doi: 10.6046/zrzyyg.2021020

引用格式:于新莉,宋妍,杨森,等.结合空间约束的卷积神经网络多模型多尺度船企场景识别[J].自然资源遥感,2021,33(4):72-81.(Yu X L,Song Y,Yang M, et al. Multi – model and multi – scale scene recognition of shipbuilding enterprises based on convolutional neural network with spatial constraints[J]. Remote Sensing for Natural Resources,2021,33(4):72-81.)

结合空间约束的卷积神经网络 多模型多尺度船企场景识别

于新莉1,宋妍1,杨森1,黄磊2,张艳杰2

(1. 中国地质大学(武汉)地理与信息工程学院,武汉 430074; 2. 国家卫星海洋应用中心,北京 100081)

摘要:船企场景识别对修复沿岸生态环境、保护水域环境以及促进船舶产业的协调发展具有现实意义,但传统方法 基于中、低层次的特征难以实现卫星遥感图像中船企的自动识别。为此,提出了结合空间约束的卷积神经网络多 模型多尺度船企场景识别方法。首先分别采用全局尺度的船企场景和局部尺度船坞(台)、厂房和船只样本训练多 个卷积神经网络模型,并进行多模型多尺度检测;进而对局部对象进行像素级定位并计算对象空间距离;最终结 合多尺度检测结果、对象标签组合方式、对象空间距离进行船企场景综合判别与提取。将此方法分别应用于中国 江苏省、日本长崎县和爱媛县周边以及韩国木浦市和巨济市周边5个典型造船密集区。结果表明,江苏省整体识 别精确度为87%,召回率为85%;日本研究区整体识别精确度为91%,召回率为87%;韩国研究区整体识别精确 度为85%,召回率为92%。实验结果表明,此方法可以较好地实现遥感船企复杂场景的识别。

关键词:卫星遥感;造船企业;卷积神经网络;多尺度;空间距离约束

中图法分类号: TP 79 文献标志码: A 文章编号: 2097 - 034X(2021)04 - 0072 - 10

0 引言

船舶产业为海洋资源开发、国防建设等提供帮助^[1],在维护海洋权益、探索海洋资源等多个领域承担着重要角色,得到世界各国尤其是沿海地区和国家的重视^[2]。结合卫星遥感图像成像范围广、重访周期短等优点,提取造船企业位置,对修复沿岸生态环境、保护水域环境以及促进船舶产业的协调发展具有现实意义。

与自然景物图像相比,卫星遥感数据具有尺度 大、场景复杂等特点。针对遥感数据中复杂场景识 别问题,部分学者结合卷积神经网络,运用多示例、 多标签^[3]等思想提取场景的多语义、多尺度特征, 取得了良好的效果^[4-7]。空间分布也作为重要特征 应用于场景识别中,多项研究表明充分利用场景的 空间布局信息能提高场景识别的准确性^[8-12]。但 此类方法需要对场景进行多次空间划分,可能导致 大型地物被过度分割产生特征丢失的情况,降低了 识别准确性[13]。

船企作为卫星影像中的一种具有高阶语义特性的复杂场景,常规遥感分类识别方法无法充分利用船企场景特征,难以取得理想的结果。本文综合了船企的多尺度、多语义特征及空间关系特征进行船企场景充分表达,提出了一种结合空间约束的卷积神经网络多模型多尺度的船企场景识别方法。

1 造船企业特征

利用卷积神经网络识别船企前应先掌握船企的 图像特征,建立船企解译标志、统计船企内部地物空 间特征并了解船企分布情况,为后续样本集的制作、 识别算法的研究提供参考。

1.1 船企解译标志

分析船企在遥感影像上的特征表现,如图1,船 企内部可以分为船坞(台)、船只和厂房,解译标志 见表1。

收稿日期: 2021-01-15;修订日期: 2021-05-13

基金项目:海洋领域融合应用示范项目(编号:2020010004)及国防科工局民用"十三五"航天预先研究项目"星载高分辨率红外高光 谱相机及应用技术"(编号:D040104)共同资助。

