

基于多源光学雷达数据融合的黄淮海平原冬小麦识别

冯权洧^{1,2} 任 燕¹ 姚晓闯^{1,2} 牛博文¹ 陈泊安¹ 赵圆圆^{1,2}

(1. 中国农业大学土地科学与技术学院, 北京 100193; 2. 自然资源部农用地质量与监控重点实验室, 北京 100193)

摘要: 遥感技术能够快速准确地获取农作物空间分布信息,为探究 2021 年黄淮海平原冬小麦空间分布信息,基于 Google Earth Engine(GEE)云平台,以 Sentinel-1 SAR 雷达影像和 Sentinel-2 光学遥感影像为数据源,通过计算极化特征、光谱特征和纹理特征,运用随机森林等 4 种机器学习方法和深度循环神经网络模型,对研究区冬小麦空间分布信息进行提取,并对比各分类器和网络架构的分类精度。结果表明,黄淮海平原冬小麦总面积约为 16 226 667 hm²,占研究区总面积的 49.17%,其中冬小麦种植面积最大的是河南省,约为 4 647 334 hm²,研究区冬小麦种植分布呈现由东向西、由南向北递减的趋势;随机森林是 4 种机器学习方法中识别精度最高的分类器,总体分类精度为 94.30%;在随机森林算法中仅使用 Sentinel-1 雷达数据总体精度为 87.38%,仅使用 Sentinel-2 光学数据总体精度为 93.95%,而融合时序 Sentinel 主被动遥感数据总体精度为 94.30%;在大范围的冬小麦分类上,深度学习模型的泛化性高于机器学习方法。

关键词: 冬小麦; 遥感分类; 多源数据融合; Google Earth Engine; 机器学习; 深度学习

中图分类号: S512.1⁺1; S127 文献标识码: A 文章编号: 1000-1298(2023)02-0160-09 OSID: 

Identification of Winter Wheat in Huang – Huai – Hai Plain Based on Multi-source Optical Radar Data Fusion

FENG Quanlong^{1,2} REN Yan¹ YAO Xiaochuang^{1,2} NIU Bowen¹ CHEN Boan¹ ZHAO Yuanyuan^{1,2}

(1. College of Land Science and Technology, China Agricultural University, Beijing 100193, China

2. Key Laboratory for Agricultural Land Quality Monitoring and Control, Ministry of Natural Resources, Beijing 100193, China)

Abstract: Current remote sensing technology can quickly and accurately obtain the spatial distribution information of crops. In order to explore the spatial distribution information of winter wheat in the Huang – Huai – Hai Plain in 2021, based on the Google Earth Engine (GEE) cloud platform. Sentinel – 1 SAR radar image and Sentinel – 2 optical remote sensing image were used as data sources, the spatial distribution information of winter wheat in the study area was extracted by computing polarization characteristics, spectral characteristics and texture characteristics, using four machine learning methods and deep learning network model. The classification accuracy of each classifier and network architecture was compared. The results showed that the total area of winter wheat in the Huang – Huai – Hai Plain was 16 226 667 hm², accounting for 49.17% of total area of the study area. The winter wheat planting area was the largest in Henan Province, accounting for 4 647 334 hm². The winter wheat planting distribution in the study area showed a decreasing trend from east to west and from south to north. Random forest was the classifier with the highest recognition accuracy among the four machine learning methods, with an overall classification accuracy of 94.30%. In the random forest algorithm, the overall accuracy of only using Sentinel – 1 radar data was 87.38%, and the overall accuracy of only using Sentinel – 2 optical data was 93.95%, while the overall accuracy of the fusion sequence Sentinel active and passive remote sensing data was 94.30%. In a wide range of winter wheat classification, the generalization of deep learning model was higher than that of machine learning.

Key words: winter wheat; remote sensing imagery classification; multi-source data fusion; Google Earth Engine; machine learning; deep learning

收稿日期: 2022 – 03 – 29 修回日期: 2022 – 05 – 26

基金项目: 国家超级计算郑州中心创新生态系统建设科技专项(201400210100)和国家自然科学基金项目(42001367)

作者简介: 冯权洧(1988—),男,副教授,博士生导师,主要从事机器学习、深度学习在农业遥感应用研究,E-mail: fengql@cau.edu.cn

通信作者: 姚晓闯(1986—),男,副教授,博士生导师,主要从事空间数据组织、管理与分析等研究,E-mail: yxc@cau.edu.cn

0 引言

冬小麦是我国三大主粮作物之一，与国家粮食安全息息相关。黄淮海平原作为我国重要的农业生产区，同时是全国最大冬小麦产区，小麦产量占全国总产量的 75%^[1]，及时精准地获取冬小麦种植面积和空间分布信息对我国农业发展具有重要意义。当前基于统计数据和遥感监测的冬小麦提取研究仍为主流研究方法，效率低且耗费大量人力成本，难以满足当前快速、准确地获取农作物信息的实际需求^[2-4]。而运用遥感可快速获取大范围地表覆被信息，在不同时空尺度的冬小麦识别方面得到了广泛应用^[5]。常用的提取冬小麦面积的方法是结合遥感影像，通过比较植被指数时序变化选择合适阈值，进而提取冬小麦的空间分布信息^[6-8]；另外，机器学习和深度学习等算法的发展也为农作物分类、冬小麦面积提取提供了新的方法^[2,9]。随着对遥感图像处理能力需求的提高，高效处理图像数据的各类云平台需求也逐渐增加。遥感云计算平台（例如：Google Earth Engine, GEE）具有强大的计算能力，能够快速分析遥感影像并及时获取大范围的农作物信息^[10-15]。由于遥感数据的工作限制和原理的不同，单一传感器难以完整反映地表覆被特征。现阶段作物识别研究多基于光学影像作为数据源，运用多时相数据增强不同地物之间的光谱可分性，能够更好地分析冬小麦的生长发育规律在遥感影像上表现出的时相变化特征，有效提高冬小麦识别和分类精度。

