文献标识码：** A **文章编号：** 1000-310X(2020)03-0386-09

DOI: 10.11684/j.issn.1000-310X.2020.03.010

A fast weak supervised detection method of small objects in sonar imagery

XU Ligang[1] ZHU Keqing[2,3] WEI Linzhe[2] WANG Peng[2]

(1 *The Navy's Wuxi Military Representative Office, Wuxi* 214061, *China*)

(2 *Institute of Acoustics, Chinese Academy of Sciences, Beijing* 100190, *China*)

(3 *University of Chinese Academy of Sciences, Beijing* 100049, *China*)

Abstract: Detection of small objects is one of the most attractive and challenging tasks in the comprehension of sonar imagery. In the paper, a fast detection method is presented under a framework of discrete cosine transform (DCT) and K-nearest neighbor (KNN) clustering. DCT is used in the generation of image fingerprint, which contributes a certain spectral sparseness to the original image; and the modified KNN model provides efficiency with a relatively low demand of labeled data. It is shown in a series of experiments that the method we proposed can reach a compromise of precision and recall rate, and achieve considerably reliable detection result on synthetic aperture sonar (SAS) images in real time imaging.

Keywords: Small object detection; Synthetic aperture sonar imagery; Weak-supervised; Discrete cosine transform; K-nearest neighbor

2019-07-08 收稿; 2019-11-28 定稿

*中国科学院青年创新促进会项目

作者简介: 徐利刚 (1985–), 男, 江苏人, 博士, 工程师, 研究方向: 主被动声呐、水声信号处理等。

†通信作者 E-mail: gman200619003@126.com

0　引言

水下图像中小目标的自动检测一直以来是声图解译领域所研究的重要课题，其成果可应用于许多水下探查任务中，包括海底打捞、小目标探测、管线追踪等。文献 [1] 中，目前较为先进的深度学习框架 YOLO 被用于声呐图像的小目标检测，通过大量的训练迭代，取得了较好的检测效果。然而在水下应用场景中，数据的获取往往需要耗费较多的时间和财力。在某些自主检测应用场景下，无论人工辅助判读还是计算机的算力，都受到了一定程度的限制。同时，由于目前对目标先验知识和抵近目标观察手段的缺少，目标的确认十分困难。在大多数情况下，研究人员所掌握的数据集相对较小，标注也不完善，很难对监督学习算法模型进行有效的训练。在这种条件下，无监督或弱监督学习方法则更为适用。Klausner 等 [2] 讨论了一种通过对海底散射建模，来预测水下目标是否存在的方法。Wang 等 [3] 将行人检测领域中的方向梯度直方图 (Histogram of oriented gradient, HOG) 和支持向量机 (Support vector machine, SVM) 用于水下目标的检测。Williams [4] 通过一组模板提取目标的阴影，并根据阴影的特征进行决策。Kim 等 [5] 基于自适应增强算法实现了基于声图的实时水下目标检测。Ma 等 [6] 利用最小二乘支持向量机方法降低了侧扫声图人工目标检测的虚警率。针对掩埋线缆目标的识别问题，Maussang 等 [7] 提出了局部统计参数融合的思想。在合成孔径声呐图像中，目标的阴影与高亮区域也是目标识别的重要特征 [8]。此外，Sawas 等 [9] 提出利用分类器级联方式可实现水下目标快速检测。

一般情况下，声图中小目标的检测主要关注两类课题：(1) 对大范围水下场景中感兴趣区域 (Regions of interests, ROI) 的提取；(2) 为每一可能含有小目标的感兴趣区域赋予正确的标签。这些课题在广义上都可归为对图像中潜在的候选区域的分类问题。为降低遍历图像的时间开销、增强对图像的概括，通常需要进行数据降维，将图像变换至某一特征域，实现对内容的抽象。

本文提出了一种基于离散余弦变换 (Discrete cosine transform, DCT) 和 K 近邻-高斯混合模型 (K-nearest Gaussian mixture model, KN-GMM) 的检测方法。检测过程共分为两个步骤，步骤 1 中，首先从声图中提取一系列包含可疑目标的候选区域，并评估置信度用于辅助决策。这类检测方法可以胜任大多数实时目标的标注任务，但对于有更高精确度要求的任务，需要采用步骤 2 的方法进一步精化结果，即通过候选框回归和决策以精确分析和筛选备选目标，这一处理有效地降低了虚警率，但在计算消耗上有所增加。本文所提出的快速检测方法框架如图 1 所示。

