



超尺度自导注意力网络的遥感船舶识别

陈会伟¹, 刘树美¹, 刘培学¹, 公茂法²

(1. 青岛黄海学院 智能制造学院, 山东 青岛 266427; 2. 山东科技大学 电气与自动化工程学院, 山东 青岛 266427)

摘要: 传统多尺度卷积神经网络因接收域有限, 难以对超尺度变化的空间目标进行建模。提出一种遥感船舶的超尺度自导注意力网络(HSSGAN)识别框架, 通过组连接的轻量级超尺度子空间模块捕获船舶的超尺度特征和尺度不变性, 使用自导注意力网络逐步细化超尺度特征图, 并在超尺度局部和全局语义之间建立长期依赖关系以增强类之间特征图的差异性。同时, 通过忽略不相关信息及聚合相关特征以增强目标船舶的识别性。实验结果表明, 与 TP-FCN、CF-SDN 和 HSF-Net 方法相比, HSSGAN 方法具有更好的识别效果, F1-Score 值为 0.966 78。

关键词: 目标识别; 遥感图像; 卷积神经网络; 超尺度特征; 子空间模块; 组连接

开放科学(资源服务)标志码(OSID):



中文引用格式: 陈会伟, 刘树美, 刘培学, 等. 超尺度自导注意力网络的遥感船舶识别[J]. 计算机工程, 2021, 47(10): 314-320.

英文引用格式: CHEN H W, LIU S M, LIU P X, et al. Remote sensing ship recognition based on hyper-scale self-guided attention networks[J]. Computer Engineering, 2021, 47(10): 314-320.

Remote Sensing Ship Recognition Based on Hyper-Scale Self-Guided Attention Networks

CHEN Huiwei¹, LIU Shumei¹, LIU Peixue¹, GONG Maofa²

(1. Institute of Intelligent Manufacturing, Qingdao Huanghai University, Qingdao, Shandong 266427, China;

2. College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao, Shandong 266427, China)

[Abstract] Due to the limited reception domain, the existing multi-scale Convolutional Neural Networks(CNN) often fail to model space targets with super-scale variation. In order to solve this problem, a Hyper-Scale Self-Guided Attention Networks (HSSGAN) recognition framework for remote sensing ships is proposed. The framework employs a lightweight super-scale subspace module connected by groups to capture the super-scale feature and scale invariance of the ship. Then the super-scale feature map is refined gradually by using the self-guided attention, and a long-term dependency relationship is established between the super-scale local and global semantics adaptively to enhance the difference of the feature maps between classes. In addition, irrelevant information is ignored while relevant features are aggregated, so the identifiability of the target ship can be enhanced. The experimental results show that the HSSGAN method exhibits improved recognition performance with the F1 value reaching 0.966 78.

[Key words] target recognition; remote sensing images; Convolutional Neural Network(CNN); hyper-scale features; subspace module; group connection

DOI: 10. 19678/j. issn. 1000-3428. 0058993

0 概述

随着遥感传感器的快速发展, 高空间分辨率遥感图像可以为目标物体提供更详细、更多样的几何结构描述^[1]及丰富的细节和空间结构信息。但由于物体尺度变化的复杂性, 遥感目标的识别仍然存在巨大挑战。

截止目前, 国内外研究人员设计了诸多算法以

检测和识别遥感图像中不同类型的物体^[2-3]。早期的多尺度目标检测识别算法多基于人工参与的特征设计, 如 DALAL^[4]、LOWE^[5]等利用多尺度滑动窗口生成不同尺度的图像块, 分别通过手工制作的 HOG (Histogram of Oriented Gradient) 和 SIFT (Scale-Invariant Feature Transform) 等特征描述符提取特征, 并将提取的特征输入支持向量机 (SVM) 或条件随机场

基金项目: 山东省重点研发计划(2017GGX20100, 2019GGX105001); 山东省高等学校科技计划项目(J18KB163); 青岛黄海学院重点项目(2019KJ01, 2019KJ02)。

作者简介: 陈会伟(1982—), 女, 副教授、硕士, 主研方向为控制工程、目标识别; 刘树美, 讲师、硕士; 刘培学, 副教授、硕士; 公茂法, 教授。

收稿日期: 2020-07-20 **修回日期:** 2020-09-25 **E-mail:** 83430761@qq.com

(CRF)中进行分类识别。TAO等^[6]利用改进的SIFT匹配策略对机场中的目标物体进行检测。XIAO等^[7]利用HOG的旋转不变性检测遥感图像中的飞机、汽车等大型目标。虽然手工制作的特征在特定的目标识别任务中表现良好,但其通用性和鲁棒性较差^[1],且特征的设计耗时较长。