目前融合多种传感器的多时空分辨率的光学遥感影像被用来提取农作物种植结构并取得了大量成果^[16]。基于高分辨率影像数据进行农作物提取也受到广泛关注^[17]。雷达遥感相对于光学遥感最大的优点在于可穿透云雨，不受日照影响和限制，可全天时、全天候工作。诸多学者也将雷达数据与光学数据结合用于农作物种植结构信息提取。ARIAS 等^[18]探讨了基于 Sentinel-1 时间序列数据进行农作物分类的可行性；GEE 是一个免费的遥感云计算平台，基于 GEE 云平台已进行了诸多在遥感领域的研究。SHELESTOV 等^[19]基于 GEE 云平台进行农作物分类研究，利用云平台获取大量多时相遥感数据，并采用多种主流分类器进行农作物提取，对比了各分类器的提取效果；谭深等^[10]运用融合多源遥感数据进行多云气候区的水稻分类研究，并基于机器学习中的随机森林分类方法获取 2016 年的水稻分布图，获得较好结果。

综上，本文基于 2020 年 12 月至 2021 年 5 月黄淮海平原的哨兵 1 号（Sentinel-1）SAR 雷达影像

和哨兵 2 号（Sentinel-2）光学遥感影像，分别构建各地物类型的光谱特征、极化特征和纹理特征数据，分析不同地物在以上特征中的时序曲线差异、冬小麦的生育周期以及对应的植被指数特征差异。利用 GEE 云平台，基于多种机器学习分类算法（朴素贝叶斯、支持向量机、决策树和随机森林）进行分类及精度验证，从而得出最优分类算法，并利用已选择的冬小麦的样本点，提取样本长时间序列的特征向量，基于深度学习中长短期记忆网络（Long short term memory, LSTM）和 Bi-LSTM 进行冬小麦的像元尺度的模型训练。基于实验结果，获得 2021 年黄淮海平原区域冬小麦种植面积及其时空变化特征，以期为该区域轮作休耕政策制订以及农业生产提供数据支持。

1 研究区域与数据来源

1.1 研究区概况

黄淮海平原位于北纬 32°~40°，东经 114°~121°，面积 3 × 10⁵ km²。跨越京、津、冀、鲁、豫、皖、苏 7 省市，地势平坦，河湖众多，交通便利，年均气温 13℃，年均降水量 710 mm，属温带大陆性季风气候，雨热同期，土质肥沃，主要生长的农作物为小麦、玉米、大豆等，其中冬小麦多于每年的 10 月上、中旬播种，次年 6 月收获^[20]。黄淮海平原研究区域范围划分参考文献[21]，包含北京、天津、河北、河南、山东、安徽、江苏 7 个省级行政区（图 1）。

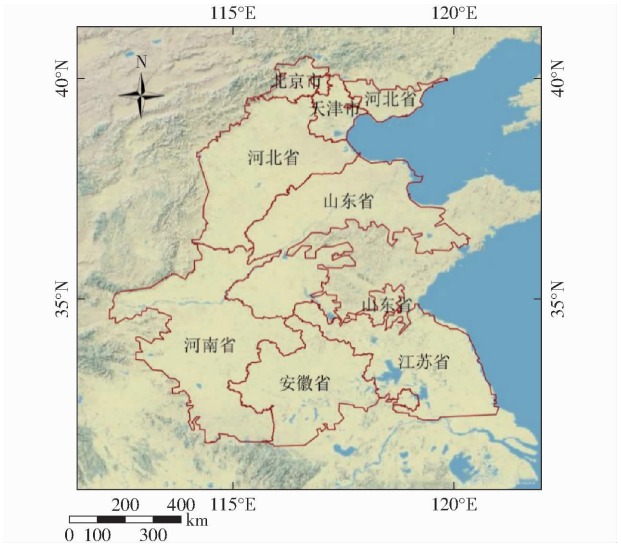


图 1 研究区范围示意图

Fig. 1 Overview of study area

1.2 数据来源与预处理

选用 Sentinel-1 SAR 数据和 Sentinel-2 光学数据进行冬小麦面积提取。Sentinel-2A 卫星覆盖 13 个光谱波段，为获得无云且包含冬小麦生长期的影像数据，通过扩大影像的时间跨度，设置时间范围

为 2020 年 12 月至 2021 年 5 月,得到黄淮海平原空间分辨率为 10 m 的 Sentinel-2 卫星影像数据,在月尺度上进行均值合成,筛选云量低于 20% 的影像。Sentinel-2 数据的裁剪、镶嵌处理通过 GEE 平台在线编写代码实现。Sentinel-1 是主动微波遥感卫星,能够全天候、全天时地对地面进行观测,具有单极化、双极化等多种极化方式^[22-23]。本研究运用 VV、VH 极化数据进行极化特征计算,并在月尺度上进行均值合成,影像空间分辨率为 10 m。

1.3 特征变量与数据集

1.3.1 特征构建

结合研究区的地理位置与农作物的物候期,选取 Sentinel-2 光学影像的光谱反射率、光谱特征和纹理特征进行计算。现有研究表明农作物遥感分类研究仅计算光谱特征会导致“同物异谱、异物同谱”的现象,影响最终分类效果,而不同地物类型在遥感影像上纹理表现差异较大。冬小麦在遥感影像上具有规则的纹理信息,易于与其他地物类型进行区分。基于此,本文运用灰度共生矩阵 (Gray-level co-occurrence matrix, GLCM) 进行纹理特征的构建,根据植被光谱曲线可知,植被在近红外 B8 波段具有最高的反射峰,植被反射率在可见光波段差异小,但近红外波段的差异明显。因此本文采用 Sentinel-2 光学影像的近红外 B8 波段计算纹理特征能够更好地区分植被等信息。为避免因选取过多纹理特征而导致信息冗余,故只选取灰度共生矩阵生成的角二矩阵 (Angular second moment, ASM)、对比度 (Contrast, CONTRAST)、相关性 (Correlation, CORR) 和熵值 4 个纹理特征变量。选择 Sentinel-1 SAR 数据的极化特征变量,共 2 个特征变量。本研究共选取 16 个特征变量,如表 1 所示。

