

一、文献综述

1 研究背景

小麦是中国乃至世界的主要粮食作物之一^[8-10],及时准确的冬小麦制图对于粮食安全以及环境可持续性都非常重要。而对小麦进行田间调查费时,劳动密集,昂贵,调查数据的质量容易受到人为错误的影响。在这种情况下高效自动的卫星遥感制图就成为了重要替代,它为从区域到局部的高时空分辨率作物制图提供了新的机会。

在此类任务中,所使用的遥感卫星主要分为两类,雷达卫星(SAR 卫星)和高光谱卫星。雷达卫星具有穿透性强的特点,因此卫星数据不容易受天气的影响。而高光谱卫星能提供更多波段信息,但容易受天气的影响,通常只会选用云量<5%的数据。

获取了遥感数据后需要建立模型对冬小麦进行分类。有三种常见的分类方法:1、基于阈值的方法;2、基于机器学习的方法;3、基于深度学习的方法。其中基于深度学习的方法是当前较为主流的方法,也是我们研究的重点方向。相比于另外两种方法,深度学习方法不需要依赖于人工进行特征工程,可以自动化的方式实现特征的提取,同时可能提取到人工难以提取到的深层次特征。在这个研究课题中,卫星数据十分丰富,我们可以构建起时序卫星序列图像,可以将多源卫星数据进行融合,甚至可以再融入统计数据。面对巨量数据,机器学习方法往往会陷入欠拟合的困境,而拥有合理模型架构的深度学习模型则会提取更多有效特征,使模型更具泛化能力。所以在这个课题中深度学习方法有着良好的应用前景,而如何寻找合理的模型架构也成为了研究者们重点关注的目标。

2 国内外研究现状

2.1 研究方向及进展

2.1.1 基于阈值的方法

通常研究者会基于卫星数据以及作物特征建立统计变量，并运用经验知识或者通过实验设立相应阈值作为作物分类的标准。这种方法的优势在于有较强的解释性，分类计算的时间与难度较低。缺点则在于需要大量时间进行人工特征提取，阈值的选定对于经验依赖性较强。下面是对一些使用该方法的具体案例的分析。

Qiu等人^[1]使用MODIS的EVI2时序数据集构建了EVE和EVL指标，并指定了相应的判定条件来对中国的华北平原与黄淮平原的冬小麦种植区进行制图。

类似的，Zhang等人^[5]设计了一个概念模型，将卫星图像的每个像素分类，其中只有三类像素含有冬小麦，即纯冬小麦像素（PA）、冬小麦与其他植被的混合像素（MA）和冬小麦和其他作物的混合像素（MB）。他们提出了三种不同的统计方法来鉴别这三种像素：1、利用Kullback-Leibler(KL)散度分析未知像素与纯冬小麦样本在NDVI时间变化特征上的相似性。确定了进一步的PA。(2)对于MA，首先使用提出的反向解混方法提取耕地成分的时间变化信息，然后如前所述从耕地成分中识别冬小麦。(3)对于只出现在PA边界上的MB，通过扩展PA创建一个掩码，并利用时间差异来识别掩码下的冬小麦。

Qu等人^[21]利用冬小麦四个关键生长阶段的多时相Landsat归一化植被指数(NDVI)数据，提出了一种新的光谱指数，称为冬小麦指数(WWI)，以突出显示和绘制冬小麦图。同时，他们采用蒙特卡洛交叉验证法推算出了区分冬小麦的WWI指数阈值。

2.1.2 基于机器学习的方法

常用的机器学习方法有随机森林、支持向量机、决策树等。随机森林分类器是一种集成分类器，它使用随机选择的训练样本和变量子集生成多个决策树。由于其分类的准确性，该分类器在遥感中应用广泛^[11]。在实际训练中，由于其

使用了 Bootstrap 抽样的方法选择特征，随机森林分类器往往不容易过拟合^[12]。而且随机森林可以对特征重要性进行排序，这是提供模型解释性的重要一步。SVM 在遥感领域特别有吸引力，因为它们即使在训练样本有限的情况下也能很好地泛化^[13]。

其中，Liu 等人^[3]利用 Landsat-8 卫星的数据构建了一个随机森林分类器用于中国华北平原冬小麦作物制图。同时，作者又构建了最大似然分类器、人工神经网络、支持向量机与随机森林分类器进行对比，发现随机森林分类器的总体精度高于最大似然分类器和人工神经网络，略低于支持向量机。但是随机森林相比于支持向量机训练速度快，调参难度低，作者认为它在该实验中是性能更好的分类器。

