



武汉大学学报(信息科学版)

Geomatics and Information Science of Wuhan University

ISSN 1671-8860, CN 42-1676/TN

《武汉大学学报(信息科学版)》网络首发论文

题目: 深度时空卷积网络支持的地块尺度作物种植类型分类
作者: 周亚男, 何金珂, 冯莉, 陈跃红, 吴田军, 张新, 骆剑承
DOI: 10.13203/j.whugis20230136
收稿日期: 2023-04-15
网络首发日期: 2023-05-09
引用格式: 周亚男, 何金珂, 冯莉, 陈跃红, 吴田军, 张新, 骆剑承. 深度时空卷积网络支持的地块尺度作物种植类型分类[J/OL]. 武汉大学学报(信息科学版).
<https://doi.org/10.13203/j.whugis20230136>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

DOI:10.13203/j.whugis20230136

引用格式：周亚男, 何金珂, 冯莉, 等. 深度时空卷积网络支持的地块尺度作物种植类型分类[J]. 武汉大学学报(信息科学版), 2023, DOI:10.13203/J.whugis20230136(ZHOU Ya'nan, HE Jinke, FENG Li, et al. Parcel-Scale Crop Type Classification Using Tile-Slice-Based Spatial-Temporal Convolutional Networks[J]. Geomatics and Information Science of Wuhan University, 2023, DOI:10.13203/J.whugis20230136)

深度时空卷积网络支持的地块尺度作物种植类型分类

周亚男¹ 何金珂¹ 冯莉¹ 陈跃红¹ 吴田军² 张新^{3,4} 骆剑承^{3,4}

1 河海大学水文水资源学院, 江苏 南京, 211100

2 长安大学理学院, 陕西 西安, 710064

3 中国科学院空天信息创新研究院, 北京, 100101

4 中国科学院大学, 北京, 101408

摘要：面向遥感影像多层次时空信息表达与作物种植类型识别的需求, 提出了一种分区分层的时空遥感作物种植类型分类方法(TSST), 该方法的核心在于地块空间的分区分层表达和深度时空分类网络。首先, 对 Sentinel-2 遥感影像地块空间进行分区和分层并构建作物类型识别的时空训练数据集和预测数据集; 然后, 构建深度时空卷积分类模型并估算预测数据集的作物类型概率; 最后, 以地块空间为约束融合地块的作物类型概率并生成最终的作物类型专题图。在法国研究区的对比与评价实验表明, 所提出方法与现有方法相比在总体精度、准确度和 F1 分数上分别取得了 3%、2%和 2%的性能提升, 验证了其在作物种植类型分类制图上的有效性, 为地块尺度遥感时间序列分析提供了一种新的思路。

关键词：Sentinel-2 遥感影像; 地块尺度; 作物类型; 时空卷积网络; 分区分层

中图分类号: P237

文献标志码: A

Parcel-Scale Crop Type Classification Using Tile-Slice-Based Spatial-Temporal Convolutional Networks

ZHOU Ya'nan¹ HE Jinke¹ FENG Li¹ CHEN Yuehong¹ WU Tianjun² ZHANG Xin^{3,4} LUO Jiancheng^{3,4}

1 Collage of Hydrology and Water Resources, Hohai University, Nanjing 211100, China

2 School of Sciences, Chang'an University, Xi'an 710064, China

3 Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China

4 University of Chinese Academy of Sciences, Beijing 101408, China

Abstract: Objectives: Parcel-based crop classification using multi-temporal satellite images plays a vital role in precision agriculture. However, exploring the multi-scale spatial information for identification of crop types from remote sensing images is a significant challenge. **Methods:** In this study, a tile-slice-based spatial-temporal method was developed for parcel-scale crop (type) classification using multi-temporal Sentinel-2 images. Central to this approach is the combined use of tile-slice-based representation of parcels in images and a deep spatial-temporal convolutional network. Firstly, spatial tiles were sliced according to crop types and parcels, to produce a tile-slice-based training sample dataset and a tile-slice-based prediction dataset. Then, the deep spatial-temporal convolutional network was established to estimate crop-type probabilities for tile-slice-based prediction dataset. Finally, tile-slice-based probabilities were fused under parcel polygons, to generate the final crop type maps. **Results:** The proposed

收稿日期: 2023-04-15

基金项目: 第三次新疆综合科学考察(2021xjkk1305); 国家重点研发计划(2021YFB3901301); 国家自然科学基金(42071316)。

第一作者: 周亚男, 副教授, 研究方向为高分辨率遥感与时序遥感。zhouyn@hhu.edu.cn

通讯作者: 张新, 博士, 研究员。zhangxin@radi.ac.cn

method is further discussed and validated through parcel-based time-series crop classifications in the France study area with multi-temporal Sentinel-2 images. The classification results demonstrated great improvements in accuracy scores (0.03, 0.02, 0.02 in overall accuracy, precision and F1, respectively) over comparison methods. **Conclusions:** Through experiments and discussions, we concluded that: (1) the proposed tile-slice-based spatial-temporal method is effective for parcel-based crop classification using Sentinel-2 image sequences, (2) the spatial division and slice in parcel-scale representation are benefit more for crop type identification of larger and spatially concentrated parcels, (3) the proposed TSST-Net is lightweight but powerful for crop classification. The above experiments and their conclusions will provide new ideas for parcel-based agricultural remote sensing time series analysis.

Key Words: Sentinel-2 image; parcel-scale; crop type mapping; spatial-temporal convolutional networks; spatial division and slicing

遥感具有覆盖范围广、重复观测周期短、成像受限少和成本低等优势，被成功应用于区域乃至全球的资源调查、作物长势监测、粮食安全预警等农业生产与管理领域^[1-4]。作物种植类型的遥感监测与制图是开展精准农业应用（诸如作物长势与粮食估产）的基础和关键。然而受限于作物生长的相似性和单次遥感观测的局限性，很难用单时相卫星影像实现对多种类型作物（特别是同季作物）的精确识别。因此，探索和表达多时相遥感影像中蕴含的时间序列信息成为遥感作物种植类型识别与制图的有效方法^[5-7]。围绕多时相遥感时序信息的提取与分析，研究人员提出了诸如统计学模型、（传统）机器学习模型、深度学习模型等方法^[8-10]。这些方法从遥感观测的时间维度出发，重点研究其序列修复、时序成分分解、时序平滑、时序分类等时间序列分析技术^[11]，提升遥感作物类型制图的性能和应用领域。

