
A FAST FOURIER CONVOLUTIONAL DEEP NEURAL NETWORK FOR ACCURATE AND EXPLAINABLE DISCRIMINATION OF WHEAT YELLOW RUST AND NITROGEN DEFICIENCY FROM SENTINEL-2 TIME-SERIES DATA

A PREPRINT

Yue Shi

School of Electronic and Electrical Engineering,
The University of Leeds,
Leeds, UK
y.shi1@leeds.ac.uk

Liangxiu Han

Department of Computing and Mathematics
Manchester Metropolitan University
Manchester M1 5GD, UK
L.han@mmu.ac.uk

Pablo González-Moreno

Department of Forest Engineering
University of Córdoba
Córdoba, Spain

Darren Dancy

Department of Computing and Mathematics
Manchester Metropolitan University
Manchester M1 5GD, UK

Wenjiang Huang

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

Zhiqiang Zhang

School of Electronic and Electrical Engineering,
The University of Leeds,
Leeds, UK

Yuanyuan Liu

Department of Computer Science,
The University of Manchester,
Manchester, UK

Mengning Huang

Beijing University of Technology
Beijing, China

Hong Miao

College of Mechanical Engineering,
Yangzhou University,
Yangzhou, China

Min Dai

College of Mechanical Engineering,
Yangzhou University,
Yangzhou, China

July 3, 2023

ABSTRACT

Accurate and timely detection of plant stress is essential for yield protection, allowing better-targeted intervention strategies. Recent advances in remote sensing and deep learning have shown great potential for rapid non-invasive detection of plant stress in a fully automated and reproducible manner. However, the existing models always face several challenges: 1) computational inefficiency and the misclassifications between the different stresses with similar symptoms; and 2) the poor interpretability of the host-stress interaction. In this work, we propose a novel fast Fourier Convolutional Neural Network (FFDNN) for accurate and explainable detection of two plant stresses with similar symptoms (i.e. Wheat Yellow Rust And Nitrogen Deficiency). Specifically, unlike the existing CNN models, the main components of the proposed model include: 1) a fast Fourier convolutional block, a newly fast Fourier transformation kernel as the basic perception unit, to substitute the traditional

convolutional kernel to capture both local and global responses to plant stress in various time-scale and improve computing efficiency with reduced learning parameters in Fourier domain; 2) Capsule Feature Encoder to encapsulate the extracted features into a series of vector features to represent part-to-whole relationship with the hierarchical structure of the host-stress interactions of the specific stress. In addition, in order to alleviate over-fitting, a photochemical vegetation indices-based filter is placed as pre-processing operator to remove the non-photochemical noises from the input Sentinel-2 time series. The proposed model has been evaluated with ground truth data under both controlled and natural conditions. The results demonstrate that the high-level vector features interpret the influence of the host-stress interaction/response and the proposed model achieves competitive advantages in the detection and discrimination of yellow rust and nitrogen deficiency on Sentinel-2 time series in terms of classification accuracy, robustness, and generalization.

Keywords Deep learning · Classification · Sentinel-2 · Winter wheat · Yellow rust · Nitrogen deficiency

1 Introduction

The plant stress caused by unfavorable environmental conditions (e.g. a lack of nutrients, insufficient water, disease, or insect damage), if left untreated, will lead to irreversible damage and decreases in plant production. Early accurate detection of plant stress is essential to be able to respond with appropriate interventions to reverse stress and minimize yield loss. Recent advances in remote sensing with enhanced spatial, temporal and spectral capacities, combined with deep learning have offered unprecedented possibilities for rapid, non-invasive stress detection in a fully automated and reproducible manner [Ji et al., 2018, Wang et al., 2020]. Currently, the deep learning models have been proven effective in remote sensing time-series analysis of plant stresses [Golhani et al., 2018, Rehman et al., 2019]. 1D-CNN and 2D-CNN with convolutions were applied either in the spectral domain or in the spatial domain [Scarpa et al., 2018, Kussul et al., 2017]. In addition, 3D-CNNs were also used across spectral and spatial dimensions [Li et al., 2017, Hamida et al., 2018]. These models do not consider temporal information. Meanwhile, temporal 1D-CNNs were proposed to handle the temporal dimension for general time series classification [Wang et al., 2017] and RNN-based models to extract features from multi-temporal observation by leveraging the sequential properties of multi-spectral data and combination of RNN [Kamilaris and Prenafeta-Boldú, 2018] and 2D-CNNs where convolutions were applied in both temporal and spatial dimensions [Zhong et al., 2019]. These preliminary works highlight the importance of temporal information which can improve the classification accuracy performance. Despite the existing works are encouraging, it suffers several limitations: 1) Over-fitting and uncertainty caused by noisy data involved in the remote sensing time-series; 2) Computing inefficiency and inaccuracy caused by the convolutional operations that are applied to all layer, particularly with the increase of size of images and the kernel. In particular, for multi-plant stresses classification, similar symptoms always lead to confusion during classification as most of the local features are extracted from the neighbor time steps. Therefore, a more effective denoise operator and larger receptive fields for the extraction of the global biological responses at various time-scale are highly desired.

One solution is to pre-filter the photochemical information from satellite time-series and change the domain through Fourier transform to model the part-to-whole relationship between the photochemical features and specific plant stress in the frequency domain. This is because the convolution operation in the spatial domain is the same as the point-by-point multiplication in the Fourier domain. According to the Fourier theory, Fourier transform provides an effective perception operation with non-local receptive field. Unlike existing CNNs where a large-sized kernel is used to extract local features, Fourier transforms with a small-sized kernel can capture global information. Therefore, the Fourier kernel has a great potential in replacing the traditional convolutional kernel in remote sensing time-series analysis without any additional effort [Yi et al., 2023]. For example, Chen et al. [2023] designed a Fourier domain structural relationship analysis framework to exploits both modality-independent local and nonlocal structural relationships for unsupervised change detection. However, the existing Fourier operators can only be sparsely inserted into the deep learning network pipeline due to their expensive computational cost. Therefore, the Fast Fourier Transformation (FFT) is an effective way to extract the global feature responses from satellite image time-series [Nguyen et al., 2020]. For example, Awujoola et al. [2022] proposed a multi-stream fast Fourier convolutional neural network (MS-FFCNN) by utilizing the fast Fourier transformation instead of the traditional convolution, it lowers the computing cost of image convolution in CNNs, which lowers the overall computational cost. Lingyun et al. [2022] designed a spectral deep network combining Fourier Convolution (FFC) and classifier by extending the receptive field. Their results demonstrated that the features around the object provide the explainable information for small objects detection.

Although the effectiveness of Fourier-based convolution has been proved by many studies, few studies have done in the multiple plant stress detection from remote sensing data. In this work, we have proposed a novel fast Fourier convolutional deep neural network (FFCDNN) for accurate and early efficient detection of plant stress with initial focus on wheat yellow rust (*Puccinia striiformis*) and nitrogen deficiency. The proposed model significantly reduce the

computing cost with improved accuracy and interpretability. Specifically, a newly fast Fourier transformation (FFT) kernel is proposed as the basic perception unit of the network to extract the stress-associated biological dynamics with various time-scale; and then the extracted biological dynamics are encapsulated into a series of high-level featured vectors representing the host-stress interactions specific to different stresses; finally, a non-linear activation function is designed to achieve the final decision of the classification. The proposed model has been evaluated with ground truth data under both the controlled and natural conditions.

The rest of this work is organized as follows. Section 2 introduces related works on existing methods of multiple plant disease classification; Section 3 details the proposed approach. Section 4 presents the material and experiment details. Section 5 illustrates the experimental evaluation results. Finally, Section 6 concludes the work.

2 The related work

2.1 Plant photochemical information filter from satellite images

The newly launched satellite sensors (e.g. Sentinel-2, worldview-3, etc.) provide the promising EO dataset for improved plant photochemical estimation [Xie et al., 2018]. Wherein, leaf chlorophyll content (LCC), canopy chlorophyll content (CCC), and leaf area index (LAI) are the most popular remotely retrievable indicators for detecting and discriminating plant stresses [Elarab et al., 2015, Haboudane et al., 2004]. Among these indicators, the LCC time-series is a key biochemical dynamics for the stresses-associated foliar component changes without (or partly) the effects from soil background and canopy structure. Estimating LCC requires the remote sensing indicators that are sensitive to the LCC but, at the same time, is insensitive to LAI and background effects [Elarab et al., 2015]. On the other hand, the LAI is one of the critical biophysics-specific proxies in characterizing the canopy architecture variations that response to the apparent symptom caused by specific stress [Li et al., 2018]. By contrast, CCC is determined by the LAI and LCC, expressed per unit leaf area, which remains multicollinearity with them and hard to be used in separating the stresses-induced biochemical changes from the biophysical impacts. Therefore, the LCC and LAI are regarded as a pair of independent variables for filtering the biochemical information between the different plant stresses [Shi et al., 2017a, Zhang et al., 2012].

Regarding to the filter methods, by using the reflectance in red-edge regions, there are two methods used in LAI and LCC estimation for minimizing saturation effect and soil background associated noises: 1) the vegetation index method [Haboudane et al., 2002, Li et al., 2018] and 2) the radiative transfer models (RTM) [Darvishzadeh et al., Sehgal et al., 2016]. For example, Clevers and Gitelson [2013] tested and compared the performance of the red-edge chlorophyll index (CIred-edge) and green chlorophyll index (CIGreen) on the Sentinel-2 bands, their results indicated that the setting of Sentinel-2 bands is well positioned for deriving these indices on LCC estimation. Punalekar et al. [2018] combined developed a PROSAIL-based model to estimate LAI and biomass on the Sentinel-2 bands, and the yielded LAI values are in agreement with the ground truth LAI measurements. However, the simply use of the remotely estimated LAI and LCC is hard to represent the non-linear host-stress interactions of the plant stresses.

2.2 Plant stress detection methods

Currently, there are two types of methods widely used in extracting the interpretable agent features for the plant stresses from satellite imagery, including the biological methods and the deep learning-based methods.

