

doi: 10. 6046/zrzyyg. 2022307

引用格式: 蒋卓然,周鑫鑫,曹伟,等. 基于协同判读机制的养殖蟹塘遥感智能检测方法[J]. 自然资源遥感,2023,35(3):25 – 34. (Jiang Z R,Zhou X X,Cao W,et al. Intelligent detection of crab ponds using remote sensing images based on a cooperative interpretation mechanism[J]. Remote Sensing for Natural Resources,2023,35(3):25 – 34.)

基于协同判读机制的养殖蟹塘遥感智能检测方法

蒋卓然¹,周鑫鑫^{2,3,4},曹伟⁵,王亚华^{1,3,4},吴长彬^{1,3,4}

(1. 南京师范大学地理科学学院,南京 210023; 2. 南京邮电大学地理与生物信息学院,南京 210023;
3. 南京师范大学虚拟地理环境教育部重点实验室,南京 210023; 4. 江苏省
地理信息资源开发与利用协同创新中心,南京 210023;
5. 南京国图信息产业有限公司,南京 210000)

摘要: 挖塘养蟹是耕地“非粮化”行为的一种,若及时发现制止,将对国家粮食安全造成危害。为了应对这一行为在遥感智能解译工作中所存在的人工判读量大、核查效率不足的挑战,提出了一种基于协同判读机制的养殖蟹塘遥感智能检测方法,该方法集成了 HRNet 分割网络和 Swin – Transformer 分类网络模型,并进一步介入人工核查,提高检测精度和工作效率。以江苏省南京市高淳区为研究区域进行了实验,结果表明,提出的基于协同判读机制的耕地“非粮化”遥感智能检测方法能够自动筛去 83. 4% 的待检测图斑,最终识别精度为 0. 972,可在大幅降低识别难度与人工核查工作量的同时提高检测精度,为实现准确高效的蟹塘等“非粮”地物检测提供可靠的解决思路。

关键词: 协同判读机制; HRNet; Swin – Transformer; 蟹塘检测; 非粮化

中图法分类号: TP 79 **文献标志码:** A **文章编号:** 2097 – 034X(2023)03 – 0025 – 10

0 引言

大闸蟹是我国的一大典型经济作物,其具有可观的经济收益和庞大的消费市场。据中商产业研究院《2020 年大闸蟹行业市场规模及产业布局情况分析》,2019 年大闸蟹产业规模突破亿元,年增速约 20%,且 2020 年我国大闸蟹市场规模或将接近 1 500 亿元。与此同时,据中国渔业统计年鉴数据显示,随着人工养殖技术的成熟和养殖规模的壮大,2019 年我国大闸蟹养殖产量达 77. 87 万 t,2020 年产量增至 80. 6 万 t,同比增长 3. 51%。因此,大闸蟹已成为养殖行业的重要经济作物之一。然而,由于蟹塘占地面积广、养殖环境要求苛刻等原因,现有耕地上挖塘养蟹行为日益增多,这不仅对耕作层造成破坏,而且会导致地区粮食产量下滑,妨害耕地可持续发展,对地区粮食产量与安全产生不良影响。因此,利用遥感影像对违规养蟹行为进行实时监测监管,是解决耕地变更蟹塘这一“非粮化”^[1–3]问题

的重点与难点。

目前,基于遥感影像进行地物检测的传统手段主要有基于像元和面向对象 2 类方法^[4–8],二者的实质均为通过提取目标地物的纹理、波段、物候等特征来分割特定的地物目标。然而,传统监督分类方法存在着同谱异物与同物异谱^[9–10]等问题,且当目标存在混合像元等复杂特征时,识别精度下降明显,检测效果较差。而深度学习作为近年来比较热门的研究方法,其在遥感影像的地物判读中已经有了广泛的应用^[11–16],且其精度和速度均高于传统解译方法。然而,由于样本质量参差不齐以及模型结构自身的局限性,神经网络检测出来的结果往往存在误差,具体表现为误判、漏判和图斑形状异常等,使得网络的检测结果必须经过人工核查才能投入使用。但在影像图幅较大的情况下,核查会因图斑数量庞大而耗费较多的人力和时间,这使得整个监测流程的效率下降。因此,为应对传统遥感智能解译工作中存在的人工判读量大、核查效率不足的挑战,亟需提出一种新的监测识别方法,在引入人工核查机制的同时兼顾人

收稿日期: 2022 – 07 – 27; 修订日期: 2022 – 08 – 22

基金项目: 国家自然科学基金项目“不动产统一登记驱动下的地籍混合维度空间数据表达模型研究”(编号: 41471318)、南京邮电大学引进人才科研启动基金项目“服务设施空间配置优化算法研究”(编号: NY221143)和虚拟地理环境教育部重点实验室开放基金项目“移动数据驱动下服务设施空间配置量子优化算法研究”(编号: 2021VGE02)共同资助。

第一作者: 蒋卓然(1998 –),女,硕士研究生,主要从事深度学习影像解译等方面的研究。Email: 201302090@njnu. edu. cn。

通信作者: 吴长彬(1977 –),男,教授,主要从事时空数据模型等方面的研究。Email: wuchangbin@njnu. edu. cn。

工判读效率,提高整个监测流程的速度和精度。

针对上述问题,本研究提出了一种基于协同判读机制的蟹塘遥感智能检测方法,由影像分割网络 HRNet(high-resolution net)^[17]进行影像识别以提高检测速度,再将分类网络 Swin-Transformer^[18]作为图斑筛选工具,辅以人工核查,从整体上提高挖塘养蟹等耕地“非粮化”行为的监测效率。

1 研究方法

本研究以存在大面积蟹塘的江浙地区为研究区域,提出了一种基于协同判读机制的蟹塘遥感智能检测方法:首先基于研究区域内的蟹塘影像构建蟹塘样本,作为蟹塘影像分割网络 HRNet 和置信度评