近年来，随着对地观测空间分辨率的提升，卫星影像越来越能够精确表达种植地块的几何边界与地块内纹理等高层次信息。因此以种植地块为分析对象、开展地块尺度的遥感时间与空间分析^[12, 13]成为当前遥感作物精准制图的主流方法（较传统的基于像元的作物类型分类），挖掘作物种植地块内的空间信息成为提升遥感作物种植类型制图的关键^[6]。相对于基于像元（光谱特征）的作物类型分类方法，面向对象的分析技术能够有效提取影像多尺度分割对象的形状、纹理、上下文关系等高级语义特征，扩展作物监测的特征维度，提升遥感作物种植类型制图的精度^[14, 15]。但此类方法往往受制于前期影像多尺度分割的效果，而难以获取精准的地块对象，增大了地块特征表达的不确定性，限制了作物类型制图精度的提升^[16, 17]。

深度学习的迅速发展为农业遥感作物分类提供了新的思路^[18-20]。考虑到遥感时序观测与视频之间的相似性，很容易将现有语义分割、视频分析等深度学习方法移植和扩展到遥感作物类型识别应用上^[21, 22]。尽管此类方法取得了较好的分类效果，但其大都建立在理想的标记样本基础之上，即作物类型标注对样本空间的全覆盖。这一假设显然不符合农业遥感野外采样的实际^[5]。并且地块的不规则和深度卷积也将会造成作物特征信息在地块边缘的扩散和混淆，降低了地块尺度（特别是面积较小的地块）作物类型识别的效果^[16]。为此，研究人员借鉴面向对象影像分析思路，将地块尺度作物类型识别任务分解为地块提取与类型识别两大步骤^[3, 23, 24]。在类型识别中，尝试通过构建地块内简单的光谱统计特征、或者利用预训练卷积网络提取影像的多层次语义特征^[13]等途径，构建地块的多维时序特征，进而利用深度循环网络挖掘时序特征中长短时间依赖信息，实现作物类型识别^[25]。尽管这类研究取得了较传统方法更好的制图效果，但在影像空间特征分析上，尚未脱离人工设计空间高层次特征的处理模式、也未能构建端到端的多层次空间特征表达模型^[13, 26]。

总体来讲，当前深度学习方法在遥感作物分类应用中仍存在以下不足：（1）缺乏对地块内部高层次空间信息的挖掘，仅计算了地块内光谱和纹理等人工设计特征，（2）缺乏面向遥感时空作物类型识别任务的端到端深度网络模型。针对上述问题，本文提出一种分区分层的遥感时空分类网络模型（tile-slice-based spatial-temporal network, TSST），其核心在于对地块空间的分区分层表

达和深度时空分类网络。相比于已有方法，本文的贡献主要表现为两个方面：（1）提出了一种适用于地块尺度遥感影像时空分析的分区分层框架，（2）构建了一个轻量且高效的作物类型识别深度网络模型。进而，利用 Sentinel-2 时序遥感影像、通过法国研究区的地块尺度作物种植类型分类实验，分析和讨论本文方法的有效性。

1 研究区与数据

1.1 研究区

本文研究区（图 1）位于法国东北部勃艮第-弗兰奇-孔泰地区（BFC）的科特迪瓦省（首府为 Dijon），中心经纬度为 $5^{\circ}1'E$ 和 $47^{\circ}17'N$ ，覆盖面积约 $5\ 000\ km^2$ 。研究区属于受大陆影响的温带海洋性气候、平均气温在 $6.8\sim 16.1^{\circ}C$ 之间、年平均降水量为 $740\ mm$ ，非常适合于小麦、油菜、葡萄和草等植被生长。

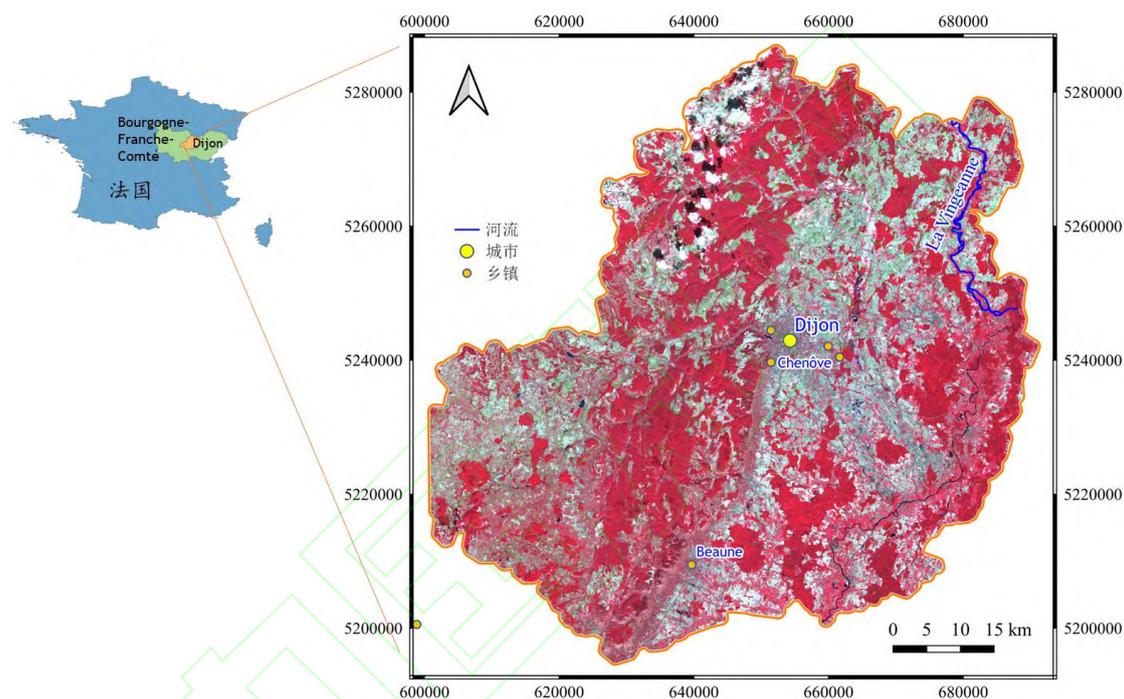


图 1 法国研究区的整体概况（2020 年 8 月 31 日获取 Sentinel-2 影像的近红外-蓝-绿波段合成）

Fig.1 The study area is in the northeast of France. The right-hand section provides an overview of the near-infrared-blue-green Sentinel-2 image acquired in August. 31, 2020