随着深度学习模型在自然语言处理、图像分割等领域的广泛应用,深度学习技术被研究人员应用于多尺度遥感目标识别和检测任务中以减少人工参与特征设计带来的误差^[8]。如HAN等^[9]使用Fast-RCNN检测高分辨率遥感图像中地理空间目标。CHEN等^[10]结合转移学习方法,利用单一结构CNN和有限的训练样本以实现飞机识别。DENG等^[11]采用级联整流线性单元(ReLU)和inception模块的CNN来捕获图像多尺度特征,并通过多个中间层生成类目标区域。ZHONG等^[12]提出一种位置敏感平衡(Position-Sensitive Balancing,PSB)框架,有效解决特征提取中的平移不变性。DING等^[13]采用密集卷积网络(Dense Convolution Networks)、多尺度表示和多种改进方案以增强VGG-16的特征提取能力,提高了检测精度。WANG等^[14]提出一种多尺度的视觉注意网络以捕获多尺度空间的上下文语义,并突出目标区域的特征。

虽然上述的多尺度目标检测方法提高了目标检测或识别的精度,但这些多尺度结构多是建立在单一结构的卷积层上,受到卷积层数的限制且缺乏丰富的接收域,导致捕获的多尺度特征难以对极端尺度变化的

目标对象进行建模,影响高空间分辨率遥感图像中超大和极小目标的识别精度。但是,如果仅通过叠加单一结构的卷积层来增强特征提取能力,就会随着网络层数的增加而出现拟合和梯度消失、爆炸等问题。

为缓解高空间分辨率遥感图像中目标对象的极端尺度变化问题,本文提出一种超尺度自导注意力网络(Hyper-Scale Self-Guided Attention Networks, HSSGAN)的目标识别框架。该网络在卷积层中构建多个轻量级多尺度结构,从多个卷积子空间组中捕获丰富的多尺度特征,根据组之间的连接将多尺度特征进行组合,形成极端尺度特征,即小尺度和超尺度特征。利用自导注意力网络逐层细化极端特征图,在极端尺度的局部和全局语义之间建立长期依赖关系,并解决冗余信息的重复使用问题以提高遥感图像中目标船舶的识别精度。

1 超尺度自导注意力网络的识别框架

高空间分辨率遥感图像中目标物体的结构通常在尺度、纹理和形状上变化较大,且容易表现出类内和类间多样性。而传统卷积神经网络的接受域有限,产生的局部特征难以对超尺度的目标物体进行详细描述。为解决这些问题,提出一种超尺度自导注意力网络的遥感船舶识别框架,利用超尺度子空间组捕获目标船舶的局部和全局语义,并逐步细化超尺度特征图,自适应地建立长期依赖关系,滤除冗余信息以提高识别精度。HSSGAN的网络结构如图1所示。