第一作者:于新莉(1997-),女,硕士,主要研究方向为深度学习、海岸带多源遥感监测等。Email: yu_xinli@ cug. edu. cn。

通信作者: 宋 妍(1980-),女,博士,副教授,主要从事摄影测量、遥感变化检测、海岛与海岸带遥感等研究。Email: songyan@ cug. edu. cn。





表1 船企内部解译标志

Tab. 1 Interpretation marks within the shipbuilding enterprise

地物 类型	解译标志		
船 坞 (台)	是指修造船的专用 场地,坞式或台式 建筑物,主要分布 在水陆交界的位置		
船只	位于船坞(台)内 或停靠在码头附 近,从内部纹理可 以明显看出还未构 建完整,正在舾装、 涂装		
厂房	蓝色、灰色或白色的矩形建筑,排列整齐,有规则纹理, 比普通房屋面积大,周围存在材料 堆放		

1.2 船企内部空间特征分析

根据船企解译标志、船企空间分布知识^[2]和中 国船舶工业年鉴资料^[14],基于资源三号卫星遥感影 像进行全国船企矢量提取。进一步,依据船企矢量 提取结果,对船企内部地物的面积和地物之间的空 间距离进行统计。

1) 船企建设面积统计。船企面积分布差异较 大,最大的船企面积达 6.8×10⁶ m²,而最小的船企 面积小于 1×10⁴ m²,占地面积在 7×10⁵ m² 以下的 船企数目占 85%。

2) 内部地物占地面积统计。对船坞(台)、厂 房、船只3 类重要船企内部地物进行分析与统计: 其中船坞(台)占地面积在5×10⁴ m² 以内的占 97%;在5×10⁴~1×10⁵ m²之间的为3%。厂房面 积在5×10⁴ m² 以内占97%,而5×10⁴~1×10⁵ m²之 间为3%。

3) 内部地物空间距离统计。船坞(台) 是船企

内最核心的地物,能够反映船企的生产能力,是船企 解译中具有唯一性的判别标志。因此,本文在建立 内部地物空间关系时,以船坞(台)为基准,对船坞 (台)与船只,船坞(台)与厂房的质心间的距离进行 近邻分析,计算厂房、船只与船坞(台)质心之间的 最近距离。

船坞(台)与厂房距离小于400 m的占93%,距 离大于400 m占比7%;船坞(台)与船只之间的距 离小于400 m的占92%,距离大于400 m的占8%。

2 遥感图像船企场景识别理论

2.1 基准网络介绍

本文选用经典残差网络(ResNet)作为基准网络。残差学习单元(图 2)分为两个分支,分别为残差函数 F(x)和输入的恒等映射 x,经过残差学习后的期望输出为 H(x) = F(x) + x,该段网络需要学习的目标函数即为 F(x) = H(x) - x。这种近路连接的方式在一定程度上降低了学习难度,不需要引入其他参数增加网络训练的计算负担,同时可以让梯度信息被有效利用和传播下去,较好地避免了网络



图 2 残差学习单元 Fig. 2 Residual learning unit

层数越来越深带来的梯度弥散现象^[15]。

2.2 多模型多尺度检测

船企内部地物之间分布密集,且形状不规则,因 此不同类别样本存在重叠和相似。多个具有相似特 征的样本一起训练时,会产生特征竞争现象,导致公 用特征被某一类别独占,特征难以被充分利用,无法 达到预期的类别检测效果。

本文提出多模型多尺度船企场景识别方法。该 方法基于全局尺度场景样本训练得到船企场景模 型,在局部尺度下,分别采用船坞(台)样本、船只样 本和厂房样本单独训练多个对象模型,对同一场景 采用训练后的多个模型进行检测,避免类别得分被 平均而导致漏检,如图3所示。



图 3 多模型检测流程 Fig. 3 Multi – model detection process

2.3 多模型多尺度场景识别

本文采用全局尺度模型进行船企场景初检测, 获取船企候选区域,并在候选区域内开展局部对象 检测,得到各局部对象的粗略位置和语义标签。为 提升场景识别的准确性,结合弱监督学习方法精化 船坞(台)、船只和厂房在图像中的位置分布,计算 各地物间的空间距离。最终结合多尺度模型检测结 果、局部对象标签组合方式以及空间距离条件进行 船企场景的综合识别。

