

基于遥感多参数和 CNN-Transformer 的冬小麦单产估测

王鹏新^{1,2} 杜江莉^{1,2} 张悦^{1,2} 刘峻明³ 李红梅⁴ 王春梅⁵

(1. 中国农业大学信息与电气工程学院, 北京 100083;

2. 农业农村部农机作业监测与大数据应用重点实验室, 北京 100083;

3. 中国农业大学土地科学与技术学院, 北京 100193;

4. 陕西省气象局, 西安 710014; 5. 中国科学院空天信息创新研究院, 北京 100094)

摘要: 为了提高冬小麦单产估测精度,改善估产模型存在的高产低估和低产高估等现象,以陕西省关中平原为研究区域,选取旬尺度条件植被温度指数(VTCI)、叶面积指数(LAI)和光合有效辐射吸收比率(FPAR)为遥感特征参数,结合卷积神经网络(CNN)局部特征提取能力和基于自注意力机制的Transformer网络的全局信息提取能力,构建CNN-Transformer深度学习模型,用于估测关中平原冬小麦产量。与Transformer模型(R^2 为0.64, RMSE为465.40 kg/hm², MAPE为8.04%)相比, CNN-Transformer模型具有更高的冬小麦单产估测精度(R^2 为0.70, RMSE为420.39 kg/hm², MAPE为7.65%),能够从遥感多参数中提取更多与产量相关的信息,且对于Transformer模型存在的高产低估和低产高估现象均有所改善。基于5折交叉验证法和留一法进一步验证了CNN-Transformer模型的鲁棒性和泛化能力。此外,基于CNN-Transformer模型捕获冬小麦生长过程的累积效应,分析逐步累积旬尺度输入参数对产量估测的影响,评估模型对于冬小麦不同生长阶段的累积过程的表征能力。结果表明,模型能有效捕捉冬小麦生长的关键时期,3月下旬至5月上旬是冬小麦生长的关键时期。

关键词: 冬小麦; 作物估产; 遥感多参数; 卷积神经网络; Transformer模型

中图分类号: TP79

文献标识码: A

文章编号: 1000-1298(2024)03-0173-10

OSID:



Yield Estimation of Winter Wheat Based on Multiple Remotely Sensed Parameters and CNN-Transformer

WANG Pengxin^{1,2} DU Jiangli^{1,2} ZHANG Yue^{1,2} LIU Junming³ LI Hongmei⁴ WANG Chunmei⁵

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

2. Key Laboratory of Agricultural Machinery Monitoring and Big Data Applications,

Ministry of Agriculture and Rural Affairs, Beijing 100083, China

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

4. Shaanxi Provincial Meteorological Bureau, Xi'an 710014, China

5. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China)

Abstract: In order to improve the accuracy of winter wheat yield estimation and the phenomena of underestimation of high yield and overestimation of low yield that exist in yield estimation models, the Guanzhong Plain in Shaanxi Province, China was taken as the study area, and the vegetation temperature condition index (VTCI), leaf area index (LAI) and fraction of photosynthetically active radiation (FPAR) at the ten-day interval were selected as remotely sensed parameters, and a deep learning model was proposed for estimating winter wheat yield by combining the local feature extraction capability of convolutional neural network (CNN) and the global information extraction capability of Transformer network based on the mechanism of self-attention. Compared with the Transformer model (R^2 was 0.64, RMSE was 465.40 kg/hm², MAPE was 8.04%), the CNN-Transformer model had higher accuracy in estimating winter wheat yield (R^2 was 0.70, RMSE was 420.39 kg/hm², MAPE was 7.65%), which can extract more yield-related information from the multiple remotely sensed parameters, and improved the underestimation of high yield and overestimation of low yield which existed in the Transformer model.

收稿日期: 2023-08-06 修回日期: 2023-09-13

基金项目: 国家自然科学基金项目(42171332)

作者简介: 王鹏新(1965—),男,教授,博士,主要从事定量遥感及其在农业中的应用研究, E-mail: wangpx@cau.edu.cn

The robustness and generalization ability of the CNN – Transformer model were further validated based on the five-fold cross-validation method and the leave-one-out method. In addition, based on the CNN – Transformer model, the cumulative effect of the winter wheat growth process was revealed, the impact of gradually accumulating ten-day scale input information on yield estimation was analyzed, and the ability of the model to characterize the accumulation process of winter wheat at different growth stages was assessed. The results showed that the model can effectively capture the critical period of winter wheat growth, which was from late March to early May.

Key words: winter wheat; yield estimation; multiple remotely sensed parameters; convolutional neural network; Transformer model

0 引言

冬小麦作为我国主要的粮食作物之一,其生长状态监测和产量信息的准确把握对于国家粮食安全政策制定具有重要价值,同时,也对市场价格调控以及对外粮食贸易等产生重要影响^[1-2]。相对于传统的作物估产方法,遥感技术具有非接触、大范围监测、长时间序列观测等优势,能够提供更准确、高效的作物长势监测和产量估测信息,为农业管理和粮食安全提供有力支持^[3]。因此,基于遥感的作物估产方法已经成为获取大区域尺度粮食产量信息的重要方式。

作物长势与单产密切相关,涉及作物的生长发育和多种环境因素的影响,如光、温、水、土、肥等因子,而应用遥感技术表征的与作物长势密切相关的监测指标是这些因子综合作用的结果^[4]。因此,目前在区域尺度的作物长势监测和产量估测研究中,多种遥感参数结合的方法被广泛应用。基于遥感数据进行作物长势监测的参数,如叶面积指数(Leaf area index, LAI)、归一化植被指数(Normalized difference vegetation index, NDVI)和光合有效辐射吸收比率(Fraction of photosynthetically active radiation, FPAR)等对提高作物长势监测的准确性发挥了重要作用。且随着高时间分辨率卫星数据产品(如MODIS数据)的积累发展,针对作物时序特征的研究不断增多。其中, LAI和 FPAR可表征在长时序条件下不同时间段内作物的生长状态和光合作用能力,是表征农情和进行作物产量估测的重要参数。条件植被温度指数(Vegetation temperature condition index, VTCI)是王鹏新等^[5]基于 NDVI和地表温度(Land surface temperature, LST)的散点图呈三角形区域分布的基础上提出的。VTCI既考虑了某一研究区域内 NDVI的变化,又考虑了在 NDVI值相同的条件下 LST的变化,能够定量化地表征作物在生长过程中受到的水分胁迫,已成功用于陕西省关中平原干旱监测及作物单产估测、预测等。因此,为进一步实现准确的作物长势监测和产量估测,可以选择

