Rice Growth Monitoring Based on 3D-CNN Satellite Remote Sensing Data

Yindong Li, Yang Chen and Wang Wang

July 17, 2024

Rice Growth monitoring based on 3D-CNN satellite remote sensing data

Li Yindong,Chen Yang ,Wang Wang

July 17, 2024

Li Yindong
Geely University of China
liyindong@bgu.edu.cn

Chen Yang
Geely University of China
chenyang@bgu.edu.cn

Wang Wang
Geely University of China
wangwang@bgu.edu.cn

Abstract

Rice is the most important food crop in China. Timely and accurate acquisition of large-area rice planting area and growth information is of great significance to China's food security. Satellite remote sensing can sensitively respond to the development of large-area rice plants and changes in soil moisture, and is an important means of rice growth monitoring[1]. Remote sensing research on rice growth monitoring mainly focuses on three major focuses: extraction of rice planting area, inversion of rice physiological parameters, and identification of rice maturity, and the research methods mainly include mathematical analysis, machine learning, and multi-source collaboration. In response to the problem that traditional remote sensing vegetation indices are difficult to accurately monitor large-area rice growth and have low accuracy, this study proposes a 3D-CNN neural network model that uses hyperspectral satellite remote sensing data sets to monitor the growth of rice in China's main rice planting areas[2]. The model uses three-dimensional convolutional neural networks (3D-CNN) and temporal convolutional neural networks (TCN) to process the spatiotemporal information and spectral information of rice satellite remote sensing images. The model first extracts spatiotemporal features through multiple three-dimensional convolutional layers, then performs spatial feature analysis by compressing and reducing the dimensionality of the output features through two-dimensional convolutional layers, and finally, the high-level feature maps are flattened and category prediction is performed through fully connected layers[3]. At the same time, a new loss function is introduced in the neural network model to eliminate the impact of the imbalance of rice yield label distribution. Finally, the new model is verified through the prediction of rice yield data in China. The results are compared with the main deep learning methods in use. The experimental results show that the method proposed in this paper can provide better predictive performance than other competitive methods.

Key words: Rice ,Chlorophyll,Hyperspectral remote sensing,3D-CNN,Growth monitoring

1 Introduction

Chlorophyll in plant leaves is the most important pigment in photosynthesis. It uses solar radiation and absorbs carbon dioxide from the atmosphere to perform photosynthesis, producing oxygen and organic matter[4]. Therefore, monitoring the level of chlorophyll content can be used to monitor the growth status of plants. In the process of rice production management, monitoring the chlorophyll content level of rice can effectively reflect the growth condition of rice, especially monitoring the chlorophyll content level of the rice canopy is of great guiding significance for the assessment of rice growth and yield prediction[5].

Satellite remote sensing technology can accurately, quickly, and non-destructively monitor the content levels of biological indicators such as chlorophyll in rice leaves. Establishing a correlation model between chlorophyll content and the nutritional status and growth trend of rice can effectively predict the yield of rice[6].

Utilizing satellite remote sensing technology, one can monitor the growth of rice by analyzing data such as the chlorophyll content, leaf area index, and aboveground biomass of the rice canopy[7]. By assessing the levels of nitrogen, phosphorus, and potassium in the rice paddies, one can predict the nutrient and water conditions necessary for rice cultivation[3]. Combining the analysis of rice growth and nutrient conditions with meteorological data allows for the prediction of rice yield. This approach provides a powerful tool for precision agriculture, enabling more informed decisions regarding crop management and resource allocation[8].

This study processes satellite remote sensing images of rice to monitor its growth. Through literature research, it has been found that the main methods for monitoring the growth of rice based on satellite remote sensing are traditional inversion methods. For example, a composite rice growth monitoring model is constructed using the spectral vegetation index - leaf area nitrogen index[9]. Growth monitoring models are built based on nitrogen-sensitive bands, such as the Ratio Vegetation Index (RVI) and the Green Normalized Difference Vegetation Index (GNDVI). The Iterative Self-Elimination Partial Least Squares (ISE-PLS) method is used to select rice growth-sensitive bands between 400 and 930 nm by gradually eliminating them, and a rice growth monitoring model is established. Multiple period correlation analysis is conducted using rice leaf nitrogen parameters and canopy spectrum parameters to establish a rice growth monitoring model for leaf nitrogen accumulation and leaf area nitrogen index[10]. There are several issues with traditional satellite remote sensing monitoring of rice growth: it cannot be used for large-area monitoring; obtaining satellite images is difficult; specialized knowledge of plant spectrum is required[11]; the generalization ability of the established models is average; and the model accuracy is relatively low[12]. This study uses deep learning methods to perform pixel interpretation of satellite remote sensing data, thereby achieving monitoring and yield estimation of rice growth over a large area[13].