表 1 中,归一化差异植被指数 (Normalized difference vegetation index, NDVI) 反映植被生长状态,通过计算近红外和红光波段之间的差异来识别植被。归一化水体指数 (Normalized difference water index, NDWI) 用于水体识别,对波段进行归一化差值处理,其对植被冠层的液态水含量变化敏感。增强型植被指数 (Enhanced vegetation index, EVI) 的计算常用于植被茂密区域的计算,可减少大气和土壤的影响,反映地表植被特征。绿色归一化差异植被指数 (Green normalized difference vegetation index, GNDVI) 用于计算叶片叶绿素浓度。绿叶指数 (Green leaf index, GLI) 反映研究植被的颜色变化信息。优化土壤调节植被指数 (Optimization soil-adjust vegetation index, OSAVI) 能够反映冬小麦的生长状况信息^[24-28]。

表 1 特征变量		
Tab.1 Characteristic variable		
传感器	特征类别	特征变量
Sentinel-1	极化特征	VV
		VH
Sentinel-2	光谱特征	红波段
		绿波段
		蓝波段
		近红外波段 NIR
		归一化植被指数 NDVI
		归一化水体指数 NDWI
		增强型植被指数 EVI
		绿色归一化差异植被指数 GNDVI
		绿叶指数 GLI
		优化土壤调节植被指数 OSAVI
	纹理特征	角二矩阵
		对比度
		相关性
		熵值

1.3.2 不同地物类型间的时序植被指数差异分析

综合分析文献[15]和研究区的地物类型分布特征,为便于进行不同地物的植被指数差异比较,共选取冬小麦和非冬小麦两大类样本点,非冬小麦地物类型由水体、不透水面、其他(植被+裸地)3类地物组成,结合遥感影像绘制各类地物在12月到次年5月的NDVI变化曲线(图2)。研究区冬小麦从1月到3月的NDVI逐月增加,变化趋势明显有别于其他地物类型,说明该阶段冬小麦绿色加深,根据研究区冬小麦物候信息可知该阶段为返青期。3—4月冬小麦NDVI偏高,为冬小麦孕穗期。4—5月NDVI植被指数呈现明显下降趋势,但总体指数值仍较高,该时期冬小麦叶绿素含量下降,5月以后的冬小麦NDVI植被指数与其他地物类型较为相似,不易进行农作物提取研究。基于特征值差异越大,分类精度越高的原则,选取返青期影像作为选择样本的基础影像,数据日期为2021年2月1日至3月

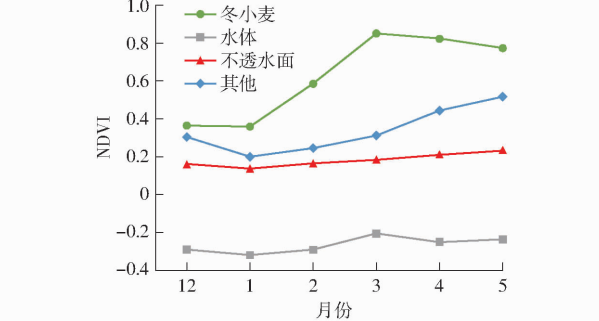


图 2 2020 年 12 月—2021 年 5 月研究区各地物类型
时序 NDVI 变化曲线

Fig.2 Time series NDVI changes of various feature type
in study area from December 2020 to May 2021

1 日。

1.3.3 样本数据集

本文分类目标作物为冬小麦,通过分析遥感影像的地表覆被信息和易混作物信息,选取 20 000 个样本点,综合文献[2]中的训练集测试集划分比例和本研究样本数据量以便于更准确地反映模型效能,从中随机抽取 30% 作为测试样本,70% 作为训练样本,获得研究区训练样本集及测试样本集(表 2)。验证数据来自各省 2020 年统计年鉴发布的冬小麦播种面积数据。根据 1.1 节可知黄淮海平原非 7 省完整区域,故以黄淮海平原包括的各个市的小麦播种面积为统计数据验证提取结果。

基于 2021 年 2 月 1 日至 3 月 1 日的 Sentinel-1

表 2 样本数据集
Tab.2 Sample data sets

数据类别	冬小麦	水体	不透水面	其他(植被+裸地)	合计
训练数据集数量	5 272	2 433	3 389	2 906	14 000
测试数据集数量	2 260	1 043	1 452	1 245	6 000
总计	7 532	3 476	4 841	4 151	

影像和 Sentinel-2 影像,制作了各类地物的目视解译标志,如表 3 所示。

2 研究方法

2.1 总体研究技术路线

本研究技术路线图如图 3 所示。首先利用 GEE

表 3 目视解译标志
Tab.3 Visual interpretation mark

类型	判读要点	图例位置	Sentinel-1影像		Sentinel-2影像	
冬小麦	冬小麦在真彩色影像中为绿色或深绿色,在假彩色影像中为红色,主要在平原地区的农村周边呈现规则的矩形分布	(118.276°E, 33.397°N) (118.261°E, 33.390°N)				
			(VV极化)	(真彩色合成)	(VH极化)	(假彩色合成)
水体	水体包括湖泊、河流等,在真彩色影像呈现绿色,在假彩色影像中为蓝色或深蓝色	(118.629°E, 33.400°N) (118.621°E, 33.394°N)				
			(VV极化)	(真彩色合成)	(VH极化)	(假彩色合成)
不透水面	不透水面包括各类建筑和道路,纹理清晰,成片集中分布	(118.204°E, 33.477°N) (118.196°E, 33.471°N)				
			(VV极化)	(真彩色合成)	(VH极化)	(假彩色合成)
其他(植被+裸地)	其他植被、裸地包含林地、其他山体植被、裸地等类型,裸地在真彩色和假彩色影像上均呈现土黄色	(118.141°E, 33.451°N) (118.134°E, 33.446°N)				
			(VV极化)	(真彩色合成)	(VH极化)	(假彩色合成)

云平台获取多时相的遥感影像,构建地物类型的光谱特征、极化特征和纹理特征,基于多种机器学习分类算法和深度循环神经网络对冬小麦样本点进行训练和精度验证,得出 2021 年黄淮海平原区域的冬小麦种植信息。

2.2 分类算法

随机森林(Random forest, RF)是利用若干个决策树对样本进行训练预测的一种分类方法,由多个 Bagging 集成学习技术训练得到的决策树组成,对输入的数据集进行判别与分类,同时给出各个变量的重要性评分,评估变量在分类中所起的作用。该方法能够在大数据集上运行,具有极好的准确率,能够较为准确地评估各个特征在分类中的重要性,解决