VTCI、LAI和 FPAR作为遥感参数,综合利用其表征的冬小麦生长状况,开展基于多参数的作物长势信息提取及单产估测研究。

目前在基于遥感的估产方法中,主要包括基于将遥感观测数据与作物生长模型相结合的数据同化方法^[6]和基于统计回归的经验估产方法^[7]。与数据同化方法相比,基于统计回归的经验估产方法不依赖于复杂的模型和数据同化过程,降低了计算成本和数据处理的复杂性,被广泛应用于区域尺度的长势监测和产量估测研究。同时,考虑到作物长势和产量之间的关系通常表现为非线性,而基于统计回归经验方法建立的非线性回归模型能够反映植被生长状态遥感参数与农作物产量之间的关系,进而实现更为精确的产量估测研究。因此,使用非线性回归模型方法进行作物长势监测和单产估测更受关注,如支持向量回归(Support vector regression, SVR)、随机森林(Random forest, RF)等^[8-9]。然而,作物生长是一个时序相关的过程,传统的机器学习方法,如 SVR和 RF,无法直接捕捉变量之间的时序关系,也无法有效地表征作物生物物质在不同时间点与环境变量之间的累积过程。而深度学习作为一种先进的机器学习方法,通过多层次的非线性变换和学习来提取数据的高级特征,能够有效地捕捉变量之间的时序关系,进而更好地处理时序相关的数据,在利用遥感数据进行作物长势监测和产量估测方面展现出巨大的潜力^[10]。卷积神经网络(Convolutional neural network, CNN)是一种深层前馈神经网络,通过局部感知野的方式能够有效地捕捉输入数据的局部特征,在作物单产估测中得到广泛关注^[11]。循环神经网络(Recurrent neural network, RNN)和长短期记忆(Long short-term memory, LSTM)神经网络对于提取时间序列数据的特征具有相对优势。目前,国内外学者已经开展了许多基于 LSTM和遥感数据的作物长势监测和产量估测研究^[12-13]。但随着时间步长的迭代, RNN、LSTM存在梯度消失爆炸、顺序时间处理以及长期记忆能力不足的缺点,这些问题影响了处理长序列

数据时的有效性^[14-15]。Transformer^[16] 相较于 RNN、LSTM 在处理序列数据时具有更好的长期记忆能力,其通过自注意力机制(Self-attention)和位置编码,能够捕捉序列中不同位置之间的长期依赖关系,充分利用计算资源,实现对整个序列的并行计算,已广泛应用于自然语言处理、农业等领域^[17-18]。ONOUFRIOU 等^[19]提出了一种基于 Transformer 的多时间线时间序列提取方法,结果表明,该架构在预测草莓产量方面优于 CNN、RNN 和 LSTM 等神经网络。JIN 等^[20]提出一种基于 Transformer 的 T2T (Transformer to high temperature) 模型,结果表明, T2T 模型相比于其他模型对次季节高温概率预测更可靠。CNN 具有强大的局部特征提取能力,而 Transformer 能够捕获长时间依赖性,因此,一些研究将 CNN 和 Transformer 结合起来进行应用。LIANG 等^[21]提出了一种融合 CNN 和 Transformer 的混合 CNN - Transformer 网络用于遥感变化检测,结果表明,混合模型表现出优异的性能。此外,由于训练样本不均衡等原因,估产模型在学习过程中会更多地关注样本数量较多的区间,从而影响对较少样本区间的估测效果,导致简单的估产模型容易造成低产高估或高产低估的问题^[13]。而基于 CNN 和 Transformer 的深度学习方法,具有综合提取局部和全局信息的优势进而在区域产量估测方面显示出潜力,但尚未得到充分挖掘,需进一步探索。

本文以陕西省关中平原为研究区域,通过选取与冬小麦长势密切相关的旬尺度时间序列 VTCI、LAI 和 FPAR 作为输入参数,构建一种结合 CNN 和 Transformer 的具有时间序列记忆优势的端到端的深度学习估产模型,实现关中平原冬小麦单产的估测。此外,为进一步提高估产模型的可解释性,本研究尝试利用时间序列参数和深度学习方法相结合的方式,表征冬小麦生长过程的累积效应,分析逐步累积旬尺度输入参数对产量估测的影响。

1 研究区与数据

1.1 研究区域概况

研究区域关中平原位于陕西省中部的渭河流域,西起宝鸡市,东至潼关县,南接秦岭,北至黄土高原,西窄东宽,东西长约 400 km,平均海拔约 500 m,地理范围为 106°22' ~ 110°24'E, 33°57' ~ 35°39'N (图 1)。关中平原为温带大陆性季风气候,年平均温度为 6 ~ 13℃,降水集中于 6—9 月,年平均降水量为 550 ~ 700 mm。该地区地势平坦、土壤肥沃,十分适宜耕作^[22]。种植模式主要为在灌溉区域夏玉米和冬小麦轮作,在旱作区域种植冬小麦。其中冬小

麦一般在 10 月上旬播种,次年 4 月下旬进入抽穗期,6 月上旬收获。根据关中平原冬小麦的实际生长状况,将冬小麦的主要生育期划分为返青期(3 月上旬—3 月中旬)、拔节期(3 月下旬—4 月中旬)、抽穗—灌浆期(4 月下旬—5 月上旬)和乳熟期(5 月中旬—5 月下旬)。本研究聚焦于冬小麦的 4 个生育时期,并采用旬时间尺度进行分析研究。

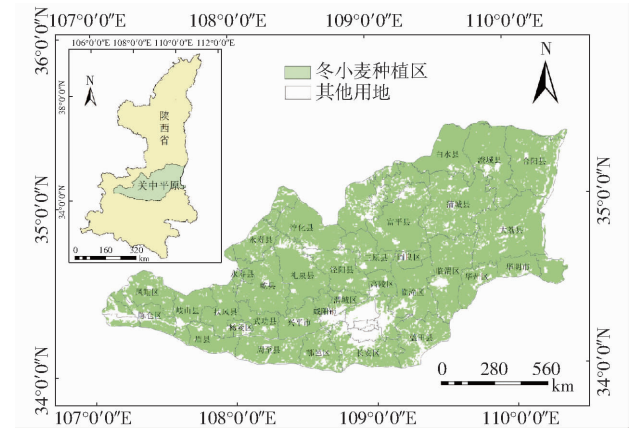


图 1 研究区位置及冬小麦种植区

Fig.1 Location of study area and winter wheat planting areas

1.2 数据来源与预处理

1.2.1 数据来源

选取的遥感数据来源于美国国家航空航天局(National Aeronautics and Space Administration, NASA)地球观测系统数据和信息系统网站(<https://ladsweb.modaps.eosdis.nasa.gov/search/>),包括覆盖研究区域的 MODIS 的地表温度产品(MYD11A1)、地表反射率产品(MYD09GA)、叶面积指数及光合有效辐射吸收比率产品(MCD15A3H)以及地表覆盖类型产品(MCD12Q1)。所采用的 2011—2021 年关中平原县域尺度冬小麦单产数据,来源于陕西省统计局发布的统计年鉴(<http://tjj.shaanxi.gov.cn/tjsj/ndsj/tjnj/>)。针对关中平原地区的 24 个县(区),选用旬尺度的 VTCI、LAI 和 FPAR 遥感参数以及冬小麦单产数据进行相关研究。