1.2 实验数据

面向地块尺度的作物种植类型识别任务，实验收集并构建了研究区作物种植地块矢量和 Sentinel-2 时间序列影像数据集。选用法国国家森林与地理信息研究所（IGN）发布的 2019 年全国作物种植的地理信息（RPG，包含种植地块矢量及种植类型），并提取研究区的作物种植地块多边形和类型数据。原始 RPG 数据包含 23 组、共 328 个作物类型。结合研究区的实际情况，合并类别并选用冬小麦（winter wheat, WWT）、冬大麦（winter barley, WBR）、冬油菜（winter rapeseed, WRP）、冬小黑麦（winter triticale, WTT）、春大麦（spring barley, SBR）、玉米（corn, CON）、大豆（soy, SOY）、向日葵（sunflower, SFL）、葡萄（grapes, GRA）、苜蓿（alfalfa, AFF）、草（grass, GRS）、和休耕地（fallow, FLW）等 12 个作物种植类别。其中，保留了两个极少数类别（向日葵和冬小黑麦，占比分别为 1.21% 和 2.10%），以测试作物分类模型的性能（也即实际作物类型识别中的类别不平衡问题）^[27]。最终获得研究区共约 53 400 个地块，并随机选择 20% 的地块作为标记样本。

考虑到研究区秋播作物的生长阶段也主要集中于来年春季，实验选用拍摄于 2019 年 3 月 1

日至 2019 年 9 月 30 日之间的 48 幅 Sentinel-2 影像（其 Path/Row 编码为 T32TFN）来监测作物的生长过程。受成像天气条件的影响，38 幅 Sentinel-2 影像受到云/影遮挡，平均云/影覆盖率为 53%；而 03 月 29 日、05 月 13 日、05 月 23 日、06 月 02 日、06 月 17 日、06 月 27 日、08 月 21 日、08 月 26 日、09 月 15 日、09 月 20 日的影像为无云影像。

2 研究方法

本文提出了一种基于 Sentinel-2 时间序列影像的地块尺度作物种植类型分类与制图方法（技术流程如图 2 所示），整体包括：（1）数据预处理、（2）空间分区与分层、（3）深度时空卷积分类模型和（4）地块尺度的类型合并等四大步骤。

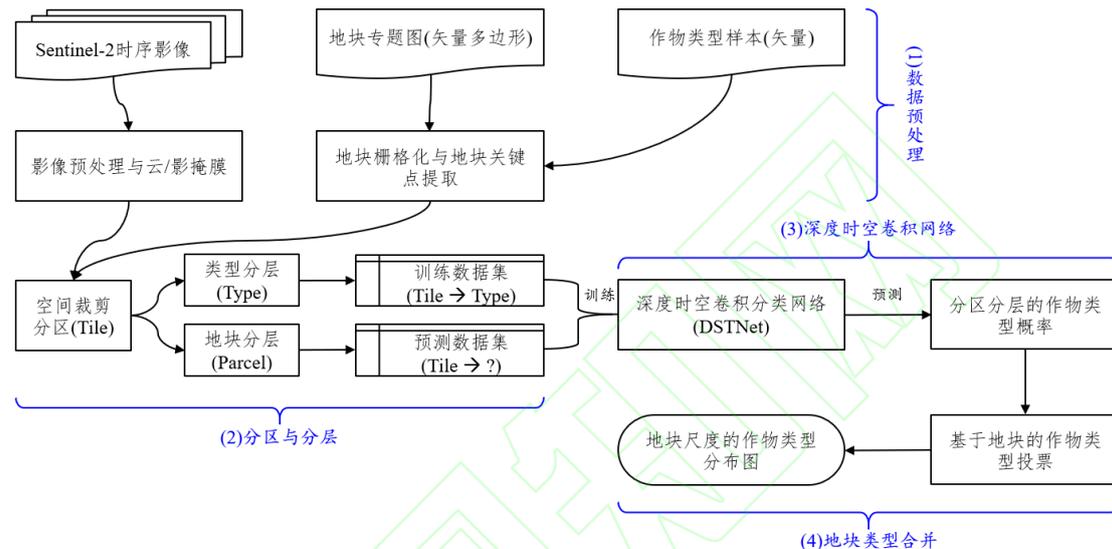


图 2 遥感时空融合的地块尺度作物种植类型分类流程

Fig.2 Flowchart of parcel-based crop mapping using spatial-temporal Sentinel-2 images

首先，数据预处理主要包括 Sentinel-2 影像大气校正、云/影检测与掩膜，和 Sentinel-2 影像与 RPG 地块数据之间的地理配准。然后，以地块关键点为中心进行地块空间（瓦片）分区、据瓦片作物类型或地块标识进行分层，建立训练数据集和预测数据集。接着，构建深度时空卷积分类模型、估算预测数据集的作物类型概率。最后，以地块为空间约束合并分区分层的作物类型概率，生成地块尺度的作物类型分类图。

2.1 数据预处理

1) 时间序列 Sentinel-2 影像

首先，利用 Sen2Cor 算法对下载的 L1C 级 Sentinel-2 影像进行大气校正，生成 L2A 级地表反射率产品，以保证不同时相、不同区域获取的 Sentinel-2 影像具有一致的反射率，增强大区域、多时相作物生长监测的可靠性^[28]。然后，提取 Sentinel-2 影像的 10 m 空间分辨率波段（波段 2、3、4 和 8）和 20 m 空间分辨率波段（波段 5、6、7、8A、11 和 12），并利用最近邻插值算法将 20 m 波段重采样为 10 m。最终，生成空间分辨率为 10 m、含有 10 个波段的 Sentinel-2 影像产品。

Sen2Cor 算法生成的影像质量场景分类栅格(quality scene classification, QSC)标记了 Sentinel-2 影像中云/影遮挡、数据缺失和过饱和等的像元区域，可用来构建 Sentinel-2 影像的掩膜。首先，目视判读 QSC 栅格并重点修正云/影边缘处的 QSC 错误分类。然后，重分类 QSC 栅格为二值掩膜，其中一类为“被污染”(contaminated)像元(含云/影遮挡、数据缺失和过饱和的像元)，另一类为“干净”(clean)像元。最后，将二值掩膜空间叠加到 Sentinel-2 影像、并标记“被污染”像元的 Sentinel-2 影像值为缺失值(也即 No Data)。需要注意的是,为构建时相规则的遥感时间序

列，完全“被污染”的影像也将保留在影像序列中。

2) 地块与作物类型样本

按照 6:2:2 的比例将 RPG 标记样本随机划分为训练样本集、验证样本集和测试样本集。其中，训练样本集和验证样本集分别用于调整模型的参数和超参数，测试样本集用于评估分类模型泛化性能。