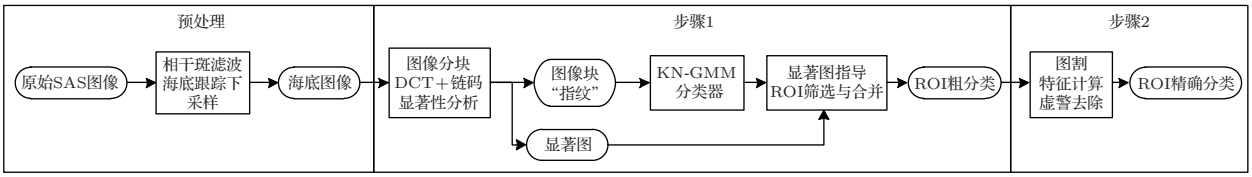


图 1　本文所提出的快速检测方法框架

Fig. 1　Framework of proposed fast detection method

1　小目标快速检测

1.1　图像块的指纹提取

步骤 1 主要完成一系列候选边界框的获取，对声图进行粗略标注。每一帧原始实时图像将被划分为统一尺寸的图像块，算法将预测每个图像块中是否有目标存在。该预测主要基于谱的稀疏性特征和数据集中学习到的知识。最后，相邻的候选图像块将被合并以形成粗略的边界框。

步骤 1 首先需要对图像进行一系列预处理操作。对于合成孔径声呐 (Synthetic aperture sonar, SAS) 图像，由于相干斑噪声的存在，需要进行自适应滤波处理。接着，通过引入图像均衡和增强缓解图像中不同位置散射强度的差异。除此之外，还需使用海底跟踪方法预先将海底和水体区域分离，检测时对二者分别处理，再对图像进行下采样并分成

小的图像块。对图像进行下采样的过程可表示为

$$\hat{\boldsymbol{I}} = \text{downsample}(g(\boldsymbol{I})) = \begin{bmatrix} \hat{\boldsymbol{I}}_{11} & \cdots & \hat{\boldsymbol{I}}_{1n} \\ \vdots & \cdots & \vdots \\ \hat{\boldsymbol{I}}_{m1} & \cdots & \hat{\boldsymbol{I}}_{mn} \end{bmatrix}, \quad (1)$$

其中，$\boldsymbol{I} \in R_{A\times B}$ 表示原始图像，$\hat{\boldsymbol{I}} \in R_{a\times b}$ 表示下采样图像，$\boldsymbol{I}_{ij} \in R_{S\times S}$ 为 $m \times n$ 图像块，$g(\cdot)$ 是自适应相干斑抑制器。

图像块的网格划分将保留一定的重叠范围 (通常为50%)，以减少边界效应的影响。为了使图像块具有相同的尺寸，对图像的最后一行或一列一般会采用更高的重叠量。换言之，m、n 可通过式 (2) 计算：

$$m = 2\left\lceil \frac{a}{S} \right\rceil - 1, \quad n = 2\left\lceil \frac{b}{S} \right\rceil - 1. \quad (2)$$

在声呐图像中，每个像素在物理上有固定的距离和方位分辨率。考虑到这一特点，图像块尺寸 S 的确定需要参考整幅图像的尺寸以及分辨率，满足每个图像块具有恒定的长度 L。设 r 为声呐图像的平均分辨率，约束条件可表达为

$$S = 2^{\left\lceil \log_2 \frac{aL}{Ar} \right\rceil}, \quad (3)$$

这里，为便于DCT的计算，强制 S 为2的整数幂。

接下来，算法将预测每个图像块中是否存在目标，这时面临的主要问题就是特征描述子和分类器的选择。本文采用的特征描述子基于DCT和链码 (chain coding)，如图2所示。DCT生成一个图像块的频谱能量图，其纹理类似于人类指纹，可作为图像块的唯一标识。主成分位于能量图的左上角，代表了稳定的低频区域。因此，从能量图中取出一个4×4矩阵，并用特定的链码对其进行排序，以保留纹理图案：

$$v_{ij} = \omega \cdot \text{chaincode}\left[\text{DCT}\left(\hat{\boldsymbol{I}}_{ij}\right)\right], \quad (4)$$