1.2.2 冬小麦种植区域提取

根据国家统计局(<http://www.stats.gov.cn/>)提供的冬小麦种植面积统计数据,近 10 年来我国冬小麦种植面积一直保持相对稳定,波动较小。关中平原冬小麦种植面积占关中平原总种植面积的 90% 以上,文献[13]研究表明,耕地数据可以作为区域产量估测研究的掩膜。因此本研究采用 MODIS 土地覆盖类型产品(MCD12Q1)提取冬小麦种植区,该产品主要通过监督决策树分类方法提取土地覆盖类型,提供 6 种不同的土地覆盖分类方案,本研究采用国际地圈生物圈计划(IGBP)的分类方

案,通过 MRT(MODIS reprojection tools)进行预处理并将提取的冬小麦种植区与关中平原行政边界矢量图叠加,得到关中平原县(区)的冬小麦种植区分布图(图 1)。

1.2.3 VTCI 时间序列数据生成

选取研究区 2011—2021 年每年 3—5 月空间分辨率为 1 000 m,时间分辨率为 1 d 的 MYD11A1 和 MYD09GA 产品,采用 MRT 对 MODIS 数据进行存储格式转换、投影转换、重采样、影像拼接和裁剪等预处理,得到每年关中平原日 NDVI 和日 LST。运用最大值合成法(Maximum value compositing, MVC),分别生成旬尺度的 NDVI 和 LST 最大值合成产品。基于多年某一句的 NDVI 和 LST 最大值合成产品,采用最大值合成法分别生成多年的旬 NDVI 和 LST 的最大值合成产品。基于每年 3—5 月旬尺度的 LST 最大值合成产品,逐像素取最小值,生成多年旬尺度 LST 最大-最小值合成产品。通过多年旬尺度 NDVI 和 LST 最大值合成产品确定研究区域 VTCI 的热边界,通过多年旬尺度 NDVI 最大值合成产品和多年旬尺度 LST 最大-最小值合成产品确定 VTCI 的冷边界,以此计算旬时间尺度的 VTCI^[23],计算式为

$$VTCI = \frac{L_{\max}(N_i) - L(N_i)}{L_{\max}(N_i) - L_{\min}(N_i)} \tag{1}$$

其中 $L_{\max}(N_i) = a + bN_i \tag{2}$

$L_{\min}(N_i) = a' + b'N_i \tag{3}$

式中 VTCI——条件植被温度指数

N_i ——某一像素的 NDVI 值

$L(N_i)$ ——研究区域内某一像素的 NDVI 值为 N_i 时的地表温度

$L_{\max}(N_i)$ 、 $L_{\min}(N_i)$ ——研究区域内当 NDVI 等于 N_i 时所有像素地表温度最大值和最小值,称热边界和冷边界

a 、 b 、 a' 、 b' ——待定系数,由研究区域 NDVI 和 LST 的散点图近似获得

基于生成的旬时间尺度的 VTCI 影像,叠加关中平原县(区)的冬小麦种植区分布图,取各县域内冬小麦种植区所包含像素的 VTCI 平均值作为该县该年该旬的 VTCI 值。

1.2.4 LAI 和 FPAR 时间序列数据生成

选取 2011—2021 年每年 3—5 月的 MODIS MCD15A3H 产品,用于获取 LAI 和 FPAR。该产品空间分辨率为 500 m,时间分辨率为 4 d。由于受云、雨天气状况等因素的影响,导致某些时间段 LAI 和

FPAR 出现不连续以及波动的特征,难以反映植被 LAI 和 FPAR 的真实变化情况。其时间序列也不能体现相应的季节性变化规律,因此应用上包络线 Savitzky Golay 滤波对 LAI 和 FPAR 数据进行平滑去噪处理。经过滤波处理后的 LAI 和 FPAR 曲线变化平稳,与冬小麦的实际生长规律相符合^[24]。

为使 LAI 与 VTCI 和 FPAR 具有相同的取值范围,对滤波后的 LAI 进行归一化处理。基于滤波后的 4 d 尺度的 LAI、FPAR 数据,通过逐像素取各旬所包含的多时相 LAI 和 FPAR 的最大值作为各旬的 LAI 和 FPAR。根据关中平原行政区划图和冬小麦种植区分布图,分别计算各县(区)内所包含所有像素的 LAI 和 FPAR 的平均值,并将其作为该县该年该旬的 LAI 和 FPAR。由于 VTCI 与 LAI 和 FPAR 的空间分辨率不同,故采用最邻近法将 1 000 m 分辨率的 VTCI 重采样为 500 m 空间分辨率。此外, VTCI、LAI 和 FPAR 在 0~1 之间,而冬小麦单产在 2 000~7 000 kg/hm²之间,因此,为了使各参数和单产数据处于同一数量级,将冬小麦单产数据归一化。

2 研究方法

2.1 Transformer 模型

Transformer 模型目前在机器翻译、文本生成任务中应用较为广泛,该模型整体结构包含输入、编码器(Encoder)、解码器(Decoder)和输出 4 部分(图 2),编码器和解码器有 M 个相同的块。输入部分包括输入向量和目标向量及位置编码,在机器翻译中,输入向量、目标向量分别代表源文本和目标文本中词汇的向量表示。

在 Transformer 的编码器结构中,由于自注意力机制是无状态的,不能直接捕捉到输入数据之间的顺序信息,因此需要加入位置编码,提供输入数据在序列中相对位置的信息,位置编码可表示为^[16]

$$PE(p, 2i) = \sin\left(\frac{p}{10\,000^{\frac{2i}{d}}}\right) \tag{4}$$

$$PE(p, 2i + 1) = \cos\left(\frac{p}{10\,000^{\frac{2i}{d}}}\right) \tag{5}$$

式中 p ——特征的位置

i ——特征的维度序号

d ——特征的维度

$PE(\cdot)$ ——位置编码

编码器包括多头注意力机制(Multi-head attention)、残差连接和层归一化(Residual connection and layer normalization)与前馈神经网络(Feed-forward network, FFN)。多头注意力机制是由多个自注意力机制组成,自注意力机制的输入为图 2 中