2.2 地块关键点与空间分区

空间分区是利用深度学习进行影像时空分析的首要步骤。通过空间分区将大区域影像（实验中 Sentinel-2 影像的空间尺度约为 $9\,000 \times 9\,000$ 像素）划分为多个影像瓦片，以缓解深度学习模型对显存等计算资源的需求；同时空间分区生成的影像瓦片也是深度学习模型的处理对象。图 3 展示了基于地块关键点的空间分区流程。首先，以 Sentinel-2 影像为参考、将地块多边形转换为栅格对象；然后，提取栅格对象的骨架线以及骨架线交点（骨架点）、利用距离变化计算栅格对象的中心点，共同组成地块的关键点集；最后，以地块关键点为中心、裁剪 $N \times N$ 像元的影像区域为该关键点对应的影像瓦片。一个地块多边形可能生成多个关键点，因此其也将对应多个影像瓦片。

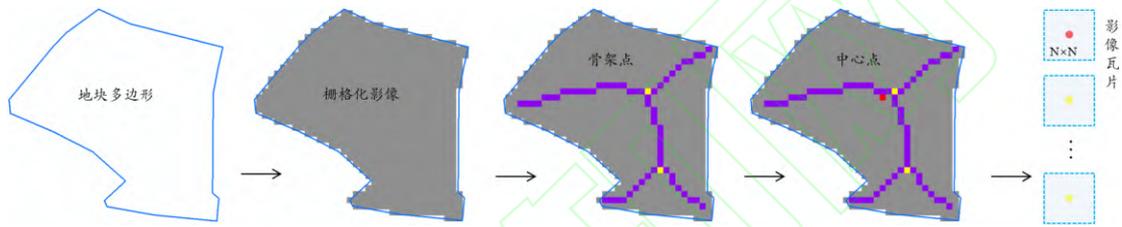


图 3 地块关键点提取与影像空间分区示意图

Fig.3 Key points of parcels and spatial division of images

影像瓦片尺寸 N 是一个关键参数；过大的 N 将导致影像瓦片区域内含有过多的类别或地块，后续分层处理将产生更高的数据缺失率；而过小的 N 将限制深度卷积的感受野，难以捕捉到地块影像的多层次信息。为此，在综合考虑研究区地块尺度、影像空间分辨率以及数据缺失率等因素的基础上，设置 N 为 32。

2.3 类别/地块分层

类别/地块分层是利用深度学习进行遥感作物时间序列分析的关键。在影像空间分区中，本文提出如图 4 所示的类被/地块分层流程。在基于作物类别的分层中，分别提取影像瓦片区域内每种类型像元的区域并标记其他像元区域为空白（影像像元值置零）；在地块分层中，则依据地块编码（比如地块序号）提取编码相同的像元区域、并将其他区域置零。在构建作物类型识别训练样本中，由于地块类别已知，因此将对瓦片区域进行类别分层，每层生成一个标记样本。在制作预测样本中，则对瓦片区域进行地块分层，每个地块分层对应一个预测样本。通过类别/地块分层，能够有效排除不同类别或地块之间的相互干扰，获取“纯净”的样本特征数据。

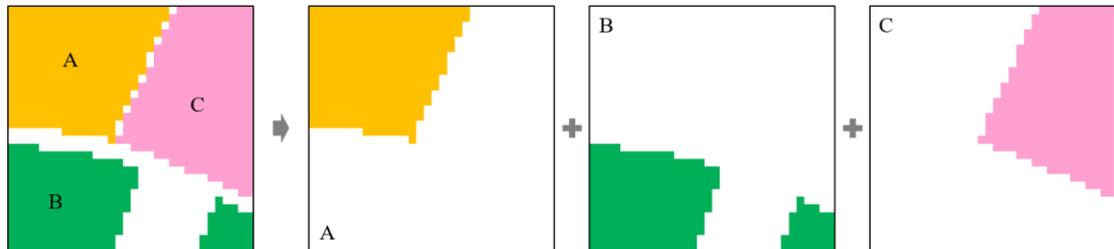


图 4 基于作物类别/地块的分层示意图。

Fig.4 Parcel slice based on crop types or parcel coding

实验过程中也注意到：瓦片分层中的空白区域（数据缺失）可能会导致模型训练过程不稳定、

模型训练不收敛甚至训练失败等问题。为此，尝试采用随机填充（random selection）和正态填充（normal selection）两种方法对空白区域进行模拟填充，以期改善模型训练过程。在随机填充中，从有效像元中随机选取像元、并将其影像像元值填充到缺失位置；在正态填充中，假定有效像元值服从正态分布、并计算其均值与标准差，进而在均值的 3 倍标准差范围内随机生成像元值、并填充到缺失位置。并且，考虑到深度学习模型对数据缺失具有一定的容忍性，实验中只有当影像瓦片序列的平均缺失率大于 0.6 时，才会模拟填充空白区域降低其数据缺失率到低于 0.6。

2.4 TSST-Net 网络模型

三维卷积神经网络能够学习影像序列的时空联合信息。本文为避免直接使用三维卷积所带来的巨大计算资源消耗，受 Pseudo-3D (P3D)^[29]、ResNet3D^[30]等三维模型的启发，将常规的三维卷积分解为二维空间卷积和一维时间卷积，提出了如图 5 所示的 TSST-Net 网络模型。整体上，网络包括空间嵌入模块（spatial embedding block, SEB）、稠密空间连接模块（dense spatial block, DSB）、注意力模块（channel-spatial attention, CSA）、向量化模块（vector block, VB）、和基于 TCN（temporal convolutional network）的分类模块等 5 个子模块。SEB 模块对输入时序影像进行空间包裹生成输入影像的高维特征表达；DSB 模块用于提取影像的多层次空间信息；CSA 模块学习影像在空间和波段（属性）维度的注意力；VB 模块将序列影像特征转化为多维时序信息；TCN 模块实现从多维时序特征到作物类别的映射，并输出作物类别概率。