这里，DCT(·) 代表图像块的离散余弦变换，chaincode(·) 代表链码，亦是DCT矩阵的查找表。

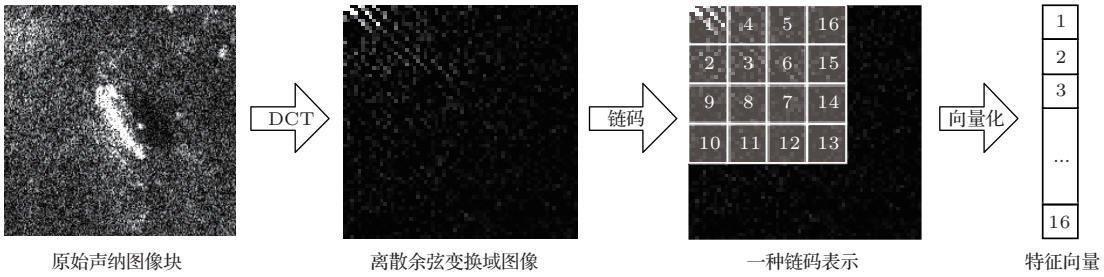


图2 图像块指纹生成示例

Fig. 2 An example of the generation of image block fingerprints

同时，除DCT指纹外，算法还将计算每个图像块的标准差、动态范围和熵，以生成显著图，作为下一步图像指纹分类的正则项。

$$\begin{cases} \sigma_{ij} = \sqrt{\dfrac{1}{S^2}\displaystyle\sum_{k=1}^{S^2}\left(\hat{\boldsymbol{I}}_{ij}^{k} - E\left(\hat{\boldsymbol{I}}_{ij}\right)\right)^2}, \\ d_{ij} = \dfrac{1}{S^2}\displaystyle\sum_{k=1}^{S^2}\left(\max\left(\hat{\boldsymbol{I}}_{ij}^{k}\right) - \min\left(\hat{\boldsymbol{I}}_{ij}^{k}\right)\right)^2, \\ h_{ij} = -\dfrac{1}{N}\displaystyle\sum_{k=1}^{N} p_k^{(ij)} \log\left(p_k^{(ij)}\right). \end{cases} \quad (5)$$

1.2 K近邻-高斯混合模型

将图像块的指纹输入基于KN-GMM的分类器。经典的K-近邻 (K-nearest neighbor, KNN) 模型通过在距离度量 (如欧几里得或曼哈顿距离) 下找到 k 个最近的训练样本，并以多数的标签作为测试样本的标签。由于算法直接采用向量化的图像指纹作为分类器的输入，在没有在特征空间中进行更高级别的抽象，因此在处理小样本问题和含有错误标签数据问题时保留了很好的稳健性。然而，这种简化也导致了一些潜在的问题：(1) 计算量将随着数据集的扩充而增加；(2) 当特征维数增加时，距离度量变得不再有效；(3) 算法不适用于类别不均衡的数据集。

为适用于对较大数据集和类别不均衡数据集的分类任务，对经典KNN模型进行改进，在算法框架中引入预处理 k 均值聚类和GMM以增强该模型对数据的概括能力。

首先，假设某一特定类别的数据服从高斯混合

分布。设数据集为 $\boldsymbol{X}=(x_1,x_2,\cdots,x_N)$，标签为 $\boldsymbol{Y}=(y_1,y_2,\cdots,y_N)$，可被描述为

$$p(\boldsymbol{X}|y_i)=\sum_{l=0}^{L}\alpha_l p(\boldsymbol{X}|\mu_l,y_i)$$
$$=\sum_{l=0}^{L}\alpha_l\exp\left(-\frac{1}{2}(x-\mu_l)^{\mathrm{T}}\sum_{l}^{-1}(x-\mu_l)\right),\quad(6)$$

其中，T 表示转置，这里，强制 α_l 为统一的值以应对数据集中类别不平衡的特点。

不同于经典 KNN 模型采用欧几里德距离度量去寻找 k 个最近邻样本，改进算法通过求解 K 近邻-高斯混合期望向量 $(\mu_1,\mu_2,\cdots,\mu_K)$ 来预测测试样本的标签。标签的预测值 $\hat{y}$ 通过可贝叶斯方程来表示：

$$\hat{y}=\arg\max_{y}p(y|x)$$
$$=\arg\max_{y_i}p(x|y_i)\,p(y_i)$$
$$\approx\arg\max_{y_i}p(y_i)\sum_{k=1}^{K}p(x|\mu_k,y_i)\,p(\mu_k).\quad(7)$$

这里，从 KN-GMM 中计算出的概率可作为图像块的置信度得分：

$$C(x)=p(y_i)\sum_{k=1}^{K}p(x|\mu_k,y_i)\,p(\mu_k).\quad(8)$$

GMM 的训练过程如下：首先，进行 K-均值聚类。具有某一类别标记的数据被分成 L 个聚类以形成 GMM 模型 L 个原始分量。对于每个聚类，估计期望 μ_l 和标准差 $\sum_l$ 来初始化 GMM。接着，通过最大期望 (Expectation maximization, EM) 算法来迭代训练 GMM。GMM 用于预测训练样本，那些预测不正确的样本会反过来修正 GMM 的参数值。直至没有训练样本被错误预测时，迭代结束。KN-GMM 的训练过程如图 3 所示。

对数据进行手动标注时，一般采用矩形框对图像中出现目标的位置进行标注。换言之，一个声图片段的标注中将包含一系列矩形框的坐标和它们所对应的类别标签。在弱监督场景下，未标注区域将被视作“疑似负样本”。对这些未标注或标注不全的样本，将首先借助基于式 (5) 所示的局部统计特征的显著图这一外部正则化条件，找出在局部统计先验下的 L2 范数最大似然标签：

$$y_{\mathrm{MLE}}=\arg\max p(\boldsymbol{X};)$$
$$\propto\arg\min_{y_i}\|\boldsymbol{\Lambda}(X-X_{y_i})\|_2,\quad(9)$$

