

# 基于对抗域适应的红外舰船目标分割

高子航<sup>1</sup>, 刘兆英<sup>1</sup>, 张 婷<sup>1</sup>, 李玉鑑<sup>2</sup>

(1. 北京工业大学信息学部, 北京 100124; 2. 桂林电子科技大学人工智能学院, 桂林 541004)

**摘要:** 为了提高红外舰船目标的分割准确率, 提出一种基于对抗域适应的红外舰船目标分割方法, 其中有标注的可见光舰船图像为源域, 没有标注的红外舰船图像为目标域。为了解决两个域之间的风格差异问题, 本文依次对源域的可见光图像进行灰度化和白化预处理, 将其转换为具有目标域风格的图像。对于目标域的红外图像, 使用去噪网络进行优化; 接着, 为了解决判别网络视野受限问题, 设计基于空洞卷积的判别网络; 最后, 针对目标域预测图像置信度低问题, 将目标域预测图像的信息熵加入到对抗损失中。在可见光和红外舰船图像组成的数据集上的实验结果高于现有方法, 证明了本文方法的有效性。

**关键词:** 域适应; 目标分割; 生成对抗学习; 红外舰船图像; 信息熵

**中图分类号:** TP183      **文献标志码:** A

## Infrared Ship Target Segmentation Based on Adversarial Domain Adaptation

GAO Zihang<sup>1</sup>, LIU Zhaoying<sup>1</sup>, ZHANG Ting<sup>1</sup>, LI Yujian<sup>2</sup>

(1. Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China; 2. School of Artificial Intelligence, Guilin University of Electronic Technology, Guilin 541004, China)

**Abstract:** To improve the segmentation accuracy of infrared ship target, we present an adversarial domain adaptation network for infrared ship target segmentation (ISADA), where the labeled visible ship images are used as the source domain and the unlabeled infrared ship images as the target domain. To address the issue of style difference between the two domains, we preprocess the visible images of the source domain in turn with graying and whitening to convert them into the images with the style of the target domain. For the infrared images in the target domain, we optimize them with a denoising network. Furthermore, to solve the matter of limited receptive field of the discriminative network, we design a discriminative network based on atrous convolution. Finally, for the problem of low confidence of the target domain prediction images, the information entropy of the target domain prediction images is added to the adversarial loss. The experimental results on the datasets composed of visible and infrared ship images is superior than the state-of-the-art methods, which demonstrates the effectiveness of the proposed method.

**Key words:** domain adaptation; object segmentation; generative adversarial learning; infrared ship images; information entropy

**基金项目:** 国家自然科学基金(61906005, 62166002, 62176009); 北京市教育委员会科技计划(KM202110005028); 北京工业大学交叉科学研究院(2021020101); 北京工业大学国际科研合作种子基金(2021A01)。

**收稿日期:** 2022-05-08; **修订日期:** 2022-08-26

## 引 言

中国拥有庞大的海域面积,舰船作为海上运输的重要载体,在海防安全中发挥着重要作用。海洋监测离不开高效的舰船目标分割<sup>[1]</sup>。如何快速、准确地获取舰船目标,实现舰船目标的精确分割,是海洋监测中迫切需要解决的问题<sup>[2]</sup>。红外热成像技术在热点区域监控、伪装目标揭露以及军事侦察等方面具有重要应用<sup>[3]</sup>。红外成像凭借着被动接收辐射变化、远距离拍摄、隐蔽性强以及全天候工作等特点,在复杂海况下有着良好的实用性。因此,红外舰船图像在海洋监测中有着广泛的应用。

通常红外舰船目标图像的信噪比低、边界模糊,这为红外图像的精确像素级标注带来困难<sup>[4]</sup>。与此同时,可见光舰船图像的信噪比高、边界清晰,在晴朗天气条件下获得的可见光图像包含更多信息,目标结构特征更为明显,且易获得其标注图像。因此,将可见光图像训练的模型应用到红外图像数据集上是一种有效解决红外图像标注困难的方法。通常做法是使用有标注可见光图像训练分割模型,然后将其应用于红外图像。然而,由于两个数据集域中的数据分布存在较大域偏移<sup>[5]</sup>,目前的高准确率分割模型并不能保证在红外图像上有良好的分割性能<sup>[6]</sup>。随着迁移学习的发展,用于解决域偏移问题的无监督域适应方法备受研究者青睐。该方法通过学习源域和目标域之间的特征分布实现域对齐。

现有的无监督域适应方法按照对齐空间的不同,可以分为基于特征空间的无监督域适应和基于像素空间的无监督域适应方法。基于特征空间的无监督域适应方法是指在特征空间直接对齐源域和目标域的特征减小域偏移<sup>[7]</sup>。基于像素空间的无监督域适应方法是指在像素空间中利用图像风格转换等技术减小两个域视觉差异<sup>[8]</sup>。现有的大多数方法是在两个域图像同为可见光场景中提出的,鲜有针对可见光图像为源域,红外图像为目标域的情况展开。这是因为一方面两个域视觉差异过大,传统的图像迁移方法不能保证风格迁移后图像的质量;另一方面,基于对抗的域适应方法对两个域的预测结果进行训练,通过对抗损失函数对齐它们的特征分布<sup>[9]</sup>。这种对齐方式仅仅是两个域的全局分布对齐,并未考虑不同类别间的对齐程度,在整个训练过程中,所有像素点被赋予的权重是一致的,已经对齐的大多数像素点会造成网络误判已收敛,从而影响分割性能。