估网络 Swin-Transformer 的训练数据;继而借助已训练完成的 HRNet 模型对整体研究区域中的蟹塘地块进行检测,并优化检测结果。为解决误检、漏检等产生的错误地块识别问题,创新性地提出了利用 Swin-Transformer 模型计算通过 HRNet 模型检测得到的蟹塘地块的置信度评估数值,并构建针对蟹塘地块进行判别的人机协同判读机制。人机协同判读机制是指先由机器进行自动判读分类,分出确切地块和模糊地块,再由人工对被分为模糊地块的部分进行核查的机制。基于蟹塘置信度评估数值与蟹塘筛选阈值,精确区分确切地块和模糊地块,降低人机交互时间成本,最后利用筛选出的蟹塘地块对初始检测结果进行优化,得到准确的蟹塘识别结果。方法的总体流程如图 1 所示。

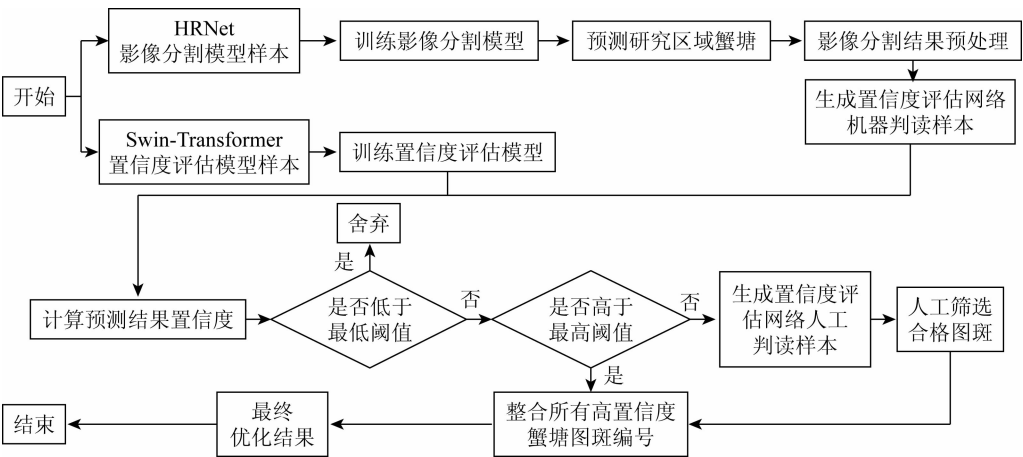


图 1 技术路线图

Fig. 1 Technology roadmap

1.1 影像分割网络 HRNet

为确保高分辨率蟹塘影像的分割精度与效率,选择语义分割网络 HRNet 作为影像分割网络。HRNet 是目前最为先进的语义分割网络之一,其精度在 COCO 数据集中的关键点检测、姿态估计和多人姿态估计这 3 项任务中都刷新了纪录,拥有着巨大的潜力。该网络能在整个任务过程中都保持高分辨率,通过不同分辨率的分支自信息融合以减少随着网路的加深而造成的信息损失。

在以往的影像分割网络中,对不同分辨率特征

层的连接都是使用串联,而 HRNet 将不同分辨率的特征层进行并联,并在此基础上对不同分辨率的特征层进行特征融合,以获取更全面的特征信息(图 2)。图 3 为 HRNet 的特征融合示意图,特征融合对于同分辨率的特征层会直接进行复制,对于需要升分辨率的特征层则使用双线性上采样和 1×1 卷积将所融合的层统一到同一通道,而对于需要降分辨率的特征图则使用跨步 3×3 卷积操作。在进行特征融合时,经过转换的不同特征层会以相加的形式进行融合。

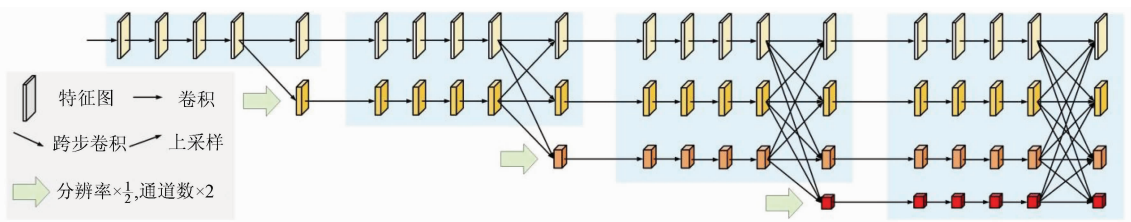


图 2 HRNet 网络结构^[17]

Fig. 2 HRNet network structure^[17]

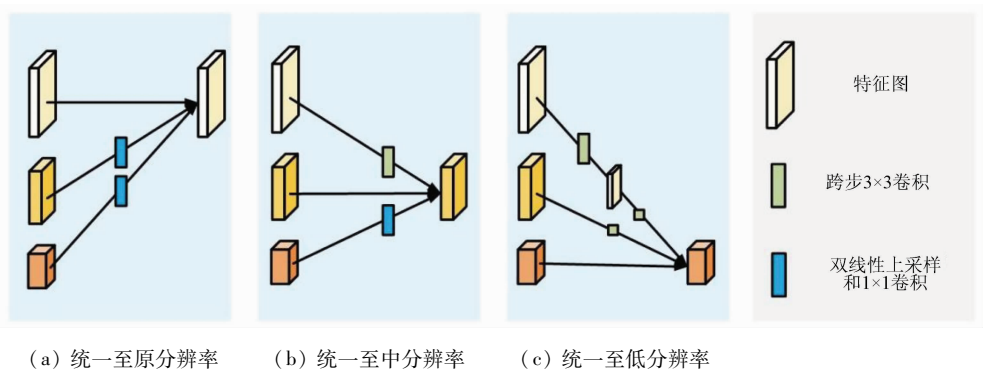


图3 特征融合示意图^[17]
Fig.3 Schematic diagram of feature fusion^[17]

本文使用 HRNet 网络所生成的影像分割模型对研究区域内的蟹塘进行检测,以得到一个初步的蟹塘识别结果。

1.2 检测结果预处理

由于影像分割模型精度不能达到百分之百,识别结果通常会存在少量误判、漏判甚至图斑形状异常等问题,此外,整幅影像在检测过程中经过了裁剪和拼接,这使得检测结果还会存在拼接缝隙和碎斑。因此,需要对影像分割的检测结果进行初步的优化,优化过程分为缝隙填充、碎斑处理和空洞填补。