Zhou 等人^[4]也使用随机森林和支持向量机建立起模型对冬小麦进行制图。与 Liu 的团队不同的是，他们选择了多种数据来源：SAR 卫星 Sentinel-1 数据和高光谱卫星 Landsat-8 数据，并且发现将两种数据结合作为模型输入能够提高模型准确度。

在 Tian 等人^[22]的研究中，他们提出了一种合成多时相 NDVI 的方法，以便使用 Landsat-7 和 -8 以及 Sentinel-2 光学图像绘制冬季作物种植区。该算法根据冬季作物的两个时间窗口——低 NDVI 值和高 NDVI 值，将多时态 NDVI 合成为三个关键值，然后利用决策树分类方法在像素级别进行冬季作物分类。

2.1.3 基于深度学习的方法

深度学习是近年来较为流行的方法。其特点在于用自动化的深度网络学习代替了人工特征工程，特征提取的过程摆脱了对作物种植经验的依赖。同时其灵活的网络架构能够更好地利用各种统计数据集，提取深层次特征，建立起泛化能力更强的模型。下面是一些应用深度学习方法的案例。

Zhong 等人^[2]提出了一个基于卷积神经网络的模型架构来对美国堪萨斯州和德克萨斯州北部的冬小麦进行制图。更重要的是，他们在模型中融入了统计数据，为我们研究提供了新的思路

由于深度学习网络具有可迁移性，一些在应用于其他作物的网络仍然有着借鉴意义。Wei 等人^[6]在研究水稻制图问题上，提出了 adaptive U-net 结构，并使用 2017 年与 2018 年的 Sentinel-1 卫星数据进行训练，在 2019 年的数据上进

行测试来证明网络的泛化能力。同时，作者也通过实验来可视化网络在训练的过程中所提取的特征，使得网络更具有解释性。

Xu 等人^[7]提出了一种名为 DeepCropMapping (DCM) 的深度学习方法基于具有注意力机制的长短期记忆结构，通过整合多时相和多光谱遥感数据，用进行大规模动态玉米和大豆制图。他们还将 DeepCropMapping 的网络与 Transformer 网络、随机森林分类器、多层感知机网络进行对比，结果表明它拥有与 Transformer 相当的性能，并在本地测试中优于随机森林分类器和多层感知机网络。DCM 模型在空间转移分析中以 95% 的置信区间显著优于其他三个模型。结果证明了 DCM 模型从时间序列中学习可泛化特征的能力。

Zhong 等人^[19]使用 Landsat 增强植被指数 (EVI) 时间序列对夏季作物进行分类。他们设计了两种类型的深度学习模型：一种基于长短期记忆 (LSTM)，另一种基于一维卷积 (Conv1D) 层。还测试了三个广泛使用的分类器以进行比较，包括 XGBoost、随机森林和支持向量机。该研究表明，基于 Conv1D 的深度学习框架在多时态分类任务中提供了一种有效且高效的时间序列表示方法。同时，他们也对模型进行了解释性分析。他们通过可视化不同层上的激活来检查基于 Conv1D 的模型的行为。该模型通过以分层方式检查各种尺度的形状来使用 EVI 时间序列。结果表明模型的较底部 Conv1D 层捕捉小规模的时间变化，而上层则专注于整体季节性模式。

2.1.4 深度学习时空模型

在深度学习领域，时空模型是最贴切于本次课题要求的，因为它们可以同时提取空间和时间的信息，从而获得更深层次的时空信息。同时这些模型都具有良好的迁移性，可能在本课题中也具有良好的性能。近年来，深度学习领域研究者们也发现了一批表现优异的时空模型。

首先是基于 RNN 的时空模型。2015 年 Shi 等人^[23]提出的 Conv-LSTM 结构将传统 LSTM 中加入卷积操作，开创了时空模型的先河。他们也从 seq2seq^[24]的 Encoder-Decoder 结构中获得启发，提出了 Encoding-Forecasting 结构来解决基于雷达图的气象预测问题。2017 年 Shi 团队^[25]对原始 GRU 进行改良，提出了 TrajGRU 进一步提升了性能。在 Conv-LSTM 之后，还有很多优秀的时空模型被提出。如 Wang 团队提出的 PredRNN^[26]和 PredRNN++^[27]，它们都是基于