2.3.1 全局尺度初检测

为缩小船企检测范围,首先基于全局尺度船企 场景模型在研究区进行初步滑窗检测,获取船企候 诜区域(图4)。



(a) 全局初检

(b) 初检细节展示 多模型多尺度船企识别流程 图 4

(c)局部对象检测

(d) 局部对象检测结果

Fig. 4 Multi - model and multi - scale shipbuilding enterprise recognition process

2.3.2 局部尺度精检测

在全局尺度初检测的基础上,将场景切割为 多个局部图像块,以此为依据开展局部尺度精检 测(如图4(c)所示,分别沿着图中箭头方向进行 滑窗检测)。最后提取出船坞(台)、船只、厂房对 象的局部图像块(如图4(d)所示,红色框为船坞 (台),黄色框为厂房,蓝色框为船只)。此时,船企 可视为由多个局部图像块的标签所组成的高阶语 义场景。

2.3.3 空间距离约束

将船企内部地物空间距离作为约束条件进一步 提高船企场景的识别精度。为了获得较为精确的空 间距离,应获取船企内部对象的位置信息。目前采 用卷积神经网络获取位置信息的成熟算法主要基于 强监督学习思想[16],需要提前大量标注目标位置的 训练样本。但对于船企这一复杂人工岸线,其具有 语义丰富、内部构造复杂、地物边界不清晰、形状不 规则等特点,为船企场景样本加上范围和位置标注 框的人力成本较高且具有主观性。如果标注者缺乏 船企场景的专家知识,则难以保证目标框选的准确 性。基于船企提取的实际需求以及样本标注的困难 性,考虑到方法的适用性和实用性,强监督学习的目 标检测并不是最合适的方法。

本文采用弱监督学习中的不精确监督方法^[17], 用图像级别标注的样本,借助卷积神经网络深层特 征的定位特性提取目标大致位置。采用梯度加权类 激活映射^[18](Grad – CAM)可视化每类对象的像素 级位置。基于深层特征图的定位特性,Grad – CAM 的每个特征图对指定类别的权重计算见公式(1):

$$\alpha_k^c = \frac{1}{z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} , \qquad (1)$$

式中: z 为该特征图中像元数量; y° 为对应类别 c 的 分数; A_{ij}^{k} 为第 k 个特征图中(i,j)所在位置的像素 值; $z = m \times n, 1 < i < m, 1 < j < n_{\circ}$ 在获取指定类别 c对输出的所有特征图的权重后, 对权重和特征图进 行加权平均即可得到 Grad - CAM。但为了限制结 果中只包含对指定类别有正影响的像元点,在加权 后的结果上增加了一个 ReLU 层,公式为:

$$L_{\text{Grad-CAM}}^{c} = ReLU\left(\sum_{k} \alpha_{k}^{c} A^{k}\right) \quad (2)$$

Grad - CAM 的优点在于可以不改变网络结构 来实现可视化,且应用十分便捷,可以直接扩展到任 意训练完成的卷积神经网络中,对网络的训练过程 没有要求。

基于深层网络的定位特性和 Grad - CAM 方法 将对象所在的大致像素位置提取出来,开展空间距 离约束。具体过程如图 5 所示,左图为船企初步识 别结果。对其内部对象进行检测,分别检出厂房 (红色框)、船坞(台)(蓝色框)、船只(绿色框),对 检测出的局部对象进行可视化,获取类别响应区域 中红色像元所在位置,据此进行船企内部不同地物 的空间约束。



图 5 局部对象空间关系约束 Fig. 5 Local object spatial relationship constraints

2.4 场景综合评判和识别

结合场景多尺度标签得分、局部对象标签组合、 局部对象空间距离约束3个决策条件对船企场景进 行综合评判,并提取满足3类条件的区域,具体为:

1)多尺度标签得分。通过调低船企场景置信 度阈值尽可能提取候选场景以避免漏检,实验中该 阈值设置为0.8。设置较高的局部对象检测阈值进 一步减少局部对象错检,实验中该阈值为0.95。置 信度高于两种阈值的区域,符合船企场景的多尺度 标签得分条件。

2)局部对象标签组合条件。局部对象组合方 式会为场景分配不同类别标签,只有当场景内同时 存在船坞(台)、船只和厂房标签,认为其满足船企 标签组合条件。 3)局部对象空间距离约束条件。计算船坞 (台)-船只和船坞(台)-厂房的空间距离,依据船 企内部地物空间统计规律,当面积和地物之间的距 离结果符合时,则认为满足对象空间距离条件。

3 船企场景识别流程

实验中计算机软硬件参数为: CPU 为 i7 - 8700,显卡为 RTX2070,操作系统为 Ubuntu16,框架 为 Pytorch1.1.0,编程语言为 Python。

长江沿岸是我国船企集群程度最高的区域,因此,选择我国长江流域江苏段(范围 E118°27′~ 121°57′,N31°38′~32°22′,总面积6307 km²)为第 一个实验区;采用2018年资源三号融合后正射影 像进行船企识别。为进一步验证方法的适用性, 依据空间分辨率2m左右的卫星影像,分别在日 本和韩国选取造船业发达的区域进行识别实验。 日本研究区为造船业较为发达的长崎县和爱媛县 周边部分区域。其中,长崎县区域(简称日本研究 区一)位于 E129°42′~130°1′,N32°39′~33°7′; 爱媛 县区域(简称日本研究区二)位于 E132°54′~133° 45′,N33°53′~34°21′。2个研究区总占地面积为 5 472 km²。韩国研究区为木浦市和巨济市周边区 域。木浦市区域(简称韩国研究区一),位于 E126°13′~126°30′,N34°33′~34°55′;巨济市区 域(简称韩国研究区二),位于E128°24′~129°3′, N34°47′~35°12′。韩国2个研究区总占地面积为 3 738 km²。

船企场景识别流程如图 6 所示,接下来对船企 场景识别的整个实验流程进行介绍。





3.1 样本集制作与影像预处理

首先进行样本集制作,充分利用现有基准数据 集 AID^[19],NWPU - RESISC45^[20],选择常见土地覆 盖类型作为背景负样本。采用两类卫星遥感数据 (资源三号和 Google Earth 2 m 分辨率遥感影像), 依靠人工目视解译识别船企并制作船企正样本,包 含全局船企场景样本、局部船坞(台)、厂房、船只样 本。船企样本集制作充分考虑全国不同地区、不同 格局、不同规模的船企的差异性,依据各地物大小不 一,制作的样本大小不同,在输入网络时,进行处理, 大小统一为 224 × 224。样本数量如表 2 所示。对 待检测试验区影像进行预处理,由于船企位于水陆 交界区域,因此依据水边线进行缓冲区提取,剔除影像中大部分陆地区域。

表 2 样本数量

Tab. 2Number of samples

样本类型	船企场景	船坞(台)	厂房	船只
样本数量/个	1 032	1 104	1 300	1 008

3.2 多模型多尺度训练

 1)样本划分与增强。样本的选取需要保证训 练样本的数量,同时应考虑不同地域样本的差异 性和多样性。为识别江苏省船企,训练模型时采 用除江苏省外的全国船企样本集。国内外船企厂 区内部地类具有高度相似性,因此为识别日本和 韩国船企,训练模型时采用全国船企样本。在划 分训练和测试样本时,按照8:2的比例随机划分。 对样本进行了随机翻转和随机旋转操作,并进行 色调变化调整。

2)网络训练参数设置与结果。基础网络采用 ResNet50,通过对在 ImageNet 数据集上训练收敛的 ResNet50网络进行微调更新网络相关参数。运用 基于动量的随机梯度下降方法,动量因子0.9,批处 理大小为64,训练过程中迭代次数为200,采用交叉 熵损失函数。学习率初始设置为0.0001,并根据迭 代次数和接近收敛程度调整学习率,每次调整将学 习率减小10倍。表3为网络训练收敛后,船企场 景、船坞(台)、厂房、船只4类地物的分类精确度、 召回率、漏警率和虚警率结果。