缝隙填充和碎斑处理需使用形态学中的膨胀和

腐蚀操作^[19]。膨胀和腐蚀是使用卷积运算求局部最大值和最小值的操作,其中,膨胀可使栅格像元边界向外扩张,腐蚀则可使栅格像元边界向内收缩。通过先膨胀后腐蚀可以填补检测结果栅格图像中的缝隙,先腐蚀后膨胀则可以消除图像中的碎斑,这2个过程分别称为闭运算和开运算^[20-21]。

空洞填补是对转为矢量图层后存在空洞的图斑进行填补,找出图斑内部存在的岛并对其进行填充,即可获得形状完整的蟹塘地块。

1.3 置信度评估网络 Swin - Transformer

Swin - Transformer 网络结构如图 4 所示。

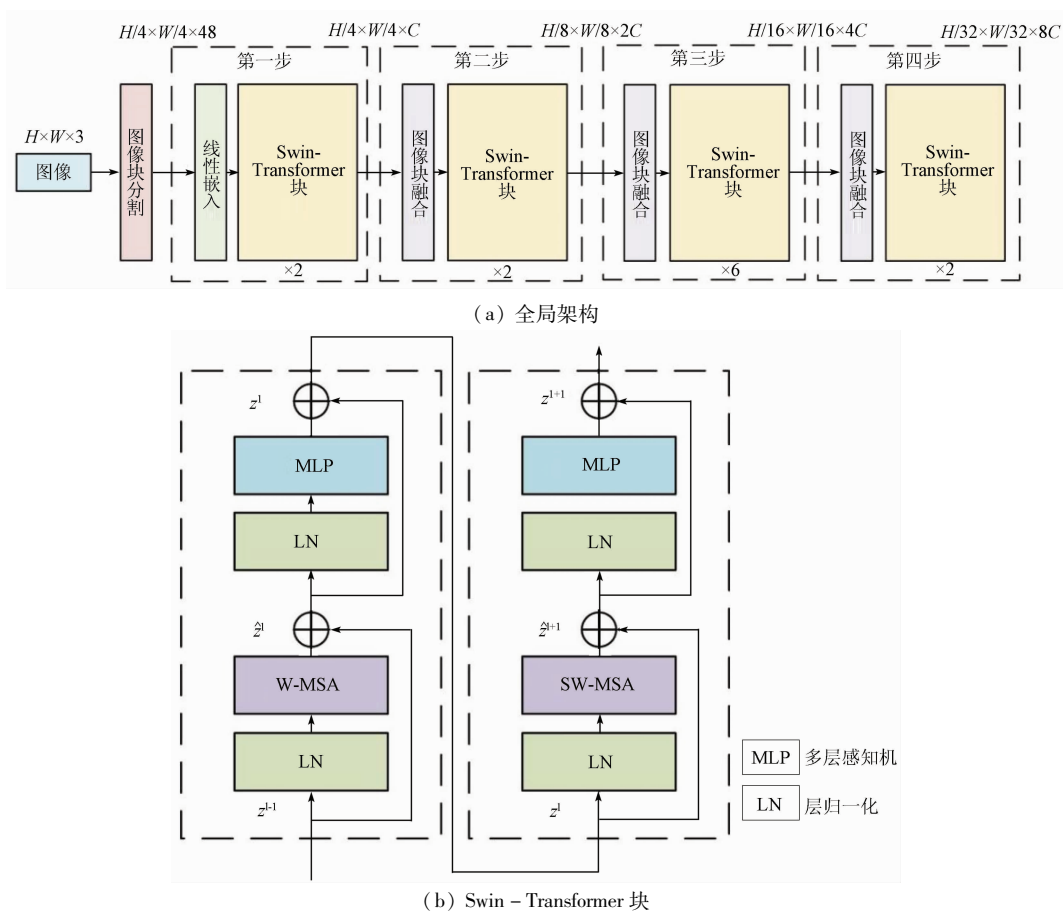


图4 Swin - Transformer 网络结构^[18]
Fig.4 Swin - Transformer network structure^[18]

由于影像光谱特征、分辨率和地貌谱系的多样性等因素影响,识别出的蟹塘地块往往精度不一,顾及蟹塘监测监管的时效性与精度要求,需提高深度学习模型检测结果的置信度标准,因此集成置信度评估网络 Swin - Transformer,检测每个蟹塘地块的置信度数值,并将置信度高的地块图斑作为地类明确的确切图斑,将置信度低的作为模糊图斑,以优化 HRNet 模型检测结果。

Swin - Transformer 衍生于 Transformer^[22] 计算机视觉任务网络,该网络使用 self - attention 机制替代原先卷积层,并引入 ResNet^[23] 跨层连接的机制,在扩大感受野的同时兼顾了精度的提升,是目前精度最高的分类网络之一。图 4 中 Swin - Transformer 的网络结构可划分为 2 个部分:①全局架构,包含图像块分割、线性嵌入、Swin - Transformer 块和图像块融合;②Swin - Transformer 块结构,包含 W - MSA 和 SW - MSA 2 个结构块。图中 H, W 和 C 分别为图像的高度、宽度和通道数; z^i 和 $\hat{z}^i$ 分别为第 i 个结构块的输出和中间输出结果。

在全局架构中,Swin - Transformer 网络首先用 Patch Partition 将输入的图像分切成没有重叠的图像块,并通过线性嵌入将分块后的图像映射到指定维度;继而基于图像块融合模块联结相邻的图像块,实现降采样的过程,最后通过 Swin - Transformer 块实现多头注意力机制,提取每个的特征。

在 Swin - Transformer 块中,W - MSA 模块主要将计算区域控制在每个当前窗口之内,以确保计算的低时间复杂度,提高网络的整体计算效率;SW - MSA 模块则负责在控制窗口数量的情况下,增加各个窗口间的信息交流,弥补了 W - MSA 模块因窗口不重合从而缺乏信息交流的缺陷。