2.2.1 Biological methods

Studies have shown that biological models can be used to map within-field crop stresses variability [Zhou et al., 2021a, Ryu et al., 2020]. This is possible because the infestation of crop stresses often leads plants to close their stomata, decreasing canopy stomatal conductance and transpiration, which in turn raises foliar biophysical and biochemical variations [Tan et al., 2019]. However, plant stress involves the complicated biophysical and biochemical responses, which demands the unique biological index only partially represent the crop stress. For instance, leaf area index (LAI) is a direct indicator of plant canopy structure features [Ihuoma and Madramootoo, 2019]. The stressed plants will lead to fluctuations on plants LAI time-series with different patterns, which will raise the higher radiations of a stressed crop [Ballester et al., 2019]. Jiang et al. [2020] proposed two LAI-derived soil water stress functions in order to quantify the effect of soil water stress on the processes of leaf expansion and leaf senescence caused by the stresses. Their results showed that the LAI-based model is sensitive to the stress-derived leaf expansion. Zhu et al. [2021] developed a vegetation indices derived model from the observed hyperspectral data of winter wheat to detect the plant salinity, the results show that the salt-sensitive blue, red-edge, and near-infrared wavebands have great performances on the detection of the plant salinity stress.

Unlike the LAI, the Photochemical associated indices directly account for leaf physiological changes such as photosynthetic pigment changes [Gerhards et al., 2019]. Photochemical reflectance is the dominant factor determining leaf reflectance in the visible wavelength (400 – 700 nm), with chlorophyll considered the most relevant photochemical index for crop stress diagnosis [Zhou et al., 2021b]. Under prolonged infestations, leaf chlorophyll content often decreases, leading to a reduction in green reflection and an increase in blue and red reflections. The spectral radiation characteristics between the red and near-infrared regions are sensitive to leaf and canopy chlorophyll content. The ratio of red and near-infrared has shown a strong sensitivity to the crop stress associated chlorophyll content changes [Ryu et al., 2020]. Cao et al. [2019] compared the feasibility of the leaf chlorophyll content, net photosynthesis rate, and maximum efficiency of photosystem on the detection crop heat stress, their findings suggest that the maximum efficiency of photosystem was the most sensitive remote sensing agent to heat stress and had the ability to indicate the start and end of the stress at the slight level or the early stage. Shivers et al. [2019] used the visible-shortwave infrared (VSWIR) spectra to model the non-photosynthetic vegetation and soil background from the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), their finding revealed the increase in temperature residuals is highly consistent with the infestation of crop stresses.

2.2.2 Machine/deep learning-based methods

Despite many studies have been focusing on crop stress detection using the biological characteristics, most of the applications require self-adjusted algorithms to improve the robustness and generalization of the model for the complicated nature conditions. Among the crop stress detection techniques, machine learning and deep learning have played a key role. For machine learning approaches, supervised models have been proved effective in data mining from the training dataset [Kaneda et al., 2017]. The data flow in the machine learning models includes feature extraction, data assimilation, optimal decision boundary searching, and classifiers for stress diagnosis. Wherein, supervised learning is deal with the classification issues by representing the labeled samples. Such models aim to find the optimal model parameters to predict the unlabelled samples [Harrington, 2012].

Deep learning has many neural layers which transforms the sensitive information from the input to output (i.e., healthy or stressed). The most applied perception neural unit is convolutional neural unit in crop stress detection [Fuentes et al., 2017, Krishnaswamy Rangarajan and Purushothaman, 2020]. Generally, the convolutional neural unit consists of dozens of layers that process the input information with convolution kernel. In the area of crop stress detection, deep learning contributed significantly to the analysis of plant stress high-level features [Jin et al., 2018]. In crop stress image classification, the multi-source images are usually used as input to extract the stress dynamics during their developments, and a diagnostic decision is used as output (e.g., healthy or diseased) [An et al., 2019, Cruz et al., 2019]. Barbedo [2019] developed a convolutional deep learning model to classify individual lesions and spots on plant leaves. This model has been successfully used in the identification of multiple diseases, the accuracy obtained in this model was 12% higher than that traditional models. Lin et al. [2019] applied a convolutional kernel based U-Net to segment cucumber powdery mildew-infected cucumber leaves, the proposed binary cross entropy loss function is used to magnify the loss of the powdery mildew stressed pixels, the average accuracy for the powdery mildew detection reaches 96.08%.

2.3 Interpretability of deep learning-based models

Although the deep learning models have been successfully applied for vegetation stress monitoring applications, most of existing deep learning-based approaches have difficulty in explaining plant biophysical and biochemical characteristics due to their black-box representations of the features extracted from intermediate layers [Shi et al., 2020]. Thus, the interpretability of deep models has become one of the most active research topics in the remote sensing based crop stress diagnosis, which can enhance and improve the robustness and accuracy of models in the vegetation monitoring applications from the biological perspective of target entities [Too et al., 2019, Brahimi et al., 2019].

Recently, the model interpretability used to disclose the intrinsic learning logic for detection and discrimination of plant stresses has received growing attention [Lillesand et al., 2015]. In other words, the interpretability that illustrates the performance of the model on characterizing the specific host-stress interaction guarantees the generalization ability of the model for the practise usages. Among the existing models, the visualization of the feature representation is the most direct method for improving the model interpretability. For example, Behmann et al. [2014a] proposed an unsupervised model for early detection of the drought stress in barley, wherein, the intermediate features produced by this model highly related with the sensitive spectral bands for drought stress. Another way to improve the interpretability of deep learning models is to construct the network architecture which can bring the network an explicit semantic meaning. For example, Shi et al. [2020] developed a biologically interpretable two-stage deep neural network (BIT-DNN) for detection and classification of yellow rust from the hyperspectral imagery. Their findings demonstrate that the BIT-DNN has great advantages in terms of the accuracy and interpretability.

2.4 Fast Fourier Transformation

The traditional receptive field used in convolution operations only conduct to the central regions to extract the local features related to the interested targets. This limits the necessity of large convolutional kernel on global feature extraction. Recently, there is an increasing interest in applying Fourier transform to deep nural networks to capture global feature. as mentioned in the introduction section, Fourier transform provides an effective perception operation with non-local receptive field. Unlike existing CNNs where a large-sized kernel is used to extract local features, Fourier transform with a small sized kernel is able to capture global information. For example, [Rippel et al., 2015] proposed Fourier transformation pooling layer that performs like principle component extraction by constructing the representation in the frequency domain. [Chi et al., 2019] proposed a integrated the Fourier transforms into a series of convolution layers in the frequency domain.

Fast Fourier Transformation-based deep learning models use the time-frequency analysis methods to extract the low-frequency host-stress interaction by limiting the high-frequency noises in the frequency-domain space [Ashourloo et al., 2016, Behmann et al., 2014b, Jakubauskas et al., 2002, Mahlein et al., 2017]. FFT is an useful harmonic analysis tool, which has been widely used in reconstruction of vegetation index time-series [Roy and Yan, 2018], curve smoothing [Bradley et al., 2007, Shao et al., 2016], and ecological and phenological applications [Jakubauskas, 2002, Jong et al., 2011, Sakamoto et al., 2005]. FFT maps the satellite time-series signals into superimposed sequences of cosines waves (terms) with variant frequencies, each component term accounting for a percentage of the total variance in the original time-series data [Jakubauskas et al., 2002]. This process facilitates the recognition of subtle patterns of interest from the complex background noises, which degrade the spectral information required to capture vegetation properties[Huang et al., 2018, Shanmugapriya et al., 2019]. For example, El Jarroudi et al. [2017] used the fast Fourier transform (FFT) method to characterize temporal patterns of the fungal disease on winter wheat between the observation sites, and then achieved the fungal disease monitoring and forecasting at the regional level. Our work advances above-mentioned research front via designing an novel fast Fourier convolutional operation unit that simultaneously uses spatial and temporal information for achieving global feature extraction during the learning process.

3 The proposed Fast Fourier Convolution Deep Neural Network (FFCDNN)

To address the challenge of the misclassification for the different plant stresses with similar symptoms, we propose a novel Fast Fourier Convolution operator to efficiently implement non-local receptive field and fuses the extracted biological information with various time-scales in frequency domain, and then, a new deep learning architecture is developed to retrieve the host-stress interaction and achieve the high accurate classification. In this section, we describe the main framework of the proposed Fast Fourier Convolution Deep Neural Network (FFCDNN), in the context of multiple plant stresses discrimination from the agent-based biological dynamics.

3.1 The network architecture of the proposed FFCDNN

Fig.1 depicts the main framework of the proposed FFCDNN for multiple crop stresses discrimination, in the context of Sentinel-2 derived biological agents (i.e. VI_{LAI} and VI_{LCC}). To be specific, a branch structure is designed to respectively pre-filter the biochemical dynamics represented by VI_{LAI} and VI_{LCC} time-series. For each of the branch, the Fourier kernel is set as the same size as the input size of VI_{LAI} and VI_{LCC} time-domain (time-series) patches, with a size of $k \times k \times K^{(1)}$, then, the Fourier kernel is point-wised multiplied by the input biological agent patches. After the Fourier convolution is performed, the ReLU function is implemented to calculate the VI_{LAI} and VI_{LCC} time-series magnitude in frequency-domain containing stress-associated biological responses, and the activation feature map, with a size of $k \times k \times K^{(2)}$, is conducted with Fourier Pool layer to highlight the most important stress information and down sampling the feature map.