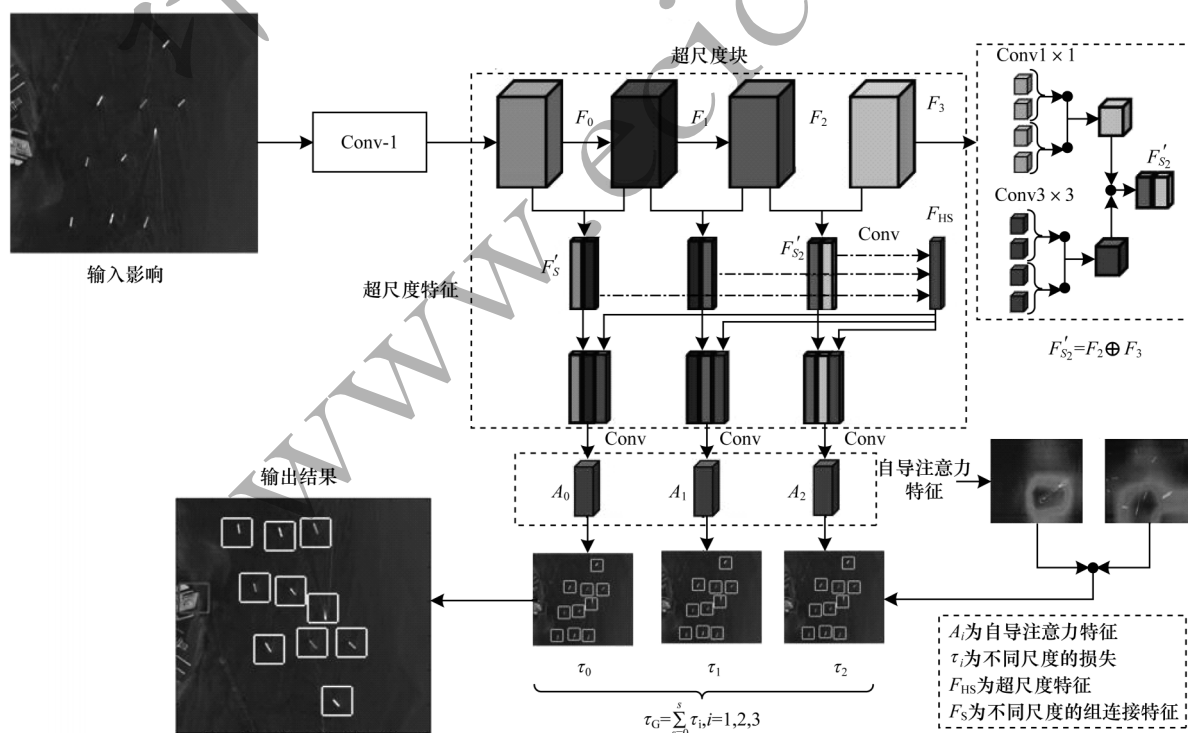


图1 HSSGAN的网络结构

Fig.1 Network structure of HSSGAN

在图1中, Conv1×1表示卷积核大小为1×1, Conv3×3表示卷积核大小为3×3。该超尺度自导注意力的遥感船舶识别框架主要包含超尺度特征提取模块和自导注意力模块2个部分。超尺度特征提取模块主要被用于提取船舶目标的不同尺度特征,其中包括极小尺度和超大尺度。而自导注意力模块主要由通道注意力和位置注意力2个模块组成,被应用于不同尺度的特征中以实现不同尺度特征的逐步细化、去除冗余信息、增强相关特征的特征能力,使特征对不同类型的目标具有更优异的区分能力。

1.1 超尺度块

超尺度块^[15]是由多个不同尺度的卷积层通过组连接构成的轻量级子空间模块,能够增强超尺度特征的表示能力。与传统卷积层通过提取多尺度特征以增强表征能力^[16-17]不同,组连接为2个卷积层中组特性之间的连接,通过实现层间信息复用使该2层能获得更多细节信息。在卷积接收域有限的情况下,此方法可提高同尺度卷积特征的利用率,并增加层间细节信息的传递。

假设初始输入为 x_i ,且 $x_i \in \{x_1, x_2, \dots, x_n\}$,经过不同尺度的单一卷积层可得到 $C_K(x_i)$,其中: K 为卷积核的大小,且满足 $K=1, 3, 5, 7$ 。组连接复合函数为 $\gamma(\cdot)$,尺度特征 $F_i \in \{F_1, F_2, \dots, F_n\}$,尺度块内的组连接计算如式(1)所示:

$$\begin{cases} F_i = \sum_{g=1}^{M-1} \text{concat}(\gamma^g(x_i), \dots, \gamma^{M-1}(x_i)) \\ \gamma(x_i) = \sum_{g=1}^M C_K^g(x_i) \odot C_K^{g+1}(x_i) \end{cases} \quad (1)$$

其中: M 表示同尺度卷积块数量; \odot 表示点积拼接; $\text{concat}(\cdot)$ 表示简单拼接。

在不增加卷积层接收域的情况下,捕获的卷积特征按照组连接方式可提高特征的复用率,弥补有限接收域带来的局限性。继续对不同尺度卷积层捕获特征 F_i 并使用组连接方式,可得到超尺度的特征图 F_{HS} ,计算公式如式(2)所示:

$$\begin{cases} F_{HS} = \text{Conv} \sum_{g=1}^{M-2} (F'_g(x_i), \dots, F'_{M-2}(x_i)) \\ F'_g(x_i) = \sum_{g=1}^{M-2} F_g(x_i) \odot F_{g-1}(x_i) \end{cases} \quad (2)$$

其中: $F'(x_i)$ 表示不同尺度的组连接特征。

由于相邻尺度组之间具有很强的关联性,对其进行组连接能有效增强特征的表征能力,且可实现

不同尺度信息组之间的互补,有利于遥感目标的准确定位。

1.2 自导注意力模块

超尺度块虽然能有效捕获目标船舶的超尺度信息,但是对图像进行编码时,容易忽略超端尺度上的细节信息以及重复使用不相关的冗余信息,导致相关信息的表征效果弱化。为解决此问题,引入自导注意力网络(Self-Guided Attention),对这些超尺度特征进行逐步细化有助于编码局部和全局语义,从而自适应地在局部和全局语义之间建立长期依赖关系,通过消除冗余信息以突出相关信息对目标船舶的表征。由于低层级特征包含大量的纹理细节信息,因此将超尺度特征 F_{HS} 与不同尺度特征 F'_S 相结合,使不同尺度特征映射的低层信息和高层语义相互补充,并将其输入到自导注意力模块中,能生成细化的注意力特征 A_S 。计算公式如式(3)所示:

$$A_S = G_{\text{AttModule}}^S(\text{Conv}(F'_S, F_{HS})) \quad (3)$$

其中: $G_{\text{AttModule}}^S(\cdot)$ 表示不同尺度的自导注意力模块; $\text{Conv}(\cdot)$ 表示卷积操作。

自导注意力模块由位置注意力(Position Attention, PA)^[18-19]和通道注意力(Channel Attention, CA)^[20]构成,位置注意力能获得丰富的上下文表示,而通道注意力可被看作是特定类的响应,能加强不同语义信息之间的关联性。

设输入到注意力的特征图为 $F \in \mathbb{R}^{C \times W \times H}$,其中: C 表示通道维度; W 表示宽维度; H 表示高维度。

PA是特征图 F 经过卷积块传递并重构生成新的特征图 $F_0^{\text{PA}} \in \mathbb{R}^{(W \times H) \times C'}$,利用上述同样操作得到 $F_1^{\text{PA}} \in \mathbb{R}^{C' \times (W \times H)}$ 的过程,2个特征图的空间注意力系数 $\alpha_S^{\text{PA}} \in \mathbb{R}^{(W \times H) \times (W \times H)}$ 如式(4)所示:

$$\alpha_{i,j}^{\text{PA}} = \frac{\exp(F_{0,i}^{\text{PA}}, F_{1,j}^{\text{PA}})}{\sum_{i=1}^T \exp(F_{0,i}^{\text{PA}}, F_{1,j}^{\text{PA}})}, T = W \times H \quad (4)$$

其中: $\alpha_{i,j}^{\text{PA}}$ 表示第 i 个位置对第 j 个位置的影响。再次对输入特征重构得到 $F_2^{\text{PA}} \in \mathbb{R}^{C \times (W \times H)}$ 。位置注意力对应的特征图 $F^{\text{PA},j}$ 如式(5)所示:

$$F^{\text{PA},j} = \lambda \sum_{i=1}^T \alpha_{i,j}^{\text{PA}} F_{2,i}^{\text{PA}} + F_j \quad (5)$$

通过位置注意力,进一步聚合了超尺度特征图的全局上下文语义。

CA是对输入特征图的通道进行重构,其注意力系数 $\alpha_S^{\text{CA}} \in \mathbb{R}^{C \times C}$ 如式(6)所示:

$$\alpha_{i,j}^{CA} = \frac{\exp(F_{0,i}^{CA}, F_{1,j}^{CA})}{\sum_{i=1}^{C \times C} \exp(F_{0,i}^{CA}, F_{1,j}^{CA})} \quad (6)$$

其中: $F_0^{CA} \in \mathbb{R}^{(W \times H) \times C}$; $F_1^{CA} \in \mathbb{R}^{C \times (W \times H)}$ 。CA对应的特征图 $F^{CA,j}$ 如式(7)所示:

$$F^{CA,j} = \kappa \sum_{i=1}^C \alpha_{i,j}^{CA} F_{2,j}^{CA} + F_j \quad (7)$$

通过CA增加了超尺度特征对不同类之间差异性的区分能力。

由于自导注意力模块被应用在多个超尺度特征图上,因此自导损失为多个模块的损失之和,定义为 τ_G ,如式(8)所示:

$$\tau_G = \sum_{s=1}^S \tau_i \quad (8)$$

其中: τ_i 表示不同超尺度注意力图的引导损失。通过自导注意力模块,可对超尺度的特征图进行逐步细化,在局部和全局语义之间建立长期依赖关系,并增强不同类之间特征图的差异性。

2 实验结果与分析

2.1 数据来源和评价指标

该数据集由Kaggle提供,共4 000张80像素×80像素的RGB高空间分辨率遥感图像,其中1 000张为有船舶图像,3 000张无船舶图像。为验证本文所提HSSGAN方法的可行性,对原始数据集进行了旋转、平移等增强处理,并将数据集随机划分为训练集、验证集和测试集。其中,20%的图像用于训练,并从中随机抽取其中10%进行5倍交叉验证。剩余80%的图像用于测试。

为进一步验证该方法的正确性,采用精准率(Precision, P)、召回率(Recall, R)和F1-Score值($F1$)等多种评测指标进行验证。计算公式如式(9)所示:

$$\begin{cases} P = \frac{T_{TP}}{T_{TP} + F_{FP}} \\ R = \frac{T_{TP}}{T_{TP} + F_{FN}} \\ F1 = \frac{2 \times P \times R}{P + R} \end{cases} \quad (9)$$