其中，$\boldsymbol{X}=(\sigma,d,h)^{\mathrm{T}}$ 为未标注块的局部统计量，式 (9) 中采用 L2 范数即为默认该局部统计矢量与各类中心之间的距离符合高斯分布。式 (9) 中权重参数矩阵 $\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_\sigma,\lambda_d,\lambda_h)$ 是一个对角阵，用于调整 3 种局部统计特征的偏好，本文方法直接按照直观理解将其赋为固定值，将三者归一化。也可考虑使用前述的 GMM 模型对这三个参数进行学习。在最大似然标签的监督下，这些未标注或标注不全的样本就可以转化为一般的监督学习场景。

通常，将一个 200 m×400 m 大小、含有目标的声图片段分割成 5 m×5 m、包含 50% 重叠的图像块，所有和标注矩形框相交的图像块都将被视作正样本，这样便可以得到大约包含 6000 个样本的数据集。这种量级的数据足以训练 KN-GMM，可保证在弱监督学习应用场景下，检测系统具有较为可靠的性能。

考虑增量学习的情况，新加入的数据样本将同样首先进行 K-均值聚类，然后用于修正 GMM 模型的参数值。

最后，通过对比置信度的预测矩阵和由每个图像块的标准差、动态范围和熵生成的显著图，剔除上述两个值与其他样本具有较大差异的候选图像块。采用 KN-GMM 和显著图指导的预测过程如图 4 所示。

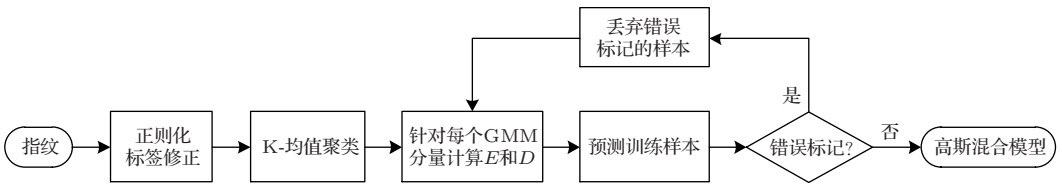


图 3　KN-GMM 的训练过程

Fig. 3　Procedure of training KN-GMM

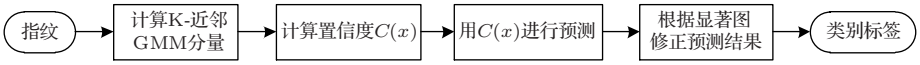


图 4　采用 KN-GMM 和显著图指导的预测过程

Fig. 4　Procedure of predicting using KN-GMM and a guidance saliency map

1.3　ROI 回归及决策

如应用场景对准确度有更高的要求，可采用步骤 2 进一步处理。步骤 2 中首先对图割算法进行了改进，使之可以适用于图像中 3 种标签部分的分割；对于每个边界框，采用这种改进算法将框内的图像分割为背景、疑似目标的高亮和阴影部分。图割方法使用高斯混合模型对三者建模，算法将预先在 ROI 上进行 K-均值聚类以初始化背景、阴影和高亮部分，再迭代执行高亮/背景和阴影/背景分割。在高亮/背景分割中，阴影和真实的背景均被认为是背景，同理，在阴影/背景分割中，高亮和真实的背景被认为是背景。基于阴影部分总是比高亮部分稍暗一些的先验知识，均值较高的阴影高斯分量和均值较低的高亮高斯分量将进行标签交换。同时，参考侧扫声呐成像几何关系，位于高亮部分近端的阴影将会被直接丢弃。

完成对 ROI 的图像分割后，一些没有检测到前景或是前景中只有很少分散像素点的 ROI 将会被丢弃。接着计算高亮部分的最小外接矩形来估计目标的尺寸，并通过对比高亮部分以及其形态学膨胀邻域计算目标强度。经上述处理后，绝大多数虚警将会被排除。ROI 回归和决策过程示例如图 5 所示。

最后，步骤 2 将对经虚警滤除后仍然保留的边界框进行回归。通过计算高亮部分的中心像素，边界框将被重新定位，使目标位于边界框的中心，以获取更为精确的 ROI，步骤 2 处理流程如图 6 所示。图像分割的引入提升了算法对弱监督场景的适应能力：首先，步骤 1 中所有和标注区域相交的图像块都将被视作正样本，步骤 2 中图像分割提供了对目标范围的精确描述，标注不够确切的边界框将被重新定位。