Subsequently, the VI_{LAI} and VI_{LCC} feature maps are sent to the hierarchical structure of the class-capsule blocks in order to build the part-to-whole relationship and to generate the hierarchical vector features for representing the high-level stress-pathogen interaction. Finally, a decoder is employed to predict the classes based on the length and direction of the hierarchical vector features in the feature space. The detailed information for the model blocks is described below:

a. Plant photochemical information filter

In this study, an agent-based photochemical information pre-filter is set as the pre-processing operator for the input satellite time-series. Based on the benchmark study of the existing vegetation agent models for LAI and LCC estimation shown in Appendix A, we use the Weighted Difference Vegetation Index (WDVI)-derived LAI, defined as VI_{LAI} , and TCARI/OSAVI-derived LCC, defined as VI_{LCC} , as the optimal plant photochemical information pre-filter on

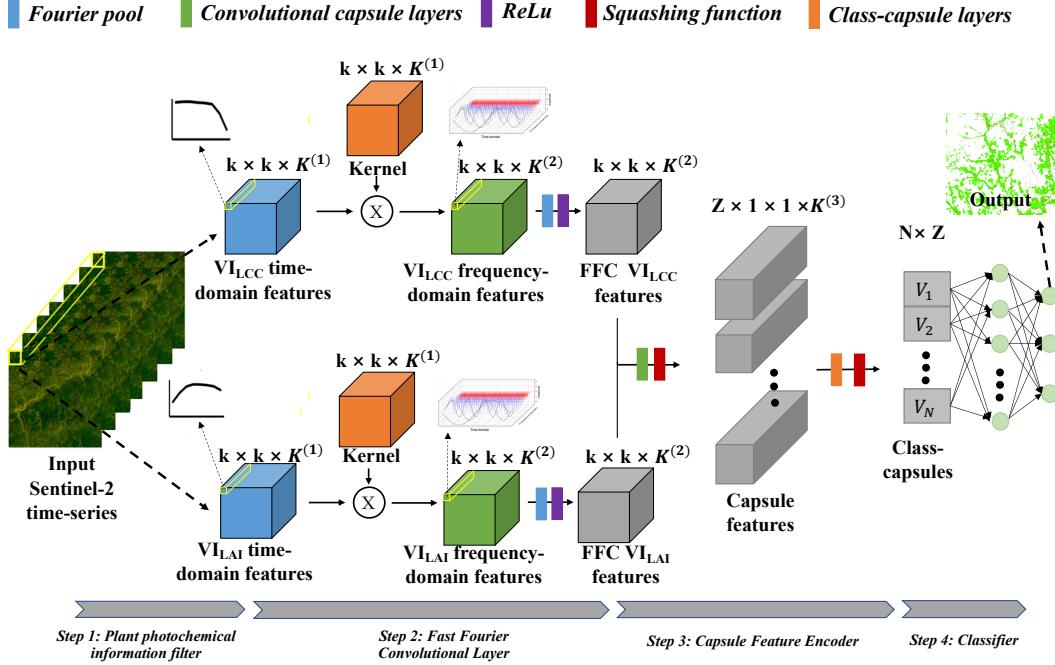


Figure 1: The workflow of the FFCDNN framework for the discrimination of multiple plant stresses from Sentinel-2 time-series

Sentinel-2 bands. And then, the VI_{LAI} and VI_{LCC} time-series will be used as the biological agents of the plant canopy structure and plant biochemical state in the follow analysis.

b. Fast Fourier Convolutional Layer

The input biological agent (i.e. VI_{LAI} or VI_{LCC}) dynamics extracted from the Sentinel-2 time-series can be viewed as a sample patch with $k \times k$ pixel vectors. Each of the pixels represents a class with $K^{(1)}$ time-series channels. Then, the 3D patches with a size of $k \times k \times K^{(1)}$ are extracted as the input of the fast Fourier convolution layer.

The fast Fourier convolution is used to decomposes the biological agent time-series into a series of frequency components with various time-scales based on the fast Fourier transform. Mathematically, FFT decomposes the original time-series signal $f(t)$ to the frequency domain by the linear combination of trigonometric functions as follows:

$$F(\varpi) = \int_{-\infty}^{+\infty} f(t) e^{-i\varpi t} dt \quad (1)$$

Where ϖ is frequency and $F(\varpi)$ is the Fourier coefficient with frequency ϖ , i is the unit of imaginary number. It is customary to use a discrete form as follow:

$$F(x)_{k \times k} = \frac{1}{K^1} \sum_{n=0}^{K^1-1} x_n e^{-\frac{2\pi x n i}{K^1}} \quad (2)$$

Where $x = 0, 1, 2, \dots, N - 1$, and N is the length of time series.

Among the frequency-domain components of the biological agents of VI_{LAI} and VI_{LCC} dynamics, the low-frequency components always indicate the soil background or phenological characteristics of the ground entities. The high-frequency region generally represents environmental noises, such as landcover variations or illumination inconsistency. Therefore, considering the infestation and development of yellow rust and nitrogen deficiency is a continuous biological process on the proxies of VI_{LAI} and VI_{LCC} , we hypothesize that the medium-frequency region represents the yellow rust and nitrogen deficiency associated VI_{LAI} and VI_{LCC} fluctuations. Thus, the yellow rust and nitrogen deficiency associated responses can be characterized from the background and environmental noises by an optimized activation function. In this study, the ReLU activation function is implemented to calculate the VI_{LAI} and VI_{LCC} time-series magnitude in the medium-frequency region, and the activation feature map, with a size of $k \times k \times K^{(2)}$, is conducted

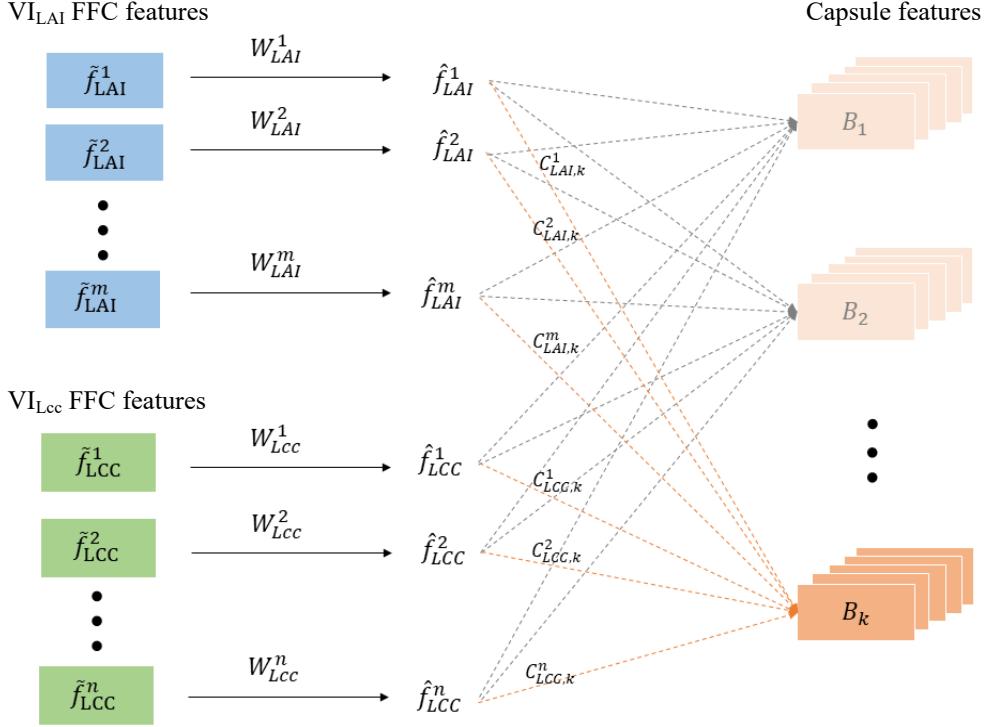


Figure 2: The dynamic routing optimization between the FFC scalar features and the capsule vector features

with Fourier Pool layer to extracted the sensitive VI_{LAI} and VI_{LCC} response in frequency-domain, and output the fast Fourier convolution (FFC) features .

c. Capsule Feature Encoder

Considering the host-stress interaction of the plant stresses is a comprehensive progress that co-represented by the various biological agents. Therefore, modelling the part-to-whole relationship is the most significant evidences for detection and discrimination of plant stresses. We develop a capsule feature encoder to rearrange the extracted VI_{LAI} and VI_{LCC} FFC features, which are the scalar features, into the joint capsule vector features. Thees joint vector features represent the hierarchical structure of the VI_{LAI} and VI_{LCC} responses to the specific plant stress. It's noteworthy that the extracted VI_{LAI} and VI_{LCC} scalar FFC features themselves respectively represent the biophysical and biochemical response to the plant stress development. Therefore, the joint vector features have great performance to characterize the intrinsic entanglement of host-stress interaction. In order to optimize the learning process between the the FFC scalar features and the capsule vector features, and dynamic routing algorithm is introduced as shown in Fig. 2.

Specifically, the VI_{LAI} and VI_{LCC} FFC features, $\{\bar{f}_{LAI}^1, \bar{f}_{LAI}^2, \dots, \bar{f}_{LAI}^m, \bar{f}_{LCC}^1, \bar{f}_{LCC}^2, \dots, \bar{f}_{LCC}^n\}$, are firstly normalized by using the normalization weights $W \in \{W_{LAI}^1, W_{LAI}^2, \dots, W_{LAI}^m, W_{LCC}^1, W_{LCC}^2, \dots, W_{LCC}^n\}$. This step smooths the feature values and makes them obey a normal distribution. In addition, this normalization operation is helpful for retraining the vanishing gradients in the back-propagation progress. After that, the normalized FFC features, $\{\hat{f}_{LAI}^1, \hat{f}_{LAI}^2, \dots, \hat{f}_{LAI}^m, \hat{f}_{LCC}^1, \hat{f}_{LCC}^2, \dots, \hat{f}_{LCC}^n\}$, are rearranged into K^3 capsule features with the coupling coefficients of c . Here, c is a series of trainable parameters that encodes the part-whole relationships between the FFC scalar features and the capsule vector features. The translation and orientation of the capsule vector feature represents the class-specific hierarchical structure characteristics in terms of VI_{LAI} and VI_{LCC} responses in frequency-domain, while its length represents the degree a capsule is corresponding to a class. To measure the length of the output vector as a probability value, a nonlinear squash function is used as follow:

$$\check{u}_m = \frac{\|u_m\|^2}{1 + \|u_m\|^2} \cdot \frac{u_m}{\|u_m\|} \quad (3)$$

wherein, $\check{u}_m^{(l)}$ is the scaled vector of \mathbf{X}_{out}^2 . This function compresses the short vector features to zero and enlarge the long vector features a value close to 1. The final output is denoted as $\mathbf{X}_{out}^3 \in \mathbb{R}^{Z \times 1 \times 1 \times K}$.

Finally, the K^3 capsule features will be weightily combined into Z class capsules and the final outputs are the class-wised biological composed feature = $\{V_1, V_2, \dots, V_Z\}$. In this study, Z is 3 because of three interested classes (i.e. healthy wheat, yellow rust, and nitrogen deficiency).

d. Classifier

Based on the characteristics of the class-capsule feature vectors, a classifier is defined to achieve the final detection and discrimination. This classifier is composed by two layers: an activation layer and a classification layer.