其中: T_{TP} 表示真阳性,即正确识别数; F_{FP} 表示假阳性,即错误识别次数; F_{FN} 表示假阴性。

2.2 实验参数设置

使用Adam优化函数对网络进行训练,学习率为0.000 5,批处理设置为64,迭代次数为250次。为证

明超尺度模块的有效性,对不同超尺度块的识别效果进行了验证,即设置为Hyper-1、Hyper-2、Hyper-3和Hyper-4。

为验证所提HSSGAN方法的优越性,与如下相关方法进行了比较。

1) HSF-Net^[21]。该方法称之为嵌入深度多尺度特征的光学遥感船舶检测,主要是将不同尺度的特征映射到同一尺度空间,并使用端对端网络进行训练检测。

2) CF-SDN^[22]。该方法为提高特征的表征能力,融合了不同层次的特征,由粗到精的对船舶进行了检测。

3) TP-FCN^[23]。该方法主要对不同深度的层分配不同任务,深层提供检测功能,浅层补充捕获特征信息,以实现精确定位,可有效缓解精确定位与特征表征能力之间的权衡问题。

2.3 结果分析

2.3.1 不同超尺度块间的比较

为验证本文所提HSSGAN方法的有效性,使用不同超尺度块在Kaggle实验数据集上进行测试验证。实验结果如表1所示,其中:Hyper-1表示1个超尺度块;Hyper-2表示2个超尺度块;Hyper-3表示3个超尺度块,即本文所提HSSGAN方法;Hyper-4表示4个超尺度块,以此顺推。

表1 不同超尺度块的实验结果

Table 1 Experimental results of different super-scale blocks

方法	精准率	召回率	F1-Score值
Hyper-1	0.897 41	0.853 02	0.868 55
Hyper-2	0.957 34	0.956 06	0.956 69
Hyper-3	0.961 58	0.973 01	0.966 78
Hyper-4	0.942 70	0.936 74	0.939 58
Hyper-5	0.930 14	0.910 02	0.919 97
Hyper-6	0.927 01	0.909 06	0.918 44
Hyper-7	0.921 44	0.903 11	0.912 18
Hyper-8	0.915 07	0.897 74	0.906 32
Hyper-9	0.908 87	0.894 32	0.901 53
Hyper-10	0.895 21	0.877 25	0.886 14
Hyper-11	0.879 40	0.859 21	0.869 18

从表1中可看出,随着超尺度块的增加,识别精度也随之增加。而当超尺度块为4,即Hyper-4时,识别精度开始下降,其F1值为0.939 58。而Hyper-11

的识别精度最低为0.869 18,因为较少的超尺度块不能有效地捕获遥感船舶的细节信息,且不能获得足量的判别特征,导致影响船舶的识别精度。当超尺度块足够多时,可能会增加冗余信息的使用率及增加网络参数量,导致网络出现拟合现象,从而使识别精度下降。

HSSGAN方法迭代200次的精准率、损失率和混淆矩阵如图2所示。图2(a)表示训练和验证集的准确率(Accuracy);图2(b)表示训练和验证集的损失率;图2(c)表示测试集的混淆矩阵。