了决策树的性能瓶颈的问题,具有良好的并行性和扩展性。朴素贝叶斯(Naive Bayes, NB)分类算法是贝叶斯分类算法中最为简单和常见的分类方法,该方法假设每一个输入的样本数据集变量均为相互独立的,极大简化了算法复杂性,因此针对实际的数据中存在属性或类别之间存在关联的情况,其分类准确率会降低^[30]。支持向量机(Support vector machine, SVM)是一种按监督学习方式对数据进行二分类模型,通过寻找最优的边距超平面对输入的样本数据进行分割,是一个监督式学习的方法^[31]。该方法适用小样本学习,计算的复杂性不取决于样本的空间维数,具有较好的鲁棒性,提高了模型的泛化能力,但解决多分类问题较为困难,模型效果好坏

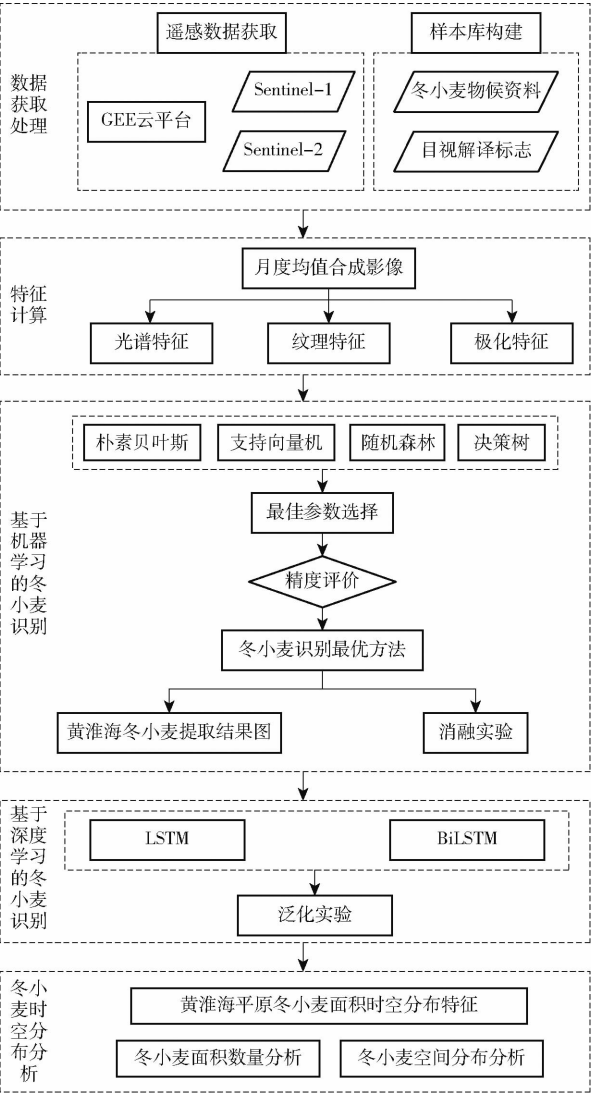


图3 技术路线图

Fig.3 Flow chart of proposed method

与参数和核函数的选择关系密切。决策树(Decision tree, DT)划分方法简单,运用递归思想,以信息熵为度量构造熵值下降最快的树,到叶子节点处熵值为0,是用概率分析的方法^[32]。

本研究基于GEE平台运用Sentinel-1 SAR雷达影像和Sentinel-2光学影像进行月度均值合成,基于返青期的光学遥感影像选择监督分类的样本点。随机森林的参数设置上,进行Grid Search调参,寻找分类的最佳参数,通过调参结果与冬小麦提取的精度、与统计数据的误差以及计算效率,将决策树的数量设置为80,决策树最大深度为7。随机森林模型通过调用GEE平台接口ee.Classifier.smileRandomForest实现,朴素贝叶斯、支持向量机、决策树模型均基于SKlearn框架构建,编程语言为Python。支持向量机算法参数设置惩罚系数为1.0,核函数类型为‘rbf’,gamma值为1/13;决策树模型采用CART决策树分类,分裂评价指数为基尼系数(Gini);朴素贝叶斯算法选用高斯朴素

贝叶斯,该方法不需要进行参数设置。

2.3 精度评价方法

冬小麦的识别精度结果通常采用两种精度评价方法。常用的遥感分类结果的验证方法为混淆矩阵(Confusion matrix)。分类结果评价精度的指标主要包括总体分类精度(Overall accuracy, OA), Kappa系数以及各个类别的制图精度(Producer’s accuracy, PA)、用户精度(User’s accuracy, UA)。Kappa系数是可以衡量总体分类精度的指标,综合了用户精度和生产者精度两个精度的评价参数。取值范围在-1和1之间,值越大图像分类精度越高,说明分类效果越好。

2.4 深度循环神经网络

2.4.1 LSTM网络

循环神经网络(Recurrent neural network, RNN)是一种用于处理序列数据的神经网络。长短时记忆网络(Long short term memory network, LSTM)是一种改进之后的深度循环神经网络,可以解决RNN无法处理长距离的依赖的问题,相比RNN结构,LSTM循环结构能够在长序列中展示更好的结果。该网络的每个循环结构有2个输出,由输入门、遗忘门、输出门和单元状态组成,不断保持一个持久的单元状态进行传递,决定信息是否遗忘或继续传递^[33]。LSTM网络结构如图4所示。图中 x_t 是当前时刻模块的输入值, h_t 是当前模块的输出值, i_t 、 o_t 、 f_t 分别对应输入门、输出门和遗忘门, g_t 为当前时刻的记忆。输入门决定当前网络的输入有多少内容保存到单元状态 c_t 中,遗忘门主要用于计算上一时刻单元状态 c_{t-1} 保留多少内容到当前时刻状态 c_t ,而输出则控制当前状态 c_t 有多少内容会输出到下一阶段LSTM模块中(即 h_t)。