Specifically, The activate function is defined as:

$$\hat{V}_h = \frac{\|V_h\|^2}{1 + \|V_h\|} \cdot \frac{V_h}{\|V_h\|} \quad (4)$$

Where, V_h is the class-capsule features corresponding to class $h \leq Z$. $\|\cdot\|$ indicates the operator of 1-norm. In fact, the orientation of the \hat{V}_h represents the instantiation parameters of the biological responses for the class h , and the length represents the membership that the feature belongs to class h . And then, an argmax function is used to achieve the final classification by seeking the largest length of \hat{V}_h . The argmax function is defined as:

$$\underset{h}{\operatorname{argmax}} O_{i,j}^5 = \{h | \forall g : \|V_g\| < \|V_h\|\} \quad (5)$$

4 Materials and experiments

In this study, we use nitrogen deficiency and the yellow rust as the study cases for model testing and evaluation. In order to comprehensively test and evaluate the classification accuracy, robustness and generalization of the proposed model, we collected two types of the data: 1) the high-quality labelled dataset under the controlled field conditions, 2) the ground survey dataset under the natural field conditions. The former is used for training and optimizing the proposed model, the latter is used for testing and evaluating the generalization and transferability of the well-trained model in the actual application cases. The detailed information is described as below:

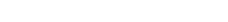
4.1 Study sites

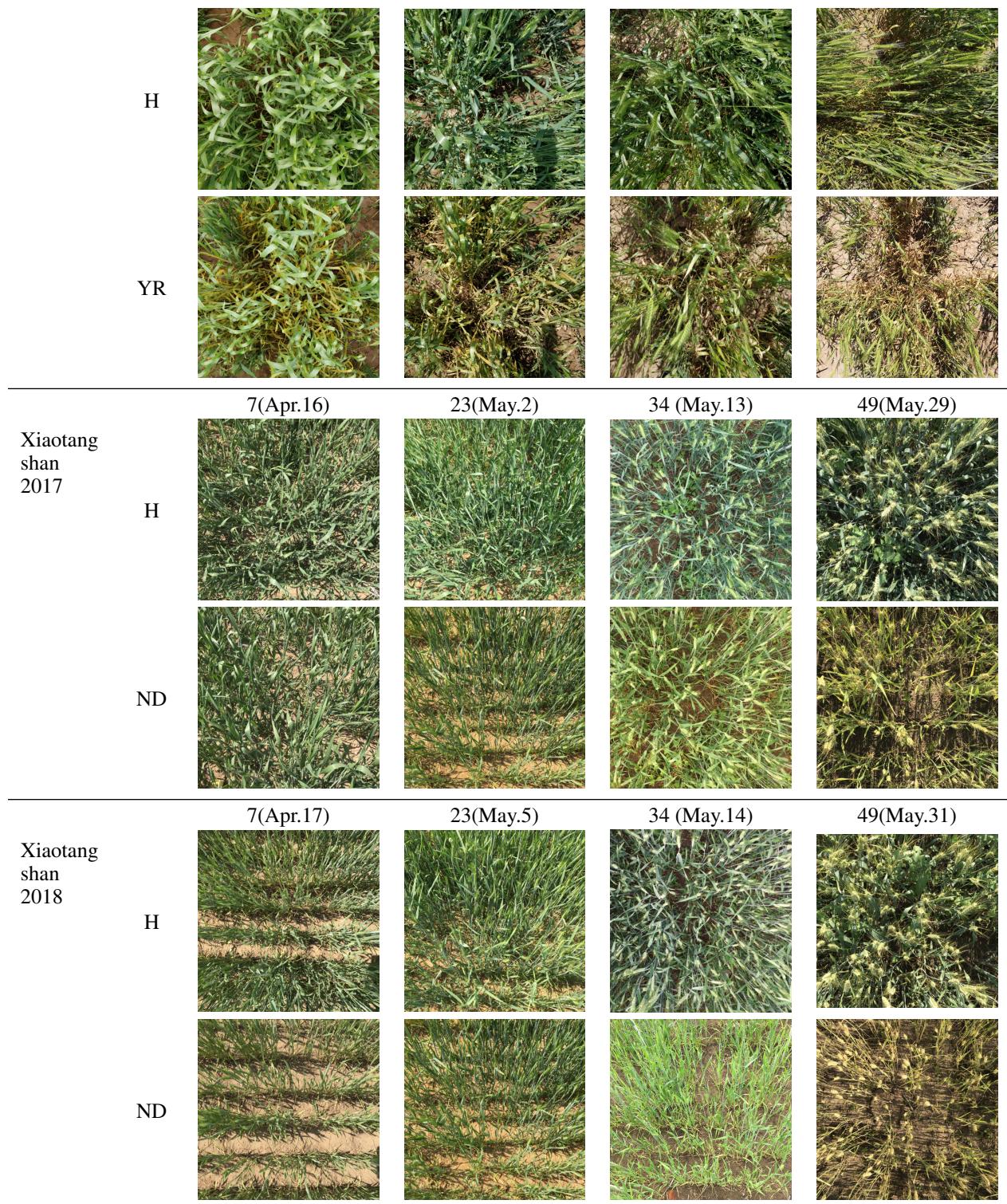
To avoid the fungus contamination on the other groups, we respectively carried out two independent experiments under similar environmental conditions recording continuous in-situ observations of: a) yellow rust infestation from 20th April to 25th May 2017 at the Scientific Research and Experimental Station of Chinese Academy of Agricultural Science (39°30'40"N, 116°36'20"E) in Langfang, Hebei province and b) nitrogen deficiency at the National Experiment Station for Precision Agriculture (40°10'6"N, 116°26'3"E) in Changping District, Beijing, China. The measurement strategies focused on eight key wheat growth stage (i.e. jointing stage, flag leaf stage, heading stage, flowering stage, early grain-filling stage, mid grain-filling stage, late grain-filling stage, and harvest stage). The detailed observation dates and the canopy photos were listed in Table 1. The same experiments were repeated from 18th April to 31th May, 2018.

For the yellow rust experiment, we used the wheat cultivar ‘Mingxian 169’ due to its susceptibility to yellow rust infestation. There was a control group and two infected groups of yellow rust (two replicates of inoculated treatment) were applied. Each field group occupied 220m² of field campaigns in which there were 8 planting rows. For the control group, a total of 8 plots (one plot in each row) with an area of 1m² were symmetrically selected in the field for hyperspectral observations and biophysical measurements. For the disease groups, the concentration levels of 5mg100⁻¹ml⁻¹ and 9mg100⁻¹ml⁻¹ spores solution was implemented to generate a gradient in infestation levels, 8 plots were applied for sampling in each replicate, respectively. All treatments applied 200kg·ha⁻¹ nitrogen and 450m³·ha⁻¹ water at the beginning of planting.

For the nitrogen deficiency experiment in Changping, the popular wheat cultivars, ‘Jingdong 18’ and ‘Lunxuan 167’, were selected. There were two replicates field groups with same nitrogen treatment were applied. Each field group occupied 600m² of field campaigns in which three fertilization levels were used in 21 planting rows of field land (7 rows per treatment) at the beginning of planting, 0kg · ha⁻¹ nitrogen (deficiency group), 100kg · ha⁻¹ nitrogen (deficiency group), and 200kg · ha⁻¹ nitrogen (control group). Similarly to Langfang, all treatments received 450m³ · ha⁻¹ water at planting.

Table 1: The state of vegetation at each measurement dates.

Location (year)	Type	Day after treatment (DAT)			
		7(Apr.20)	14(Apr.27)	23 (May.6)	27(May.10)
Langfang 2017	H				
					
	YR				
					
Langfang 2018	H				
					
	YR				
					



Note: H=healthy YR=yellow rust, ND=nitrogen deficiency

4.2 The simulation of Sentinel-2 bands

The simulated Sentinel-2 bands are regarded as the pure spectral signatures without the effects of atmosphere conditions. For this purpose, the reflectance and transmittances of the sampling plots were firstly collected using an ASD FieldSpec spectroradiometer (Analytical Spectral Devices, Inc., Boulder, CO, USA). In each plot, 10 scans were taken at 1.2m above the wheat canopy. The spectroradiometer was fitted with a 25° field-of-view bare fiber-optic cable, and operated in the 350 – 2,500nm spectral region. The sampling interval was 1.4nm between 350 and 1,050nm, and 2nm between 1,050 and 2,500nm. A white spectral reference panel (99% reflectance) was acquired once every 10 measurements to minimize the effect of possible difference in illumination. Only the bands in the range of 400~1000nm were adopted in this study in order to match the visible-red edge-near infrared bands of Sentinel-2 and avoid bands below 400nm and above 1000nm that were affected by noises [Shi et al., 2017b]. In order to keep radiance consistence, the sampling was conducted at the same period of time between 11:00 and 13:30 local time under a cloud-free sky.

Subsequently, we integrated the field canopy hyperspectral data with the sensor's relative spectral response (RSR) function to simulate the multispectral bands of Sentinel-2. The formula is given as:

$$R_{sentinel-2} = \frac{\int_{\lambda_{start}}^{\lambda_{end}} R_{ground}(\lambda) \cdot RSR(\lambda) dx}{\int_{\lambda_{start}}^{\lambda_{end}} RSR(\lambda) dx} \quad (6)$$

where $R_{sentinel-2}$ is the simulated multispectral channel of Sentinel-2 sensor, λ_{start} and λ_{end} represent the beginning and ending reflectance wavelength of Sentinel-2's corresponding channel, respectively, and R_{ground} is the ground truth canopy hyperspectral data, RSR is the relative spectral response of Sentinel-2 sensor (<https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/document-library/>), both of the R_{ground} and RSR are the functions of wavelength.