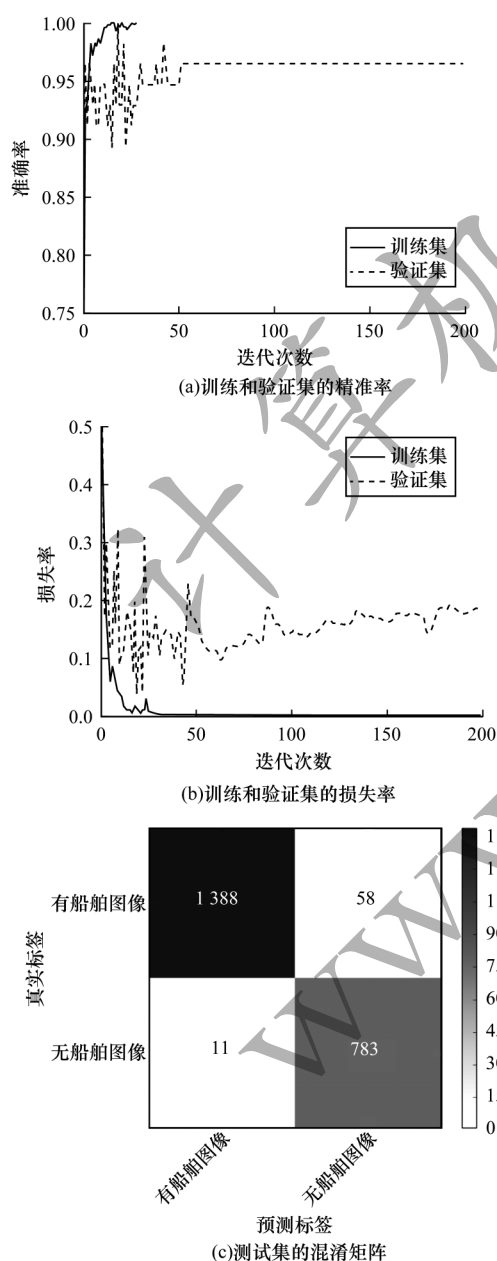


图2 HSSGAN的精准率、损失率和混淆矩阵
Fig.2 Accuracy rate, obfuscation matrix and loss rate of HSSGAN

从图2中可看出,随着迭代次数的增加,精准率先增加然后处于稳定状态,损失率则下降一定程度后处于稳定状态。

HSSGAN方法对不同尺度船舶的检测效果如图3所示。图3(a)~图3(c)的左部分为自导注意力对不同尺度变化船舶的热力图,右部分为识别结果。

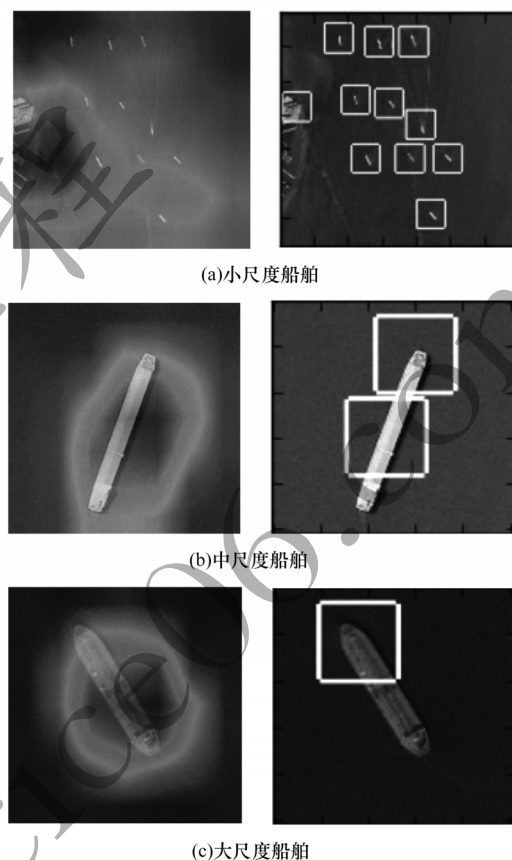


图3 不同尺度的识别效果

Fig.3 Recognition effects of different scales

由图3可看出,超尺度块能有效获取遥感船舶的超尺度特征,即对极端尺度的船舶特征具有较强的敏感性,特别是对尺度较小的船舶,如图3(a)所示。这是因为超尺度块捕获了遥感船舶的高层语义,同时也融合了低层级信息。

自导注意力网络可对超尺度块的特征图逐步细化,并在局部和全局语义间建立长期依赖关系,同时进一步丰富上下文语义,强调相关特征的关联。该功能尤其对尺度较大的遥感船舶非常明显,如图3(b)和图3(c)所示。

2.3.2 与相关方法的比较

为验证HSSGAN方法的优越性,与相关方法进行了对比,实验结果如表2所示。

表 2 与相关方法对比的实验结果

Table 2 Experimental results compared with relevant methods

方法	精准率	召回率	F1-Score 值
HSF-Net ^[21]	0.952 84	0.949 78	0.951 27
CF-SDN ^[22]	0.921 56	0.944 05	0.929 74
TP-FCN ^[23]	0.831 95	0.831 94	0.833 93
HSSGAN	0.961 58	0.973 01	0.966 78

从表 2 中可以看出,本文所提方法 HSSGAN 在精准率、召回率和 F1-Score 值上均取得最优效果,分别为 0.961 58、0.973 01 和 0.966 78。这是因为超尺度块有效捕获了遥感船舶的超尺度特征,从多个尺度对目标船舶进行了描述,同时自导注意力网络使这些特征更加精细,让相关特征受到了更多关注。

相比于 TP-FCN 方法,CF-SDN 和 HSF-Net 方法的识别效果更好,F1-Score 值分别提高了 0.095 81 和 0.117 34。因为这 2 种方法都采用了多尺度策略,从不同尺度对图像中的目标对象进行了描述,对比结果也证明了多尺度策略的有效性。