基于以上分析,本文提出一种基于对抗域适应的红外舰船目标分割方法(Infrared ship target segmentation based on adversarial domain adaptation, ISADA)。首先,对源域可见光图像依次使用灰度化和白化方法进行预处理;对于目标域红外图像,使用去噪网络去除其噪声并平滑图像,减小两个域的外观差异;然后,将预处理后的两个域图像输入到分割网络并预测其分割结果;其次,设计并使用空洞卷积扩大判别网络的感受野;最后,计算目标域图像的熵值并作为该图像的置信度水平加入到对抗损失中,增大难分类像素点图像的权值,提高红外舰船目标的分割精度。

## 1 相关工作

基于特征空间对齐的域适应图像分割方法通常依靠生成对抗思想训练网络。例如,Hoffman等<sup>[10]</sup>从全局和类别两个方面考虑,通过将域适应技术加入到图像分割中,提出基于全卷积模块的无监督域适应方法。Sankaranarayanan等<sup>[11]</sup>将神经网络获取的中间特征投影到图像空间,并添加辅助分类损失函数用于网络训练。Tsai等<sup>[12]</sup>以两个域分割输出的空间布局、局部内容和上下文信息存在相似性为动机,在特征输出空间进行对抗训练减小域偏移。Luo等<sup>[13]</sup>考虑到不同特征的权重问题,通过对各个特征赋予不同权重来学习两个域的特征分布,提出了一种基于不同特征级的域适应方法。Vu等<sup>[14]</sup>通过直接和间接最小化目标域图像预测熵值减小域偏移。Chen等<sup>[15]</sup>使用最大平方损失并引入图像级权重比例因子缓解域偏移问题。杨佳林等<sup>[16]</sup>为了降低图像分辨率的影响,使用不同层的分辨率特征作为特征提取部分的分支,以保存图像更多的细节信息。Chen等<sup>[17]</sup>使用目标导向蒸馏模块模仿真实图像的训练

特征,并在预测输出的不同区域分别进行对抗训练以提升网络的分割性能。

基于像素空间对齐的域适应图像分割方法通常在域图像输入分割网络前减小外观差异。例如, Hoffman等<sup>[18]</sup>提出基于图像转换网络的域适应模型。在转换网络中引入语义一致性损失保留原图像内容,并加入更多语义信息以减小外观差异。Hong等<sup>[19]</sup>在源域添加一个条件生成器用于学习将源域图像特征转换为真实图像特征以及它们的判别网络。Wu等<sup>[20]</sup>通过将源域图像转换为目标域的风格,并对齐生成网络中不同维度的通道特征以提升分割性能。Zhu等<sup>[21]</sup>提出保守损失函数惩罚极大极小情况,使网络通过梯度上升和下降学习域不变特征。Wu等<sup>[22]</sup>提出新的图像转换网络共同转换两个域图像以减小两个域外观差异,并在分割网络中引入伪标签作为监督无标注的图像。

上述方法分别从特征空间和像素空间两方面降低域偏移提升目标域分割性能,且两个域数据集图像均为信息丰富的可见光图像。本文提出一种基于对抗域适应的红外舰船目标分割方法,与现有模型相比,在像素空间使用更简单、更易操作的图像处理方法代替原有需要训练的图像转换网络预处理源域图像,并使用快速、灵活的去噪网络去除红外图像噪声。