Crop or vegetation growth monitoring using satellite remote sensing imagery typically involves two methods[14]. The first method aggregates spectral bands into vegetation indices that represent the physical properties of vegetation, enabling vegetation growth monitoring[15]. The second method directly utilizes raw multi-temporal images for vegetation growth monitoring. Guerschman et al[16]. used multi-temporal Landsat TM data for land cover classification, and the results showed that using raw images achieved higher accuracy than using the Normalized Difference Vegetation Index (NDVI)[17]. Spectral, spatial, and temporal features are the foundations for extracting crop growth information from remote sensing[18]. Seasonality is one of the most prominent characteristics of crops, and multi-temporal remote sensing is an effective way to monitor crop growth dynamics and classification[19]. Shallow machine learning algorithms such as Support Vector Machines (SVM) and Random Forest (RF) have a limited number of non-linear transformation combinations and are greatly affected by feature engineering (FE), resulting in poor resolution of complex heterogeneous features in images[16]. Deep learning is considered a breakthrough technology in the fields of machine learning and data mining (including remote sensing). Due to its hierarchical feature representation, high efficiency, and end-to-end automated learning, it has gradually become the mainstream algorithm in the field of image pattern recognition[20]. Convolutional Neural Networks (CNN) are one of the most successful network structures in deep learning methods, and studies have shown that CNNs perform better than other models in most image classification problems[21]. For multi-temporal remote sensing images or time series NDVI, 3D CNNs are particularly suitable for extracting dynamic features of crop growth and are superior to mainstream methods such as 2D CNN, SVM, and nearest neighbor classification[22]. A comparative study of CNN, Recursive Neural Network (RNN), and hybrid neural networks (CNN+RNN) based on multi-spectral time series data concluded that the most effective method is the hybrid configuration network. Li et al[23]., drawing on the transformer structure in natural language processing (NLP) knowledge to explore multi-time series patterns, proposed a hybrid model CNN-transformer that significantly improves the accuracy of crop classification. Gadiraju et al[24]. proposed a multimodal deep learning scheme that jointly uses spatial, spectral, and phenological features to identify crop types, reducing the prediction error by 60

One-dimensional CNNs provide an effective and efficient method for long time series remote sensing image crop type identification. Jie Yi et al[25]. believe that Long Short-Term Memory (LSTM) networks have a clear advantage in classifying crops using multi-source remote sensing data fusion of time series NDVI[26]. The main advantage of deep learning is its ability to effectively approximate highly complex problems without the need for prior feature engineering. Remote sensing images can provide dynamic or temporal information[27]. Although significant progress has been made in the theory, technical methods, and practical applications of crop remote sensing , 2D CNNs lack the ability to accurately extract three-dimensional features[28]. The information extracted in the third dimension (i.e., the temporal dimension) is averaged and folded into a scalar, thus not fully exploiting the characteristics of this dimension. The structural design of 3D convolution is very suitable for spatiotemporal representation[29]. However, 3D CNNs have high computational complexity, many parameters are not easy to train, and they perform poorly when processing similar textures on multi-spectral bands, so the application of 3D CNNs in vegetation growth monitoring is relatively rare. In response to the insufficient use of time series remote sensing information in the growth monitoring process of crops, the similar expression of features in medium-resolution images, and most studies focusing on a small number of crop categories, this paper proposes a deep learning algorithm model for rice growth monitoring based on 3D-CNN satellite remote sensing data[30]. It explores the optimization process of the model and analyzes the role of spatial and spectral information in the model, providing new ideas for monitoring vegetation growth with satellite remote sensing data[31].

Support Vector Machines (SVM) and Random Forest (RF) are shallow machine learning algorithms that have a limited number of non-linear transformation layers and are greatly influenced by Feature Engineering (FE)[32]. As a result, they perform poorly in distinguishing complex heterogeneous features within images. Deep learning is considered a breakthrough technology in the fields of machine learning and data mining, including remote sensing[33]. Its advantages, such as hierarchical feature representation, high computational efficiency, and end-to-end automated learning, have made it the mainstream algorithm in the field of image pattern recognition[34].

Convolutional Neural Networks (CNN) are one of the most successful network structures in deep learning methods. Studies have shown that CNNs generally perform better than other models in most image classification problems[35]. For multi-temporal remote sensing images or time series NDVI, 3D CNNs are particularly adept at extracting the dynamic features of crop growth, outperforming mainstream methods such as 2D CNNs, SVM, and nearest neighbor classification[36].