4.3 Ground truth plant parameters collection

The plant LAI and LCC were synchronously measured on the same place where the canopy spectral measurements were made. The LCC was measured by the Dualex Scientific sensor (FORCE-A, Inc. Orsay, France), a hand-held leaf-clip sensor designed to non-destructively evaluate the content of chlorophyll and epidermal flavonols. The LCC values were collected with the default unit, which were used preferentially because of the strong relationship between their digital readings and real foliar chlorophyll. Considering the canopy structure-derived multiple scattering process, the first three leaves from the top are regarded as the most effective one with maximum photosynthetic absorption rate, which not only represent the average growth state of the whole plant, but also contribute most to the canopy reflected radiation measured by our observations. Therefore, for each sampling plot, the first, second and third wheat leaves, from the top of ten randomly selected plant (30 leaves for each plot), were chosen for LCC measurements. For the LAI acquisition, the LAI-2200 Plant canopy analyzer (Li-Cor Biosciences Inc., Lincoln, NE, USA) was used in each 1m × 1m subplot.

4.4 Ground truth plant stress severity assessments

In this study, the disease index (DI) was used to measure the severity of yellow rust, and the fertilization level was used to measure the severity of nitrogen deficiency. Specifically, the disease index (DI) was calculated using the method mentioned in Zhang et al. [2012]. It's noted that, because the slight stressed ($DI < 20$) generate invisible influence on wheat yield and do not trigger enough spectral responses on the top of canopy (TOC) reflections of the 10m × 10m Sentinel-2 pixels, the samples with $DI < 20$ were labelled as "healthy wheat", otherwise they were labelled as "yellow rust". In order to guarantee the uniformed bias in each observation, all leaves were manual inspected by the same specially-assigned investigators according to the National Rules for the Investigation and Forecasting of Plant Diseases (GB/T 15795-1995). For nitrogen deficiency, three fertilization levels (i.e. $0kg\text{ha}^{-1}$, $100kg\text{ha}^{-1}$, and $200kg\text{ha}^{-1}$) were controlled in our experiments, here we labelled the fertilization levels of $200kg\text{ha}^{-1}$ as "healthy wheat", otherwise they were labelled as "nitrogen deficiency". The distribution of the collected DI of yellow rust and the fertilization levels of nitrogen deficiency are shown in Fig.3

4.5 The ground survey dataset under the natural field conditions

In order to evaluate the generalization and transferability of the proposed model in the actual applications under the natural conditions, we collected the actual Sentinel-2 time-series and the ground truth data in two different sites, the one is located in the Ningqiang county ($37^{\circ}35'51''N$, $118^{\circ}35'19''E$), Shaanxi province, 2018, and another one is located in Shunyi district ($41^{\circ}20'41''N$, $116^{\circ}24'8''E$), Beijing, 2016. In Ningqiang county, a total of 9 cloud-free Sentinel-2 images and 55 ground truth plots were collected. In Shunyi district, a total of 6 cloud-free Sentinel-2 images and 32 ground truth were collected. All of the collected Sentinel-2 images were atmospherically corrected using the SEN2COR

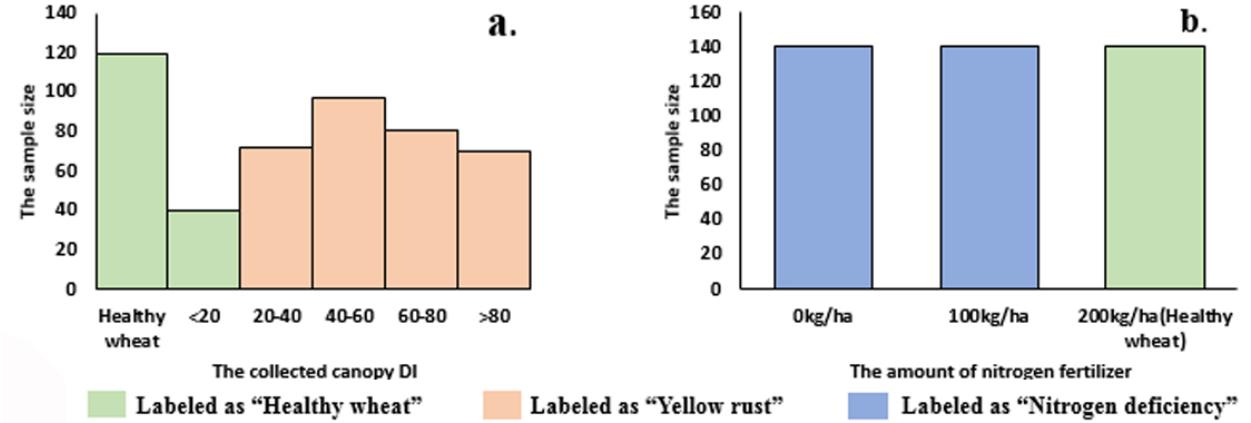


Figure 3: The distribution of the (a) collected disease index (DI) of yellow rust and (b) fertilization levels of nitrogen deficiency

procedure, converting top-of-atmosphere (TOA) reflectance into top-of-canopy (TOC) reflectance. TOC products were the result of a resampling procedure with a constant ground resampling distance of 10 m for visible and near-infrared bands (B2, B3, B4, and B8) and 20m for red-edge bands (B5, B6, B7). The spatial resolution of the red-edge bands (B5, B6, B7) was homogenized to 10m using nearest neighbour resampling. Such process was conducted in the ESA SNAP 6.0 software. The basic principle of the nearest neighbour resampling was described in Roy and Yan [2018]’s study. The overview of the sampling plots and Sentinel-2 collection are shown in Fig.4.

In both surveys, LAI and LCC values were measured by the same approaches used in the experiments under controlled field conditions. Each sample was collected in an area of about $10m \times 10m$ (for corresponding to the spatial resolution of Sentinel-2 bands), of which the centre coordinates were recorded using a GPS with differential correction (accuracy in the order of 2-5 m). The sketch of the sampled site setting is shown in Fig.5.

DIs of yellow rust were measured by the same method used in the experiments under controlled field conditions. In each plot, a plot was labeled as “yellow rust” when $DI > 20$. On the other hand, nitrogen deficiency in each plot was investigated by requesting the history of fertilizer application to the local farmers, a plot was labeled as “nitrogen deficiency” when the history of fertilizer application $< 150Kg/ha$. The statistical distribution of the labeled classes were shown in Fig.6.

5 Results and discussion

In this section, the proposed model is tested and evaluated in three different aspects, including the model performance on detecting and discriminating the yellow rust and nitrogen deficiency, computing efficiency, and the interpretability assessment.

Firstly, to test the performance of the proposed FFCDNN on detection and discrimination of yellow rust and nitrogen deficiency, two popular classification methods, thus, support vector machine (SVM) and convolutional neural network (CNN), are selected for experimental comparison and validation. Specifically, for the configuration of SVM classifier, the radial basis function (RBF) kernel is used in the SVM classification frame, and a grid-based approach proposed by Rumpf et al. [2010] is used to specify the parameter C and δ . For the CNN classifier, a network architecture with four convolutional layers and two fully connect layers proposed by Chen et al. is employed, all the hyperparameters of the CNN classifier have been optimized for the experiments. Regarding the model assessments, five widely used metrics, thus the average accuracy, producer’s accuracy, user’s accuracy, Kappa value, and computing time, are employed in this study to evaluate the classification accuracy. The definition of these matrices can be find in [Mahlein et al., 2017]

Secondly, for the interpretability assessment of the model, a post-hoc analysis is used to expose the learning process and feature representations of the data life in the proposed model. Specifically, a canonical discriminant analysis is first used to measure the intra-class distance and the separability in each learning stage of the model. The definition of the canonical discriminant analysis is described in our previous study [Shi et al., 2017a]. And then, the coefficients of

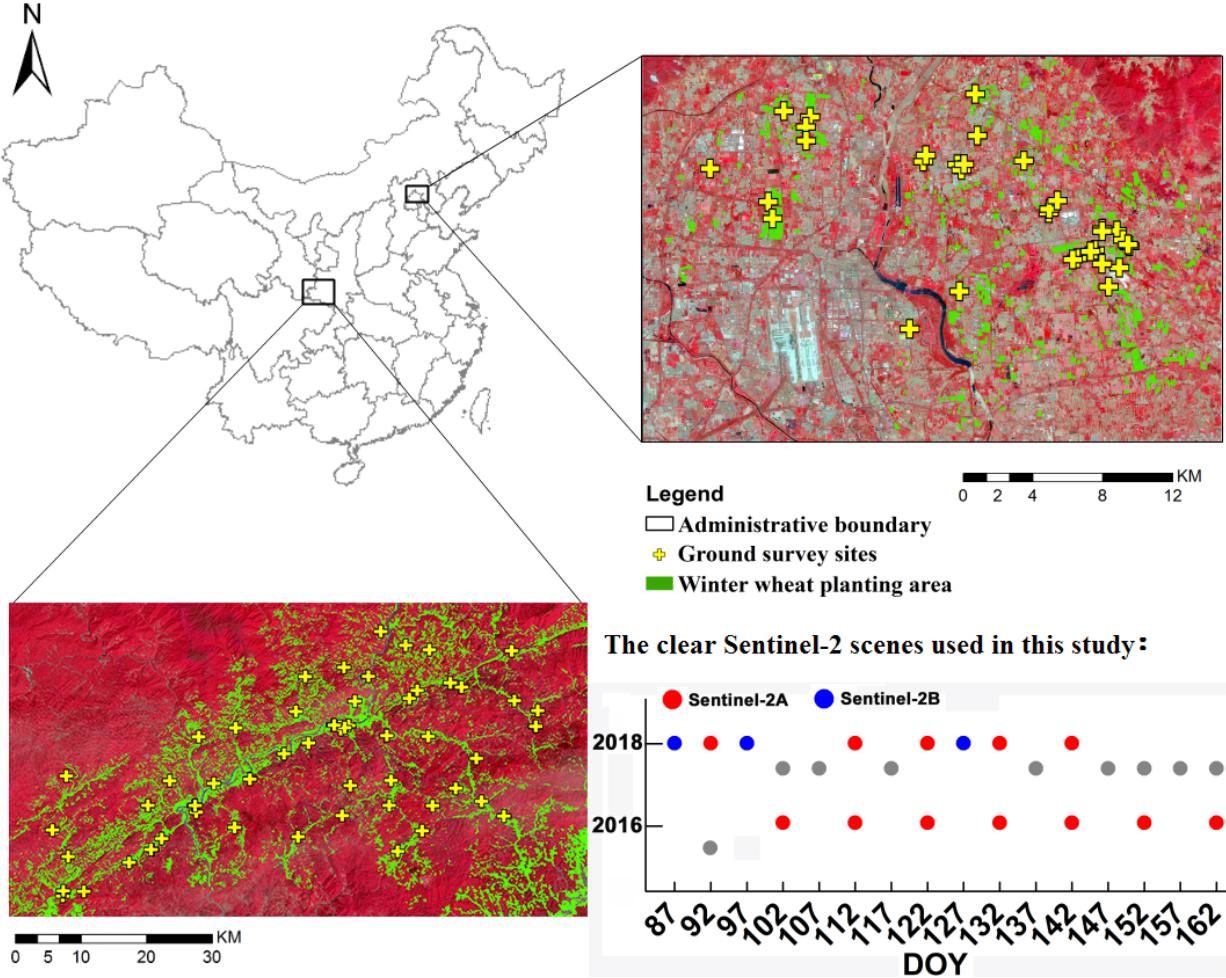


Figure 4: False-color maps of the experimental sites of Ningqiang county, Shaanxi, (bottom left) and Shunyi district, Beijing (top right). And the overview of the sentinel-2 imagery used.

determination (R^2) between the generated biological composed features and the ground-measured severity of yellow rust and nitrogen deficiency are calculated based on univariate correlation analysis.