利用 Swin - Transformer 网络良好的分类性能,对 HRNet 网络检测出的每个蟹塘地块结果输出一个置信度数值,置信度越高,对应地块为蟹塘的可能性越大。

1.4 模型精度评价指标

为了评价蟹塘检测模型的精度,采用准确率 A 、精确率 P 、召回率 R 与 F1 系数作为初步评价模型的检测精度指标。

准确率为被正确分割的样本数与所有样本数之比,其计算公式为:

$$A = \frac{TP + TN}{TP + TN + FP + FN}, \quad (1)$$

式中: TP 为检测值与真实值相同且检测值为正样本的数量; TN 为检测值与真实值相同且检测值为

负样本的数量; FP 为检测值与真实值不同且检测值为正样本的数量; FN 为检测值与真实值不同且检测值为负样本的数量。精度指标数值越大,表明被正确分割的像素数越多,模型精度越好,反之则表明分割效果较差。

精确率为检测为正样本的数量与其中真正为正样本的数量之比,反映了模型“找得对”的能力,其计算公式为:

$$P = TP / (TP + FP). \quad (2)$$

召回率为检测为正样本的数量与总体正样本的数量之比,反映了模型“找得全”的能力,其计算公式为:

$$R = TP / (TP + FN). \quad (3)$$

F1 系数为精度和召回率的调和平均数,是评价模型性能的重要指标,其计算公式为:

$$F1 = \frac{2TP}{2TP + FP + FN}. \quad (4)$$

1.5 构建人机协同判读体系

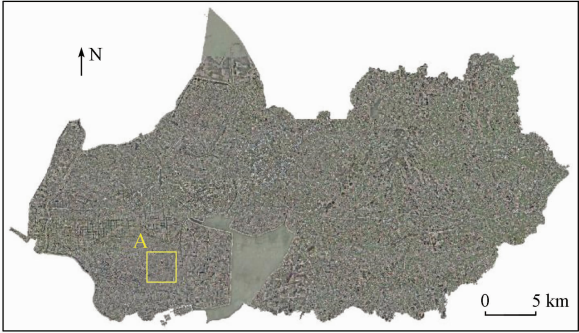
在获得所有蟹塘检测结果图斑的置信度 Cf 后,据此将检测结果分为 3 类:①当 $Cf > \alpha$ 时,图斑确定为蟹塘;②当 $Cf < \beta$ 时,图斑确定为非蟹塘;③当 $\beta < Cf < \alpha$ 时,图斑归为模糊地类。其中, α 和 β 分别为蟹塘筛选阈值的上限和下限,此处分别设为 0.8 和 0.5,被归为模糊地类的图斑需要进一步介入人工判读,根据蟹塘样本名称中带有图斑编号,对待核查的模糊地类图斑进行人工判读样本的生成,以便人工核查。待人工核查完成后,将人工核查和机器筛选合格的图斑编号进行汇总并对原始检测结果进行筛选,即可得到优化后的最终检测结果。

2 研究区及样本获取

2.1 研究区域概况

江浙地区具有水网密集、耕地破碎、蟹塘养殖面积大等特征。截至 2020 年,全国蟹塘养殖面积已超 4 600 km²,而江浙沪地区的占比达到 70%,已经成为国内养蟹较为集中的区域。作为江浙一带的养蟹大区,江苏省南京市高淳区拥有大面积的养殖蟹塘。高淳区位于 E118°41' ~ 119°21', N31°13' ~ 31°26' 之间,地处长江流域中下游,整体地势平坦,淡水资源丰富,辖区内拥有多个河流与湖泊,已形成较具规模化的螃蟹养殖产业。本研究采用了 2021 年江苏南京市高淳区的高分辨率遥感影像,空间分辨率为 0.3 m,如图 5 所示。可以看到,该区域具有大规模

的蟹塘,且蟹塘呈现沿河网聚集状分布,因此本研究区域具有典型性。



(a) 高淳区全域



(b) 区域 A 放大结果 (c) 区域 B 放大结果

图5 研究区域概况

Fig.5 Overview of the study area

2.2 样本制作

本研究涉及样本主要包括: ①HRNet 模型样本; ②Swin - Transformer 模型样本; ③人工判读样本。为确保模型的检测精度,在进行样本制作之前,需对遥感影像进行大气校正、几何纠正和图像融合等预处理,消除影像中的辐射失真和几何畸变。待预处理完成后,方可进行样本制作。

2.2.1 HRNet 模型样本

基于一景 2021 年江苏省南京市高淳区 0.3 m 空间分辨率的遥感影像进行样本制作。首先,沿影像中的蟹塘地块边界进行勾画,生成矢量蟹塘图斑

(图 6(a))。然后,对所有被矢量蟹塘图斑覆盖的地区进行影像切割,生成 256 像素 × 256 像素大小的影像切片(图 6(b)),并根据图斑覆盖位置,生成以黑色为背景、红色为地物的对应掩模(图 6(c))。最后,将配套的影像切片和掩模按 4: 1 的比例分为训练集和验证集,此时一套完整的 HRNet 训练样本制作完成。基于上述流程,共生成了 5 000 对 256 像素 × 256 像素大小的 HRNet 模型样本,其中,训练集和验证集分别为 4 000 对和 1 000 对。



(a) 矢量图斑 (b) 影像切片 (c) 掩模

图6 HRNet 样本示意图

Fig.6 Schematic diagram of HRNet sample

2.2.2 Swin - Transformer 模型样本

Swin - Transformer 模型样本分为训练样本和机器判读样本,其中,训练样本用于训练 Swin - Transformer 模型。首先,对与蟹塘纹理相近的地块进行了勾画,生成了矢量非蟹塘图斑(图 7(a))。然后,用这些矢量非蟹塘图斑与 HRNet 模型训练样本制作过程中生成的矢量蟹塘图斑分别对影像进行依边界裁剪的掩模提取,生成黑色为背景、实际影像为地物的无统一大小的影像切片(图 7(b) - (c))。最后,将蟹塘和非蟹塘的影像切片按 4: 1 的比例分为训练集和验证集,且蟹塘为正样本、非蟹塘为负样本,此时一套完整的 Swin - Transformer 模型训练样本制作完成。基于上述流程,本文共生成了 7 000 张 Swin - Transformer 模型训练样本,其中蟹塘样本为 5 000 张,非蟹塘样本为 2 000 张,训练集和验证集分别为 5 600 张和 1 400 张。