A comparison of classification performance based on multi-spectral time series data among CNNs, Recursive Neural Networks (RNN), and hybrid neural networks (CNN+RNN) concluded that the most effective method is the hybrid configuration network. This suggests that combining different neural network architectures can leverage their respective strengths to achieve better classification results in remote sensing image analysis[24].

The primary advantage of deep learning is its ability to effectively approximate highly complex problems without the need for prior feature engineering. Remote sensing images can provide dynamic or temporal information. Although significant progress has been made in the theory, technical methods, and practical applications of crop remote sensing [14-15], 2D CNNs lack the capability to accurately extract three-dimensional features. The information extracted in the third dimension, which is the temporal dimension, is averaged and collapsed into a scalar, thus not fully exploiting the characteristics of this dimension. The structural design of 3D convolution is highly suitable for spatiotemporal representation[37]. However, 3D CNNs have high computational complexity and a large number of parameters, making them difficult to train. Moreover, they perform poorly when processing classes with similar textures in multi-spectral bands, which is why the application of 3D CNNs in vegetation growth monitoring is relatively rare[38].

In response to issues such as insufficient utilization of time series remote sensing information in crop growth monitoring, similar feature expression in medium-resolution imagery, and a focus on extracting a limited number of crop categories in most studies, this paper proposes a deep learning algorithm model for rice growth monitoring based on 3D-CNN satellite remote sensing data[39]. The paper explores the optimization process of the model and analyzes the role of spatial and spectral information within the model, offering new perspectives for monitoring vegetation growth using satellite remote sensing data.

The exploration of the model's optimization process includes refining the architecture of the 3D CNN to handle the high dimensionality of data while maintaining computational efficiency[39]. It also involves finding the right balance between the spatial and temporal features that contribute to the accurate classification of crop growth stages. By addressing these challenges, the proposed model aims to enhance the capability of satellite remote sensing in providing valuable insights into crop health and yield estimation, ultimately contributing to more informed decision-making in agricultural management[7].

2 Methodology

2.1 Research Assignment

In the research, given multiple temporal remote sensing images of certain areas, the first step is to identify the rice planting areas through deep learning algorithms and calculate the area of rice cultivation. When using 3D-CNN algorithms to process satellite remote sensing data, both spectral and spatial information is processed to monitor the growth of rice[15]. Therefore, the multiple temporal remote sensing images of a county-level unit are first defined as:

.

$$I \in R_{t\times c\times h\times w} \tag{1}$$

Where, that is, I is a time series of length t, and each remote sensing image has c channels, h in height and w in width.For any at time t, use all the time sequence images before time t to learn a prediction model, and finally get the predicted crop growth at time t, which is denoted as p_i^t. This problem can be defined as[18]:

$$p_i^t = F(I_i^l, ..., I_i^{t-1}) \tag{2}$$

2.2 Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNNs) are multi-layer feedforward neural networks that excel at image and video recognition, classification, and segmentation tasks by balancing local and global features. CNNs utilize convolutional layers to extract local features from images, followed by pooling layers that reduce the spatial dimensions of the features while increasing invariance to image displacements. A CNN typically consists of multiple neural network layers, with each layer connected to the next through a set of learnable weights. Each layer processes a local portion of the image, and these portions are scanned across the entire image to capture features at different scales, both local and global[13].

Within the CNN framework, image features are generalized through alternating convolutional and pooling layers until high-dimensional features are obtained. The classification of the image is then performed by a fully connected layer at the end of the network. Additionally, multiple feature maps can exist within a single convolutional layer, and the weights of the convolutional nodes are shared across the same feature map. This setup allows the network to learn various features while keeping the number of weight parameters in the neural network manageable.The nonlinear activation function is used to enhance the nonlinearity of the feature. Specifically, the main operations performed in CNN can be summarized as follows[23]:

$$Q^l = P_s(\sigma(O^{l-1} * W^l + b^l)) \tag{3}$$

Q^l represents the input feature map of the l layer, W^l and b^l represent the weight and deviation of the layer respectively, and they convolve the input feature map by linear convolution *, () represents the nonlinear function outside the convolution layer. The maximum pooling (P_s) operation of the ss window size is then used to aggregate statistical information for features within a specific region, thus outputting the feature map O^l at layer l

2.3 3D Convolutional Neural Network (3D CNN)

3D Convolutional neural networks (3D CNNs) play an important role in the analysis of hyperspectral remote sensing data[12]. Hyperspectral remote sensing data has the characteristics of atlas integration and rich spectral information, which can provide more details and features than traditional RGB images. 3D CNNS are able to process both the spatial and spectral dimensions of an image, allowing for more efficient extraction and analysis of this data. Space-spectral feature extraction: 3D CNNs are capable of extracting both spatial and spectral features of hyperspectral images simultaneously, which is essential for understanding image content and accurate classification. Dimensionality reduction processing: Since the dimensionality of hyperspectral images is very high, dimensionality reduction using techniques such as principal component analysis (PCA) is a common practice to reduce computation and remove redundant information. This paper proposes a new CNN-based architecture, which combines spatial features and spectral information features of remote sensing images for analysis. The model architecture is shown in Figure 1[12].