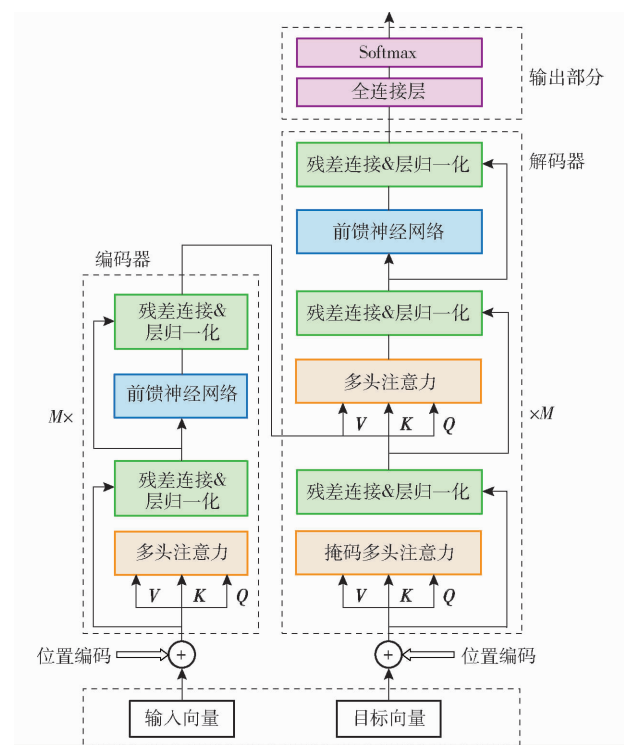


图 2 Transformer 模型结构图

Fig. 2 Structure diagram of Transformer model

的输入向量或者上一个编码器中相同层的输出。通过得到对应的查询矩阵 Q 、键矩阵 K 和值矩阵 V ，根据查询矩阵和键矩阵的相关性对值矩阵进行加权求和得到自注意力的输出结果^[16]。具体计算公式为

$$\text{Attention}(\mathbf{V}, \mathbf{K}, \mathbf{Q}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \quad (6)$$

式中 softmax ——自注意力权重的归一化函数

d_k —— Q 和 K 矩阵的列数, 即向量维度

$\text{Attention}(\cdot)$ ——自注意力机制计算函数

多头注意力机制通过将每个自注意力机制的输出进行拼接得到输出, 具体公式为

$$\text{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Concat}(h_1, h_2, \dots, h_j, \dots, h_k)\mathbf{W}_o \quad (7)$$

其中 $h_j = \text{Attention}(\mathbf{W}_j^Q \mathbf{Q}, \mathbf{W}_j^K \mathbf{K}, \mathbf{W}_j^V \mathbf{V})$

$$(j = 1, 2, \dots, k) \quad (8)$$

式中 $\text{MultiHead}(\cdot)$ ——多头注意力机制计算函数

k ——注意力头部数量

h_j ——自注意力机制输出

$\text{Concat}(\cdot)$ ——将多个单头注意力结果进行连接的函数

$\mathbf{W}_j^Q, \mathbf{W}_j^K, \mathbf{W}_j^V$ —— Q, K, V 第 j 头的线性变换权重矩阵

$\mathbf{W}_o$ ——多头注意力矩阵连接后的线性变换权重矩阵

残差连接和层归一化由两部分组成, 可表示为

$$\mathbf{Z} = \text{LayerNorm}(\mathbf{X} + \text{MultiHead}(\mathbf{X})) \quad (9)$$

$$\mathbf{I} = \text{LayerNorm}(\mathbf{Z} + \text{FeedForward}(\mathbf{Z})) \quad (10)$$

式中 $\text{LayerNorm}(\cdot)$ ——层归一化函数

$\text{FeedForward}(\cdot)$ ——前馈神经网络

$\mathbf{X}$ ——多头注意力的输入

$\mathbf{Z}$ ——多头注意力经过残差连接和层归一化后的输出

$\mathbf{I}$ ——前馈神经网络经过残差连接和层归一化后的输出

$\mathbf{X} + \text{MultiHead}(\mathbf{X})$ 和 $\mathbf{Z} + \text{FeedForward}(\mathbf{Z})$ 是残差连接部分, 用于解决多层网络训练的问题, 使网络只关注当前差异的部分。层归一化将每一层神经元的输入和输出数据的均值和方差保持不变, 加快收敛速度。

模型的解码器与编码器整体结构相似, 但在解码器的第 1 个多头注意力机制中引入了掩码操作。掩码操作能够限制注意力机制在生成目标序列时只能关注已生成的部分, 而不能访问未生成的部分, 以确保按顺序生成目标序列, 并且只能依赖于之前已生成的部分信息。输出部分包含 1 个全连接层和 1 个 Softmax 层, 其中 Softmax 层用于计算下个目标的概率。

2.2 CNN-Transformer 模型构建

本研究提出的 CNN-Transformer 模型(图 3)将 CNN 作为模型的一部分, 用于提取时间序列数据中的局部特征。由于 CNN 很难捕捉到全局信息, 而基于自注意力机制获取遥感多参数长期依赖特征的 Transformer 可以作为局部细节的补充。因此, 本研究基于 CNN 和 Transformer 构建估产模型, 从时间序列数据中自适应学习旬尺度的时间依赖关系, 更准确地从遥感多参数中提取冬小麦长势信息, 从而取得更好的产量估测结果。

本研究使用 Pytorch 构建和训练 CNN-Transformer 模型。模型由 1 个输入层、3 个一维卷积块、1 个 Transformer 层和 1 个输出层组成。每个卷积块的卷积核大小为 3, 步长为 1。时间序列 $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_t\}$ 作为模型的输入, 其中 $\mathbf{x}_t$ 表示冬小麦旬尺度遥感数据 $(VTCL_t, LAI_t, FPAR_t)$ 组成的向量。输出层作为模型最后一层估测冬小麦产量。为了更好地捕捉全局关系、进行非线性建模以及灵活的输出维度, Transformer 的解码器采用全连接层进行替代^[21]。

将 2011—2021 年每年关中平原 24 个县旬尺度 VTCL、LAI、FPAR 与其对应的每年各县的单产数据作为一组样本, 总共 264 组样本, 并对总样本进行随机划分, 80% 用于训练集, 20% 用于测试集。通过多次训练网络, 最优批量大小为 16, 训练步为 300 个