(a) 矢量图斑 (b) 蟹塘样本 (c) 非蟹塘样本

图7 Swin - Transformer 样本示意图

Fig.7 Schematic diagram of Swin - Transformer sample

而机器判读样本则是根据 HRNet 模型的检测结果进行制作的,用于对 HRNet 模型的检测结果图斑置信度进行评价。机器判读样本的制作流程与训

练样本相同,区别在于所使用的矢量图斑为 HRNet 模型的检测结果图斑。此外,机器判读样本的命名中带有对应的图斑编号,旨在与检测结果图斑和下

文所使用的人工判读样本进行关联。

2.2.3 人工判读样本

人工判读样本是针对 Swin - Transformer 检测结果中属于模糊地块的图斑设计的,旨在融合图斑所属地块附近的上下文背景信息,以方便人工判读。首先,根据 Swin - Transformer 检测结果中属于模糊地块的图斑编号,找到对应的矢量图斑;然后,根据这些对应的矢量图斑,分别计算其最小外接矩形,并对其进行边界系数扩张,得到扩大后的外接矩形框;最后,根据这些外接矩形框对影像进行裁剪,并根据图斑轮廓对影像进行地块边界绘制,生成对应的人工判读样本(图 8)。人工判读样本保留了更多地块周边上下文信息,有利于人工判读时进行广泛取证、迁移推理以辅助人工判别地块类型。



图 8 人工判读样本示意图

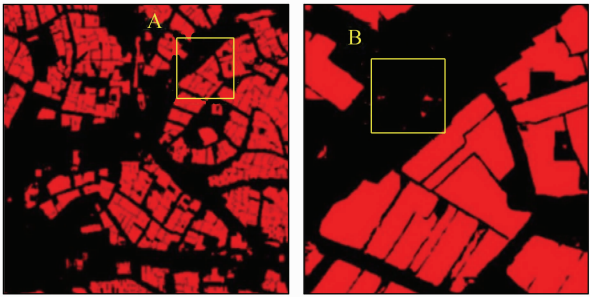
Fig. 8 Schematic diagram of manual interpretation sample

3 结果与分析

3.1 影像分割模型训练与结果

本研究利用深度学习语义分割网络 HRNet 对研究区域内的蟹塘进行模型训练解译。基于 Py-Torch 深度学习框架,搭建并运行 HRNet 深度学习网络。所有训练数据及分割代码均部署在一台 Linux 服务器上,该服务器的系统为 Ubuntu16.04.4,处理器为 Core i7 - 8700k,内存为 32 GB,显卡版本为 GTX2080Ti - 12 GB。模型训练过程中,设置优化器为 SGD,初始学习率为 1×10^{-3} ,迭代次数 epoch 为 200 次,其中学习率自适应于迭代次数,即随着迭代次数增加,学习率逐渐递减并最终归 0。基于构建的分割模型,最终获得 2021 年高淳区蟹塘检测结果,如图 9 所示。HRNet 模型可以准确分割出遥感影像中的蟹塘地块,模型准确率与 F1 系数较高,达到了 0.827 和 0.852,精确率和召回率分别为 0.837 和 0.867,分割效果较好。然而,由于蟹塘特征复杂以及曝光等成像因素影响,出现部分识别的蟹塘超出地块边界以及耕地/水域被错误识别成碎块蟹塘

的情况(图 9(b)区域 B)。这些误判会降低蟹塘检测精度,对后续“非粮化”指标计算带来负面影响。



(a) 局部区域检测结果 (b) 区域 A 放大结果

图 9 HRNet 模型蟹塘检测部分结果

Fig. 9 Partial results of crab pond prediction of HRNet model

为消除因分割产生的地块拼接缝、碎斑以及空洞等问题,在 HRNet 初步检测结果后介入拼接缝填补(闭运算)、碎斑处理(开运算)与空洞填补等优化措施。图 10(a)、(b)、(d)、(e)分别展示了经形态学开闭运算前后的检测掩模对比,此处所使用的卷积核大小为 3×3 。可以看到,经过形态学开闭运算操作后,大部分的拼接缝隙被有效填补,地块碎斑也得到了去除,蟹塘地块的拓扑结构得到有效维持与修复,且地块原本的形状并未发生改变。此外,拼接缝的填补也使得蟹塘掩模及图斑边界毛刺减少,蟹塘规则性得到保持。图 10(c)和(f)为填补矢量图斑空洞前后的结果,利用矢量图斑的拓扑关系将其中的岛全部进行填补,得到填补的矢量图斑能够实现对地块的全覆盖。

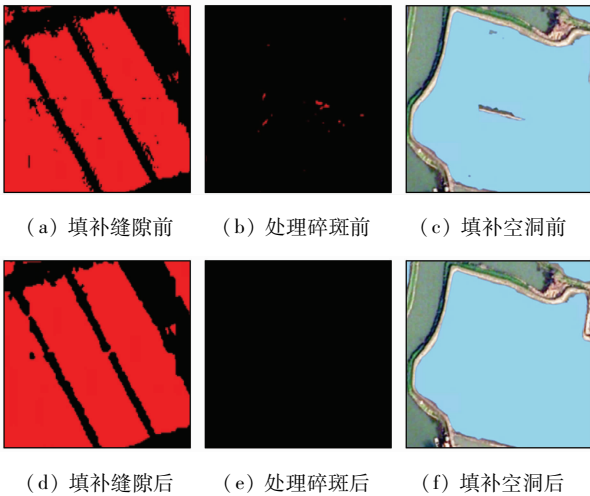


图 10 影像分割结果优化
Fig. 10 Optimization of image segmentation results

3.2 置信度评估模型训练与结果

考虑到蟹塘检测的海量数据与时效性,借助 Swin - Transformer 深度学习模型,通过训练测试,对

HRNet 模型检测得到的蟹塘图斑进行检测判读,筛选出确定为蟹塘/非蟹塘的地块。在训练的过程中,设置优化器为 Adam,初始学习率为 1×10^{-3} ,迭代次数 *epoch* 为 200 次。