5.1 Model test on detection and discrimination of the yellow rust and nitrogen deficiency

5.1.1 Experiment one: model testing on the simulated Sentinel-2 bands under controlled field conditions

The first experiment is to evaluate the performance of the proposed model on the detection and discrimination of yellow rust and nitrogen based on the simulated Sentinel-2 bands under controlled field conditions. For model training and validation, a 5-folder cross validation is employed to measure the classification accuracy, and the computing time (CT) is used to measure the computing efficiency. The comparison of the classifications of the proposed FFCDNN, SVM, and CNN is shown in Table. 2. Our results show that the proposed FFCDNN achieves the best classification in terms of accuracy and Kappa value on both of training dataset (with the overall accuracy of 92.8% and Kappa of 0.891) and validation dataset (with the overall accuracy of 87.5% and Kappa of 0.812). The misclassification mainly occurs between the healthy wheat and nitrogen deficiency. In the term of computing efficiency, although the computing time of the proposed model is not the best among the baseline, it is highly improved from the traditional convolution-based deep learning model.

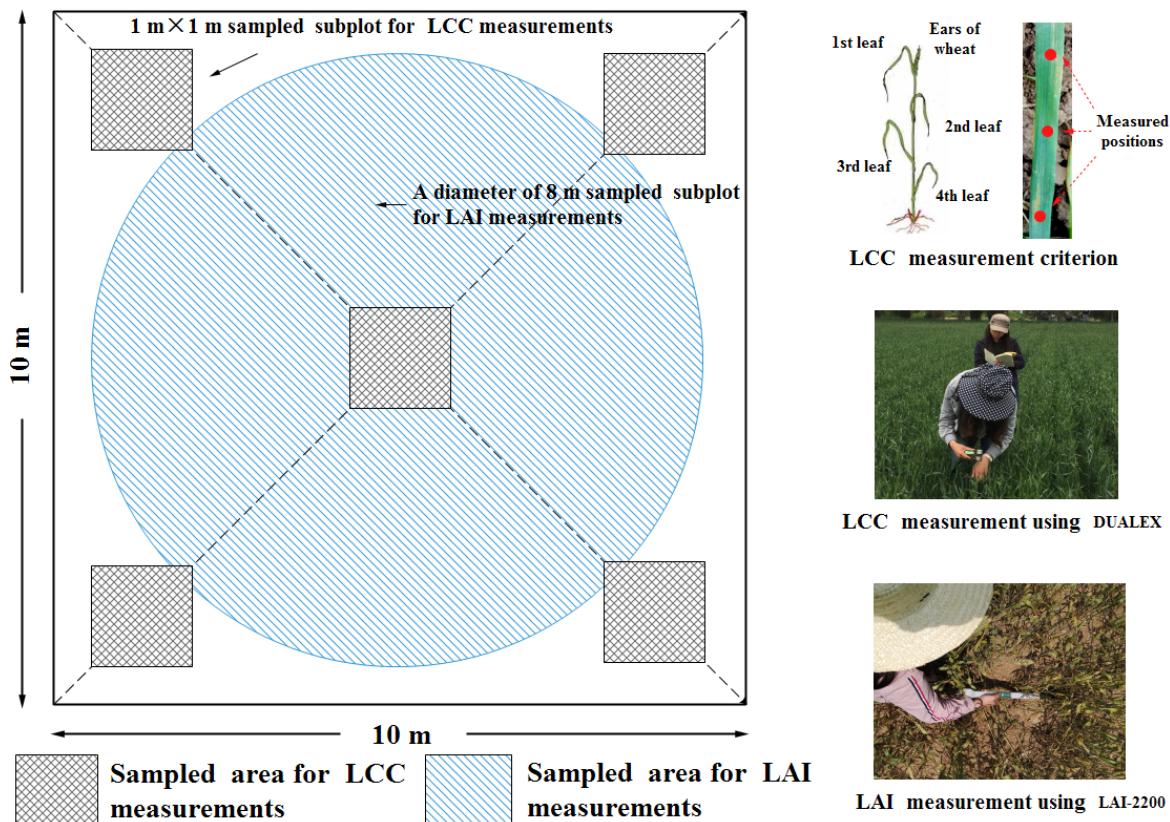


Figure 5: The measurement sketch of the synchronously ground LAI and LCC truth data collection.

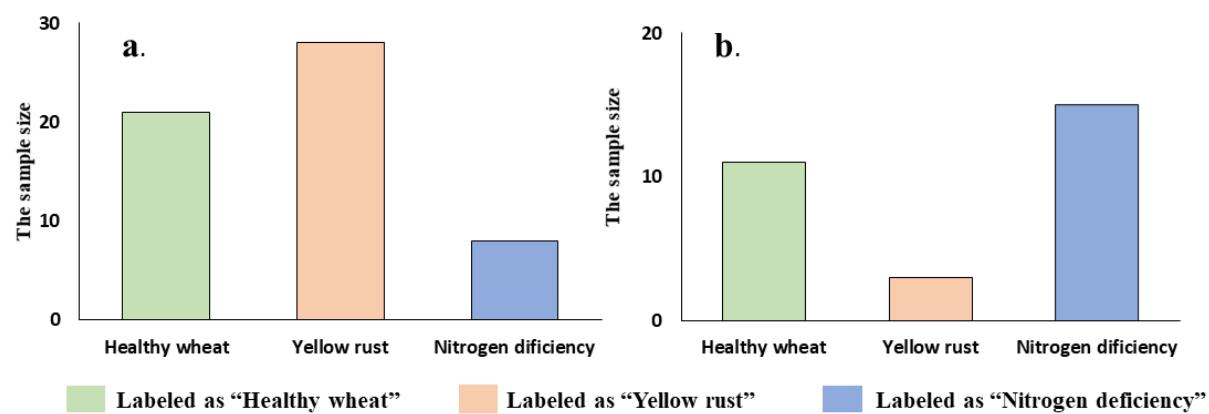


Figure 6: The distribution of the labeled classes in (a) Ningqiang and (b) Shunyi.

Table 2: Accuracy evaluation for the classification of yellow rust and nitrogen based on the simulated Sentinel-2 bands under controlled field conditions .

		Training					Validation				
SVM	Health	160	10	12	87.9	OA(%)	97	8	15	80.8	
	YR	6	169	8	92.3		7	106	13	84.1	
	NS	14	13	148	84.6	88.3	16	14	84	73.7	79.7
	P(%)	88.9	88	88.1			80.8	82.8	75		
	Kappa		0.824					0.695			
	CT(s)		108.7						32.5		
		Training					Validation				
CNN	Health	165	7	11	90.2	OA(%)	101	6	10	86.3	
	YR	4	180	6	94.7		7	109	12	85.2	
	NS	11	5	151	90.4	91.9	12	13	90	78.3	83.3
	P(%)	91.7	93.8	89.9			84.5	85.2	80.4		
	Kappa		0.877					0.749			
	CT(s)		481.5						71.2		
		Training					Validation				
FFCDNN	Health	168	5	9	92.3	OA(%)	103	5	8	88.8	
	YR	5	181	7	93.8		8	115	7	88.5	
	NS	7	6	152	92.1	92.8	9	8	97	85.1	87.5
	P(%)	93.3	94.3	90.5			85.8	89.8	86.6		
	Kappa		0.891					0.812			
	CT(s)		277.4						46.2		

5.1.2 Experiment two: model test on the actual Sentinel-2 bands under natural filed conditions

The second experiment aims to further evaluate the robustness and transferability of the proposed model on the actual Sentinel-2 images under natural field conditions. For this purpose, the pre-trained models in the last section are directly used in the pixel-wise classification of yellow rust and nitrogen deficiency on the actual Sentinel-2 time-series in Ningqiang county and Shunyi district, and the ground truth samples are used as validation. The accuracy assessment of the pre-trained SVM, CNN, and FFCDNN are listed in Table. 3. Our results illustrate that the well-trained FFCDNN achieves an overall accuracy of 80.7% (Kappa = 0.69) for the model test in Ningqiang, and an overall accuracy of 79.3% (Kappa = 0.644) for the model test in Shunyi district. These accuracies are basically consistent with the results reported in experiment one. In comparison, there are evident declines could be figured out in the classification performance of SVM and CNN model, the average decline in term of overall accuracy respectively reaches 52.7% for SVM and 36.3% for CNN. Overall, these results suggest that the proposed FFCDNN provides a more stable and robust performance than the comprtitors on the classification of yellow rust and nitrogen deficiency in different field conditions.

Table 3: The accuracy assessment of the pre-trained models on actual Sentinel-2 time-series on classification of yellow rust and nitrogen deficiency in Ningqiang county of Shaanxi and Shunyi distinct of Beijing.