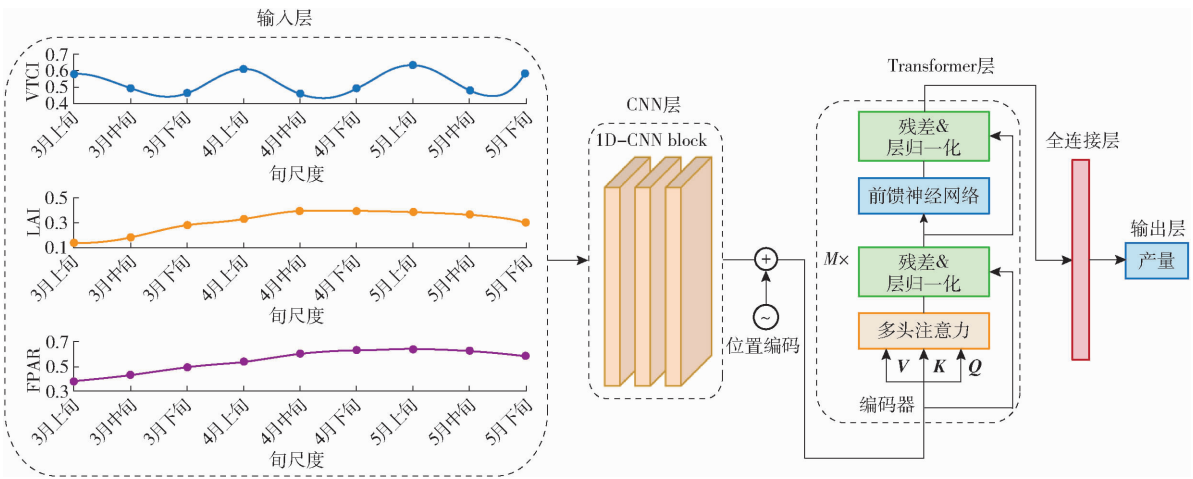


图 3 CNN-Transformer 估产模型结构图

Fig. 3 Structure diagram of CNN-Transformer yield estimation model

迭代周期 (epochs), 优化器为 Adam, 损失函数为均方误差, 初始学习率为 0.001。同时, 为了减少模型的过拟合现象并提高泛化能力, 引入 L_2 正则化项。

2.3 基线模型和模型性能评估

本研究选择 Transformer 模型作为基线模型, 构建了一个用于产量估测的 3 层深度神经网络模型, 包括 1 个输入层、1 个 Transformer 层和 1 个输出层。为了进行对比分析, 基线模型的参数设置与 CNN-Transformer 模型保持一致。为了评价模型在产量估测方面的性能, 选择均方根误差 (Root mean square error, RMSE)、平均绝对百分比误差 (Mean absolute percentage error, MAPE) 和决定系数 (Coefficient of determination, R^2) 评估测试集中模型输出与目标向量之间的关系。为了进一步评估 CNN-Transformer 模型的鲁棒性和泛化能力, 采用 5 折交叉验证方法和留一法进行验证。

为了探究冬小麦生长过程的累积效应, 通过逐步添加旬时间尺度的输入信息, 系统地观察模型在不同生长阶段的性能提升情况。量化不同时间尺度输入参数对冬小麦产量的相对重要性, 以深入了解不同阶段的输入信息对产量预测的贡献程度, 并准确识别冬小麦的关键生长阶段。

3 结果与分析

3.1 CNN-Transformer 和 Transformer 模型性能比较

基于旬尺度遥感输入参数分别构建 Transformer 和 CNN-Transformer 估产模型, 得到估测精度结果如图 4 所示。从整体上看, 与 Transformer 模型 (RMSE 为 465.40 kg/hm², MAPE 为 8.04%, R^2 为 0.64) 相比, CNN-Transformer 模型获得了更准确的估测精度 (RMSE 为 420.39 kg/hm², MAPE 为

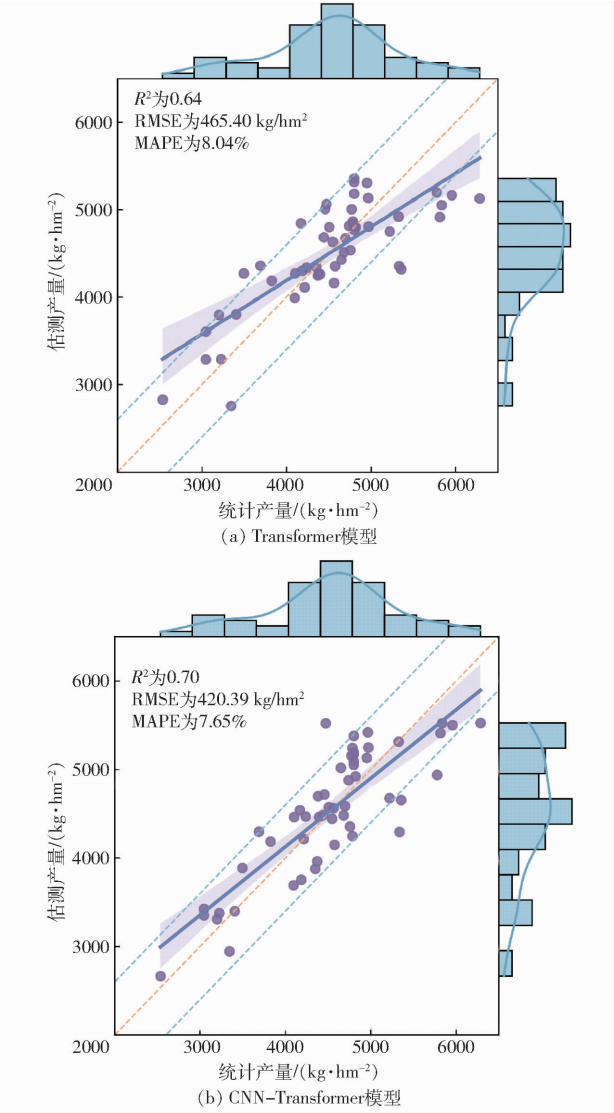


图 4 Transformer 模型和 CNN-Transformer 模型
单产估测精度对比

Fig. 4 Comparisons of yield estimation accuracy between Transformer model and CNN-Transformer model
7.65%, R^2 为 0.70)。 R^2 增加 0.06, RMSE 和 MAPE 分别减少 45.01 kg/hm² 和 0.39 个百分点。2 种模

型统计产量与估测产量的散点图大致都分布在 1:1 线附近(图 4),CNN - Transformer 模型的拟合线斜率较小,且只有 5 个县(图 4b)的统计产量与估测产量之间的绝对误差大于 600 kg/hm²(图中蓝色虚线),表明 CNN - Transformer 模型具有合理的误差,表现出较好的产量估测性能。