表 1 列出了 Swin – Transformer 模型检测的蟹塘地块精度。其中 HRNet 模型共计检测得到蟹塘图斑数量为 18 471,依据蟹塘筛选阈值可自动化判别出的蟹塘、非蟹塘和模糊地块的数量分别为 11 191, 4 212 和 3 068,其中,在判别为蟹塘的图斑中实际为蟹塘的图斑数量为 10 464,判别精确率为 0.935; 在判别为非蟹塘的图斑中实际为非蟹塘的图斑数量为

3 862,判别精确率为 0.917; Swin – Transformer 模型准确判别出的蟹塘/非蟹塘数量为 14 326,模型准确率为 0.930; 此外,模型的精确率为 0.935,召回率为 0.967,F1 系数可达到 0.941。上述结果表明, Swin – Transformer 模型可以有效区分指定阈值下的蟹塘与非蟹塘地块,且模型整体精度较好,表现较优; 此外,此方法能够自动筛选掉 83.4% 确定图斑,仅剩余 16.6% 需介入人工判读,也从侧面表明借助 Swin – Transformer 网络的自动判别,可以大幅减少需人工判别的蟹塘/非蟹塘图斑数量,提升大尺度范围内的目视判别效率。

基于 Swin – Transformer 模型推理蟹塘的结果如表 2 所示。其中第 1 列为正确检测出的蟹塘地块,这些地块均为纹理清晰、具有典型特征的完整蟹塘地块; 第 2 列为正确检测出的非蟹塘地块,可以看出,对于耕地(第 2 列 1,2,4 行)与水工建筑(第 2 列 3 行)等与蟹塘具有明显区别的非蟹塘地块,均能够准确检测; 第 3 列为检测错误的蟹塘地块,错误情形主要发生在与蟹塘内陆地纹理相近的耕地地块(第 3 列 1,2 行)、水陆交融图斑(第 3 列 3,4 行)中,这些地块特征与蟹塘特征较相似,容易对蟹塘影

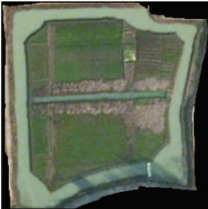
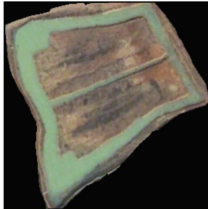
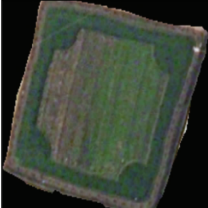
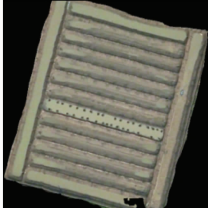
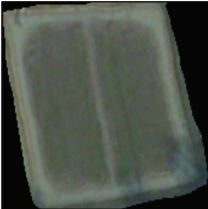
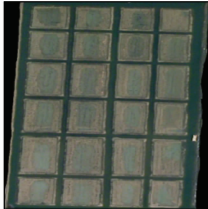
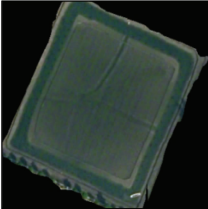
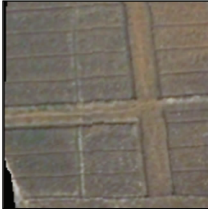
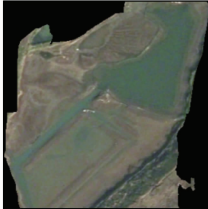
表 1 Swin – Transformer 模型检测蟹塘地块精度

Tab.1 Accuracy of Swin – Transformer model to predicts the crab pond

分类结果	真实类别		总计
	蟹塘	非蟹塘	
蟹塘	10 464	727	11 191
非蟹塘	350	3 862	4 212
总计	10 814	4 589	15 403
准确率	0.930		
精确率	0.935		
召回率	0.967		
F1 系数	0.941		

表 2 Swin – Transformer 模型推理结果

Tab.2 Reasoning results of Swin – Transformer model

序号	预测正确的蟹塘地块	预测正确的非蟹塘地块	预测错误的蟹塘地块	预测错误的非蟹塘地块
1				
				
				
				

像的智能判读产生负面影响；第 4 列为检测错误的非蟹塘地块，该类地块主要受到地块规则与色相，与常见的蟹塘地块特征存在些许差异，导致 Swin - Transformer 模型将之划分为非蟹塘。

针对基于蟹塘阈值筛选出的模糊图斑，需进一步采用人工判读，确定模糊图斑的具体归属类别。为此，设计了一套协同多任务判读系统，可辅助人工判读。借助该系统，人工判读的分类效率可达 1 000 ~ 1 200 张/h，同时能够有效改善传统判读过程中“找图斑、等加载、填属性”等效率较低的限制。经人工应用总结，在保障判读严肃性、严谨性的前提下，对模糊图斑而言，按照本实验检测出的 3 068 个模糊图斑，仅需 1 人在 3 h 内即可完成判读，判读效率大幅提升。经人工判读后，筛去 563 张非蟹塘图斑，保留 2 505 张蟹塘图斑。

综合上述实验结果，人机协同判读共筛去 4 775 张非蟹塘图斑，保留了 13 696 张蟹塘图斑，在以图斑为单位的情况下，检测准确率由 0.721 提升到了 0.972(表 3)。可见，本文提出的基于协同判读机制的蟹塘遥感智能检测方法，可有效提高监测效率，确

保蟹塘监测监管的有效性。

表 3 人机协同判读前后精度变化
Tab.3 Accuracy changes before and after collaborative interpretation mechanism