Conv-LSTM 的模型。他们对于 LSTM 中的记忆结构进行了进一步的优化，实验表明该模型在多个数据集上达到了 SOTA 水平。

其次是基于 CNN 的时空模型。2015 年 Tran 等^[28]提出了基于 3d 卷积的 C3D 模型能够很好地提取时空信息，在视频预测领域有着优异的表现。之后在 2017 年，I3D^[29]作为 C3D 的优化被提出，它解决了 3D 卷积参数过多计算复杂度高的问题。2019 年，Wang^[30]将 3D 卷积与 LSTM 结合起来，提出了 E3D-LSTM，在视频预测中也达到了 SOTA 水平。同时作者也指出该模型可以在除了视频预测外的其他时空问题中发挥作用。

2.2 存在问题

遥感数据中存在样本不平衡的问题。往往遥感图像中是小麦的像素比不是小麦的像素要少很多，会影响模型判别小麦的准确率。针对这个问题，研究者也选择了几种解决方法：1、尽量选择小麦比例较大的卫星数据。2、改变损失函数，增加小麦像素的权重。但是这两种解决方法都不能普适性地解决样本不平衡的问题。

卫星图像数据会有缺失，天气（云量）也会影响图像数据的可用性，这对于利用时间序列图像构建模型的方法有较大影响。为了提高模型的泛化能力，将多源卫星数据融合是一种可行的^[14]途径，这也对卫星图像的完整性提出了更高要求。

现有许多研究是基于 MODIS 卫星数据，分辨率仅为 250m，为了更精确的冬小麦制图工作，就对于卫星分辨率提出了更高要求。所以现有许多研究是基于 10m 分辨率的 Sentinel 卫星和 Landsat 卫星。

部分区域缺少有地块标签的数据进行模型训练。由于该研究主流模型是有监督模型，需要有标签数据进行训练，在缺少地块标签数据的区域进行冬小麦制图时则需要从其他区域训练模型并进行迁移，精确度并不能得到保证。

基于深度学习的作物制图的主要限制在于模型的可解释性。由于分层特征转换能力，深度神经网络通常优于人工规则设计的浅层分类器^[16]。然而，深度学习模型通常被认为是“黑匣子”，具有不可见的特征学习过程，可产生更高级别的抽象特征^[17]。解释深度学习方法的特征学习管道可能有利于我们解释复

杂的判别模式并确保模型的可靠性^[18]。所以还需要更多面向模型解释性的研究。

3 研究展望

在数据源方面，现在较多的研究使用的是中分辨率卫星数据 MODIS，为了进行更准确的小麦制图，使用更高分辨率的卫星数据是未来的趋势。Landsat 卫星、Sentinel 卫星都提供免费的 10m 分辨率的卫星数据。除此之外，多源数据融合也是未来研究的方向之一，随着模型性能越发强大，需要有更丰富的数据来源提供更丰富的特征信息。

在模型方面，随着深度学习的发展，可用的模型越来越多，也就有了更多可能的架构，不同模型架构的探究和比较也是未来研究的重中之重。

在研究地点方面，现在大部分使用有监督学习算法的研究者都倾向于选择美国作为数据来源。因为美国有 CDL 标签数据以供分类。作为全球最大的冬小麦生产国之一，中国的冬小麦产量占全球的 18% 以上。然而，中国却没有大空间范围内具有高空间分辨率的小麦分布图^[21]。所以未来的研究也会很大部分集中于如何为这些原本没有小麦分布图的国家 and 地区制作小麦分布图。