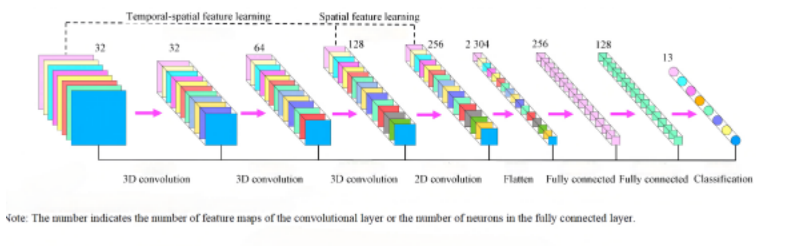

Figure 1: 3D-CNN convolutional neural network model

.

The model consists of three interconnected parts. In the first part, 3D convolution is used to perform image spatial feature analysis, extracting spatial features of remote sensing images through multiple 3D CNN layers without folding spectral information of remote sensing images. In the second part, 2D convolution is used to introduce remote sensing spectral information for analysis, and the output of 3D CNN is compressed and then transmitted to 2D CNN. Further, only important information is extracted from remote sensing spectral information, and the accuracy of remote sensing image recognition can be effectively improved by combining the spectral information[10].

2.4 Method

A new deep learning framework is used to make full use of the spatial and spectral features of satellite remote sensing images. The model consists of two parts. First, for each spectral step t(1,2.... T), the framework introduces a 3DCNN network to extract the sensory image features of the spectral features. Finally, each image is mapped to an N-dimensional input of the same size at the corresponding spectral step. Finally, a prediction layer composed of fully connected neural networks is deployed to generate predictions from the generated representation of joint features. The model structure is shown in Figure 2. The objective function of the model is expressed as follows:

$$L = \frac{1}{n} \sum_{i=1}^{n} W_i \left(p_i - y_i\right)^2 \tag{4}$$

2.5 loss function

The commonly used method in image semantic segmentation task is to use cross-entropy loss function to train the model. The formula of cross-entropy loss function is as follows:[14]

$$CE_{\text{loss}} = -\frac{1}{N} \sum_{n=1}^{N} \sum_{c=1}^{C} I_c^{(n)} \cdot \log\left(p_c^{(n)}\right) \tag{5}$$

In formula (5), N is the total number of training images in each batch, C is the set of all categories, () n ci is the unique hot label class c of the NTH image block sample in the current batch, () n cp is the softmax probability of model prediction sample n being class c. However, since this loss is calculated by summing over all pixels, it does not account well for unbalanced classes. Although the introduction of weight factor balances the importance of positive samples and negative samples, it does not solve the imbalance problem of hard to classify samples and easy to classify samples.

he focal loss function (Focalloss) can solve these difficulties by reducing the weight of easily classified samples and focusing more attention on difficult classified samples. The coke loss function is based on the standard cross entropy loss by introducing the adjustment factor, and its formula is as follows:

$$Focal_{loss} = (1 - p_c^{(n)})^\gamma * CE_{loss} \tag{6}$$

3 Materials

Based on multi-temporal Sentinel-2A and Landsat8 data, SVM, CNN, and 3D-CNN were used to extract rice seedling growth, and the results were shown in Figure 2.