在小尺度上的识别效果如图 4 所示,其中椭圆形圈表示该区域内无船舶。

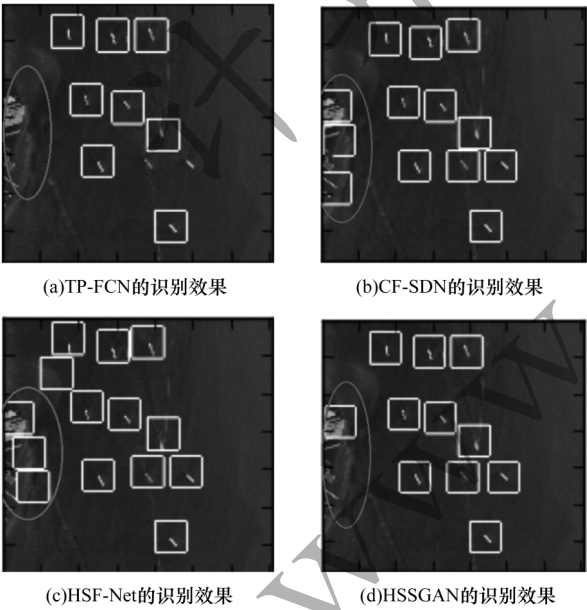


图 4 不同方法在小尺度上的识别效果

Fig.4 Recognition effects of different methods on a small scale

由图 4 可知,TP-FCN 方法虽然对椭圆形圈内的目标进行了精准识别,但对水中船舶的辨识度有待提高。CF-SDN 和 HSF-Ne 方法将红色圈内的其他目标误识别成船舶,但对水上目标进行了正确识别。

2.3.3 自导注意力模块的研究

为验证自导注意力网络对识别效果的影响,针对自导注意力网络进行了消融实验测试,结果如表 3 所示,其中:Non-guided 表示无自导注意力模块;PA-guided 表示只使用了位置注意力模块;CA-guided 表示只使用了通道注意力模块;加粗字体表示最大值。

表 3 自导注意力模块的消融实验结果

Table 3 Ablation results of self-directed attention module

方法	精准率	召回率	F1-Score 值
Non-guided	0.942 07	0.942 32	0.942 20
PA-guided	0.946 46	0.962 54	0.953 38
CA-guided	0.953 44	0.960 77	0.956 86
HSSGAN	0.961 58	0.973 01	0.966 78

从表 3 中可以看出,无自导注意力模块的 F1-Score 值仅为 0.942 20,识别效果最差。尽管位置注意力模块对船舶的识别效果有所提高,但通道注意力模块的识别效果要稍强于位置注意力模块,这证明在该数据集上通道注意力模块的贡献要高于空间注意力模块(PA-guided),即在船舶识别过程中,增加超尺度特征在类间的差异性能更好地提高识别效果,识别效果的可视化如图 5 所示。

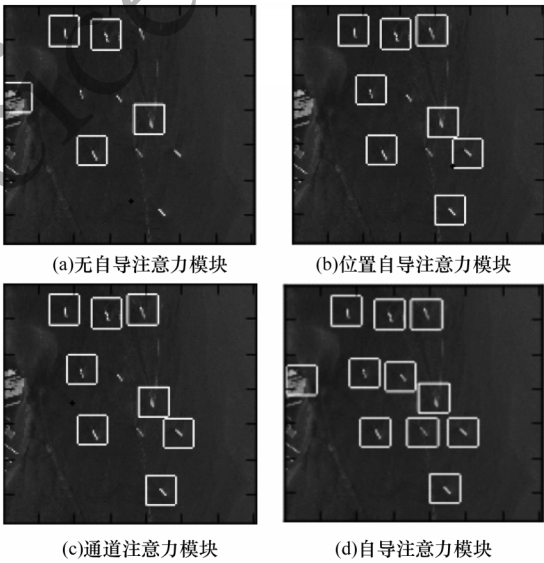


图 5 识别效果的可视化

Fig.5 Visualization of the recognition effect

由图 5 可知,图 5 (b)和图 5 (c)给出的识别效果一样,这是因为通道注意力模块的 F1 值仅比位置注意力模块提高了 0.004 44,这进一步说明了两者的共同使用的增强效果要高于单个组件,即所提的自导注意力模块细化的特征图更好地聚合了局部和全局语义。

3 结束语

本文设计超尺度自导注意力识别框架,采用超尺度的子空间组捕获遥感船舶的超尺度特征和不同层级的结构信息,使用自导注意力模块逐步聚合相关的上下文语义及滤除冗余信息,增加特征在不同类之间的差异性,并突出相关信息的表征能力。实验结果表明,所提模型在定量和定性上均优于HSF-Net、CF-SDN、TD-FCN等其他相关方法。