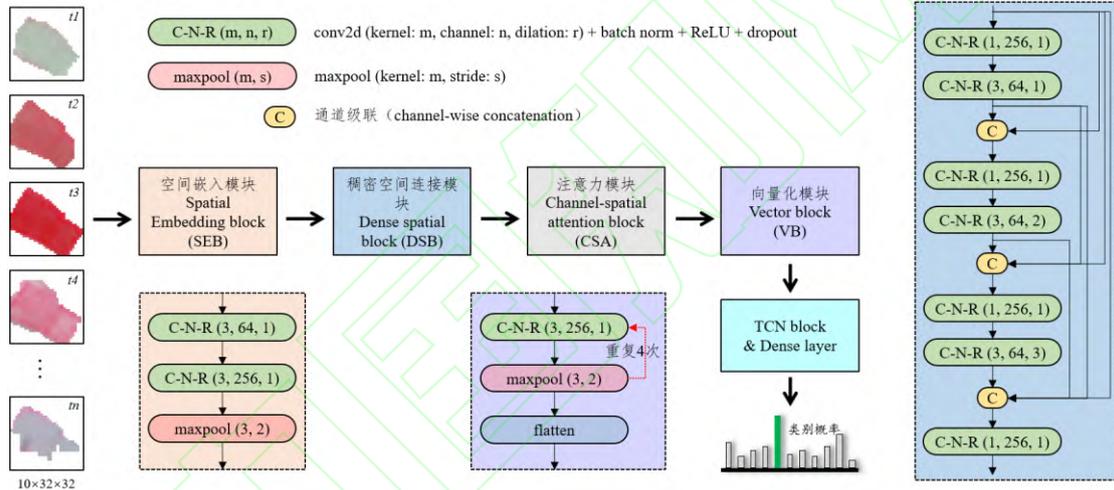


图 5 TSST-Net 网络的整体架构

Fig.5 Overall architecture of the TSST-Net

1) 多层次空间特征

TSST 网络模型利用 DSB 模块提取影像的多尺度空间信息，其实现沿用了 DSNet^[31]提出的稠密膨胀卷积模块（图 5）。具体而言，DSB 利用具有密集感受野的膨胀卷积来捕获影像空间的多尺度信息、且保持较低的模型参数量；利用稠密的通道级联（channel-wise concatenation）来获取影像波段间的多尺度特征；并在膨胀卷积层之前利用 1×1 的卷积层来阻止特征通道数的增大。

2) 时间序列分类

相对于神经网络的序列建模，TCN 网络具有多层次的感受野、稳定的梯度传播和更低的资源消耗等优点^[32]。本文选用 TCN 实现从地块时序特征到作物类型的映射。在实现上，TCN 利用膨胀卷积来捕捉序列观测的长时间依赖关系，利用残差链接实现网络的跨层信息传递。具体地，TCN 由三个连续的一维卷积层组成、然后是批量归一化（BatchNorm）、ReLU 激活函数和随机丢弃（Dropout）层。经 TCN 时序编码后的特征被展平并传递到具有批归一化、ReLU 激活函数和 Dropout 的全连接层，并经 SoftMax 函数计算作物类别概率。

3) 损失函数

针对研究区作物类型的类别不均衡性，采用多类别 focal loss 损失函数^[33]来训练时空分类模

型。

2.5 地块尺度作物类别合并

空间分区生成了地块的多个影像瓦片、也对应多了个作物类别概率。因此需要融合地块的多个类型概率以生成最终的地块作物种植类型。假设待预测瓦片 i 区域内的有效像元数（同时落入瓦片与地块多边形内）为 n_i 、瓦片 i 的作物类别概率为 $p_i=(p^1_i, p^2_i, \dots, p^T_i)$ ，其中 T 为分类体系的类别个数。借鉴集成学习中的投票思想，计算作物类别概率的加权（以 n_i 为权重）和，并将概率最高的类别作为最终的地块作物类型 Type ，公式如下：

$$\text{Type} = \arg \max_{t \in (1,T)} \left\| \sum_{i=1}^K n_i \times p_i \right\| \quad (1)$$

式中， K 为地块的关键点数。

2.6 性能评价与对比

1) 对比方法

本文选用多个对比方法来验证提出 TSST-Net 模型的有效性。对比方法分为两大类：基于像元的时序分类（pixel-based）和地块尺度（parcel-scale）的时序分类方法。其中基于像元的分类方法以像元时序光谱值为输入，利用 LSTM^[3]（long- and short-term memory）和 Transformer^[5] 模型提取其时序特征并分类；进而以地块多边形为空间约束，投票合并上述像元分类结果，生成地块尺度类别投票（parcel voting）分类结果；同时以地块为分析单元提取地块区域内光谱均值构建地块时间序列观测，并输入到 LSTM 和 Transformer 模型，生成地块尺度特征（parcel feature）的作物分类结构。为保证对比实验的公平性，所有模型均采用相同的参数设置（包括优化器、损失函数、批大小、初始学习率、早停机制等），并选用其最优训练模型进行测试。

2) 评价指标

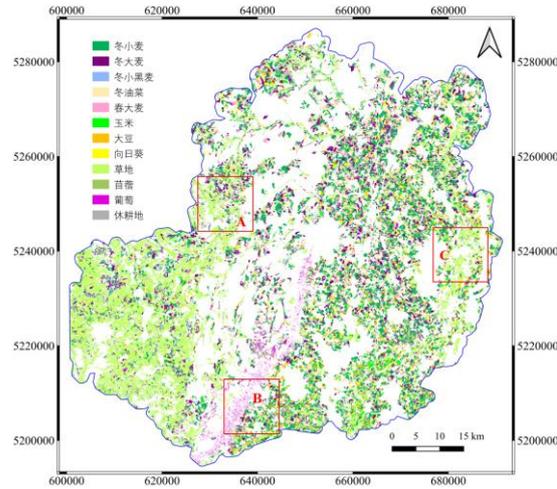
通过逐地块对比测试样本类别与分类模型预测类别，构建模型分类结果的混淆矩阵并计算总体精度（overall accuracy, OA）、准确度（ P ）、召回率（ R ）、 $F1$ 分数等指标。总体精度 OA 表达了正确分类的地块数与测试样本集中地块数之间的比值。通过公式 $P = TP / (TP + FP)$ 和 $R = TP / (TP + FN)$ 来计算准确度和召回率，其中 TP、TN、FP 和 FN 分别代表混淆矩阵中真阳性、真阴性、假阳性和假阴性的数量。准确度和召回率的调和平均数 $F1 = 2 \times UA \times PA / (UA + PA)$ 更能够表达某一类作物的分类精度。整体上来说，更高的 OA、 P 、 R 和 $F1$ 分数表示更好的分类结果，反之亦然。

3 实验与讨论

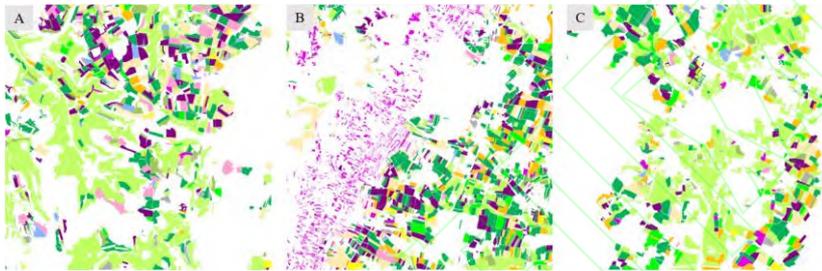
3.1 实验结果

1) 作物类型分类结果

TSST 方法生成的作物种植类型图和局部细节如图 6 所示。通过目视解译评价作物类型识别方法的整体效果。草地主要分布在研究区西南部以及 Vingeane 河两岸（局部细节 C）。葡萄主要种植在从切诺夫（Chenôve）到博恩（Beaune）之间的狭长谷底（局部细节 B），这里也是自中世纪以来葡萄栽培和葡萄酒生产的典型区域。研究区中东部的平地上主要种植着冬小麦、冬大麦、玉米等作物。



(a) 研究区作物种植类型分类全图



(b) 作物种植类型局部细节

图 6 本文 TSST 方法的作物种植类型分类结果。

Fig.6 Crop classification using the proposed TSST-Net.