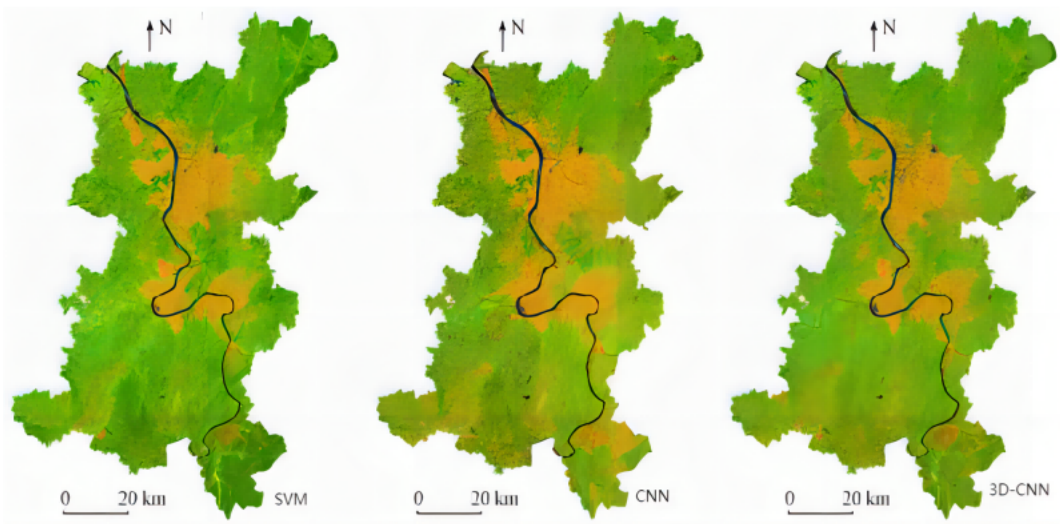

Figure 2: Rice extraction results based on three classifiers

hrough comparison and qualitative analysis of rice information in the visual interpretation results of GF-5 and GoogleEarth high-resolution remote sensing images, the distribution of rice in the classification results based on 3D-CNN is basically consistent with the real distribution, especially in urban areas, rice can be well separable from other vegetation, with less misclassification and missing. The results showed that this method could extract rice information from regions with high heterogeneity. In contrast, the classification results based on 3D-CNN are closer to the distribution of rice in the study area. The rice extraction accuracy of the three classifiers is shown in Table 4.

The OA and Kappa coefficients of the CNN model reached 96.11% and 0.94 respectively, and the UA and PA of rice were both above 95%, indicating a high recognition accuracy. In contrast, SVM and CNN based classification results are not good, OA and Kappa coefficients are 81.47%, 0.75 and 83.77%, 0.80, respectively.

In order to verify that CNN algorithm has stronger anti-interference and generalization ability in double-cropping rice extraction, this paper selected three rice typical type regions and compared and analyzed the rice extraction results of 3D-CNN, CNN and SVM classification methods (FIG. 4). It can be seen from the results that compared with SVM and RF classification algorithms, CNN can still effectively extract rice information in the regions with greater heterogeneity. However, it can be seen from the three rice farming areas that although the classification is based on multi-temporal and multi-source remote sensing data, rice is still easily misclassified by the classifier into vegetable base and other crops.

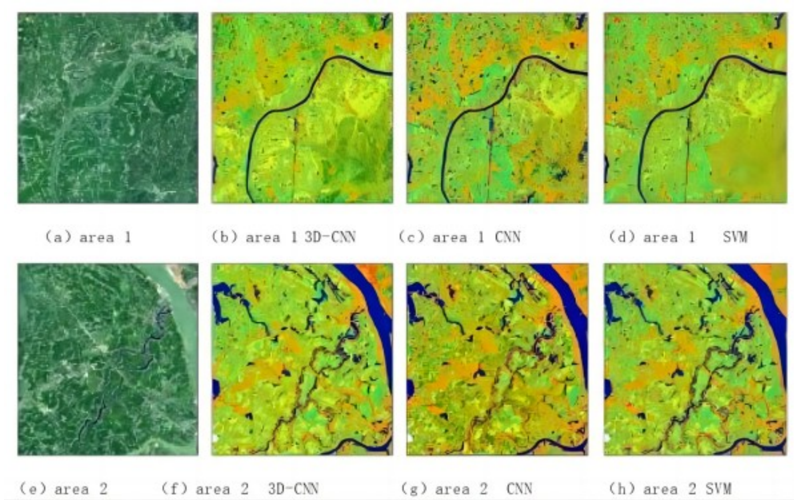

Figure 3: Classification results of 3 typical rice farming regions with different classifiers

The CNN method avoids the confusion between rice and other land species to a large extent, and the phenomenon of misclassification and misclassification is less. However, in the RF and SVM classification results, there were different degrees of misclassification, such as rice was misclassified into other crops and vegetables, and forest land was misclassified into other crops. In previous studies on rice extraction, regions with different planting density, heterogeneity and fragmentation degree of rice patches would produce pixel sets of different scales with similar spectra, leading to overfitting or over-smoothing problems in the classifier.

Therefore, when rice is extracted by different classifiers and classification methods, if the spectral similarity between samples is too high, the results will have overfitting and over-smoothing effects. In this paper, CNN adopts regularization method to avoid high-intensity overfitting to a certain extent [23]. In addition, CNN uses the convolutional layer for feature extraction, and each neuron integrates the local information of the high-level through local perception, so that the entire classification framework can obtain all the representational information in the image scene. These representation information are created by different convolution nuclei, that is, these features can enable CNN to understand the semantics of the entire scene, and are also the main reason why CNN classification results do not produce large-scale "pepper and salt phenomenon". Secondly, CNN uses the pooling layer to carry out high-level abstract expression of features and realize deep mining of images, which enables CNN to obtain better classification results than SVM and RF in areas with large land patch fragmentation.