## 2 本文提出的方法

本文提出一种基于对抗域适应的红外舰船目标分割方法,该方法的网络模型包括图像预处理、分割网络和判别网络3部分,整体网络如图1所示。

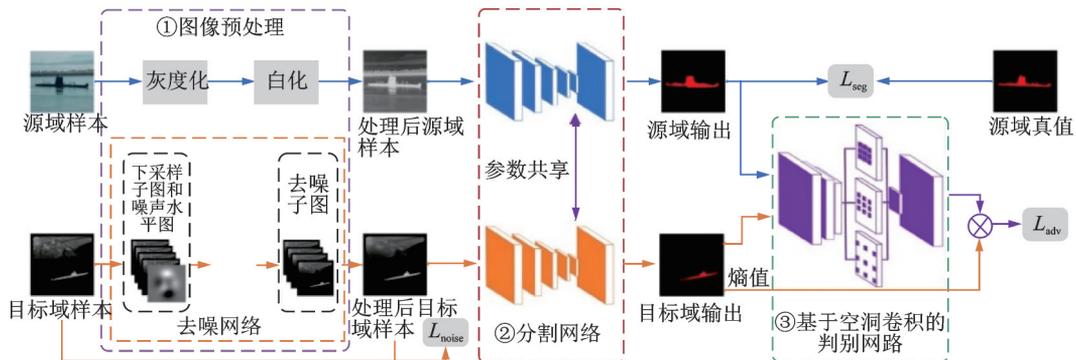


图1 整体网络结构

Fig. 1 Overall network structure

### 2.1 图像预处理

考虑到可见光图像和红外图像间的外观差异较大。本文分别对源域图像和目标域图像进行预处理以减小两个域的外观差异。为了便于后续分割网络对相同通道数的两个域数据集训练,使用图像灰度化和白化操作将源域图像转变为目标高亮的三通道灰度图而不是单通道灰度图<sup>[23]</sup>;由于红外图像的噪声分布不一,难以估计,本文使用快速、灵活的去噪网络将红外图像转变为目标无噪且清晰的三通道灰度图<sup>[24]</sup>。由于真实噪声图像通常依赖信号、非高斯和空间变化,故可以使用高斯白噪声近似模拟真实场景中的噪声图像。这是因为:(1)当没有关于噪声源的特定先验信息时,高斯白噪声是一个常见的选择;(2)噪音的来源是多分布的,难以准确表示。

在训练去噪网络之前,本文对原图像添加不同级别的高斯白噪声模拟原有图像噪声水平,生成对应的噪声图像并得到对应的噪声水平图。使用噪声水平图可以控制降噪和细节保存之间的权衡,并且不会产生视觉伪影。去噪网络结构如图2所示,网络命名为 $F$ 。

图像去噪可以从低分辨率图像中提取深层次特征恢复原有高分辨率的细节,本文选取4个大小为原图四分之一的图像作为低分辨率原图,在不影响模型性能的前提下同时提高计算速度。首先,将带有噪

声的原图下采样至4个大小为原图四分之一、通道数不变的子图,将这4个子图与相同大小的噪声水平图 $M$ 连接作为网络的输入;将其输入到连续的10个卷积块中,第1个卷积块包括1层卷积(Convolution, Conv)、修正线性单元(Rectified linear unit, ReLU)激活函数,中间卷积块的结构均相同,包括1层卷积、批归一化(Batch normalization, BN)和修正线性单元激活函数,最后1个卷积块仅包括1层卷积;最后,对卷积后的特征图上采样至原图大小得到去噪图像。整个过程中,每个卷积块输出的特征图大小不变,卷积核大小均为 $3 \times 3$ ,通道数均为96。该网络的目的是得到噪声较少且平滑的图像。

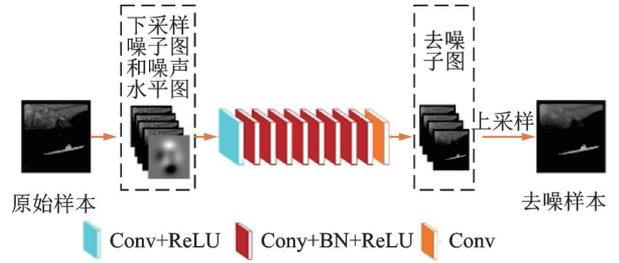


图2 去噪网络结构

Fig.2 Structure of the denoising network

## 2.2 分割网络

本节将2.1节转换后的图像分别作为新的源域和目标域输入到分割网络 $G$ 中。将两个域转换后的图像分别表示为 $\{I_s\}$ 和 $\{I_t\}$ 。本文采用以Resnet101为主干网络的DeepLab-v2分割网络,两个域图像共享其参数。该分割网络的分类器在最后1个卷积层后,由空洞金字塔池化模块构成,池化的采样率为 $\{6, 12, 18, 24\}$ ,该模块可以在不降低分辨率的情况下同时捕获多个尺度的特征和图像上下文信息。

## 2.3 判别网络

本节将介绍基于空洞卷积的判别网络模块,命名为 $D$ 。原始的判别网络存在感受视野受限问题,空洞卷积被用来保存图像经过卷积下采样过程中丢失的信息<sup>[25]</sup>。本文使用3个不同空洞率的空洞卷积替代传统全卷积判别结构。如图1所示,该判别网络是一个卷积网络,其中前3层是卷积层,卷积核大小均为 $4 \times 4$ ,通道数分别为 $\{64, 128, 256\}$ ,步长均为2;第4层是1个空洞卷积层,该层包含3个并联的空洞卷积,空洞率分别为1、3、6,卷积核大小均为 $4 \times 4$ ,通道数均为256,步长均为1;第5层是卷积层,卷积核大小为 $4 \times 4$ ,通道数为1,步长为2,接受3个空洞卷积层的相加特征作为输入;最后使用上采样操作将输出特征图恢复至输入大小返回判别结果。

## 2.4 损失函数

对于去噪网络 $F$ ,原始去噪网络的损失函数为

$$L_1(F) = -\frac{1}{2N} \sum_{i=1}^N \|F(y_i, M_i; \Theta) - x_i\|^2 \quad (1)$$