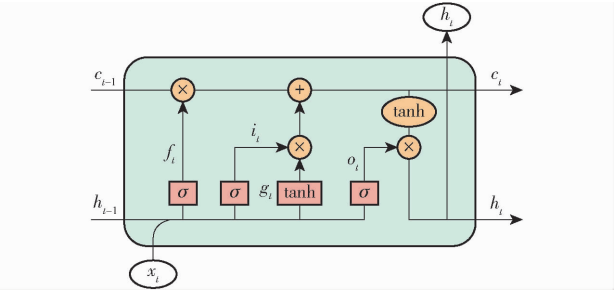


图4 LSTM网络结构

Fig.4 LSTM network structure

2.4.2 Bi-LSTM网络

双向长短记忆网络(Bi-LSTM)是循环神经网络的一种扩展形式,由前向LSTM与后向LSTM组合而成,实际的模型应用中,预测需要考虑完整的上下文信息,需要使用完整的输入序列信息,因此采用双向的循环神经网络建模上下文信息,前向LSTM和后向LSTM层输出向量可以使用相加、平均值或

连接等方式进行处理(图 5)^[34]。

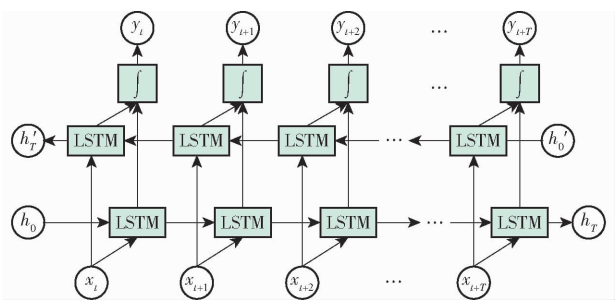


图 5 Bi-LSTM 结构图
Fig. 5 Bi-LSTM network structure

3 结果与分析

3.1 分类结果及精度分析

运用 4 种机器学习方法对研究区进行冬小麦识别,得到研究区冬小麦分类精度 OA 如表 4 所示。

表 4 不同机器学习分类器冬小麦分类精度比较
Tab.4 Comparison of accuracy of winter wheat based on machine learning classifiers in study area

特征	方法	分类精度 OA/%
光谱 + 纹理 + 极化	决策树 DT	92. 20
	朴素贝叶斯 NB	88. 20
	支持向量机 SVM	91. 40
	随机森林 RF	94. 30

由表 4 可知随机森林在 4 种机器学习方法中提
取精度最高。黄淮海平原冬小麦面积分层图和冬小
麦空间位置分布图如图 6 所示。

在 ArcMap 上对分类结果图进行多维分析处
理,对各省市的冬小麦面积进行统计。图 7 为黄淮
海平原冬小麦种植面积的县级分布分层统计图,由
浅到深的颜色表示冬小麦种植面积由少及多。对每
个研究区内的冬小麦种植面积求和,将其相加得到
黄淮海平原冬小麦种植总面积,并计算每个研究区
的种植面积在总面积中的比例,计算式为

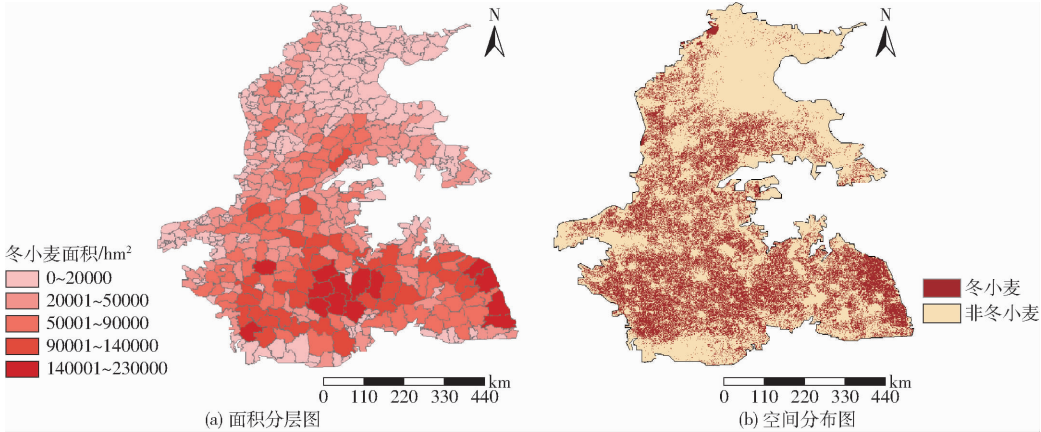


图 6 2021 年黄淮海平原冬小麦面积分层、空间分布图
Fig. 6 Stratified and spatial distribution maps of winter wheat area in Huang - Huai - Hai Plain in 2021

$$W = \sum_{i=1}^5 W_i \tag{1}$$

$$\alpha_i = \frac{W_i}{W} \times 100\% \quad (i = 1, 2, \cdots, 5) \tag{2}$$

式中 W——黄淮海平原冬小麦种植面积
W_i——研究区 i 的冬小麦种植总面积
α_i——研究区 i 冬小麦面积占总面积的比例

黄淮海平原冬小麦总面积为 16 226 667 hm²,占研究
区总面积的 49. 17%,研究区涵盖的各市冬小麦播
种面积统计数据总和约为 15 690 949. 13 hm²,为检
验提取准确性,计算相对误差为 3. 41%。5 个研究
区中,河南省的冬小麦种植面积最多,约为 4 647 334
hm²,占黄淮海平原冬小麦总种植面积的
28. 64%;江苏省次之,约为 3 427 012 hm²,占黄淮海
平原冬小麦总种植面积的 21. 12%;京津冀地区种
植面积最少,约为 1 886 401 hm²,仅占总面积的
11. 63%;山东省、安徽省的冬小麦种植面积分别占
黄淮海平原冬小麦总种植面积的 19. 77%、
18. 84%。本研究黄淮海地区的各省冬小麦提取面
积相对误差均在 5% 以内,由文献[15]可知,相对误
差在 5% 以内即为提取误差较小、提取效果较好。
黄淮海平原的冬小麦种植分布呈现一种由东向西、
由南向北递减的趋势。

3.2 消融实验

基于随机森林的消融实验是通过对比单一传感
器和多传感器遥感数据的冬小麦分类识别精度,仅
使用月度均值合成的 Sentinel - 1 SAR 数据,总体精
度为 87. 38%,Kappa 系数为 0. 763 8;仅使用月度均
值合成的 Sentinel - 2 光学数据,其总体精度为
93. 95%,Kappa 系数为 0. 972 0;而以 Sentinel - 1 和
Sentinel - 2 多源遥感数据为数据源,选择多时相影
像在月尺度上进行均值合成,得到的冬小麦分类结
果为总体精度为 94. 30%,Kappa 系数为 0. 978 4。