		Ningqiang					Shunyi				
SVM	Health	9	8	4	42.9	OA(%)	4	4	5	40	
	YR	6	14	1	66.7		3	1	4	12.5	
	NS	6	6	3	20	45.6	4	1	6	54.5	37.9
	P(%)	42.9	50	37.5			36.4	33.3	40		
	Kappa		0.891					0.812			
	CT(s)		277.4						46.2		

		Ningqiang					Shunyi				
		Health	YR	NS	U(%)	OA(%)	Health	YR	NS	U(%)	OA(%)
CNN	Health	11	5	3	57.9		5	1	4	50	
	YR	4	18	1	78.3	57.9	2	2	2	33.3	55.2
	NS	6	5	4	26.7		4	0	9	69.2	
	P(%)	52.4	64.3	50			45.5	66.7	60		
	Kappa			0.344				0.272			
	CT(s)			548.1				102.5			
		Ningqiang					Shunyi				
FFCDNN	Health	16	1	2	84.2	OA(%)	9	1	2	75	
	YR	2	24	0	92.3		0	2	1	66.7	
	NS	3	3	6	50	80.7	2	0	12	85.7	79.3
	P(%)	76.2	85.7	75			81.8	66.7	80		
	Kappa			0.69				0.644			
	CT(s)			268.7				81.6			

For the demonstration purpose, the FFCDNN-based classification map of the yellow rust and nitrogen deficiency in Ningqiang and Shunyi are respectively illustrated in Fig. 7 and Fig. 8. The spatial distributions of yellow rust and nitrogen deficiency in Ningqiang and Shunyi are consistent with our field survey. Specifically, for Ningqiang case, the yellow rust is mainly located around the river where ideal moisture is provided for the infestation and development of yellow rust (see the zoom-in window in Fig. 7), the nitrogen deficiency is distributed around the edge of the county where the high transportation cast results in the poor fertilization management. For the Shunyi case, the nitrogen deficiency mainly occurs in the edge of the field patches (see the zoom in window in Fig. 8), the yellow rust slightly occurs in the west of the study area. These monitoring results are double-checked through telephone interviews with the local plant protection department.

5.2 The interpretability assessment of the model

The interpretability is one of the important matrices that measuring bias and provide explainable reason for prediction decisions from model. In this study, the interpretability assessment mainly focus on the data life in the proposed FFCDNN model and the representations of the intermediate features.

5.2.1 The data life in the proposed FFCDNN model

In this study, two significant modules are proposed to characterize the yellow rust- and nitrogen deficiency-associated information from the Sentinel-2 time-series, thus, 1) the FFC features extraction and 2) the capsule feature generation. In order to evaluate the effects of each module on the inter-class separability, we conduct a canonical discriminant analysis to measure the clusters of the intermediate features. In the canonical discriminant analysis, the first two canonical discriminant functions are employed to establish the projective scatter plots. In addition, we gradually add the modules into the FFCDNN framework and compare their effects on classification accuracy. The visualization of the comparison is illustrated in Fig.9.

a. The base model without the characterized modules

The base model architecture without the characterized modules is similar to a multi-layer perception (MLP), thus, the VI_{LAI} and VI_{LCC} time-series produced by the biological feature retrieval layer $L^{(1)}$ will directly input into the classifier $L^{(5)}$. The inter-class separability of the time-series features is shown in the second column of Fig.9, and the overall accuracy achieved by the base model is approximately 51.7%.

b. Adding the FFC layer

In the FFCDNN, the FFC features extraction is the most important step to extract the yellow rust and nitrogen deficiency associated VI_{LAI} and VI_{LCC} frequency-domain features from the background noises. The canonical discriminant analysis indicates that, by comparison with the time-series features, the extracted frequency-domain features reveal the

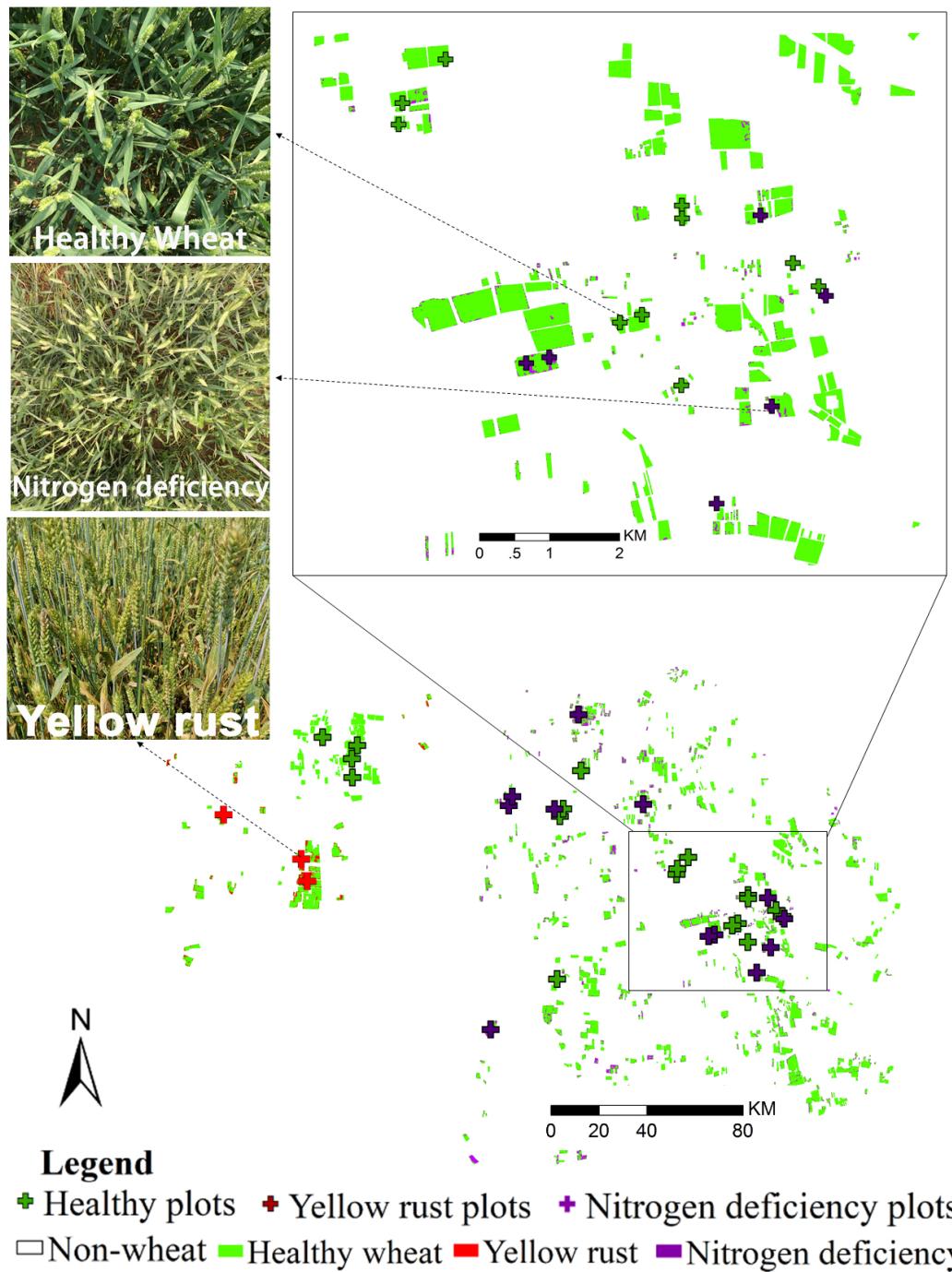


Figure 7: The occurrence monitoring and mapping of yellow rust in Ningqiang county, Shaanxi province (the window shows the zoom in for the classification on the sub-region).

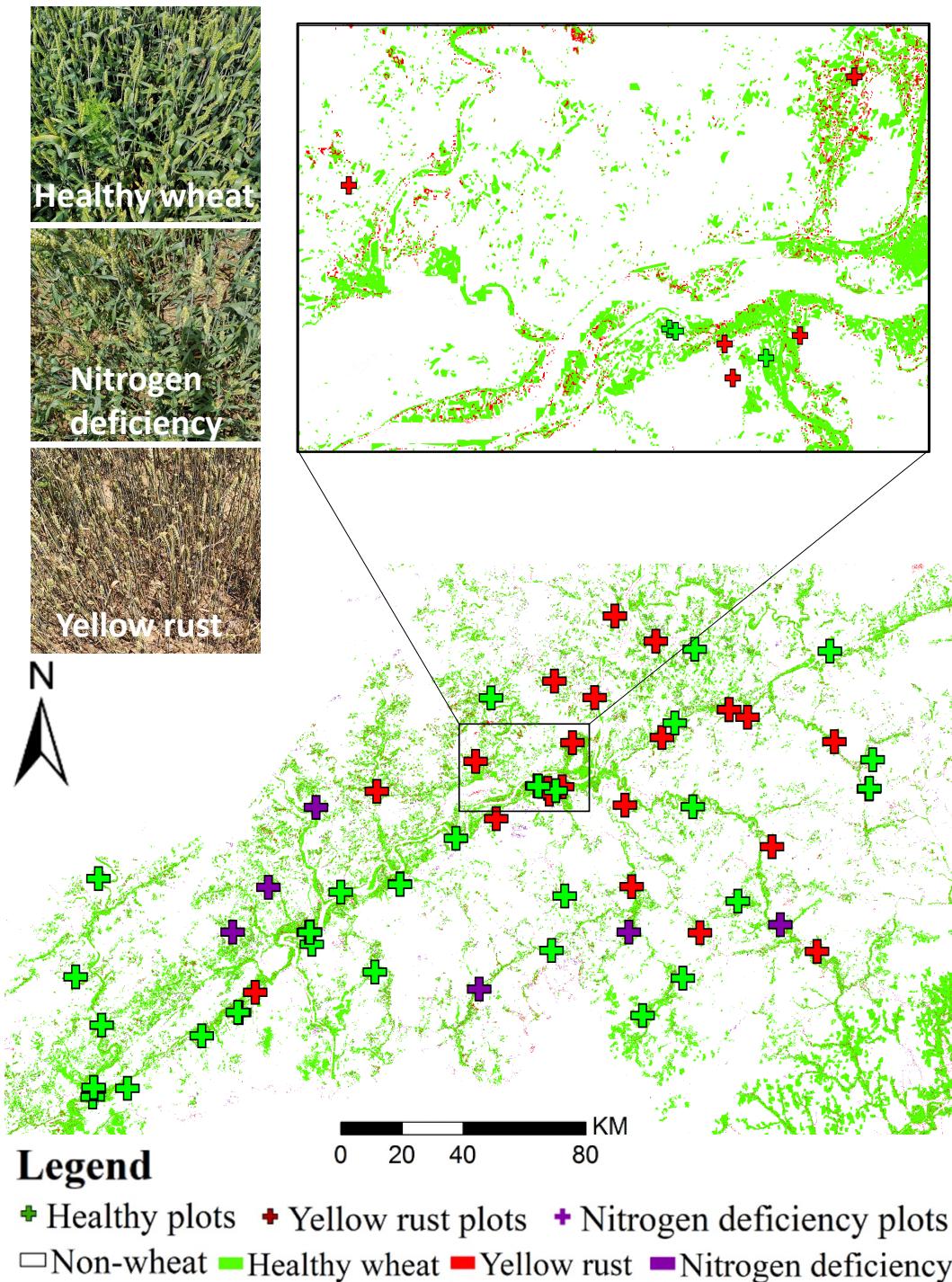


Figure 8: The detection and discrimination of yellow rust and nitrogen deficiency in Shunyi district, Beijing (the window shows the zoom in for the classification on the sub-region).