2) 精度分析与对比

为定量评价 TSST 方法的有效性，计算和对比了多种作物分类方法（包括基于像元的 LSTM 和 Transformer 方法、地块尺度类别投票合并的 LSTM-V 和 Transformer-V 方法,以及地块尺度特征特征的 LSTM-P 和 Transformer-P 方法）结果的精度指标，如表 1 所示。

表 1 作物类型分类精度评价与对比。

Tab.1 Accuracy comparison of crop type classifications

| 精度 | 基于像元 | | 地块尺度 | | | | 本文 TSST 方法 |
|----|---------|-------------|----------------|---------------|---------|----------------|------------|
| | LSTM | Transformer | 类别投票 | | 尺度特征 | | |
| | | | LSTM-V | Transformer-V | LSTM-P | Transformer-P | |
| OA | 0.756 1 | 0.769 3 | 0.808 2 | 0.791 6 | 0.800 3 | 0.801 3 | 0.833 5 |
| P | 0.545 7 | 0.561 4 | 0.657 2 | 0.673 2 | 0.658 2 | 0.694 6 | 0.718 7 |
| R | 0.642 7 | 0.660 0 | 0.744 5 | 0.711 6 | 0.745 6 | 0.720 4 | 0.743 8 |
| F1 | 0.590 3 | 0.606 7 | 0.698 2 | 0.691 9 | 0.699 2 | 0.707 3 | 0.731 0 |

注：红色字体标识了最优指标分数，粗体标识了次优指标分数。

整体上看，本文 TSST 方法取得了最好的作物分类效果。其中，OA、P、R、F1 指标分数分别达到 0.833 5、0.718 7、0.743 8 和 0.731 0。特别是：相对于次优精度指标，TSST-Net 方法在 OA、P 和 F1 分数上分别提高了 0.03、0.02 和 0.02。这也定量说明了 TSST 方法在遥感时空作物类型识别上的优势。具体而言，像元尺度的 LSTM 和 Transformer 方法取得了最低的作物分类精度，而地块尺度分类方法均取得了较好的分类结果。这是显而易见的，因为以地块多边形为空间约束的特征统计和作物类别合并能够有效消除特征提取和分类中的异常值和椒盐噪声，生成更高精度的作物类型分类结果。进而相对于类别投票和尺度特征方法，TSST-Net 方法则能够有效

挖掘和利用影像的多层次空间信息，也最终获得了最好的分类结果。实验还发现：相对于召回率得分的提升，TSST 取得了更大的精度提升。究其原因可能是，由于挖掘和利用了地块空间的高层语义信息，提升了模型的识别准确度。

3.2 分析与讨论

1) 作物类型 VS 精度提升

以 LSTM-P 方法结果为基准，分析 TSST 方法分类结果中不同类别作物识别的 $F1$ 精度及其提升（如图 7 所示）。

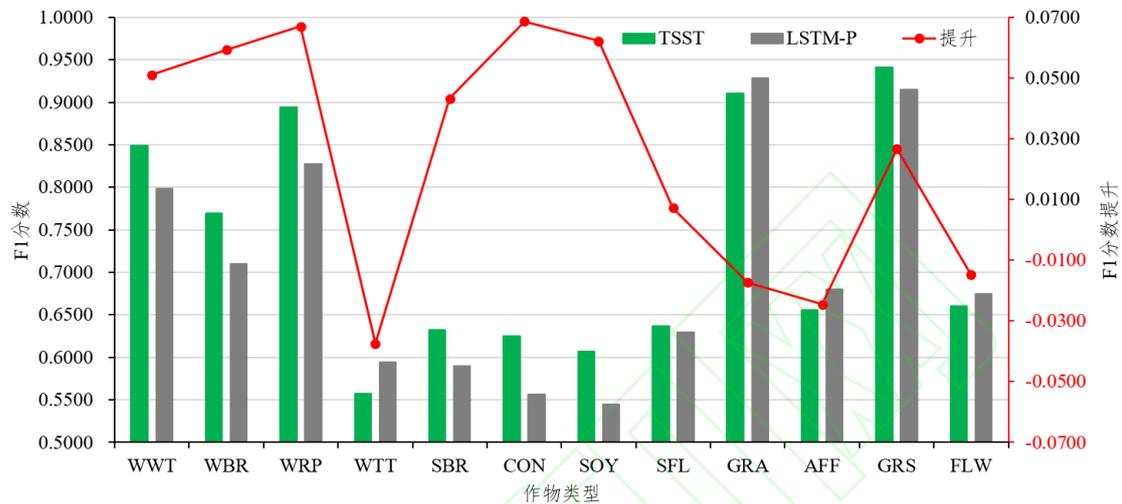


图 7 不同类型作物的识别精度及其提升.

Fig.7 Classification accuracies and improvement of different types of crops.

遥感时空作物类型识别中，不同类型作物表现出较大的性能差异。TSST 方法和基准方法 LSTM-P 均在 WRP、GRA 和 GRS 等作物上取得了相对较高的精度，而在 WTT、CON 和 SOY 作物的识别中表现较差。一方面是由于 GRA 和 GRS 作物其相对独特的生长过程；另一方面结合样本中作物类别数目来看，尽管实验采用 focal loss 来克服样本类别不平衡的问题，样本数占优的类别仍取得更好的分类精度。进一步分析本文 TSST 方法在不同作物上的性能提升。其中在 WRP、CON 和 SOY 作物上提升明显，而在 WTT、GRA、AFF 和 FLW 类被上出现了稍微的精度下滑。从作物地块的平均面积和空间分布来看，提升效果明显的多表现在地块面积大、空间分布集中的作物类别，地块面积小、分布零散的作物则表现出更差的效果。可能原因是，方法中分区分层处理将会提高影像瓦片序列中缺失值的比率、也提升了时空卷积网络学习影像空间特征的难度，尤其是面积较小且分布零散作物类型的影像瓦片，其缺失率更高、学习到特征的噪声更大、识别模式更不稳定。