由于样本量以及样本不均衡的原因,深度学习估产模型较容易出现高产低估或低产高估的现象。例如,由图 4 边缘分布直方图和概率密度曲线可看出,冬小麦统计产量服从正态分布且主要集中在 3 500 ~ 5 000 kg/hm²之间,超过 5 000 kg/hm²的高产量样本数量相对较少。当冬小麦统计产量高于 5 000 kg/hm²时,2 种模型的估产精度都有下降趋势,存在高产低估现象。但从 2 种模型对于局部(高产部分或低产部分)的产量估测精度来看,CNN - Transformer 模型在对局部的产量估测方面表现优于 Transformer 模型。对于高产部分,其统计产量与估测产量的散点多在误差线(±600 kg/hm²)以内,表明模型能够比较准确地估测高产区域的产量,改善了高产低估的现象。对于低产部分,相比于基线模型的估测结果(图 4a),CNN - Transformer 模型(图 4b)对于低产部分的估测更加贴近 1:1 线,且其分布直方图更加接近统计产量的分布直方图,说明模型能够较准确地捕捉低产量的特征。因此,从 2 种模型对于整体和局部的产量估测表现上看,CNN - Transformer 模型更能够从长时间序列中有效地提取与产量相关的作物长势信息特征,更全面地考虑作物生长过程中的时序规律,使得模型稳定性和估测结果的准确性均有所提高。

为进一步分析 2 种模型在不同产量区间估测方面的表现,对研究区不同产量区间的冬小麦估测结果使用残差箱形图进行分析比较(图 5)。根据产量分布,将冬小麦统计产量划分为 3 个范围:2 000 ~ 3 500 kg/hm²为低产量,3 500 ~ 5 000 kg/hm²为中产量,5 000 ~ 7 000 kg/hm²为高产量。残差是实际观测值与估测值之间的差异,通过分析冬小麦统计产量和估测产量间的差异评估模型的拟合优度。结果显示,对于中产量区间,2 种模型的残差基本在零上下浮动,表明 2 种模型对中产量区间的冬小麦产量变化有较好的捕捉能力。对于高产量区间的残差,正残差占主体,2 种模型均存在高产低估现象,CNN - Transformer 模型的正残差较小,对于高产量的估测效果优于 Transformer 模型。对于低产量区间的残差,Transformer 模型的残差主要呈现负值,而 CNN - Transformer 模型的残差基本分布在零附近,说明 CNN - Transformer 模型对于低产量的估测效果

较好,其估测结果更接近真实值。因此,CNN - Transformer 模型对于不同产量区间的单产估测表现均优于基线模型,进一步说明了 CNN 在局部特征提取方面的优势,能够更有效地捕捉与冬小麦产量相关的信息。

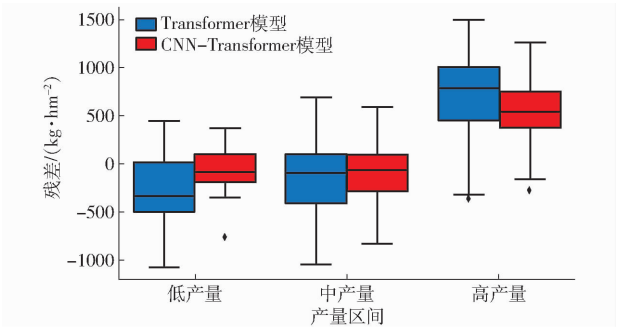


图 5 Transformer 模型和 CNN - Transformer 模型在不同产量区间的残差箱形图

Fig. 5 Box plots of yield residuals at different yield intervals using Transformer model and CNN - Transformer model

3.2 CNN - Transformer 估产模型鲁棒性分析
3.2.1 5 折交叉验证

CNN - Transformer 模型在产量估测方面表现出了良好性能,且在不同产量区间上的产量估测具有较好表现。为了验证 CNN - Transformer 模型的鲁棒性,采用 5 折交叉验证法进行验证和评估(表 1)。结果表明,在 5 个子集中, R^2 均超过 0.60, RMSE 均小于 512 kg/hm², MAPE 均低于 10%。每个子集的估产结果都表现出稳定且良好的性能,进一步验证了 CNN - Transformer 模型在冬小麦产量估测方面的可靠性和稳定性。

表 1 CNN - Transformer 估产模型的 5 折交叉验证结果
Tab.1 Five-fold cross-validation results for CNN - Transformer yield estimation model

折数	R^2	RMSE/(kg·hm ⁻²)	MAPE/%
折 1	0.61	480.73	8.24
折 2	0.63	511.11	8.24
折 3	0.70	444.03	7.48
折 4	0.66	481.15	7.89
折 5	0.62	503.06	9.13

3.2.2 留一法
为了进一步验证估产模型在时间上的鲁棒性,采用留一法进行模型性能的评估(图 6)。结果表明,2011—2021 年 CNN - Transformer 估产模型的 R^2 平均值为 0.63, RMSE 平均值为 433.54 kg/hm²。综合分析发现,除了 2013 年、2020 年以及 2021 年,其他年份产量估测精度较好,估测精度也更加集中在平均 R^2 和平均 RMSE 附近,原因可能是其产量多分布在 3 500 ~ 5 000 kg/hm² 范围内,冬小麦产量分布

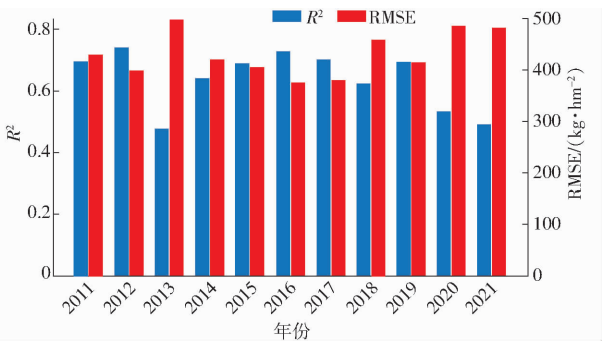


图 6 基于留一法的 CNN-Transformer 估产模型精度
Fig.6 Accuracy of CNN-Transformer yield estimation model based on leave-one-out method

均匀且相对稳定,模型在训练过程中能够有效地捕获与产量相关的特征。2013 年的估测精度明显降低,可能原因是该年份陕西省遭受了严重的干旱,平均降水量明显低于其他年份的降水量,且气温小幅升高,严重影响了冬小麦生长发育,冬小麦产量显著下降,导致模型在估测过程中没有快速捕捉到产量变化。相比于其他年份,2020 年以及 2021 年模型估测精度相对较低。在对数据进行分析时发现,2020 年和 2021 年许多县的统计产量相对于其他年份相同县的统计产量较高,使得模型在估测过程中

表 2 CNN-Transformer 模型所有年份累积旬尺度的平均 R^2 和 RMSE

Tab.2 Average R^2 and RMSE for all years of cumulative decadal scales for CNN-Transformer model									
参数	3 月上旬	3 月中旬	3 月下旬	4 月上旬	4 月中旬	4 月下旬	5 月上旬	5 月中旬	5 月下旬
R^2	0.40	0.42	0.47	0.51	0.56	0.59	0.61	0.62	0.63
RMSE/(kg·hm ⁻²)	592.12	584.53	560.86	546.75	522.34	511.69	506.32	498.63	491.75