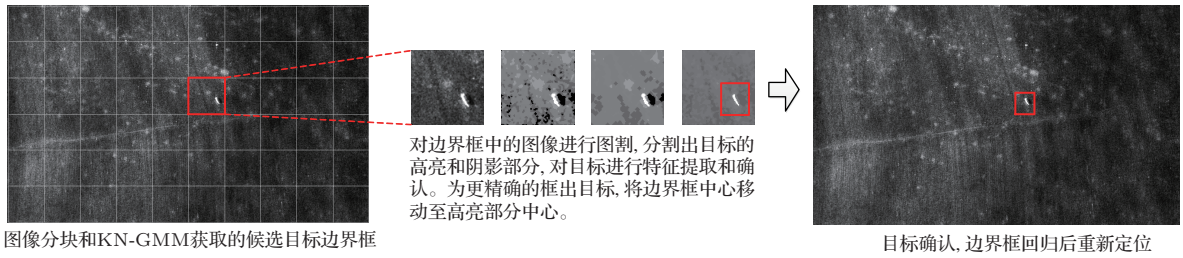


图 5　ROI 回归和决策过程示例

Fig. 5　An example of ROI regression and determination process

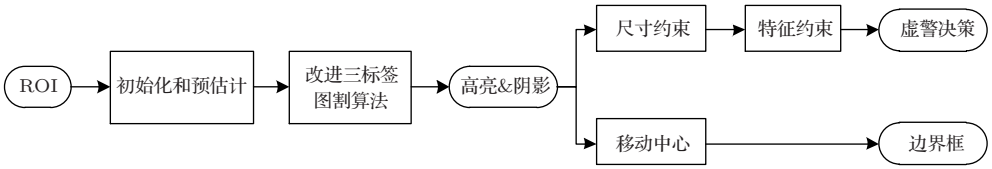


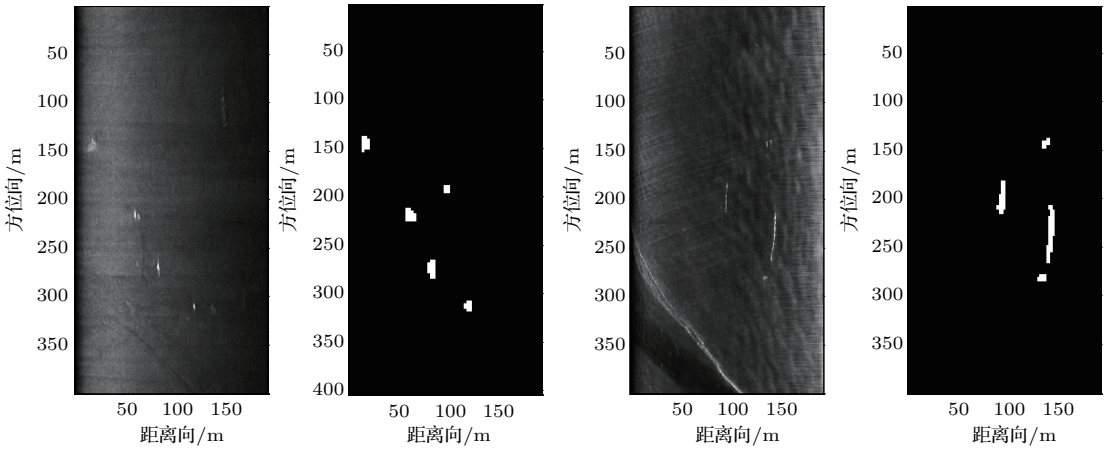
图 6　步骤 2 处理流程

Fig. 6　Procedure of Phase 2

2　试验结果

2.1　数据集

对检测效果的评估主要基于专家标注的高频 SAS 图像数据集。数据集由一系列 200 m×400 m 的连续实时 SAS 条带图像片段组成，并对条带图像片段中的小目标进行人工标注作为真实标记 (Ground truth, GT)。以在丹江口水库布放的两类仿制目标为例，如图 7 所示，图中几个目标被放置在湖底，包括 ϕ3 cm 电缆和 ϕ40 cm 圆柱体煤气罐。目标的图像及其高亮-阴影分割结果如图 8 所示，第一行为目标实物照片，第二行为目标高频 SAS 图像，第三行为目标高亮-阴影分割结果。