式中: $N$ 为训练样本总个数; $x_i$ 为第 $i$ 个样本; $y_i$ 为第 $i$ 个样本中添加一定高斯白噪声的图像; $M_i$ 为第 $i$ 个样本的噪声水平图(噪声水平为 $\sigma$ )且是均匀映射,矩阵所有元素均为 $\sigma$ ; $\Theta$ 为网络的参数。

为了使去噪图像更加平滑,加入平滑图像损失,令 $u_i = F(y_i, M_i; \Theta)$ 表示去噪图像,该损失函数为

$$L_2(F) = -\frac{1}{2N} \sum_{i=1}^N \left\| (\nabla_x(u_i - x_i))^2 + (\nabla_y(u_i - x_i))^2 \right\| \quad (2)$$

式中 $\nabla_x$ 和 $\nabla_y$ 分别为相邻像素间沿 $x$ 方向和 $y$ 方向像素值强度的梯度值。该去噪网络部分的总损失函数为

$$L(F) = L_1(F) + L_2(F) \quad (3)$$

此外,本文将分割网络和判别网络分别作为对抗网络的两部分,并进行“最大最小”训练,其中分割网络得到两个域的分割结果,判别网络对分割结果进行判定。对于判别网络 $D$ ,其损失函数为

$$L_d(D) = -\sum_{l=1}^N \sum_{h,w} (1-z) \log(D(I^l)^{(h,w,0)}) - \sum_{l=1}^N \sum_{h,w} z \log(D(I^l)^{(h,w,1)}) \quad (4)$$

式中: $z=0$ 为该样本属于源域; $z=1$ 为该样本属于目标域; $h,w$ 分别为图像的高和宽; $l$ 为两个域转换后的第 $l$ 个图像。为了能够判别真假,该网络需要不断更新迭代最小化目标函数。

对于分割网络 $G$ ,本文使用经典的交叉熵损失作为监督损失函数,表示为

$$L_{\text{seg}}(G) = - \sum_{l=1}^N \sum_{h,w} \sum_{c \in C} Y_s^{l(h,w,c)} \log(P_s^{l(h,w,c)}) \quad (5)$$

式中: $C$ 为类别总数; $P_s^l$ 为源域第 $l$ 个样本图像输入分割网络生成的预测图像; $Y_s^l$ 为第 $l$ 个样本图像对应的真值标签。

为了能够提高那些难分类点的分割准确率,本文引入信息熵进行判别。一幅图像的信息熵值越小,代表其可信度越高;信息熵值越大,表示其可信度越低。对于目标域中的预测图像,需要赋予熵值大的像素点更多的权重,提升其网络学习能力。目标域预测图像的熵值计算公式为

$$E_t = - \sum_{h,w} \sum_{c \in C} P_t^{(h,w,c)} \lg P_t^{(h,w,c)} \quad (6)$$

式中 $P_t$ 为目标域图像输入分割网络得到的预测结果。

为了使源域预测图像和目标域预测图像的特征分布更接近,本文将目标域预测图像的熵值加入对抗损失中作为其权重参数,定义对抗损失函数为

$$L_{\text{adv}}(D) = - \sum_{l=1}^N E_t \cdot \sum_{h,w} \lg(D(P_t^l)^{h,w,1}) \quad (7)$$

式中 $P_t^l$ 为目标域第 $l$ 个样本图像输入分割网络生成的预测输出图像。综合式(5,7),可以得到分割网络的损失函数,表示为

$$L_{\text{total}}(G, D) = L_{\text{seg}}(G) + \lambda_{\text{adv}} L_{\text{adv}}(D) \quad (8)$$

式中参数 $\lambda_{\text{adv}}$ 用于均衡两个损失函数。根据对抗学习思想,本文优化该生成网络的损失函数为

$$\min_G \max_D L_{\text{total}}(G, D) \quad (9)$$

即在网络训练中,最小化源域图像的分割损失,同时最大化目标图像被预测成源域图像的概率。

本文提出基于对抗域适应的红外图像目标分割算法如算法1所示。

#### 算法1 基于对抗域适应的红外图像目标分割算法

输入:源域数据集 $I_S$ ,目标域数据集 $I_T$ ,网络循环次数 $K$ ;

输出:网络训练的权重;

$i=0$

while  $i \leq K$  do:

for  $j$  in range(len( $I_S$ )):

(1) 分别读取两个数据集图像 $\{I_S\}$ 和 $\{I_T\}$ 作为网络的输入,对源域和目标域图像分别进行图像处理和去噪网络处理,利用目标域图像优化去噪损失;

(2) 将预处理之后的数据集作为新的分割网络输入,源域图像输入分割网络优化分割损失 $L_{\text{Seg}}$ 并生成对应分割图 $P_s$ ,目标域图像输入分割网络生成对应分割图像 $P_t$ ;

(3) 将两个域输出图像输入判别器优化判别损失 $L_D$ ,计算目标域输出图像熵值并优化对抗损失 $L_{\text{adv}}$ 。

end while

### 3 实验结果与分析

为验证本文方法的有效性,本节依次阐述所提模型的数据集、实验细节、评价指标、实验结果和消融实验。深度框架为Pytorch,操作系统为Window7,GPU型号为NVIDIA Tesla K40c。