参考文献

- [1] CHENG G, HAN J. A survey on object detection in optical remote sensing images[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, 117(32): 11-28.
- [2] LI K, WAN G, CHENG G, et al. Object detection in optical remote sensing images: a survey and a new benchmark[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 159(45): 296-307.
- [3] 李文斌, 何冉. 基于深度神经网络的遥感图像飞机目标检测[J]. *计算机工程*, 2020, 46(7): 268-276.
- LI W B, HE R. Remote sensing image aircraft target detection based on deep neural network[J]. *Computer Engineering*, 2020, 46(7): 268-276. (in Chinese)
- [4] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//*Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington D. C., USA: IEEE Press, 2005: 886-893.
- [5] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [6] TAO C, TAN Y, CAI H, et al. Airport detection from large IKONOS images using clustered SIFT keypoints and region information[J]. *IEEE Geoscience and Remote Sensing Letters*, 2010, 8(1): 128-132.
- [7] XIAO Z, LIU Q, TANG G, et al. Elliptic fourier transformation-based histograms of oriented gradients for rotationally invariant object detection in remote-sensing images[J]. *International Journal of Remote Sensing*, 2015, 36(2): 618-644.
- [8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Region-based convolutional networks for accurate object detection and segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 38(1): 142-158.
- [9] HAN X, ZHONG Y, ZHANG L. An efficient and robust integrated geospatial object detection framework for high spatial resolution remote sensing imagery[J]. *Remote Sensing*, 2017, 9(7): 666.
- [10] CHEN Z, ZHANG T, OUYANG C. End-to-end airplane detection using transfer learning in remote sensing images[J]. *Remote Sensing*, 2018, 10(1): 139.
- [11] DENG Z, SUN H, ZHOU S, et al. Multi-scale object detection in remote sensing imagery with convolutional neural networks[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, 145: 3-22.
- [12] ZHONG Y, HAN X, ZHANG L. Multi-class geospatial object detection based on a position-sensitive balancing framework for high spatial resolution remote sensing imagery[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, 138(23): 281-294.
- [13] DING P, ZHANG Y, DENG W J, et al. A light and faster regional convolutional neural network for object detection in optical remote sensing images[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, 141(24): 208-218.
- [14] WANG C, BAI X, WANG S, et al. Multiscale visual attention networks for object detection in VHR remote sensing images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2018, 16(2): 310-314.
- [15] ZHENG Z, ZHONG Y, MA A, et al. HyNet: hyper-scale object detection network framework for multiple spatial resolution remote sensing imagery[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 166(14): 1-14.
- [16] HAQUE M R, MISHU S Z. Spectral-spatial feature extraction using PCA and multi-scale deep convolutional neural network for hyperspectral image classification[C]//*Proceedings of the 22nd International Conference on Computer and Information Technology*. Washington D. C., USA: IEEE Press, 2019: 1-6.
- [17] SU B, LIU Y, FU J, et al. Reduction of motion artifacts in head CT exams using multi-scale convolutional neural network[EB/OL]. [2020-07-09]. https://www.researchgate.net/publication/339839413_Reduction_of_motion_artifacts_in_head_CT_exams_using_multi-scale_convolutional_neural_network.
- [18] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation[C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Washington D. C., USA: IEEE Press, 2019: 3146-3154.
- [19] 尤洪峰, 田生伟, 禹龙, 等. 基于 Word Embedding 的遥感影像检测分割[J]. *电子学报*, 2020, 16(1): 75-83.
- YOU H F, TIAN S W, YU L, et al. Embedding remote sensing image detection segmentation based on Word Embedding[J]. *Journal of Electronics*, 2020, 16(1): 75-83. (in Chinese)
- [20] LIU W, SUN Y, JI Q. MDAN-UNet: multi-scale and dual attention enhanced nested U-Net architecture for segmentation of optical coherence tomography images[J]. *Algorithms*, 2020, 13(3): 60-71.
- [21] LI Q, MOU L, LIU Q, et al. HSF-Net: multiscale deep feature embedding for ship detection in optical remote sensing imagery[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(12): 7147-7161.
- [22] WU Y, MA W, GONG M, et al. A coarse-to-fine network for ship detection in optical remote sensing images[J]. *Remote Sensing*, 2020, 12(2): 246-253.
- [23] LIN H, SHI Z, ZOU Z. Fully convolutional network with task partitioning for inshore ship detection in optical remote sensing images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(10): 1665-1669.