表 3 网络训练结果 Tab. 3 Network training results

rub.5 Retwork training results				(70)	
指标	船企场景	船坞(台)	厂房	船只	-
精确度	99.5	96.3	98.8	97.8	_
召回率	100	99.2	99.4	98.4	
虚警率	0.5	3.7	1.2	2.2	
漏警率	0	0.8	0.6	1.6	

3.3 船企场景识别实验

通过训练获得4类模型:船企场景模型、船坞 (台)模型、船只模型和厂房模型。采用上述多个模 型进行船企多模型多尺度检测,并采用弱监督学习 的方法计算场景内部地物之间的空间距离,结合多 个约束条件进行船企场景的综合评判与识别。图 7-9为不同实验区的检测过程和结果。具体步骤





(d) 区域二全局检测结果







(f) 区域二综合判别结果



Fig. 7

(a) 全局检测结果



(e) 区域二局部检测结果

图 7 江苏省研究区域结果展示

Display of the results of the study area in Jiangsu Province



(c) 综合判别结果

(b) 局部检测结果 图 8 日本研究区域结果展示 Fig. 8 Display of results of the study area in Japan



(a) 全局检测结果



(b) 局部检测结果

Fig. 9 Display of results of the study area in Korea

图 9

韩国研究区域结果展示



(c) 综合判别结果

(0%)

如下:依据1.2节的分析结果,采用2m分辨率遥 感影像进行船企识别时,将船企全局场景检测框长 度设置在400~600像素之间;局部对象检测框长 度设置在100~200像素之间。为避免图像块切割 导致场景特征被破坏,将滑窗移动步长设置的较小 (长度约为局部对象检测框的1/3)。采用非极大值 抑制方法将大面积重复的检测框删除。最后,采用 结合场景多尺度检测标签得分、对象多标签组合方 式和对象空间距离约束进行船企场景综合判别。

4 实验结果分析

4.1 江苏省结果与精度

目前多种基础网络如 AlexNet^[21] 及 VGG16^[22] 等均可以应用在本文的框架中,为了验证 ResNet50 基础网络在船企识别中的有效性,分别采用 AlexNet 和 VGG16 网络模型作为基础网络,使用本文方法对 江苏省船企进行识别,模型训练和检测参数设置根 据实际情况微调,以达到最优效果。识别结果的精 确度、召回率、虚警率和漏警率如表4 所示。可以看 出,AlexNet 和 VGG16 基础网络识别结果的虚警率 较高,对船企和其他沿江目标的区分能力有限,识别 效果不如 ResNet50 网络模型。

表 4 基于不同基础网络模型的江苏省船企识别结果 Tab. 4 Recognition results of shipbuilding enterprises

in Jiangsu Province based on different

	basic network models		
基础网络	AlexNet	VGG16	ResNet50
精确度	46.63	42.36	87
召回率	73.17	69.92	85
虚警率	53.37	57.64	13
漏警率	26.83	30.08	15

以 ResNet50 作为基础网络,使用本文方法识别 出的江苏省船企分布如图 10 所示,区域内共识别出 121 个船企,其中正检 105 个,错检 16 个,漏检 18 个 船企。识别结果细节见表 5。



图 10 江苏省船企分布的最终识别结果 Fig. 10 Final recognition results of the distribution of shipbuilding enterprises in Jiangsu Province