### 3.1 数据集

本文使用的数据集为网络获取的自制可见光-红外舰船数据集。源域包含 600 幅可见光图像,全部用于训练;目标域包含 200 幅红外图像,训练集包含 120 幅图像,测试集包含 80 幅图像。源域的可见光图像全部拥有像素级标注,而目标域红外图像仅有测试集拥有像素级标注,训练集没有。部分源域图像和目标域测试图像以及各自对应标签如图 3 所示。

网络训练中,首先对两个域数据集进行扩充。对两个域数据集分别进行水平翻转、平移、缩放和裁剪操作,最终得到源域训练数据集共 3 091 幅图像,目标域数据集共 790 幅图像,其中无标签训练数据集共 474 幅图像,测试数据集包含 316 幅图像。数据集图像大小均为 256 像素×256 像素,通道数均为 3。

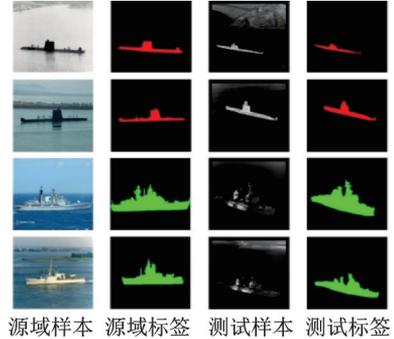


图 3 可见光和红外图像示例

Fig.3 Examples of visible and infrared images

### 3.2 实验细节和评价指标

对于去噪网络,噪声水平设定为  $\sigma=30$ ,初始学习率设为  $10^{-3}$ ,当完成 50 个循环后学习率以多项式衰减,衰减率为 0.9。分割网络使用随机梯度下降法优化参数中,动量参数为 0.9,权重衰减为  $10^{-4}$ ,初始学习率设为  $2.5 \times 10^{-4}$ [26],式(8)中的参数  $\lambda_{adv}=0.001$ 。对于判别网络,使用自适应矩估计优化器进行训练[27],动量参数分别设为 0.9 和 0.99,初始学习率设为  $2.5 \times 10^{-4}$ ,使用幂值为 0.9 的多项式衰减降低学习率,训练总轮数设置为 15 000,每 1 000 轮训练之后保存网络权重。

本文采用的评价指标是平均交并比(mean Intersection over union, mIoU)[28]。mIoU 表示的是预测区域和真实区域间交集与并集的比值,即有

$$mIoU = \frac{1}{k+1} \frac{\sum_{i=0}^k p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}} \quad (10)$$

式中: $P_{ii}$ 为分类正确的数量; $P_{ij}$ 为实际类别为  $i$  而被预测为类别  $j$  的数量; $P_{ji}$ 为实际类别为  $j$  而被预测为类别  $i$  的数量; $k+1$  表示数据集中的总类别数量。

### 3.3 实验结果

本节将本文方法与现有基于特征空间(FCN Wld<sup>[10]</sup>、MaxSquare<sup>[15]</sup>、AdaptsegNet<sup>[12]</sup>、ADVENT<sup>[14]</sup>、CLAN<sup>[13]</sup>)和基于像素空间(CycADA<sup>[18]</sup>、CGAN<sup>[19]</sup>、DANNet<sup>[22]</sup>)的域适应分割方法进行对比分析,实验结果见表 1,其可视化结果如图 4 所示。从表 1 和图 4 可知:

(1)本文方法获得了 71.23% 的平均交并比,与基于特征空间的方法对比提高了 2.49%~19.96%。此外,本文方法得到的预测图像轮廓更完整,分割区域较为准确,而对比的方法并不能准确识别图像的目标区域。

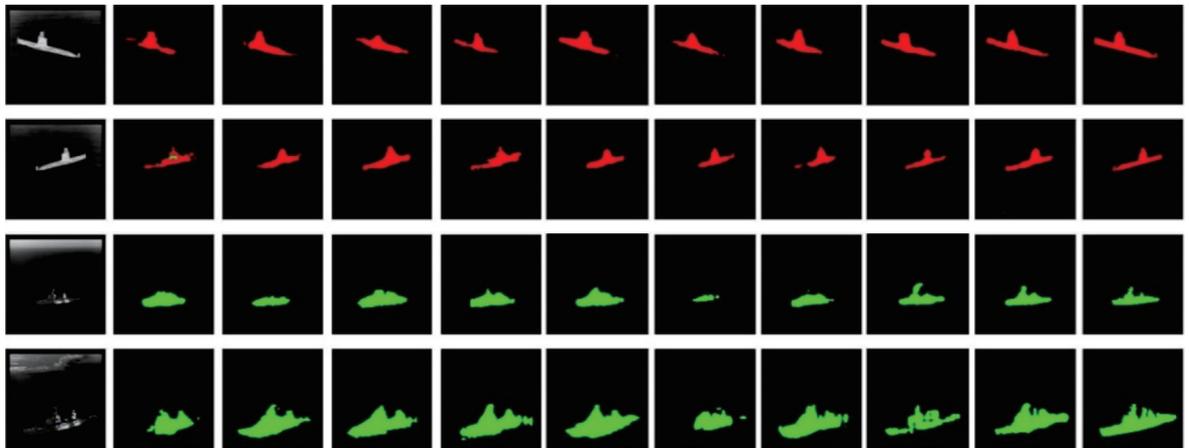
(2)与基于像素空间的方法对比,本文方法的交并比提高了 2.2%~15.87%。另外,本文方法针对性地对数据集图像分别预处理并进行特征空间对齐,有助于提高分割结果。