(a) 正样本I　(b) 正样本I的真标注　(c) 正样本II　(d) 正样本II的真标注

图 7　两类目标的正样本及其真实标记

Fig. 7　Two positive samples and their ground truth

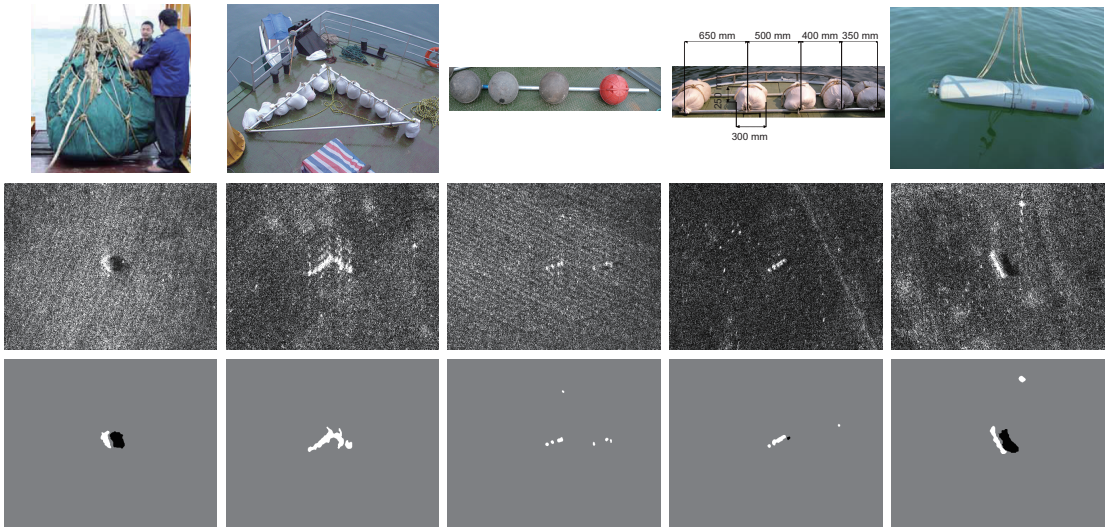


图 8　布放目标示例

Fig. 8　Examples of counterfeit targets

2.2　检测结果

采用图 7 中的样本 2(丹江口水库试验数据，包含电缆和圆柱体目标)，从中获取 15138 个图像块来训练 KN-GMM，对南中国海试验数据中圆柱体目标的检测的实例如图 9、图 10 和图 11 所示。

借助显著图指导和 ROI 回归的处理，所有疑似区域的目标均被正确标出。下面，使用不同尺寸的、从不同地点获取的数据集进一步验证算法的鲁棒性。采用 A、B、C 三个试验数据集，分别来自丹江口、千岛湖和南中国海，字母 A、B、C 后面的数字代表所使用的片段编号。数据集被标注为 0 ~ 4 等级，代表 5 个渐进的置信度，其中等级 0 代表不存在目标，等级 4 代表存在确信的人造目标。在一般场景下，3 个模型均采用高召回率约束 (即认为 2 ~ 4 等级目标均为正样本) 进行训练。采用与弱监督学习相关的工作，HOG 和 SVM 作为对比方法。

算法的精确度 (Precision, P) 和召回率 (Recall, R) 可由检测结果和真实标记统计计算得到。连通的正样本图像块被认为是同一个目标。例如图 8 中的样本 1，总共有 5 个正样本。此外，本文还选择了 F 值 (F-measure) 作为对算法的整体性评估：

$$F=\frac{\left(1+\beta^2\right)PR}{\beta^2P+R},\quad \beta=1. \tag{10}$$

如表 1 所示，本文方法可在数据集较小时获得相对较低的虚警。

表 1　不同数据集试验中得到的精确度和召回率

Table 1　Precision and recall rate of experiments on different dataset

序号	训练集	测试集	HOG + SVM			本文方法		
			精确度	召回率	F 值	精确度	召回率	F 值
1	A1	A64	0.620	0.721	0.667	0.803	1.0	0.891
2	A4	A64	0.714	0.819	0.763	0.824	1.0	0.904
3	A32	A64	0.821	0.901	0.859	0.859	1.0	0.924
4	A64	A64	0.865	0.951	0.906	0.859	1.0	0.924
5	A1	C123	0.563	0.547	0.555	0.776	0.852	0.835
6	A4	C123	0.615	0.652	0.632	0.781	0.930	0.849
7	A32	C123	0.772	0.852	0.810	0.762	0.948	0.845
8	A64	C123	0.795	0.913	0.850	0.773	0.948	0.851