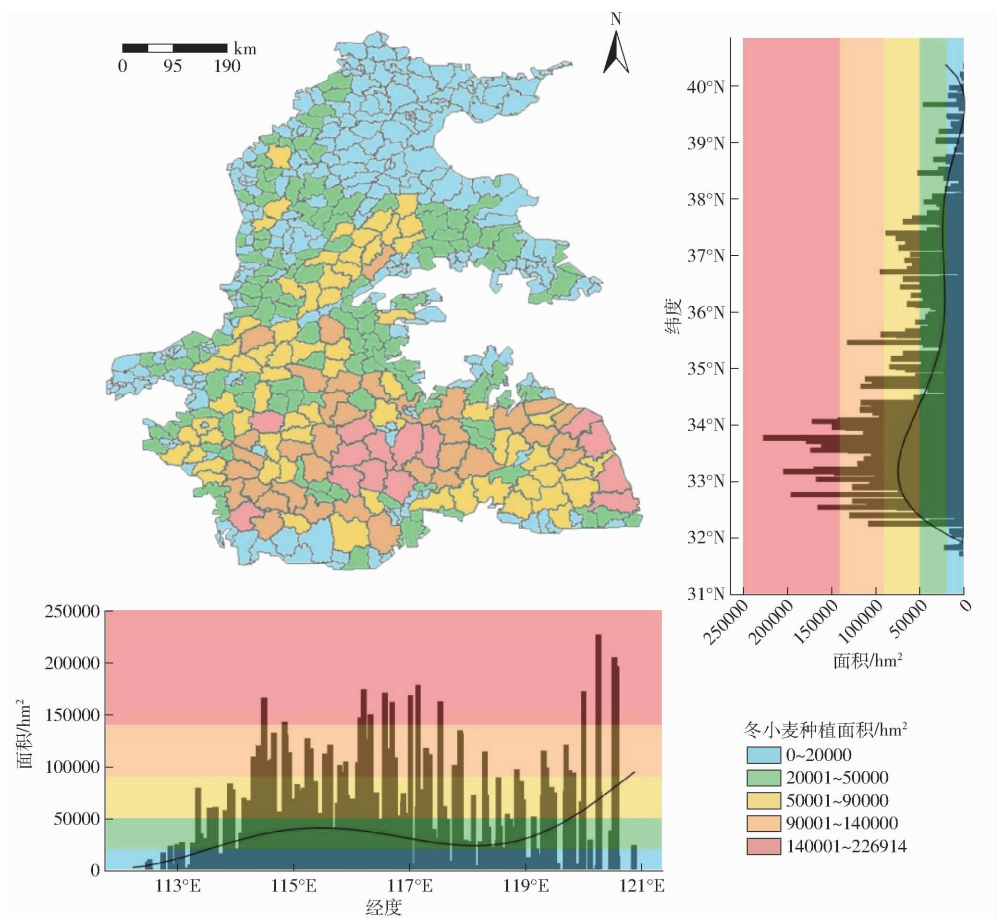


图 7 冬小麦种植面积的县级分布分层统计图

Fig. 7 Stratified statistical map of winter wheat planting area at county level

因此融合时序 Sentinel 主被动遥感数据能够提高分类总体精度和冬小麦的识别精度。

3.3 深度学习精度结果

本实验的深度学习框架为 Tensorflow 2.3.0、Keras 2.4.3。操作系统为 Ubuntu 18.04, CPU 为 Intel Core i7-9700K @ 3.60 GHz, GPU 为 NVIDIA RTX 2080ti。编程语言及版本为 Python 3.7.0, 学习率为 10^{-5} , 迭代次数为 2 000 次, 批处理大小为 1 000, 损失函数为交叉熵, 优化函数为 Adam。

基于深度学习的 LSTM 和 Bi-LSTM 网络架构进行冬小麦识别提取研究, 随机森林、LSTM 和 Bi-LSTM 模型的冬小麦识别精度分别为 94.30%、93.73%、93.87%, 表明深度学习的冬小麦识别精度略低于机器学习随机森林方法, 所以进行深度学习与机器学习的泛化性对比实验。本研究只尝试了 LSTM 系列模型, 并不能完全代表所有的深度学习模型的冬小麦分类精度均低于随机森林方法。

3.4 泛化实验结果

为探究机器学习算法对新样本的适应能力, 将具有同一规律的其他数据集输入已经训练好的网络, 判断输出结果的优劣即为模型的泛化能力。对在某一省份已训练好的网络中输入另一省份的冬小

麦样本数据集, 并输出精度, 得到基于随机森林方法泛化结果(表 5), 将 5 个区域矩阵对角线精度计算平均值, 得到在原本区域的预测结果为 99.11%; 对角线以外的其余精度的平均值代表在其他区域的预测能力, 二者对比即泛化能力, 分别为 90.98%、94.09%、92.63%、87.63%、92.44%。由此说明, 随机森林只在原本实验区域预测精度较高, 在其他区域的识别精度较低, 泛化性低。

表 5 随机森林方法泛化实验结果

Tab. 5 Experimental results of generalization of

random forest method					%
训练	预测				
	安徽	山东	江苏	京津冀	河南
安徽	98.23	88.53	92.10	73.96	91.67
山东	92.15	99.60	97.38	94.16	94.66
江苏	92.26	98.07	99.54	94.58	92.49
京津冀	86.52	93.93	91.74	99.29	90.92
河南	93.00	95.81	89.29	87.82	98.91

基于深度循环神经网络方法 Bi-LSTM 的泛化实验结果如表 6 所示, 得出深度学习在原本区域的预测精度为 96.65%, 对角线以外的其余精度的平均值代表其 在其他区域的预测能力, 分别为 93.72%、96.00%、94.39%、90.84%、95.24%。由

此说明,深度学习的泛化能力在 5 个区域均高于随机森林方法,深度学习在其他地区的泛化性高于随机森林的泛化性,随机森林只在当前实验区域识别精度较高,在其他区域的识别精度较低。