(3)本文方法的平均交并比均高于其他方法的结果,并且分割结果更为完整、准确。这是因为本文

表 1 不同方法的实验结果

Table 1 Experimental results of different methods

类别	方法	mIoU/%
特征空间	MaxSquare <sup>[15]</sup>	51.27
	FCN Wld <sup>[10]</sup>	65.85
	AdaptsegNet <sup>[12]</sup>	66.46
	ADVENT <sup>[14]</sup>	68.26
	CLAN <sup>[13]</sup>	68.74
像素空间	CycADA <sup>[18]</sup>	55.36
	CGAN <sup>[19]</sup>	67.95
	DANNet <sup>[22]</sup>	69.03
本文	ISADA	71.23



测试图像 MaxSquare FCN Wld AdaptSegNet ADVENT CLAN CycADA CGAN DANNet ISADA(Ours) 标签图像

图4 不同方法的可视化结果

Fig.4 Visualization results of different methods

方法分别从可见光和红外图像的外观和缩小两个域的差异两方面改进,在缩小两个域外观差异的基础上提高分割图像性能。

### 3.4 图像预处理对分割结果的影响

本节分析图像预处理部分对分割结果的影响,将本文方法分别与不作任何预处理、仅处理源域图像和仅处理目标域图像3种情况做对比,实验结果见表2,可视化结果如图5所示。从表2和图5可知:

(1)本文方法获得了71.23%的平均交并比,而与不对可见光和红外舰船图像做预处理时,获得了68.71%的平均交并比,提高了2.52%。这表明通过缩小可见光与红外图像的外观差异,可以提高红外舰船图像的分割结果。

(2)与仅处理源域和仅处理目标域的方法对比,本文方法的平均交并比分别提高了1.61%和1.05%。这表明同时预处理两个域图像,在对齐颜色空间基础上保留清晰的原有图像并缩小两个域外观差异,对目标域图像进行去噪处理能够改善图像质量,进而提高分割结果。

(3)同时对源域和目标域图像进行处理,缩小二者的外观差异,同时提高图像质量,有助于红外舰船目标的分割。

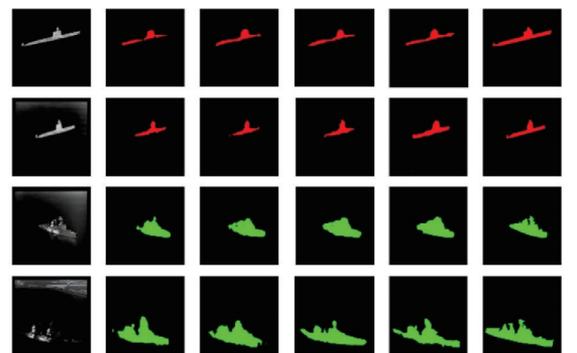
### 3.5 空洞卷积判别网络对分割结果的影响

本节分析空洞卷积判别网络对分割结果的影响,并对比不同空洞率条件下的结果。实验结果见表3,可视化结果如图6所示。从表3和图6可知:

表2 图像预处理对分割结果的影响

Table 2 Influence of image preprocessing on segmentation results

方法	mIoU/%
不使用预处理	68.71
仅处理源域	69.62
仅处理目标域	70.18
分别处理两个域	71.23



测试图像 不使用预处理 仅处理源域 仅处理目标域 分别处理两个域 标签图像

图5 图像预处理模块的消融结果

Fig.5 Ablation results of image preprocessing module

(1)与不添加空洞卷积的方法对比,添加不同空洞率的空洞卷积扩大判别网络感受野整体可以提高分割性能。例如,空洞率为{1,2,4}时获得了70.54%的平均交并比,空洞率为{1,3,5}时获得了70.97%的平均交并比,空洞率为{1,3,6}时获得了71.23%的平均交并比,这表明使用空洞卷积有助于提高红外舰船目标的分割性能。

(2)空洞率的选取不当也会造成性能下降,如空洞率为{1,2,6}时,获得了69.71%的平均交并比,比不使用空洞卷积的结果低0.71%,这是由于选取组合倍数差距较大。

(3)整体上使用空洞卷积有助于提高红外舰船目标的分割性能,然而不同的空洞率具有不同的分割结果,需要选择合适的空洞率,一种经验是尽量使空洞率的组合倍数较小。

### 3.6 添加熵值对分割结果的影响

本节分析基于熵值对抗损失对分割结果的影响,对比是否使用熵值作为损失权重的结果。实验结果见表4,可视化结果如图7所示。从表4和图7可知:

(1)当损失函数不添加熵值时,分割结果会出现舰船目标轮廓不完整或者分割区域过大的情况。当损失函数添加熵值时,使模型能够对齐舰船目标中不易分类的像素,分割结果更为准确、完整。

(2)将熵值添加至损失函数中的mIoU较之前提高1.19%,证明了该部分的有效性。

## 4 结束语

本文提出一种基于对抗域适应的红外舰船目标分割方法,将可见光舰船图像作为源域、红外舰船图像作为目标域,从像素空间和特征空间实现无监督域适应红外舰船分割。本文的贡献主要为两方面:在像素空间,通过图像预处理对齐两个域的颜色特征以缩小外观差异;在特征空间,构建基于空洞卷积的判别网络扩大其感受野并使用基于熵值的对抗损失对齐不易分类的像素以缩小两个域分布差异。实验表明了本文方法的优越性。然而,本文提出的方法在目标较小的红外舰船图像上仍然存在识别不准确、分割不完整的问题,未来考虑引入图像边缘信息细化边缘特征,并使用图像显著性模块间接提升性能。

表3 空洞卷积对分割结果的影响

Table 3 Influence of atrous convolution on segmentation results

方法	mIoU/%
不使用空洞卷积	70.42
空洞率组合1,2,4	70.54
空洞率组合1,2,6	69.71
空洞率组合1,3,5	70.97
空洞率组合1,3,6	71.23