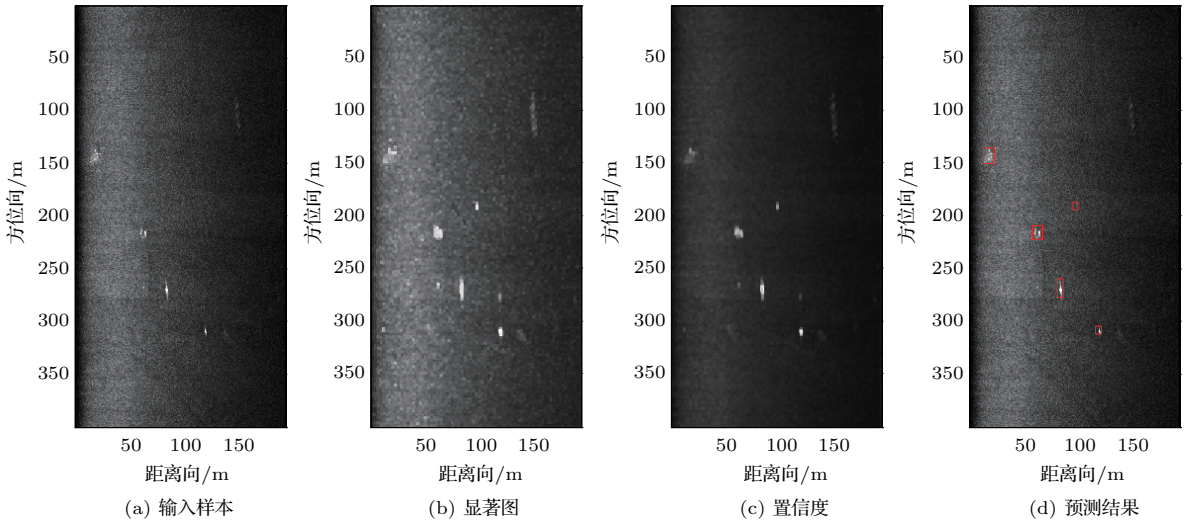


图 9　基于图像指纹和 KN-GMM 的预测结果示例 1

Fig. 9　Example 1 of predictions by image fingerprints and KN-GMM

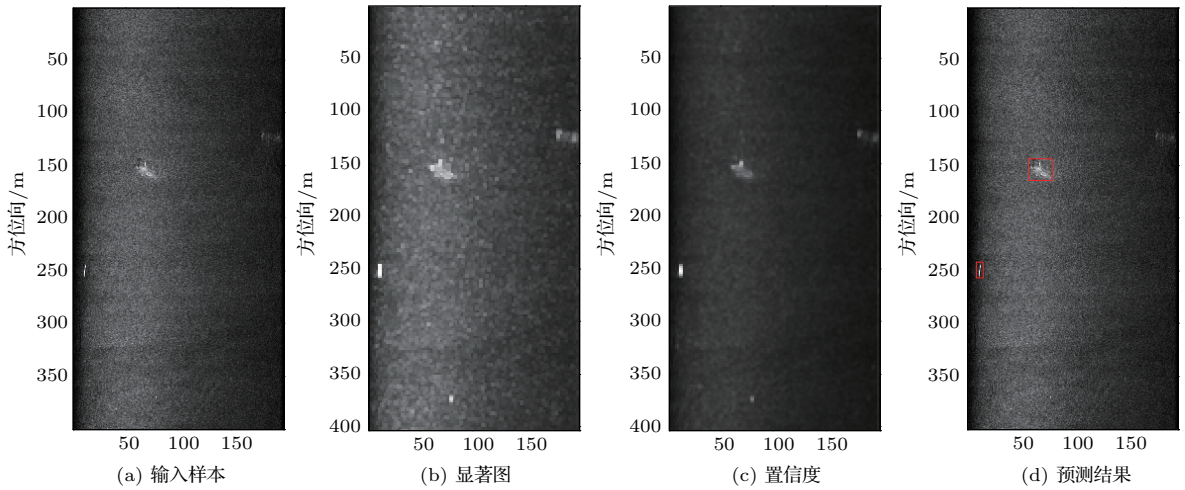


图 10　基于图像指纹和 KN-GMM 的预测结果示例 2

Fig. 10　Example 2 of predictions by image fingerprints and KN-GMM

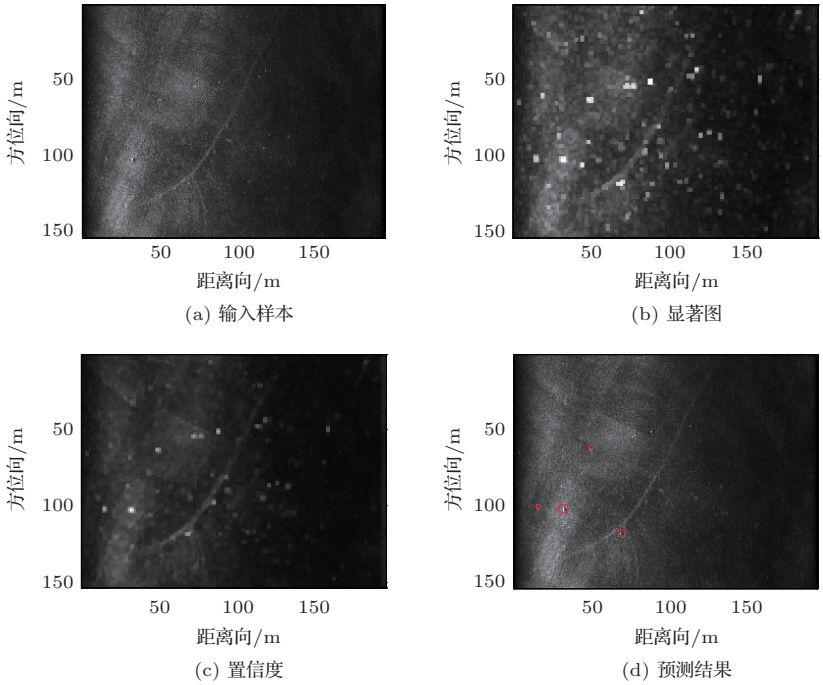


图 11　基于图像指纹和 KN-GMM 的预测结果示例 3

Fig. 11　Example 3 of predictions by image fingerprints and KN-GMM

2.3　算法性能

实时工作条件下，一般将以流水的形式逐帧对条带式侧视声图进行目标检测处理。当图像帧达到一定大小后，触发一次检测计算，并要求在下一帧相同大小的图像累积结束前完成检测。图像帧的生成时间可由沿航迹向分辨率和拖体的运动速度计算得出。通过对比目标检测的时间开销和产生该图像帧所需的测绘时间，可大致估计出算法是否满足实时性要求。若检测时间开销有较大冗余，还可选择设置一定比例的帧重叠，以改善在帧边界的目标检测效果。