4 Discussion

Previous studies on regional rice information extraction using optical remote sensing images can be roughly divided into three categories: (1) single phase image + image statistical methods, such as supervised classification (MLC and SVM, etc.), unsupervised classification (threshold method and ISODATA, etc.) or object-oriented classification; (2) Time series remote sensing images + supervised classification (DT and RF, etc.); (3) Images of special phenological period + pixel-based classification, such as using normalized water index, spectral band or vegetation index before and after water inundated or key phenological characteristics, to extract rice information based on pixels.

In this paper, 3D-CNN neural network model was used to monitor rice growth in major rice growing areas in China based on satellite remote sensing data sets. The model uses three-dimensional convolutional neural network (3D-CNN) and temporal convolutional neural network (TCN) to process spatiotemporal and spectral information of rice satellite remote sensing images. The model first extracts spatio-temporal features through multiple 3D convolution layers, then performs spatial feature analysis through 2D convolution layer after dimensionality reduction of the output features, and finally flattens the high-level feature maps and performs category prediction through the fully connected layer. At the same time, a new loss function was introduced into the neural network model to eliminate the influence of unbalanced distribution of rice yield labels. Finally, the new model is verified by the prediction of rice yield data in China. The results are compared with the mainly used deep learning methods. Experimental results show that the proposed method can provide better prediction performance than other competing methods.

5 conclusion

In the study of rice growth monitoring based on 3D-CNN satellite remote sensing data, 3D convolutional neural network (3D-CNN) combined with temporal convolutional network (TCN) was mainly used to improve the ability to capture spatio-temporal information and band information of remote sensing images. Using domestic remote sensing satellite time series data, this paper presents an effective detection method for major rice growing areas in China. The satellite remote sensing data of missing phase was obtained by spatiotemporal fusion method, and the rice growth was monitored by 3D-CNN method. 3D-CNN technology shows potential in rice growth monitoring, but there are still some challenges, such as broken plots and complex terrain, the diversity of cultivation conditions in paddy fields, and the problems of asynchronous and intercrop rotation. Future research may consider the use of radar satellite data to enhance data sources, as well as the application of spatio-temporal fusion technology to obtain long and dense time-series optical images to overcome the shortage of data sources caused by optical images susceptible to cloud and rain.

References

[1] L. S. CAI Yaotong and L. S. CAI Yaotong, "Extraction of paddy rice based on convolutional neuralnetwork using multi source remote sensing data," *Remote sensing of natural resources*, vol. 32, pp. 97–104, Dec. 2020.

[2] L. Chen, Z. Wei, and Y. Xu, "A Lightweight SpectralSpatial Feature Extraction and Fusion Network for Hyperspectral Image Classification," *Remote Sensing*, vol. 12, p. 1395, Apr. 2020.

[3] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. L. Yuille, and Y. Zhou, "TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation," Feb. 2021. arXiv:2102.04306 [cs].

[4] D. M. G. dela Torre, J. Gao, C. Macinnis-Ng, and Y. Shi, "Phenology-based delineation of irrigated and rain-fed paddy fields with Sentinel-2 imagery in Google Earth Engine," *Geo-spatial Information Science*, vol. 24, pp. 695–710, Oct. 2021. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10095020.2021.1984183.

[5] D. M. G. dela Torre, J. Gao, and C. Macinnis-Ng, "Remote sensing-based estimation of rice yields using various models: A critical review," *Geo-spatial Information Science*, vol. 24, pp. 580–603, Oct. 2021. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10095020.2021.1936656.

[6] J. Dong and X. Xiao, "Evolution of regional to global paddy rice mapping methods: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 119, pp. 214–227, Sept. 2016.

[7] J. J. Erinjery, M. Singh, and R. Kent, "Mapping and assessment of vegetation types in the tropical rainforests of the Western Ghats using multispectral Sentinel-2 and SAR Sentinel-1 satellite imagery," *Remote Sensing of Environment*, vol. 216, pp. 345–354, Oct. 2018.

[8] S. FROLKING, J. QIU, S. BOLES, X. XIAO, J. LIU, Y. ZHUANG, C. LI, and X. QIN, "Combining remote sensing and ground census data to develop new maps of the distribution of rice agriculture in China," *Combining remote sensing and ground census data to develop new maps of the distribution of rice agriculture in China*, vol. 16, no. 4, pp. 38.1–38.10, 2002. Place: Washington, DC Publisher: American Geophysical Union.