4 参考文献

- [1] Qiu B , Luo Y , Tang Z , et al. Winter wheat mapping combining variations before and after estimated heading dates[J]. *Isprs Journal of Photogrammetry & Remote Sensing*, 2017, 123(JAN.):35-46.
- [2] Zhong L , Hu L , Zhou H , et al. Deep learning based winter wheat mapping using statistical data as ground references in Kansas and northern Texas, US[J]. *Remote Sensing of Environment*, 2019, 233:111411-.
- [3] Liu J , Feng Q , Gong J , et al. Winter wheat mapping using a random forest classifier combined with multi-temporal and multi-sensor data[J]. *International Journal of Digital Earth*, 2017:1-20.
- [4] Zhou T , Pan J , Zhang P , et al. Mapping Winter Wheat with Multi-Temporal SAR and Optical Images in an Urban Agricultural Region[J]. *Sensors*, 2017, 17(6):1210.
- [5] Xiwang, Zhang, Fang, et al. Identification and mapping of winter wheat by integrating temporal change information and Kullback–Leibler divergence - ScienceDirect[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2019, 76:26-39.
- [6] Wei, Pengliang, et al. "Large-scale rice mapping under different years based on time-series Sentinel-1 images using deep semantic segmentation model." *ISPRS Journal of Photogrammetry and Remote Sensing* 174 (2021): 198-214.
- [7] Xu J , Zhu Y , Zhong R , et al. DeepCropMapping: A multi-temporal deep learning approach with improved spatial generalizability for dynamic corn and soybean mapping[J]. *Remote Sensing of Environment*, 2020, 247:111946.
- [8] Eisavi V , Homayouni S , Yazdi A M , et al. Land cover mapping based on random forest classification of multitemporal spectral and thermal images[J]. *Environmental Monitoring&Assessment*, 2015.
- [9] Upadhyay P , Ghosh S K , Kumar A . Temporal MODIS data for identification of wheat crop using noise clustering soft classification approach[J]. *Geocarto International*, 2015.
- [10] Ren J , Chen Z , Zhou Q , et al. Regional yield estimation for winter wheat with MODIS-NDVI data in Shandong, China[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2008, 10(4):403-413.
- [11] Belgiu, Dragut. Random forest in remote sensing: A review of applications and future directions[J]. *ISPRS J PHOTOGRAMM*, 2016, 2016,114(-):24-31.
- [12] Breiman L. Random Forests[J]. *Machine Learning*, 2001, 45(1):5-32.
- [13] Mountrakis G , Im J , Ogole C . Support vector machines in remote sensing: A review[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2011, 66(3):247-259.
- [14] Orynbaikyzy A , Gessner U , Conrad C . Crop type classification using a combination of optical and radar remote sensing data: a review[J]. *International Journal of Remote Sensing*, 2019, 40(17-18):1-43.
- [15] Xu, Jinfan, et al. "Towards interpreting multi-temporal deep learning models in crop mapping." *Remote Sensing of Environment* 264 (2021): 112599.
- [16] LeCun et al., 2015 Deep learning Y. LeCun, Y. Bengio and G. Hinton *Nature*, 521 (2015), pp. 436-444
- [17] Heaven, Douglas. "Why deep-learning AIs are so easy to fool." (2019): 163-166.
- [18] Xu, Jinfan, et al. "Towards interpreting multi-temporal deep learning models in crop mapping." *Remote Sensing of Environment* 264 (2021): 112599.
- [19] Zhong, Liheng, Lina Hu, and Hang Zhou. "Deep learning based multi-temporal crop classification." *Remote sensing of environment* 221 (2019): 430-443.
- [20] Qu, Chang, Peijun Li, and Chengming Zhang. "A spectral index for winter wheat mapping using multi-temporal Landsat NDVI data of key growth stages." *ISPRS Journal of Photogrammetry and Remote Sensing* 175 (2021): 431-447.
- [21] Dong, Jie, et al. "Early-season mapping of winter wheat in China based on Landsat and Sentinel images." *Earth System Science Data* 12.4 (2020): 3081-3095.

-
- [22] Tian, Haifeng, et al. "Mapping winter crops in China with multi-source satellite imagery and phenology-based algorithm." *Remote sensing* 11.7 (2019): 820.
 - [23] SHI Xingjian, Zhouong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. Convolutional lstm network: A machine learning approach for precipitation nowcasting. In *Advances in neural information processing systems*, pages 802 – 810, 2015.
 - [24] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. In *Advances in neural information processing systems*, pages 3104 – 3112, 2014.
 - [25] Xingjian Shi, Zhihan Gao, Leonard Lausen, Hao Wang, Dit-Yan Yeung, Wai-kin Wong, and Wang-chun WOO. Deep learning for precipitation nowcasting: A benchmark and a new model. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.
 - [26] Wang, Yunbo, et al. "Predrnn: Recurrent neural networks for predictive learning using spatiotemporal lstms." *Proceedings of the 31st International Conference on Neural Information Processing Systems*. 2017.
 - [27] Wang, Yunbo, et al. "Predrnn++: Towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning." *International Conference on Machine Learning*. PMLR, 2018.
 - [28] Tran, Du, et al. "Learning spatiotemporal features with 3d convolutional networks." *Proceedings of the IEEE international conference on computer vision*. 2015.
 - [29] Carreira, Joao, and Andrew Zisserman. "Quo vadis, action recognition? a new model and the kinetics dataset." *proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017.
 - [30] Wang, Yunbo, et al. "Eidetic 3d lstm: A model for video prediction and beyond." *International conference on learning representations*. 2018.