数据集 A、B、C 典型图像的处理结果如表 2 所示。所使用 CPU 配置为 i5 四代，8 GB 内存。

随着图像中疑似目标数量的增加，本文算法在步骤 2 中的时间消耗也会增加，因为对每一个疑似目标都需基于图割进行 ROI 回归和决策。由于步骤 1 的设置目的在于尽可能地减少漏检，步骤 2 则用于排除虚警，可以认为若获取的 ROI 与正样本差异越大，场景越复杂，步骤 2 的必要性就益发凸显。为满足实时性要求，算法将对尺寸超出一定范围的 ROI 进行下采样以避免图割计算时间过长，同时引入目标的属性进行预评估。所提出算法可以满足实时性要求，但不能保证可适用于最坏的情况。

表 2　本文方法在不同场景下的效果

Table 2　The performance of proposed method on different scene

数据集	A	B	C
像素	18980 × 6000	9941 × 5274	12576 × 5216
正样本	4	6	4
ROI 标注数量	9	54	4
图像帧生成时间/s	380.00	226.08	251.54
步骤 1 处理时间/s	39.33	17.26	18.93
步骤 1 和步骤 2 的处理时间/s	51.32	44.58	23.13

3　结论与展望

3.1　结论

本文提出了一种用于水下小目标检测的快速弱监督学习算法。算法包含基于 DCT 和 KNN 用于获取较高目标召回率的步骤 1 和基于图割算法用于消除虚警的步骤 2。DCT 对目标区域的海底地物散射变化进行了总体描述，KNN 模型为低可靠性数据提供了兼容性，增强了对小样本集的适应能力，基于图割的回归步骤使检测系统更加准确和高效。在合成孔径声呐的不同数据集中的试验结果证明本文

方法在一般情况下，可以满足实时性要求，同时在使用较小训练集时，可以达到相对较高的精确度和召回率。

3.2 展望

步骤 1 和步骤 2 的级联构造了一种简单框架以实现对精确度和召回率的折中，但仍存在一些局限，有待未来进一步的工作中解决：

(1) 复杂场景中的灵活性。本文方法在地貌丰富 (如从千岛湖获取的数据集 B) 的图像中测试时，步骤 1 的效果较差。在目前的框架下，只能采用独立的手动级别 1 来标注自然地形，期望通过 KN-GMM 来消除干扰。复杂场景的自适应一直是目标检测中的一个重要话题。

(2) 增量学习。KNN 模型的特点决定了其预测时间将随着数据集规模的增大而大幅增加。对 KN-GMM 进行优化可降低时间消耗，但从增量学习的角度来说，效率仍有待提高。

(3) 虚警消除步骤的优化。本文方法的步骤 2 对于简单虚警的消除略显复杂。可以使用预评估消除一些容易判断的虚警，从而避免使用图割，降低时间开销。

(4) 强化对目标细节信息的利用。本文方法主要针对颗粒度较为粗糙的分类体系，DCT 指纹特征更关注声图的局部区域特征，这样的框架保留了对地物场景的整体概括，但目标的细节信息 (如多视角特征等) 并未得到充分利用。在步骤 2 中增加对目标特征的提取和筛选将有助于提升算法对于更精细的分类问题的适应能力。

参 考 文 献

[1] Einsidler D, Dhanak M, Beaujean P. A deep learning approach to target recognition in side-scan sonar imagery[C]. OCEANS 2018 MTS/IEEE Charleston, Charleston, SC, 2018: 1–4.

[2] Klausner N, Azimi-Sadjadi M R. Non-Gaussian target detection in sonar imagery using the multivariate laplace distribution[J]. IEEE Journal of Oceanic Engineering, 2045, 40(2): 452–464.

[3] Wang X, Wang J, Yang F, et al. Target detection in colorful imaging sonar based on HOG[C]. 2018 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), Qingdao, 2018: 1–5.

[4] Williams D P. Fast target detection in synthetic aperture sonar imagery: a new algorithm and large-scale performance analysis[J]. IEEE Journal of Oceanic Engineering, 2015, 40(1): 71–92.

[5] Kim B, Yu S C. Imaging sonar based real-time underwater object detection utilizing AdaBoost method[C]. IEEE Underwater Technology，2017, 2017: 1–5.

[6] Ma N, Chia C S. False alarm reduction by LS-SVM for manmade object detection from side-scan sonar images[C]. IEEE Oceans 2007 Proceedings, 2007.

[7] Maussang F, Rombaut M, Chanussot J, et al. Fusion of local statistical parameters for buried underwater mine detection in sonar imaging[J]. EURASIP Journal on Advances in Signal Processing, 2008: 1–19.

[8] Lopera O, Dupont Y. Automated target recognition with SAS: shadow and highlight-based classification[C]. IEEE Oceans 2012 Proceedings, 2012.

[9] Sawas J, Petillot Y, Pailhas Y. Cascade of boosted classifiers for rapid detection of underwater objects[C]. EUCA 2010 Conference Istanbul, 2010.