[9] G. Wang, "Quantifying the Spatio-Temporal Variations and Impacts of Factors on Vegetation Water Use Efficiency Using STL Decomposition and Geodetector Method | Semantic Scholar."

[10] T. Jiang, T. Xu, and X. Li, "VA-TransUNet: A U-shaped Medical Image Segmentation Network with Visual Attention," in *Proceedings of the 2022 11th International Conference on Computing and Pattern Recognition*, ICCPR '22, (New York, NY, USA), pp. 128–135, Association for Computing Machinery, May 2023.

[11] M.-J. Lambert, P. C. S. Traor, X. Blaes, P. Baret, and P. Defourny, "Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's cotton belt," *Remote Sensing of Environment*, vol. 216, pp. 647–657, Oct. 2018.

[12] M.-J. Lambert, P. C. S. Traor, X. Blaes, P. Baret, and P. Defourny, "Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's cotton belt," *Remote Sensing of Environment*, vol. 216, pp. 647–657, Oct. 2018.

[13] H. Li, X. Wang, S. Wang, J. Liu, Y. Liu, Z. Liu, S. Chen, Q. Wang, T. Zhu, L. Wang, and L. Wang, "ChinaRiceCalendar seasonal crop calendars for early-, middle-, and late-season rice in China," *Earth System Science Data*, vol. 16, pp. 1689–1701, Apr. 2024.

[14] J. Li, H. Wang, A. Zhang, and Y. Liu, "Semantic Segmentation of Hyperspectral Remote Sensing Images Based on PSE-UNet Model," *Sensors*, vol. 22, p. 9678, Dec. 2022.

[15] M. Liang, H. Wang, X. Yu, Z. Meng, J. Yi, and L. Jiao, "Lightweight Multilevel Feature Fusion Network for Hyperspectral Image Classification," *Remote Sensing*, vol. 14, p. 79, Jan. 2022. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute.

[16] B. Mishra, L. Busetto, M. Boschetti, A. Laborte, and A. Nelson, "RICA: A rice crop calendar for Asia based on MODIS multi year data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 103, p. 102471, Dec. 2021.

[17] S. Liu, Y. Chen, Y. Ma, X. Kong, X. Zhang, and D. Zhang, "Mapping Ratoon Rice Planting Area in Central China Using Sentinel-2 Time Stacks and the Phenology-Based Algorithm," *Remote Sensing*, vol. 12, p. 3400, Oct. 2020.

[18] Y. Luo, Z. Zhang, Y. Chen, Z. Li, and F. Tao, "ChinaCropPhen1km: a high-resolution crop phenological dataset for three staple crops in China during 20002015 based on leaf area index (LAI) products," *Earth System Science Data*, vol. 12, pp. 197–214, Jan. 2020. Publisher: Copernicus GmbH.

[19] K. Mao, F. Gao, S. Zhang, and C. Liu, "An Information Spatial-Temporal Extension Algorithm for Shipborne Predictions Based on Deep Neural Networks with Remote Sensing ObservationsPart I: Ocean Temperature," *Remote Sensing*, vol. 14, p. 1791, Apr. 2022.

[20] J. Nalepa, L. Tulczyjew, M. Myller, and M. Kawulok, "Segmenting Hyperspectral Images Using Spectral-Spatial Convolutional Neural Networks With Training-Time Data Augmentation," July 2019. arXiv:1907.11935 [cs].

[21] J. Nalepa, M. Myller, Y. Imai, K.-i. Honda, T. Takeda, and M. Antoniak, "Unsupervised Segmentation of Hyperspectral Images Using 3D Convolutional Autoencoders," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, pp. 1948–1952, Nov. 2020. arXiv:1907.08870 [cs].

[22] B. Pan, Y. Zheng, R. Shen, T. Ye, W. Zhao, J. Dong, H. Ma, and W. Yuan, "High Resolution Distribution Dataset of Double-Season Paddy Rice in China," *Remote Sensing*, vol. 13, p. 4609, Jan. 2021. Number: 22 Publisher: Multidisciplinary Digital Publishing Institute.

[23] G. Sun, Y. Pan, W. Kong, Z. Xu, J. Ma, T. Racharak, L.-M. Nguyen, and J. Xin, "DA-TransUNet: Integrating Spatial and Channel Dual Attention with Transformer U-Net for Medical Image Segmentation," Nov. 2023. arXiv:2310.12570 [cs, eess].

[24] M. Zhang, L. Liu, Y. Jin, Z. Lei, Z. Wang, and L. Jiao, "Tree-shaped multiobjective evolutionary CNN for hyperspectral image classification," *Applied Soft Computing*, vol. 152, p. 111176, Feb. 2024.