表 5 江苏省船企识别结果细节展示

Tab. 5 Detailed results in Jiangsu Province



进一步分析,发现正确识别结果主要包含大中 型船企以及特征较为明显的小型船企和滩涂船企; 错检结果主要为港口和沿海工厂;漏检主要为特征 不显著的中小型和滩涂船企。

4.2 境外研究区域船企识别结果分析与精度评价

4.2.1 日本研究区域

日本研究区一船企识别结果如图11(a)所示,



(a) 区域一

图 11 日本研究区域的船企识别结果 tion results of shiphuilding entermines in Jan

Fig. 11 The recognition results of shipbuilding enterprises in Japan study area

共检出5处船企场景,其中4处为正检场景,1个为 错检场景。日本研究区二识别结果如图 11(b) 所 示,共检出17个检测框,其中16个为正检,1个为 错检,并存在3处漏检。对日本2个研究区船企整 体识别结果进行统计,日本实验区船企识别的精确 度为91%, 召回率为87%, 虚警率为9%, 漏警率为 13%,证明制作的船企样本数据集和多模型多尺度 方法应用于日本区域依然有效。表6中显示日本研





究区识别结果示例,漏检一个大型和两个小型船企, 发生错检为沿海工厂和港口。

4.2.2 韩国研究区域

韩国研究区识别结果如表7和图12所示。研 究区一共检出6个正检场景(图12(a))。韩国研究 区二共检出16个为正检场景,4个错检场景, 存在2个漏检船企(图12(b))。韩国研究区船企 的总体识别精确度为85%,召回率为92%。对识别



Tab. 7 Detailed display of recognition results of shipbuilding enterprises in Korean research areas





(a) 区域一

图 12 韩国研究区域的船企识别结果

Fig. 12 The recognition results of shipbuilding enterprises in Korea study area

结果分析得知,错检主要为港口,而漏检主要为小型 船厂。

通过对日韩研究区域船企识别的结果分析,可 知本文方法可以较好地对日韩区域进行船企识别。 漏检和错检主要原因为:船企内部空间距离、厂房 色调等方面与国内存在差异,如国内船企厂房主要 为棕色、白色和蓝色,而日本部分船企的厂房呈浅绿 色;同时,日韩研究区中少量船企建在岛屿上,空间 布局较为独特。

5 结论

针对船企语义丰富、场景复杂的特点,本文充分 利用了场景多尺度、多语义、空间布局信息,提出了 一种结合空间距离约束的多模型多尺度船企场景识 别方法。将该方法应用于国内外研究区域进行船企 识别,实验结果证明了本文方法的可行性和有效性, 提取的船企空间分布情况能够为政府产能调控、产 业布局提供宏观决策依据。

参考文献(References):

[1] 杨 青.长三角船舶产业集群研究[D].上海:上海海事大学, 2007.

Yang Q. Research on the shipbuilding industry cluster in the Yangtze River delta [D]. Shanghai: Shanghai Maritime University, 2007.

[2] 刘辉,史雅娟,曾春水.中国船舶产业空间布局与发展策略
 [J]. 经济地理,2017,37(8):99-107.
 Liu H,Shi Y J,Zeng C S. Spatial layout and development strategy

of the ship – related industry in China[J]. Economic Geography, 2017,37(8):99 – 107.

- Zhou Z H, Zhang M L, Huang S J, et al. Multi instance and multi – label learning[J]. Artificial Intelligence, 2012, 176(1):2291 – 2320.
- [4] Li Z L, Xu K, Xie J F, et al. Deep multiple instance convolutional neural networks for learning robust scene representations [J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58 (5): 3685 – 3702.
- [5] Liu W H, Li Y D, Wu Q. An attribute based high level image representation for scene classification [J]. IEEE Access, 2019, 7: 4629 – 4640.
- [6] Khan N, Chaudhuri U, Banerjee B, et al. Graph convolutional network for multi – label VHR remote sensing scene recognition [J]. Neurocomputing, 2019, 357:36 – 46.
- [7] Chen Z M, Wei X S, Wang P, et al. Multi Label Image Recognition with Graph Convolutional Networks [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2019.
- [8] Clement M, Kurtz C, Wendling L. Learning spatial relations and shapes for structural object description and scene recognition [J]. Pattern Recognition, 2018, 84:197-210.
- [9] Lazebnik S, Schmid C, Ponce J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories [C]// Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference, 2006.
- [10] Jiang Y, Yuan J, Gang Y. Randomized spatial partition for scene rec-

ognition[M]. Berlin:Springer Berlin Heidelberg, 2012:730-743.