2) 云/影缺失值处理

本文 TSST 方法通过空间分区与类别/地块分层减弱了类别之间时空特征的混淆，但也导致了影像瓦片更多的像元值缺失。实验也设计随机填充和正态填充两种缺失值修复方案。不同填充方案的作物分类精度如表 2 所示，No 代表不进行云/影掩膜操作，Zero 为直接置零处理，Random 和 Normal 分别为随机和正态填充处理，#epoch 为模型训练达到稳定所需的训练次数。

表 2 不同缺失值填充方法的作物分类精度。

Tab.2 Classification accuracies resulted from different reconstruction of missing values.

| 精度 | No | Zero | Random | Normal |
|-----|---------|---------|---------|---------|
| OA | 0.757 9 | 0.833 5 | 0.825 1 | 0.834 5 |
| P | 0.522 3 | 0.718 7 | 0.689 0 | 0.706 5 |
| R | 0.612 0 | 0.743 8 | 0.776 1 | 0.761 0 |

| | | | | |
|--------|---------|---------|---------|---------|
| F1 | 0.563 6 | 0.731 0 | 0.729 9 | 0.732 7 |
| #epoch | 92 | 58 | 42 | 30 |

整体上, Zero、Random、Normal 方案均取得较 No 更好的精度。其原因是: 相对于 No 方案中作物空间特征会受到污染像元的干扰, 其他方法均消除了污染像元, 并且 Random 和 Normal 方法还模拟填充了缺失值。两种填充方法并没有取得较 Zero 处理更好的分类效果。尽管两种填充方法模拟了缺失值, 但填充操作也破坏了影像的高层语义信息(如纹理等)。但两种填充操作能够有效提升分类模型的训练效率、降低模型训练所需的次数, 使得模型收敛更快。

3) 模型性能对比与分析

为评价本文构建 TSST-Net 网络的性能, 实验将其与流行的深度时空分析网络模型(ResNet-TCN 和 P3D)对比。表 3 展示多个模型在参数量、预测速度和总体精度方面的性能对比, 其中 ResNet-TCN 和 P3D 网络分别选用了 50 层和 63 的骨干网络。

表 3 时空网络模型性能对比与分析.

Tab.3 Performance comparison of spatial-temporal models for crop classification.

| 分类精度 | ResNet-TCN | P3D | DSTNet |
|------|------------|---------|---------|
| 参数量 | 23.68 M | 24.97 M | 1.70 M |
| 预测速度 | 2.21 | 2.43 | 1.19 |
| OA | 0.834 1 | 0.829 7 | 0.833 5 |

在模型分类精度上, 3 个模型取得了几乎一样的总体精度。相对于 ResNet-TCN 和 P3D 模型, TSST-Net 模型具有更小的参数量和更快的预测速度。也就是说, TSST-Net 方法以更轻量的模型取得了与媲美 ResNet-TCN 和 P3D 的作物类型识别效果。

4 结 语

充分挖掘影像中作物种植的时空信息是提升遥感作物种植类型分类和制图的有效途径。本文在地块空间的分区分层表达和深度时空卷积分类网络的基础上, 提出了一种分区分层的遥感时空作物种植类型分类方法(TSST)。基于 Sentinel-2 时序遥感影像的法国研究区地块尺度作物种植类型制图实验表明: (1) TSST 方法能够有效学习影像空间的多尺度时空语义特征并提升遥感时空作物识别的精度, 相对于现有方法在总体精度、准确度和 F1 分数上分别提高 3%、2%和 2%; (2) 面向地块特征表达的分区分层处理, 更有利于大地块面积、空间集中分布的作物类型识别; (3) TSST-Net 网络是一种轻量且高效的作物类型识别模型。上述实验及其结论将为农业遥感时间序列分析提供新的思路和方法。同时, 这些方法也能容易移植和适应于其他卫星数据和应用。

此外, 云/影等成像条件限制了光学遥感对作物生长过程的时空观测。后续研究将考虑引入和融合微波遥感观测(如 Sentinel-1 和高分 3 号卫星数据)来提升对作物生成的时间序列监测和制图能力。

参考文献

- [1] CHEN Zhongxin, REN Jianqiang, TANG Huajun, et al. Progress and perspectives on agricultural remote sensing research and applications in China. [J]. Journal of Remote Sensing, 2016, 20(5): 748-767. (陈仲新, 任建强, 唐华俊, 等. 农业遥感研究应用进展与展望 [J]. 遥感学报, 2016, 20(5): 748-767.)
- [2] WU Bingfang, ZHANG Miao, ZENG Hongwei, et al. Agricultural monitoring and early warning in the era of big data. [J]. Journal of Remote Sensing, 2016, 20(5): 1027-1037. (吴炳方, 张淼, 曾红伟, 等. 大数据时代的农情监测与预警 [J]. 遥感学报, 2016, 20(5): 1027-1037.)
- [3] ZHOU Y N, LUO J, FENG L, et al. Long-short-term-memory-based crop classification using high-resolution optical images and multi-temporal SAR data [J]. GIScience & Remote Sensing, 2019, 56(8): 1170-91.