[25] G. Sun, Y. Pan, W. Kong, Z. Xu, J. Ma, T. Racharak, L.-M. Nguyen, and J. Xin, "DA-TransUNet: Integrating Spatial and Channel Dual Attention with Transformer U-Net for Medical Image Segmentation," Nov. 2023. arXiv:2310.12570 [cs, eess].

[26] B. W. Tienin, G. Cui, R. Mba Esidang, Y. A. Talla Nana, and E. Z. Moniz Moreira, "Heterogeneous Ship Data Classification with SpatialChannel Attention with Bilinear Pooling Network," *Remote Sensing*, vol. 15, p. 5759, Dec. 2023.

[27] H. C. Tinega, E. Chen, and D. O. Nyasaka, "Improving Feature Learning in Remote Sensing Images Using an Integrated Deep Multi-Scale 3D/2D Convolutional Network," *Remote Sensing*, vol. 15, p. 3270, June 2023.

[28] F. Ujoh, T. Igbawua, and M. Ogidi Paul, "Suitability mapping for rice cultivation in Benue State, Nigeria using satellite data," *Geo-spatial Information Science*, vol. 22, pp. 332–344, Oct. 2019.

[29] A. Velazquez-Mena, H. Benitez-Perez, R. C. Rodrguez-Martnez, and R. F. Villarreal-Martnez, "DICOMIST: An methodology for Performing Distributed Computing in Heterogeneous ad hoc Networks," *INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL*, vol. 19, July 2024. Number: 4.

[30] K. Venkataramani, C. Z. Marshak, D. Bekaert, M. Simard, M. Denbina, A. L. Handwerger, and S. Chan, "Harmonizing SAR and Optical Data to Map Surface Water Extent: A Deep Learning Approach," *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*, pp. 3349–3352, July 2023. Conference Name: IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium ISBN: 9798350320107 Place: Pasadena, CA, USA Publisher: IEEE.

[31] X. Wang, X. Shi, and F. Ling, "Images difference of ASAR data for rice crop mapping in Fuzhou, China," *Geo-spatial Information Science*, vol. 13, pp. 123–129, Jan. 2010.

[32] X. Wu, J. Lu, J. Wu, and Y. Li, "Multi-Scale Dilated Convolution Transformer for Single Image Deraining," pp. 1–6, Sept. 2023.

[33] Y. Yang, W. Huang, T. Xie, C. Li, Y. Deng, J. Chen, Y. Liu, and S. Ma, "Elevation Gradients Limit the Antiphase Trend in Vegetation and Its Climate Response in Arid Central Asia," *Remote Sensing*, vol. 14, p. 5922, Nov. 2022.

[34] J. Zhang, H. Wu, Z. Zhang, L. Zhang, Y. Luo, J. Han, and F. Tao, "Asian Rice Calendar Dynamics Detected by Remote Sensing and Their Climate Drivers," *Remote Sensing*, vol. 14, p. 4189, Aug. 2022.

[35] G. Zhang, X. Xiao, J. Dong, W. Kou, C. Jin, Y. Qin, Y. Zhou, J. Wang, M. A. Menarguez, and C. Biradar, "Mapping paddy rice planting areas through time series analysis of MODIS land surface temperature and vegetation index data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 106, pp. 157–171, Aug. 2015.

[36] X. Zhang, B. Wu, G. E. Ponce-Campos, M. Zhang, S. Chang, and F. Tian, "Mapping up-to-Date Paddy Rice Extent at 10 M Resolution in China through the Integration of Optical and Synthetic Aperture Radar Images," *Remote Sensing*, vol. 10, p. 1200, Aug. 2018. Number: 8 Publisher: Multidisciplinary Digital Publishing Institute.

[37] X. Zhao, K. Nishina, H. Izumisawa, Y. Masutomi, S. Osako, and S. Yamamoto, "Monsoon Asia Rice Calendar: a gridded rice calendar in monsoon Asia based on Sentinel-1 and Sentinel-2 images," Aug. 2023.

[38] Q. Zhou and A. Ismaeel, "Integration of maximum crop response with machine learning regression model to timely estimate crop yield," *Geo-spatial Information Science*, vol. 24, pp. 474–483, July 2021.

[39] A.-X. Zhu, F.-H. Zhao, H.-B. Pan, and J.-Z. Liu, "Mapping Rice Paddy Distribution Using Remote Sensing by Coupling Deep Learning with Phenological Characteristics," *Remote Sensing*, vol. 13, p. 1360, Jan. 2021. Number: 7 Publisher: Multidisciplinary Digital Publishing Institute.