指标	人机协同判读前	人机协同判读后
蟹塘图斑数	18 471	13 696
非蟹塘图斑数	0	4 775
实际蟹塘图斑数	13 319	13 319
准确率	0.721	0.972

3.3 应用分析

基于上述理论与试验，对整个高淳区进行蟹塘检测及二次判别筛选，结果如图 11 所示。可以看出，经 Swin - Transformer 模型判别与人工判读后，蟹塘识别精度得到有效提升，且规则性与边界拓扑效果更好。经统计，研究区域内的蟹塘总面积约为 985.445 km²，通过套合三调数据，耕地总面积约为 960.687 km²，其中被蟹塘占用的耕地面积约为 18.192 km²，占比约 1.9%，表明高淳区未出现大规模挖塘养蟹等“非粮”行为，并且未对当地耕地造成破坏。

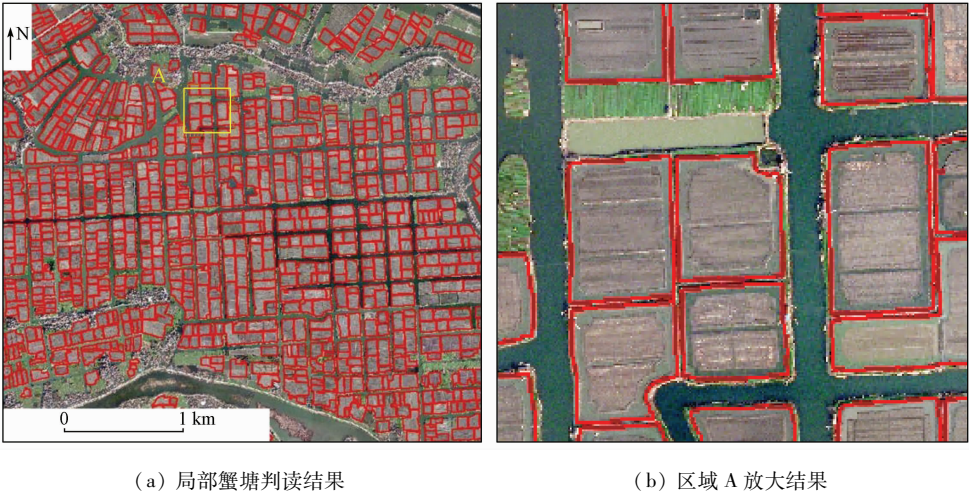


图 11 最终蟹塘判读结果

Fig. 11 Final interpretation results of crab pond

4 结论

为开展对蟹塘的遥感智能解译工作，并应对工作后期存在的人工判读量大、核查效率不足的挑战，实现遥感智能化监测的实时动态开展，本研究创新性地提出一种基于协同判读机制的蟹塘遥感智能检测方法，通过 HRNet 深度学习模型初步检测蟹塘地块，初步检测结束后借助 Swin - Transformer 模型得到检测蟹塘地块的置信度数值，最终通过人机协同判读机制完成对蟹塘地块的二次筛查。

通过对 1 景高淳区遥感影像进行语义分割与图像分类，蟹塘识别的总体精度可以达到 0.972，研究试验表明，通过 Swin - Transformer 模型的分类筛查与人工核查，能够快速、有效和精准地识别大范围区域内蟹塘等典型“非粮化”地物，自动判别比例高达 83.4%，能够大幅减少人工判别的样本数量，提高整体的目视判别效率。

本研究提出的基于协同判读机制的蟹塘遥感智能检测方法能够有效提高蟹塘的识别效率，降低人工核检的工作量，为提高此类耕地“非粮化”行为的检测速度和精度提供了一种新的思路和方法。然

而,本研究也存在一些不足之处,譬如仅对发生误判的图斑进行了去除,而对于漏检的图斑则依然需要进行人工勾画填补,因此下一步拟探寻新的修正机制,进一步提升整个业务流程的效率和可行性。

参考文献(References):

[1] 杨瑞珍,陈印军,易小燕,等. 耕地流转中过度“非粮化”倾向产生的原因与对策[J]. 中国农业资源与区划,2012,33(3):14-17.
Yang R Z, Chen Y J, Yi X Y, et al. The causes and countermeasures on the excessive “non – food” tendency in the process of land transfer[J]. Chinese Journal of Agricultural Resources and Regional Planning, 2012, 33(3): 14 – 17.

[2] 王 勇,陈印军,易小燕,等. 耕地流转中的“非粮化”问题与对策建议[J]. 中国农业资源与区划,2011,32(4):13-16.
Wang Y, Chen Y J, Yi X Y, et al. The non – grain problem in the process of land transfer and the countermeasures[J]. Chinese Journal of Agricultural Resources and Regional Planning, 2011, 32(4): 13 – 16.

[3] 匡远配,刘 洋. 农地流转过程中的“非农化”、“非粮化”辨析[J]. 农村经济,2018(4):1-6.
Kuang Y P, Liu Y. Distinguishing “non – agriculturalization” and “non – grainification” in the process of rural land transfer[J]. Rural Economy, 2018(4): 1 – 6.

[4] 李小江,孟庆岩,王春梅,等. 一种面向对象的像元级遥感图像分类方法[J]. 地球信息科学学报,2013,15(5):744-751.
Li X J, Meng Q Y, Wang C M, et al. A hybrid model of object – oriented and pixel based classification of remotely sensed data[J]. Journal of Geo – Information Science, 2013, 15(5): 744 – 751.

[5] 张继贤,顾海燕,杨 懿,等. 高分辨率遥感影像智能解译研究进展与趋势[J]. 遥感学报,2021,25(11):2198-2210.
Zhang J X, Gu H Y, Yang Y, et al. Research progress and trend of high – resolution remote sensing imagery intelligent interpretation [J]. National Remote Sensing Bulletin, 2021, 25(11): 2198 – 2210.

[6] Blaschke T. Object based image analysis for remote sensing[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2010, 65(1):2-16

[7] Ma L, Li M, Ma X X, et al. A review of supervised object – based land – cover image classification[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2017, 130:277-293.