3 月下旬至 5 月上旬,随着时序累积以及更多的生长信息输入到模型中, R^2 提升 0.14, RMSE 减少 54.54 kg/hm²,表明模型能够有效地捕捉到冬小麦生长状态的变化,更好地拟合作物生长的时序规律,进一步说明这一时期的输入信息对于产量估测尤为重要,是冬小麦生长的关键时期。3 月下旬至 4 月中旬是冬小麦的拔节期,是根系的建立和发育、茎秆的延伸和加粗、叶片的绿化和扩展的关键阶段。此时冬小麦对土壤中水分的吸收利用最为迫切,正常的绿叶生长能够对冬小麦促进干物质的积累,从而对产量的形成起到重要的作用。4 月下旬至 5 月上旬,模型的估测精度也有所提升,原因是随着冬小麦生长发育阶段的推进,逐渐从拔节期转变为抽穗-灌浆期,此时冬小麦进入了生殖生长的后期,植株开始抽穗,籽粒开始灌浆,光合作用能力达到最高点,植株能够更有效地吸收和利用光能和养分,促进籽粒的膨大和充实,对冬小麦产量的形成起到了关键作用。

在冬小麦生长后期(5 月中旬至 5 月下旬),模

难以准确地拟合高产量数据,从而导致了估测精度下降。

3.3 单产估测的旬尺度累积效应评估

为了评估模型对于冬小麦不同生长阶段累积过程的表征能力,利用构建的 CNN-Transformer 模型分别对不同旬的累积时间尺度进行单产估测(表 2)。将 2011—2021 年 3 月上旬至 5 月下旬中每一旬所有年份的遥感参数作为输入,依次累积输入到 CNN-Transformer 模型中进行单产估测。结果表明,模型估测能力在不同生长阶段受到累积旬输入参数的影响程度存在差异。随着不断累加旬时间尺度信息,模型估测能力逐渐提高。模型在 3 月下旬之前的估测精度较低,原因是在冬小麦生长前期,表征作物生长和环境的信息相对较少,因此模型无法充分捕捉到作物生长特征;另一原因是当前冬小麦处于返青期,其根、茎、叶等生长关键过程尚未完全展开,因此对水分等环境因素的敏感性较低,导致模型预测能力较弱。且这一时期的水分胁迫对产量的影响可能呈现滞后效应,不会立即对作物生长造成显著影响,而是会在后几句产生影响^[25]。

型估测精度提升较缓慢, R^2 仅提升 0.01,原因可能是随着冬小麦的生长,其对水分的需求相对减少。此时,冬小麦已经建立了相对稳定的根系结构和水分调节机制,逐渐适应了环境中的水分胁迫。因此,即使在较为干旱的环境中,冬小麦也表现出一定程度的耐受性。单产与水分供应的关系不再那么紧密,导致模型对水分因素的敏感性减弱,从而估测精度的提升速度变慢。

4 讨论

本研究通过选取旬尺度遥感参数(VTCI、LAI、FPAR),结合 CNN 的局部特征提取能力和 Transformer 的全局时序信息提取能力,构建了 CNN-Transformer 估产模型,用于估测关中平原县级冬小麦产量。结果表明,CNN-Transformer 模型在单产估测方面具有可靠性与适用性。VTCI、LAI 和 FPAR 是表征作物生长状况的关键指标,在决定作物产量形成过程中起着重要的作用。在本研究中,使用 VTCI、LAI 和 FPAR 作为模型的输入参数以

提取作物的生长状况信息,对冬小麦进行长势监测和单产估测研究。然而,影响冬小麦产量的因素很多,气象数据和土壤数据是农作物生长和产量形成的重要驱动因素之一,未来可以考虑添加气象数据和土壤数据作为输入参数。另一方面,本研究使用的遥感数据均来源于MODIS,具有较高的时间分辨率,但空间分辨率相对较低,单一卫星遥感数据难以同时具备高空间分辨率和高时间分辨率。未来可选择较高的空间分辨率影像与MODIS数据进行时空数据融合,利用深度学习方法进一步提高县域尺度乃至田块尺度冬小麦单产估测精度。

为了提高深度学习模型的可解释性和理解参数对作物产量的影响,本研究采用CNN-Transformer结构来捕捉冬小麦生长过程的累积效应。通过逐步累加不同句时间尺度输入参数,分析模型精度的提高程度,并进一步识别冬小麦关键生长阶段,在一定程度上提升了模型的可解释性。在农业领域中,了解每个输入参数在不同生长阶段对产量的影响至关重要,这有助于农业生产者制定种植实践和减少环境影响等决策。尽管本研究通过逐步输入参数的方式分析模型精度的提高程度,但仍难以定量评估各个参数对产量估测的相对重要性。未来的研究可以采用一些方法来分析特征的重要性,例如去除某一自变量的分析法或注意力机制方法识别对产量影响

最大的输入参数。

5 结论

(1)相比于Transformer模型,CNN-Transformer模型能从VTCI、LAI和FPAR遥感参数中更有效地学习与冬小麦生长和产量相关的特征,模型 R^2 提升0.06, RMSE和MAPE分别减少45.01 kg/hm²和0.39个百分点。此外,CNN-Transformer模型对于Transformer模型存在的高产低估和低产高估现象均有所改善,该模型在捕捉高产冬小麦的生长特征和单产估测方面表现较好。

(2)CNN-Transformer估产模型表现出较好的鲁棒性,在不同产量区间和不同年份表现出较好的泛化能力和稳定性,表明模型能够适应不同年份冬小麦生长环境的变化,捕捉到冬小麦生长过程中的共性特征,不仅仅局限于特定年份的情况,且能够捕捉不同产量水平下的冬小麦生长特征和变化趋势。

(3)在冬小麦的不同生长阶段,使用CNN-Transformer结构能够有效地捕获冬小麦生长过程中的累积效应。随着输入时序数据的不同,累积每句的输入参数对产量的影响程度不同。其中,3月下旬至5月上旬从遥感多参数中提取的信息对于产量估测至关重要,是冬小麦生长的关键时期。