中河南省的冬小麦种植面积最多,约为 4 647 334 hm²;黄淮海平原的冬小麦种植分布呈现一种由东向西、由南向北递减的趋势。

(2)河南省的冬小麦种植面积最多,占黄淮海平原冬小麦总种植面积的 28.64%;京津冀地区种植面积最少,仅占总面积的 11.63%。

(3)随机森林在 4 种机器学习方法中识别精度最高,总体分类精度为 94.30%,其他机器学习方法的总体分类精度分别为决策树 92.20%,朴素贝叶斯 88.20%,支持向量机 91.40%。

(4)融合时序 Sentinel 主被动遥感数据能够提高分类总体精度和冬小麦的识别精度。

(5)在大范围的冬小麦分类上,深度学习模型的泛化性高于机器学习方法。随机森林泛化性较低,并只在当前实验区域识别精度较高,在其他区域的识别精度较低。

表 6 深度学习的 Bi-LSTM 泛化实验结果

Tab.6 Experimental results of generalization of Bi-LSTM for deep learning network %

训练	预测				
	安徽	山东	江苏	京津冀	河南
安徽	96.78	94.77	95.42	87.34	96.35
山东	94.07	97.73	95.40	93.22	95.47
江苏	94.27	96.90	97.07	90.62	94.27
京津冀	92.80	95.82	92.67	94.99	94.85
河南	93.75	96.50	94.06	92.16	96.69

4 结论

(1)黄淮海平原冬小麦总面积为 16 226 667 hm²,其

参 考 文 献

[1] MENG Q, SUN Q, CHEN X, et al. Alternative cropping systems for sustainable water and nitrogen use in the North China Plain[J]. Agriculture, Ecosystems and Environment, 2012, 146(1):93-102.

[2] 何昭欣, 张森, 吴炳方, 等. Google Earth Engine 支持下的江苏省夏收作物遥感提取[J]. 地球信息科学学报, 2019, 21(5): 752-766.

HE Zhaoxin, ZHANG Miao, WU Bingfang, et al. Extraction of summer crop in Jiangsu based on Google Earth Engine[J]. Journal of Geo-information Science, 2019, 21(5): 752-766. (in Chinese)

[3] KOBAYASHI N, TANI H, WANG X, et al. Crop classification using spectral indices derived from Sentinel-2A imagery[J]. Journal of Information and Telecommunication, 2020, 4(1): 67-90.

[4] KUMAR L, MUTANGA O. Google Earth Engine applications since inception: usage, trends, and potential[J]. Remote Sensing, 2018, 10(10): 1509.

[5] MA T, LIU Q, SUN H. Application of multi-source remote sensing technology in land use classification[J]. Bulletin of Surveying and Mapping, 2018(8): 56-61.

[6] 杨慧宇, 王征强, 白建军, 等. 基于多特征提取与优选的冬小麦面积提取[J]. 陕西师范大学学报(自然科学版), 2020, 48(1): 40-49.

YANG Huiyu, WANG Zhengqiang, BAI Jianjun, et al. Winter wheat area extraction based on multi-feature extraction and feature selection[J]. Journal of Shannxi Normal University (Natural Science Edition), 2020, 48(1): 40-49. (in Chinese)

[7] 张莎, 张佳华, 白雲, 等. 基于 MODIS-EVI 及物候差异免閾值提取黄淮海平原冬小麦面积[J]. 农业工程学报, 2018, 34(11): 150-158.

ZHANG Sha, ZHANG Jiahua, BAI Yun, et al. Extracting winter wheat area in Huanghuaihai Plain using MODIS-EVI data and phenology difference avoiding threshold[J]. Transactions of the CSAE, 2018, 34(11):150-158. (in Chinese)

[8] 王利民, 刘佳, 姚保民, 等. 基于 GF-1 影像 NDVI 年度间相关分析的冬小麦面积变化监测[J]. 农业工程学报, 2018, 34(8): 184-191.

WANG Limin, LIU Jia, YAO Baomin, et al. Area change monitoring of winter wheat based on relationship analysis of GF-1 NDVI among different years[J]. Transactions of the CSAE, 2018, 34(8): 184-191. (in Chinese)

[9] 程千, 徐洪刚, 曹引波, 等. 基于无人机多时相植被指数的冬小麦产量估测[J]. 农业机械学报, 2021, 52(3):160-167.

CHENG Qian, XU Honggang, CAO Yinbo, et al. Grain yield prediction of winter wheat using multi-temporal UAV based on multispectral vegetation index[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(3):160-167. (in Chinese)

[10] 谭深, 吴炳方, 张鑫. 基于 Google Earth Engine 与多源遥感数据的海南水稻分类研究[J]. 地球信息科学学报, 2019, 21(6): 937-947.

TAN Shen, WU Bingfang, ZHANG Xin. Mapping paddy rice in the Hainan Province using both Google Earth Engine and remote sensing images[J]. Journal of Geo-information Science, 2019, 21(6): 937-947. (in Chinese)

[11] PHAN T, KUCH V, LEHNERT L. Land cover classification using Google Earth Engine and random forest classifier-the role of image composition[J]. Remote Sensing, 2020, 12(15): 2411.

[12] LIU P, HU G, CHEN Y, et al. High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine platform[J]. Remote Sensing of Environment, 2018, 209: 227-239.

[13] LUO C, QI B, LIU H, et al. Using time series Sentinel-1 images for object-oriented crop classification in Google Earth Engine[J]. Remote Sensing, 2021, 13(4): 561.

[14] 马战林, 刘昌华, 薛华柱, 等. GEE 环境下融合主被动遥感数据的冬小麦识别技术[J]. 农业机械学报, 2021, 52(9):