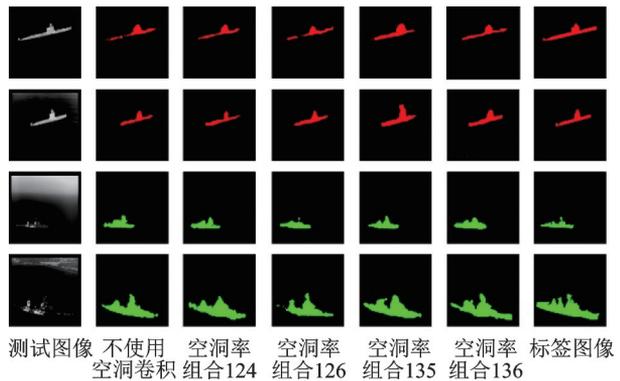


图6 空洞卷积模块的消融结果

Fig.6 Ablation results of atrous convolution module

表4 基于熵值对抗损失对分割结果的影响

Table 4 Influence of entropy-based adversarial loss on segmentation results

方法	mIoU/%
不添加熵值	70.04
添加熵值	71.23

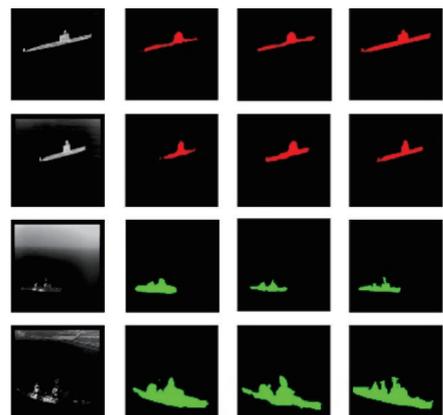


图7 基于熵值的消融结果

Fig.7 Ablation results based on entropy

## 参考文献:

- [1] LUO H J, WU K S, GUO Z W, et al. Ship detection with wireless sensor networks[J]. *Journal of IEEE Transactions on Parallel and Distributed Systems*, 2012, 23(7): 1336-1343.
- [2] JIN D R, BAI X Z. Distribution information based intuitionistic fuzzy clustering for infrared ship segmentation[J]. *Journal of IEEE Transactions on Fuzzy Systems*, 2020, 28(8): 1557-1571.
- [3] TANAKA K, IKEYA N, TAKATANI T, et al. Time-resolved far infrared light transport decomposition for thermal photometric stereo[J]. *Journal of IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(6): 2075-2085.
- [4] CORDTS M, OMRAN M, RAMOS S, et al. The cityscapes dataset for semantic urban scene understanding[C]// *Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV: IEEE Computer Society, 2016: 3213-3223.
- [5] SHIMODAIRA H. Improving predictive inference under covariate shift by weighting the log-likelihood function[J]. *Journal of Statistical Planning and Inference*, 2000, 90(2): 227-244.
- [6] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs[J]. *Journal of Computer Science*, 2014(4): 357-361.
- [7] GANIN Y, VICTOR S L. Unsupervised domain adaptation by backpropagation[C]// *Proceedings of the 32th International Conference on Machine Learning*. Lille: International Machine Learning Society, 2015: 1180-1189.
- [8] BOUSMALIS K, SILBERMAN N, DOHAN D, et al. Unsupervised pixel-level domain adaptation with generative adversarial networks[C]// *Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI: IEEE Computer Society, 2017: 95-104.
- [9] GOODFELLOW I J, POUGET-ABADIE J, MIRAZ M, et al. Generative adversarial nets[C]// *Proceedings of the 28th Conference and Workshop on Neural Information Processing Systems*. Montreal: NIPS Foundation, 2014: 2672-2680.
- [10] HOFFMAN J, WANG D, YU F, et al. FCNs in the Wild: Pixel-level adversarial and constraint-based adaptation[J]. *Journal of Computer Science*, 2016. DOI:10.48550/arXiv.1612.02649.
- [11] SANKARANARAYANAN S, BALAJI Y, JAIN A, et al. Learning from synthetic data: Addressing domain shift for semantic segmentation[C]// *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake city: IEEE Computer Society, 2018: 3752-3761.
- [12] TSAI Y, HUNG W, SCHULTER S, et al. Learning to adapt structured output space for semantic segmentation[C]// *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE Computer Society, 2018: 7472-7481.
- [13] LUO Y, ZHENG L, GUAN T, et al. Taking a closer look at domain shift: Category-level adversaries for semantics consistent domain adaptation[C]// *Proceedings of the 31th IEEE Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE Computer Society, 2019: 2502-2511.
- [14] VU T, JAIN H, BUCHER M, et al. ADVENT: Adversarial entropy minimization for domain adaptation in semantic segmentation[C]// *Proceedings of the 31th IEEE Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE Computer Society, 2019: 2512-2521.
- [15] CHEN M, XUE H, CAI D. Domain adaptation for semantic segmentation with maximum squares loss[C]// *Proceedings of the 31th IEEE Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE Computer Society, 2019: 2090-2099.
- [16] 杨佳林. 结合域适应和生成对抗网络的语义分割算法[D]. 太原: 太原理工大学, 2021.  
YANG Jialin. Semantic segmentation algorithm based on domain adaptation and generative adversarial network[D]. Taiyuan: Taiyuan University of Technology, 2021.
- [17] CHEN Y, LI W, GOOL L. ROAD: Reality oriented adaptation for semantic segmentation of urban scenes[C]// *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE Computer Society, 2018: 7892-7901.
- [18] HOFFMAN J, TZENG E, PARK T, et al. CyCADA: Cycle-consistent adversarial domain adaptation[C]// *Proceedings of the 35th International Conference on Machine Learning*. Stockholm: International Machine Learning Society, 2018: 1994-2003.

- [19] HONG W, WANG Z, YANG M, et al. Conditional generative adversarial network for structured domain adaptation[C]// Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE Computer Society, 2018: 1335-1344.
- [20] WU Z, HAN X, LIN Y, et al. DCAN: Dual channel-wise alignment networks for unsupervised scene adaptation[C]// Proceedings of the 15th IEEE Conference on European Conference on Computer Vision. Munich: IEEE Computer Society, 2018: 535-552.
- [21] ZHU X, ZHOU H, YANG C, et al. Penalizing top performers: Conservative loss for semantic segmentation adaptation[C]// Proceedings of the 15th IEEE Conference on European Conference on Computer Vision. Munich: IEEE Computer Society, 2018: 587-603.
- [22] WU X, WU Z, GUO H, et al. DANNet: A one-stage domain adaptation network for unsupervised nighttime semantic segmentation[C]// Proceedings of the 33th IEEE Conference on Computer Vision and Pattern Recognition. Nashville, TN: IEEE Computer Society, 2021: 15764-15773.
- [23] ZHANG L, SHEN Y, LI H Y. VSI: A visual saliency-induced index for perceptual image quality assessment[J]. Journal of IEEE Transactions on Image Processing, 2014, 23(10): 4270-4281.
- [24] ZHANG K, ZUO W M, ZHANG L. FFDNet: Toward a fast and flexible solution for CNN-based image denoising[J]. IEEE Transactions on Image Processing, 2018, 27(9): 4608-4622.
- [25] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[C]// Proceedings of the 4th International Conference on Learning Representations. Caribe Hilton, San Juan: ICLR, 2016.
- [26] BOTTOU L. Large-scale machine learning with stochastic gradient descent[C]// Proceedings of the 19th International Conference on Computational Statistics. Paris France: COMPSTAT, 2010: 177-186.
- [27] KINGMA D P, BA J. Adam: A method for stochastic optimization[C]// Proceedings of the 3th International Conference on Learning Representations. Caribe Hilton, San Juan: ICLR, 2015.
- [28] BEHERA S K, RATH A K, SETHY P K. Fruits yield estimation using faster R-CNN with mIoU[J]. Journal of Multimedia Tools and Applications, 2021, 80(12): 19043-19056.

## 作者简介:



高子航(1996-),男,硕士研究生,研究方向:深度学习、计算机视觉,E-mail: gaozihang@emails.bjut.edu.cn。



刘兆英(1986-),女,博士,副教授,研究方向:模式识别、深度学习等,E-mail: zhaoying.liu@bjut.edu.cn。



张婷(1986-),通信作者,女,博士,讲师,研究方向:模式识别、深度学习等,E-mail: zhangting@bjut.edu.cn。



李玉鑑(1968-),男,博士,教授,研究方向:模式识别、机器学习等,E-mail: liyujian@bjut.edu.cn。

(编辑:刘彦东)