参 考 文 献

- [1] 赵龙才,李粉玲,常庆瑞. 农作物遥感识别与单产估算研究综述[J]. 农业机械学报, 2023, 54(2): 1-19.
ZHAO Longcai, LI Fenling, CHANG Qingrui. Review on crop type identification and yield forecasting using remote sensing [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(2): 1-19. (in Chinese)
- [2] 吴炳方,张森,曾红伟,等. 全球农情遥感速报系统20年[J]. 遥感学报, 2019, 23(6): 1053-1063.
WU Bingfang, ZHANG Miao, ZENG Hongwei, et al. Twenty years of CropWatch: progress and prospect[J]. Journal of Remote Sensing, 2019, 23(6): 1053-1063. (in Chinese)
- [3] 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.
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. (in Chinese)
- [4] 王鹏新,田惠仁,张悦,等. 基于深度学习的作物长势监测和产量估测研究进展[J]. 农业机械学报, 2022, 53(2): 1-14.
WANG Pengxin, TIAN Hui ren, ZHANG Yue, et al. Crop growth monitoring and yield estimation based on deep learning: state of the art and beyond[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(2): 1-14. (in Chinese)
- [5] 王鹏新,龚健雅,李小文. 条件植被温度指数及其在干旱监测中的应用[J]. 武汉大学学报(信息科学版), 2001, 26(5): 412-418.
WANG Pengxin, GONG Jianya, LI Xiaowen. Vegetation temperature condition index and its application for drought monitoring [J]. Geomatics and Information Science of Wuhan University, 2001, 26(5): 412-418. (in Chinese)
- [6] ZHANG Y, WANG P, TANSEY K, et al. Enhanced feature extraction from assimilated VTCI and LAI with a particle filter for wheat yield estimation using cross-wavelet transform[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023, 16: 5115-5127.
- [7] LI L, WANG B, FENG P, et al. Crop yield forecasting and associated optimum lead time analysis based on multi-source environmental data across China[J]. Agricultural and Forest Meteorology, 2021, 308: 108558.
- [8] 黎锐,李存军,徐新刚,等. 基于支持向量回归(SVR)和多时相遥感数据的冬小麦估产[J]. 农业工程学报, 2009, 25(7): 114-117.

- LI Rui, LI Cunjun, XU Xingang, et al. Winter wheat yield estimation based on support vector machine regression and multi-temporal remote sensing data[J]. Transactions of the CSAE, 2009, 25(7): 114 – 117. (in Chinese)
- [9] 刘峻明, 和晓彤, 王鹏新, 等. 长时间序列气象数据结合随机森林法早期预测冬小麦产量[J]. 农业工程学报, 2019, 35(6): 158 – 166.
- LIU Junming, HE Xiaotong, WANG Pengxin, et al. Early prediction of winter wheat yield with long time series meteorological data and random forest method[J]. Transactions of the CSAE, 2019, 35(6): 158 – 166. (in Chinese)
- [10] REICHSTEIN M, CAMPS-VALLS G, STEVENS B, et al. Deep learning and process understanding for data-driven earth system science[J]. Nature, 2019, 566(7743): 195 – 204.
- [11] YANG Q, SHI L, HAN J, et al. Deep convolutional neural networks for rice grain yield estimation at the ripening stage using UAV-based remotely sensed images[J]. Field Crops Research, 2019, 235: 142 – 153.
- [12] LIN T, ZHONG R, WANG Y, et al. DeepCropNet: a deep spatial-temporal learning framework for county-level corn yield estimation[J]. Environmental Research Letters, 2020, 15(3): 034016.
- [13] TIAN H, WANG P, TANSEY K, et al. A deep learning framework under attention mechanism for wheat yield estimation using remotely sensed indices in the Guanzhong Plain, PR China[J]. International Journal of Applied Earth Observation and Geoinformation, 2021, 102: 102375.
- [14] BENGIO Y, SIMARD P, FRASCONI P. Learning long-term dependencies with gradient descent is difficult[J]. IEEE Transactions on Neural Networks, 1994, 5(2): 157 – 166.
- [15] YU W, GONZALEZ J, LI X. Fast training of deep LSTM networks with guaranteed stability for nonlinear system modeling[J]. Neurocomputing, 2021, 422: 85 – 94.
- [16] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30: 5998 – 6008.
- [17] ACHEAMPONG F A, NUNOO-MENSAH H, CHEN W. Transformer models for text-based emotion detection: a review of BERT-based approaches[J]. Artificial Intelligence Review, 2021, 54(8): 5789 – 5829.
- [18] LIU M, CHAI Z, DENG H, et al. A CNN – Transformer network with multiscale context aggregation for fine-grained cropland change detection[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 4297 – 4306.
- [19] ONOUFRIOU G, HANHEIDE M, LEONTIDIS G. Premonition net, a multi-timeline Transformer network architecture towards strawberry tabletop yield forecasting[J]. Computers and Electronics in Agriculture, 2023, 208: 107784.
- [20] JIN W, ZHANG W, HU J, et al. Transformer for sub-seasonal extreme high temperature probabilistic forecasting over eastern China[J]. Theoretical and Applied Climatology, 2023, 151(1 – 2): 65 – 80.
- [21] LIANG S, HUA Z, LI J. Hybrid Transformer – CNN networks using superpixel segmentation for remote sensing building change detection[J]. International Journal of Remote Sensing, 2023, 44(8): 2754 – 2780.
- [22] 李艳, 王鹏新, 刘峻明, 等. 基于条件植被温度指数的冬小麦主要生育时期干旱监测效果评价——Ⅱ. 改进的层次分析法和变异系数法组合赋权[J]. 干旱地区农业研究, 2014, 32(1): 236 – 239.
- LI Yan, WANG Pengxin, LIU Junming, et al. Evaluation of drought monitoring effects in the main growing stages of winter wheat by using the vegetation temperature condition index—Ⅱ. Improved analytic hierarchy process and variation coefficient method[J]. Agricultural Research in the Arid Areas, 2014, 32(1): 236 – 239. (in Chinese)
- [23] 孙威, 王鹏新, 韩丽娟, 等. 条件植被温度指数干旱监测方法的完善[J]. 农业工程学报, 2006, 22(2): 22 – 26.
- SUN Wei, WANG Pengxin, HAN Lijuan, et al. Further improvement of the approach to monitoring drought using vegetation and temperature condition indexes from multi-years' remotely sensed data[J]. Transactions of the CSAE, 2006, 22(2): 22 – 26. (in Chinese)
- [24] 王鹏新, 苟兰, 李俐, 等. 基于时间序列叶面积指数傅里叶变换的作物种植区域提取[J]. 农业工程学报, 2017, 33(21): 207 – 215.
- WANG Pengxin, XUN Lan, LI Li, et al. Extraction of planting areas of main crops based on Fourier transformed characteristics of time series leaf area index products[J]. Transactions of the CSAE, 2017, 33(21): 207 – 215. (in Chinese)
- [25] 张悦, 王鹏新, 张树誉, 等. 基于双参数和 Morlet 多时间尺度特性的冬小麦单产估测[J]. 农业机械学报, 2021, 52(10): 243 – 254.
- ZHANG Yue, WANG Pengxin, ZHANG Shuyu, et al. Yield estimation of winter wheat based on two parameters and Morlet multi-scale characteristics[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(10): 243 – 254. (in Chinese)