[8] 张 裕,杨海涛,袁春慧. 遥感图像分类方法综述[J]. 兵器装备工程学报,2018,39(8):108-112.
Zhang Y, Yang H T, Yuan C H. A survey of remote sensing image classification methods[J]. Journal of Ordnance Equipment Engineering, 2018, 39(8): 108 – 112.

[9] 魏飞鸣,李小文,顾行发,等. 基于形状参数的遥感图像“同谱异物”目标区分[C]//第十四届全国图象图形学学术会议论文集,2008:477-481.
Wei F M, Li X W, Gu X F, et al. Shape – based classification of “spectral similar” objects in remote sensing image processing [C]// Proceedings of the 14th National Conference on Image Graphics, 2008: 477 – 481.

[10] 秦其明. 遥感图像自动解译面临的问题与解决的途径[J]. 测绘科学,2000,25(2):21-24.
Qing Q M. The problem and approach in the auto – interpretation of remote sensing imagery [J]. Science of Surveying and Mapping, 2000, 25(2): 21 – 24.

[11] 李国清,柏永青,杨 轩,等. 基于深度学习的高分辨率遥感影像土地覆盖自动分类方法[J]. 地球信息科学学报,2021,23(9):1690-1704.
Li G Q, Bo Y Q, Yang X, et al. Automatic deep learning land cover classification methods of high – resolution remotely sensed images [J]. Journal of Geo – Information Science, 2021, 23(9): 1690 – 1704.

[12] 王立国,王丽凤. 结合高光谱像素级信息和 CNN 的玉米种子品种识别模型[J]. 遥感学报,2021,25(11):2234-2244.
Wang L G, Wang L F. Variety identification model for maize seeds using hyperspectral pixel – level information combined with convolutional neural network [J]. National Remote Sensing Bulletin, 2021, 25(11): 2234 – 2244.

[13] Zhang L P, Zhang L F, Du B. Deep learning for remote sensing data: A technical tutorial on the state of the art[J]. IEEE Geoscience and Remote Sensing Magazine, 2016, 4(2): 22-40.

[14] 王华俊,葛小三. 一种轻量级的 DeepLabv3+ 遥感影像建筑物提取方法[J]. 自然资源遥感,2022,34(2):128-135. doi:10.6046/zrzyyg.2021219.
Wang H J, Ge X S. Lightweight DeepLabv3+ building extraction method from remote sensing images[J]. Remote Sensing for Natural Resources, 2022, 34(2): 128 – 135. doi: 10. 6046/zrzyyg. 2021219.

[15] 刘万军,高健康,曲海成,等. 多尺度特征增强的遥感图像舰船目标检测[J]. 自然资源遥感,2021,33(3):97-106. doi:10.6046/zrzyyg.2020372.
Liu W J, Gao J K, Qu H C, et al. Ship detection based on multi – scale feature enhancement of remote sensing images[J]. Remote Sensing for Natural Resources, 2021, 33(3): 97 – 106. doi: 10. 6046/zrzyyg. 2020372.

[16] 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. doi:10.6046/zrzyyg.2021020.
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. doi: 10. 6046/zrzyyg. 2021020.

[17] Sun K, Xiao B, Liu D, et al. Deep high – resolution representation learning for human pose estimation [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019:5693-5703.

[18] Liu Z, Lin Y, Cao Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]// Proceedings of the IEEE/ CVF International Conference on Computer Vision(CVPR), 2021: 10012-10022.

[19] 戴青云,余英林. 数学形态学在图象处理中的应用进展[J]. 控制理论与应用,2001,18(4):5.
Dai Q Y, Yu Y L. The advances of mathematical morphology in image processing [J]. Control Theory and Applications, 2001, 18

(4):5.

[20] 刁智华,赵春江,吴 刚,等. 数学形态学在作物病害图像处理中的应用研究[J]. 中国图象图形学报,2010,15(2):194 – 199.

Diao Z H,Zhao C J,Wu G ,et al. Application research of mathematical morphology in image processing of crop disease[J]. Journal of Image and Graphics,2010,15(2):194 – 199.

[21] Yang K,Zeng L B,Wang D C. A fast arithmetic for the erosion and dilation operations of mathematical morphology[J]. Computer Engineering and Applications,2005,41(34):54 – 56.

[22] Vaswani A,Shazeer N,Parmer N,et al. Attention is all you need [C]// Proceedings of the 31st International Conference on Neural Information Processing Systems,2017:6000 – 6010.

[23] He K,Zhang X,Ren S,et al. Deep residual learning for image recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2016.

Intelligent detection of crab ponds using remote sensing images based on a cooperative interpretation mechanism

JIANG Zhuoran¹, ZHOU Xinxin^{2,3,4}, CAO Wei⁵, WANG Yahua^{1,3,4}, WU Changbin^{1,3,4}

(1. School of Geography, Nanjing Normal University, Nanjing 210023, China; 2. School of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications, Nanjing 210023, China; 3. Key Lab of Virtual Geographic Environment, Ministry of Education, Nanjing Normal University, Nanjing 210023, China; 4. Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China; 5. Nanjing Guotu Information Industry Co., Ltd., Nanjing 210000, China)

Abstract: Digging ponds to raise crabs is a non – grain behavior of cultivated land, endangering national food security. However, the intelligent interpretation of remote sensing images targeting this behavior faces challenges such as laborious manual interpretation and low verification efficiency. Based on a cooperative interpretation mechanism, this study proposed an intelligent method for detecting crab ponds using remote sensing images. This method, integrating the HRNet segmentation network and the Swin – Transformer classification network models and combining manual verification, improved the detection accuracy and work efficiency. The application results of this method to Gaochun District, Nanjing City, Jiangsu Province show that the method for intelligent detection can automatically determine 83.4% of the spots for detection, with final identification accuracy of 0.972. The method proposed in this study can significantly reduce the identification difficulty and manual verification workload while improving the detection accuracy. Therefore, this study will provide a reliable solution for the accurate and efficient detection of non – grain surface features such as crab ponds.

Keywords: cooperative interpretation mechanism; HRNet; Swin – Transformer; crab pond detection; non – grain (责任编辑: 陈 理)