- [11] Weng C Q, Wang H X, Yuan J S, et al. Discovering Class Specific Spatial Layouts for Scene Recognition [J]. IEEE Signal Processing Letters, 2017, 24(8):1143 – 1147.
- [12] He K M, Zhang X Y, Ren S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37 (9):1904-1916.
- [13] Xie L, Lee F F, Liu L, et al. Scene recognition: A comprehensive survey[J]. Pattern Recognition, 2020, 102:107205.
- [14] 船舶工业年鉴编委会.中国船舶工业年鉴 2018[Z].[s.1]:船 舶工业年鉴编委会,2018.
 Editorial Committee of Yearbook of Shipbuilding Industry. Yearbook of China shipbuilding industry 2018[Z].[s.1]:Editorial Committee of Yearbook of Shipbuilding Industry,2018.
- [15] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas:2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2016: 770-778.
- [16] 颜 荔. 基于卷积神经网络的遥感图像飞机目标识别研究
 [D]. 合肥:中国科学技术大学,2018.
 Yan L. Research on airplane target recognition in remote sensing images based on convolutional neural network[D]. Hefei:University of Science and Technology of China,2018.
 [17] Zhou Z H A brief introduction to weakly supervised learning[1].
- $[\,17\,]\,$ Zhou Z H. A brief introduction to weakly supervised learning[J]. National Science Review, 2018, 5(1):44–53.
- [18] Selvaraju R R, Cogswell M, Das A, et al. Grad CAM: Visual explanations from deep networks via gradient – based localization[J]. International Journal of Computer Vision, 2020, 128(2):336 – 359.
- [19] Xia G S, Hu J W, Hu F, et al. AID: A benchmark data set for performance evaluation of aerial scene classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(7):3965 – 3981.
- [20] Cheng G, Han J W, Lu X Q. Remote sensing image scene classification: Benchmark and state of the art [J]. Proceedings of the IEEE, 2017, 105(10): 1865 - 1883.
- [21] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6):84 - 90.
- [22] Karen S, Andrew Z. Very deep convolutional networks for large scale image recognition [J]. arXiv, 2014.

Multi – model and multi – scale scene recognition of shipbuilding enterprises based on convolutional neural network with spatial constraints

YU Xinli¹, SONG Yan¹, YANG Miao¹, HUANG Lei², ZHANG Yanjie²

School of Geography and Information Engineering, China University of Geosciences, Wuhan 430074, China;
 National Satellite Ocean Application Service, Beijing 100081, China)

Abstract: The scene recognition of shipbuilding enterprises is of practical significance for the restoration of the coastal ecological environment, the protection of water environment, and the promotion of the coordinated development of shipbuilding enterprises. However, it is difficult to realize the automatic recognition of shipbuilding enterprises from satellite remote sensing images based on traditional medium – and low – level features. Therefore,

this paper proposes a multi – model multi – scale scene recognition method of shipbuilding enterprises based on a convolutional neural network with spatial constraints and the steps are as follows. Firstly, train multiple convolutional neural network models using the samples of global – scale shipbuilding enterprise scenes and local – scale docks (slipways), workshops, and ships individually, and conduct multi – model multi – scale detection. Then, locate local – scale objects at a pixel level and calculate the spatial distance of the objects. Finally, conduct comprehensive judgment and extraction of the shipbuilding enterprise scenes according to the multi – scale detection results, the combination method of object tags, and the spatial distance of objects. The method was applied to five typical shipbuilding intensive areas in Jiangsu Province, China, the surrounding areas of Nagasaki and Ehime prefectures, Japan, and Mokpo and Geoje cities, South Korea. As a result, the overall recognition accuracy and recall rate were 87% and 85%, respectively in Jiangsu Province, were 91% and 87%, respectively in the study area in Japanese, and were 85% and 92%, respectively in the study area in South Korean. The experimental results show that this method can realize the effective recognition of the complex scenes of shipbuilding enterprises based on remote sensing images.

Keywords: satellite remote sensing; shipbuilding enterprises; convolutional neural networks; multi - scale; spatial distance constraint

(责任编辑:李瑜)