- 195 – 205.
- MA Zhanlin, LIU Changhua, XUE Huazhu, et al. Identification of winter wheat by integrating active and passive remote sensing data based on Google Earth Engine platform[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 195 – 205. (in Chinese)
- [15] 周珂, 柳乐, 张俨娜, 等. GEE 支持下的河南省冬小麦面积提取及长势监测[J]. 中国农业科学, 2021, 54(11): 2302 – 2318. ZHOU Ke, LIU Le, ZHANG Yanna, et al. Area extraction and growth monitoring of winter wheat in Henan Province supported by Google Earth Engine[J]. Scientia Agricultura Sinica, 2021, 54(11): 2302 – 2318. (in Chinese)
- [16] KUSSUL N, LAVRENIUK M, SKAKUN S, et al. Deep learning classification of land cover and crop types using remote sensing data[J]. IEEE Geoscience and Remote Sensing Letters, 2017, 99: 1 – 5.
- [17] 甄晓菊, 张雪红, 吴国明, 等. 基于 Sentinel – 2A NDVI 时间序列数据的冬小麦识别[J]. 江苏农业科学, 2019, 47(16): 239 – 245.
- ZHEN Xiaoju, ZHANG Xuehong, WU Guoming, et al. Winter wheat identification based on Sentinel – 2A NDVI time series data[J]. Jiangsu Agricultural Sciences, 2019, 47(16): 239 – 245. (in Chinese)
- [18] ARIAS M, CAMPO-BESCÓS M, ÁLVAREZ-MOZOS J. Crop classification based on temporal signatures of Sentinel – 1 observations over Navarre Province, Spain[J]. Remote Sensing, 2020, 12(2): 278.
- [19] SHELESTOV A, LAVRENIUK M, KUSSUL N, et al. Exploring Google Earth Engine platform for big data processing: classification of multi-temporal satellite imagery for crop mapping[J]. Frontiers in Earth Science, 2017, 5: 17.
- [20] 王学, 李秀彬, 谈明洪, 等. 华北平原 2001—2011 年冬小麦播种面积变化遥感监测[J]. 农业工程学报, 2015, 31(8): 190 – 199.
- WANG Xue, LI Xiubin, TAN Minghong, et al. Remote sensing monitoring of changes in winter wheat area in North China Plain from 2001 to 2011[J]. Transactions of the CSAE, 2015, 31(8): 190 – 199. (in Chinese)
- [21] REN S, GUO Bin, WU Xi, et al. Winter wheat planted area monitoring and yield modeling using MODIS data in the Huang – Huai – Hai Plain, China[J]. Computers and Electronics in Agriculture, 2021, 182: 106049.
- [22] INGLADA J, VINCENT A, ARIAS M, et al. Improved early crop type identification by joint use of high temporal resolution SAR and optical image time series[J]. Remote Sensing, 2016, 8(5): 362.
- [23] 周涛, 潘剑君, 韩涛, 等. 基于多时相合成孔径雷达与光学影像的冬小麦种植面积提取[J]. 农业工程学报, 2017, 33(10): 215 – 221.
- ZHOU Tao, PAN Jianjun, HAN Tao, et al. Planting area extraction of winter wheat based on multi-temporal SAR data and optical imagery[J]. Transactions of the CSAE, 2017, 33(10): 215 – 221. (in Chinese)
- [24] DONG C, ZGAO G, QIN Y, et al. Area extraction and spatiotemporal characteristics of winter wheat-summer maize in Shandong Province using NDVI time series[J]. PLoS ONE, 2019, 14(12): 226508.
- [25] 贺原惠子, 王长林, 贾慧聪, 等. 基于随机森林算法的冬小麦提取研究[J]. 遥感技术与应用, 2018, 33(6): 1132 – 1140.
- HE Yuanhuizi, WANG Changlin, JIA Huicong, et al. Research on extraction of winter wheat based on random forest[J]. Remote Sensing Technology and Application, 2018, 33(6): 1132 – 1140. (in Chinese)
- [26] KHAMDMOV R, SALIEV E, RAKHMANOV K. Classification of crops by multispectral satellite images of Sentinel – 2 based on the analysis of vegetation signatures[J]. Journal of Physics: Conference Series, 2020, 1441(1): 012143.
- [27] 化国强, 肖靖, 黄晓军, 等. 基于全极化 SAR 数据的玉米后向散射特征分析[J]. 江苏农业科学, 2011, 39(3): 562 – 565.
- HUA Guoqiang, XIAO Jing, HUANG Xiaojun, et al. Corn backscattering characteristics analysis based on full polarimetric SAR data[J]. Jiangsu Agricultural Sciences, 2011, 39(3): 562 – 565. (in Chinese)
- [28] 姜亚珍, 张瑜洁, 孙琛, 等. 基于 MODIS – EVI 黄淮海平原冬小麦种植面积分带提取[J]. 资源科学, 2015, 37(2): 417 – 424.
- JIANG Yazhen, ZHANG Yujie, SUN Chen, et al. MODIS – EVI based winter wheat planting information extraction of zoning on Huanghuaihai Plain[J]. Resources Science, 2015, 37(2): 417 – 424. (in Chinese)
- [29] 李旭青, 刘世盟, 李龙, 等. 基于 RF 算法优选多时相特征的冬小麦空间分布自动解译[J]. 农业机械学报, 2019, 50(6): 218 – 225.
- LI Xuqing, LIU Shimeng, LI Long, et al. Automatic interpretation of spatial distribution of winter wheat based on random forest algorithm to optimize multi-temporal features[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(6): 218 – 225. (in Chinese)
- [30] ZHANG X, LIU J, QIN Z, et al. Winter wheat identification by integrating spectral and temporal information derived from multiresolution remote sensing data[J]. Journal of Integrative Agriculture, 2019, 18(11): 2628 – 2643.
- [31] 罗桓, 李卫国, 景元书, 等. 基于 SVM 的县域冬小麦种植面积遥感提取[J]. 麦类作物学报, 2019, 39(4): 81 – 88.
- LUO Huan, LI Weiguo, JING Yuanshu, et al. Remote sensing extraction of winter wheat planting area based on SVM[J]. Journal of Triticeae Crops, 2019, 39(4): 81 – 88. (in Chinese)
- [32] 林雪, 彭道黎, 黄国胜, 等. 结合多尺度纹理特征的遥感影像面向对象分类[J]. 测绘工程, 2016, 25(7): 22 – 27.
- LIN Xue, PENG Daoli, HUANG Guosheng, et al. Object-oriented classification with multi-scale texture feature based on remote sensing image[J]. Engineering of Surveying and Mapping, 2016, 25(7): 22 – 27. (in Chinese)
- [33] ALEX S. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network[J]. Physica D-Nonlinear Phenomena, 2020, 404: 132306.
- [34] HUGO C, OSMAR A, OSMAR L, et al. Rice crop detection using LSTM, Bi – LSTM, and machine learning models from Sentinel – 1 time series[J]. Remote Sensing, 2020, 12(16): 2655.