- [4] DONG Jinwei, WU Wenbin, HUANG Jianxi, et al. State of the Art and Perspective of Agricultural Land Use Remote Sensing Information Extraction[J]. Journal of Geo-information Science, 2020, 22(4): 772-783. (董金玮, 吴文斌, 黄健熙, 等. 农业土地利用遥感信息提取的研究进展与展望 [J]. 地球信息科学学报, 2020, 22(4): 772-783.)
- [5] RUßWURM M, KÖRNER M. Self-attention for raw optical satellite time series classification [J]. ISPRS journal of photogrammetry and remote sensing, 2020, 169: 421-35.
- [6] WU Zhifeng, LUO Jiancheng, SUN Yingwei, et al. Research on Precision Agricultural based on the Spatial-temporal Remote Sensing Collaboration[J].Journal of Geo-information Science, 2020, 22(4): 731-742. (吴志峰, 骆剑承, 孙莹伟, 等. 时空协同的精准农业遥感研究 [J]. 地球信息科学学报, 2020, 22(4): 731-742.)
- [7] QIAN Lisha, JIANG Hao, CHEN Shuisen, et al. Extracting field-scale crop distribution in Lingnan using spatiotemporal filtering of Sentinel-1 time-series data[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE),2022,38(5):158-166. (钱丽沙, 姜浩, 陈水森, 等. 基于时空滤波 Sentinel-1 时序数据的田块尺度岭南作物分布提取 [J]. 农业工程学报, 2022, (5): 158-166.)
- [8] ZHANG X, ZHOU Y N, LUO J. Deep learning for processing and analysis of remote sensing big data: A technical review [J]. Big Earth Data, 2021: 1-34.
- [9] CAI Y, GUAN K, PENG J, et al. A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach [J]. Remote sensing of environment, 2018, 210: 35-47.
- [10] GARIOUDA, VALERO S, GIORDANO S, et al. Recurrent-based regression of Sentinel time series for continuous vegetation monitoring [J]. Remote Sensing of Environment, 2021, 263: 112419.
- [11] ZHOU Y N, WANG S, WU T, et al. For-backward LSTM-based missing data reconstruction for time-series Landsat images [J]. GIScience & Remote Sensing, 2022, 59(1): 410-30.
- [12] WANG S, ZHOU Y N, YANG X, et al. BSNet: Boundary-semantic-fusion network for farmland parcel mapping in high-resolution satellite images [J]. Computers and Electronics in Agriculture, 2023, 206: 107683.
- [13] ZHOU Y N, LUO J, FENG L, et al. DCN-based spatial features for improving parcel-based crop classification using high-resolution optical images and multi-temporal SAR data [J]. Remote Sensing, 2019, 11(13): 1619.
- [14] ZHOU Y, LI J, FENG L, et al. Adaptive Scale Selection for Multiscale Segmentation of Satellite Images [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, 10(8): 3641-51.
- [15] LI Qianjing, LIU Jun, MI Xiaofei, et al. Object-oriented crop classification for GF-6 WFV remote sensing images based on Convolutional Neural Network[J]. National Remote Sensing Bulletin. 2021, 25(2):549-558. (李前景, 刘珺, 米晓飞, 等. 面向对象与卷积神经网络模型的 GF-6 WFV 影像作物分类 [J]. 遥感学报, 2021, 25(2): 549-558.)
- [16] LUO Jiancheng, WU Tianjun, WU Zhifeng, et al. Methods of Intelligent Computation and Pattern Mining based on Geo-parcels[J].Journal of Geo-information Science, 2020, 22(1): 57-75. (骆剑承, 吴田军, 吴志峰, 等. 地理图斑智能计算及模式挖掘方法研究 [J]. 地球信息科学学报, 2020, 22(1): 57-75.)
- [17] XU L, MING D, ZHOU W, et al. Farmland extraction from high spatial resolution remote sensing images based on stratified scale pre-estimation [J]. Remote Sensing, 2019, 11(2): 108.
- [18] GONG Jianya. Chances and Challenges for Development of Surveying and Remote Sensing in the Age of Artificial Intelligence[J]. Geomatics and Information Science of Wuhan University, 2018, 43(12): 1788-1796. (龚健雅. 人工智能时代测绘遥感技术的发展机遇与挑战 [J]. 武汉大学学报(信息科学版), 2018, 43(12): 1788-1796.)
- [19] LUO Jiancheng, WU Tianjun, XIA Liegang. The Theory and Calculation of Spatial-spectral Cognition of Remote Sensing[J]. Journal of Geo-information Science, 2016, 18(5): 578-589. (骆剑承, 吴田军, 夏列钢. 遥感图谱认知理论与计算 [J]. 地球信息科学学报, 2016, 18(5): 578-589.)
- [20] ZHANG Bing. Remotely Sensed Big Data Era and Intelligent Information Extraction[J]. Geomatics and Information Science of Wuhan University, 2018, 43(12): 1861-1871. (张兵. 遥感大数据时代与智能信息提取 [J]. 武汉大学学报(信息科学版), 2018, 43(12): 1861-1871.)
- [21] YANG L, HUANG R, HUANG J, et al. Semantic segmentation based on temporal features: Learning of temporal-spatial

- information from time-series SAR images for paddy rice mapping [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 60: 1-16.
- [22] WEI P, CHAI D, LIN T, et al. Large-scale rice mapping under different years based on time-series Sentinel-1 images using deep semantic segmentation model [J]. *ISPRS journal of photogrammetry and remote sensing*, 2021, 174: 198-214.
- [23] HAN Yanxin, MENG Jihua. A review of per-field crop classification using remote sensing[J]. *Remote Sensing for Land & Resources*, 2019, 31(2): 1-9. (韩衍欣, 蒙继华. 面向地块的农作物遥感分类研究进展[J]. *自然资源遥感*, 2019, 31(2): 1-9.)
- [24] YANG Yingpin, WU Zhifeng, LUO Jiancheng, et al. Parcel-based crop distribution extraction using the spatiotemporal collaboration of remote sensing data[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2021, 37(7): 166-174. (杨颖频, 吴志峰, 骆剑承, 等. 时空协同的地块尺度作物分布遥感提取[J]. *农业工程学报*, 2021, 37(7): 166-174.)
- [25] SUN Y, LI Z-L, LUO J, et al. Farmland parcel-based crop classification in cloudy/rainy mountains using Sentinel-1 and Sentinel-2 based deep learning [J]. *International Journal of Remote Sensing*, 2022, 43(3): 1054-73.
- [26] ZHANG Qiankun, MENG Jihua, REN Chao. Crop classification based on two-dimensional representation and CNN model from remote sensing[J]. *National Remote Sensing Bulletin*, 2022, 26(7): 1437-1449. (张乾坤, 蒙继华, 任超. 构建地块二维表征及 CNN 模型的作物遥感分类 [J]. *遥感学报*, 2022, 26(7): 1437-1449.)
- [27] RÜßWURM M, PELLETIER C, ZOLLNER M, et al. Breizhcrops: A time series dataset for crop type mapping. *ISPRS [J]. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2020, 43: B2-2020.
- [28] SONG C, WOODCOCK C E, SETO K C, et al. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? [J]. *Remote sensing of Environment*, 2001, 75(2): 230-44.
- [29] QIU Z, YAO T, MEI T. Learning spatio-temporal representation with pseudo-3d residual networks; proceedings of the proceedings of the IEEE International Conference on Computer Vision, F, 2017 [C].
- [30] HARA K, KATAOKA H, SATOH Y. Can spatiotemporal 3d cnns retrace the history of 2d cnns and imagenet?; proceedings of the Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, F, 2018 [C].
- [31] DAI F, LIU H, MA Y, et al. Dense scale network for crowd counting; proceedings of the Proceedings of the 2021 international conference on multimedia retrieval, F, 2021 [C].
- [32] BAI S, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling [J]. *arXiv preprint arXiv:180301271*, 2018.
- [33] LIN T-Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection; proceedings of the Proceedings of the IEEE international conference on computer vision, F, 2017 [C].