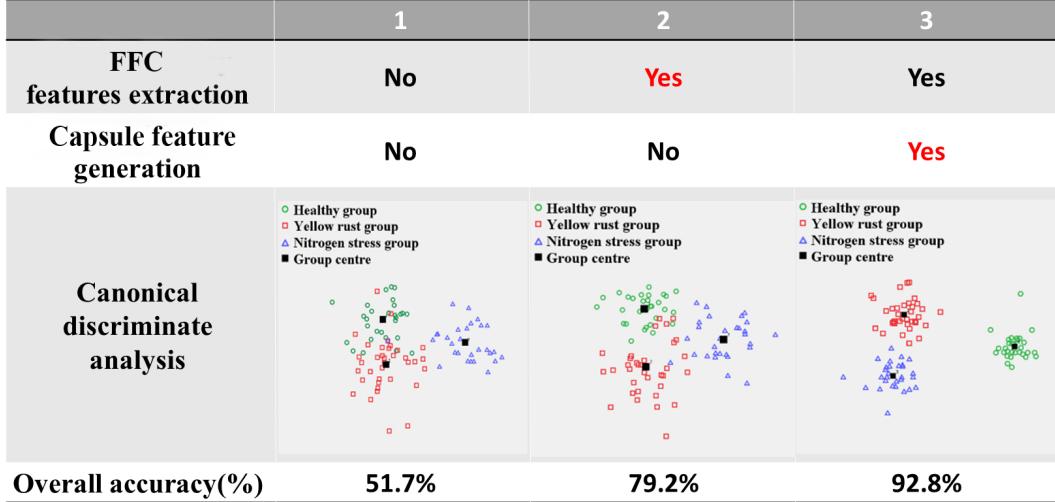


Figure 9: The visualization of the comparison for showing the effects of each module in FFCDNN on the canonical discriminate analysis and overall accuracy. Each column is a model with the modules in the top. The red highlight the main difference of the current model with the previous one.

greater clusters between the different classes (the third column of Fig.9), and the overall accuracy reaches approximately 79.2%.

c. Adding the capsule feature encoder

The capsule feature encoder is the most intelligent part of the proposed FFCDNN, which encapsulates the extracted scalar biological features into the vector features with the explicit biological representation of the target classes. The evident clusters and class edges can be figured out in the canonical projected scatter plot (the firth column of Fig.9), and the final overall accuracy reaches 92.8%.

5.2.2 The representations of the intermediate features

The primary contribution of this study is to model the part-to-whole relationship between the Sentinel-2-derived biological agents (i.e. VI_{LAI} and VI_{LCC}) and the specific stresses, by encapsulating the scalar FFC features into the low-level class-associated vector structures. The philosophy behind the biologically composed features is that the vector features provide a hierarchical structure to represent the entanglement of the VI_{LAI} and VI_{LCC} fluctuations associated with yellow rust and nitrogen deficiency, and provide evidence for detection and discrimination of yellow rust and nitrogen deficiency.

The coefficients of determination (R^2) between the components of the generated biological composed features and the ground-measured severity of yellow rust and nitrogen deficiency are calculated based on univariate correlation analysis (see Fig.10). It's noted that, according to Nyquist theorem, the maximum frequency component after FFT is 26 HZ, thus, the dimensionality of the generated biological composed features will be less than 52. Our results illustrate that, for yellow rust, both the VI_{LAI} and VI_{LCC} frequency features located in the low-frequency regions (2-4 Hz) highly relate with the severity levels of yellow rust, which means the host-pathogen interaction of yellow rust may induce the chronic impacts on the VI_{LAI} and VI_{LCC} fluctuation. These findings are in agreement with the biophysical and pathological characteristics of yellow rust that were reported in our previous study Shi et al. [2018]. For nitrogen deficiency, the associated VI_{LAI} fluctuations are mainly located in the frequency regions of 5-15 Hz, and the associated VI_{LCC} fluctuations are mainly located in the frequency regions of 6-13 Hz. Which means the nitrogen deficiency may give rise to a more acute VI_{LAI} and VI_{LCC} responses than that of yellow rust on the Sentinel-2 time-series. For instance, as reported in Behmann et al. [2014b], the occurrence of nitrogen deficiency in green plants is associated with the poor photosynthesis rates, and further lead to abnormal LAI and LCC (i.e. reduced growth and chlorotic leaves).

6 Conclusion

The proposed FFCDNN model differs from existing approaches to the detection and discrimination of multiple plant stresses in the following three aspects: 1) Our model primarily considers plant biochemical information specific to the

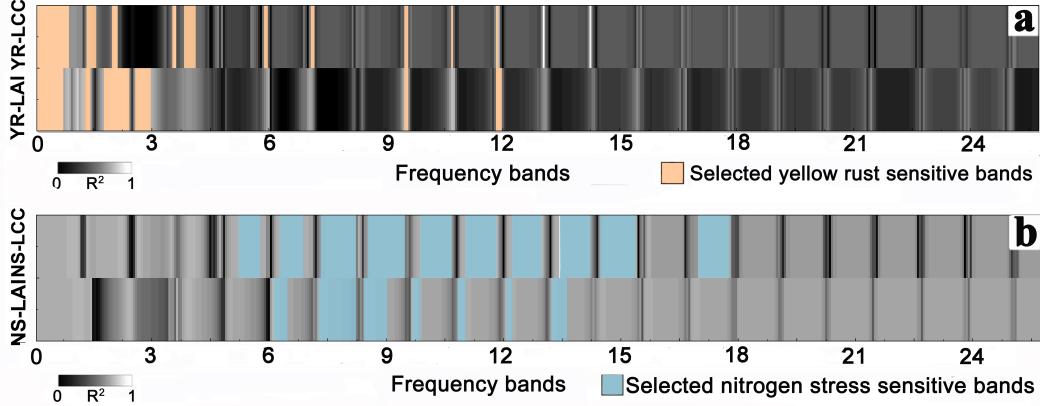


Figure 10: The coefficients of determination (R^2) between the components of the generated biological composed features and the ground measured severity of (a) yellow rust and (b) nitrogen deficiency.

stresses. 2) The proposed Fast Fourier convolution kernel represents the first attempt to use the FFT-based kernel in a deep neural network for biological dynamic extraction from the Sentinel-2 time-series. 3) The well-designed capsule feature encoder demonstrates excellent performance in modeling the part-to-whole relationship between the extracted biological dynamics and the host-stress interaction. These three characteristics improve the interpretability of our model for decision making, akin to human experts.

However, two challenges persist in the practical use of the proposed implementation. Firstly, the performance of our model is inherently limited by the accurate extraction of the biochemical pre-filter. The Sentinel-2 based VI_{LAI} and VI_{LCC} estimations struggle to represent the real LAI and LCC values accurately, leading to the mis-estimation of the biological dynamics of specific stresses. Secondly, errors from the gap conditions and the co-registration of Sentinel-2 imagery introduce uncertainty in the modeling processes. These are primary reasons for the performance decline in the practical application of the FFCDNN. Future research will investigate whether integrating information provided by multi-source satellites into the FFCDNN framework could compensate for the LAI and LCC estimations and gap-related error, thereby further improving accuracies in detecting and discriminating yellow rust and nitrogen deficiency.

In conclusion, modeling the biochemical progress of specific plant stress is a key factor that influences the effectiveness of deep learning applications in the remote sensing detection and discrimination of multiple plant stresses. In this study, we proposed the FFCDNN model to analyze the stress-associated VI_{LAI} and VI_{LCC} biological responses from Sentinel-2 time-series to achieve multiple plant classification at the regional level. Comparisons with state-of-the-art models reveal that the proposed FFCDNN exhibits competitive performance in terms of classification accuracy, robustness, and generalization ability.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

YS planned the study, designed the field experiments, developed the algorithm, and drafted the manuscript. LXH and DD reviewed, edited, conducted interviews and supervised the manuscript and lead the revision. PGM and WJH prepared and conducted interviews, reviewed and edited the manuscript and conducted interviews. ZQZ, YYL and MN provided literature reviews, HM and MD reviewed and edited the manuscript. All authors improved the manuscript by responding to the review comments. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by BBSRC (BB/R019983/1), BBSRC (BB/S020969/1), and Jiangsu Provincial Key Research and Development Program-Modern Agriculture (Grant No. BE2019337) and Jiangsu Agricultural Science and Technology Independent Innovation (Grant No. CX(20)2016).

Acknowledgments

The authors would like to thank Dr. Bo Liu for providing the field for our experiments in Langfang in this study.

References

- Shunping Ji, Chi Zhang, Anjian Xu, Yun Shi, and Yulin Duan. 3d convolutional neural networks for crop classification with multi-temporal remote sensing images. *Remote Sensing*, 10(1):75, 2018.
- Senzhang Wang, Jiannong Cao, and Philip Yu. Deep learning for spatio-temporal data mining: A survey. *IEEE transactions on knowledge and data engineering*, 2020.
- Kamlesh Golhani, Siva K Balasundram, Ganesan Vadmalai, and Biswajeet Pradhan. A review of neural networks in plant disease detection using hyperspectral data. *Information Processing in Agriculture*, 5(3):354–371, 2018. ISSN 2214-3173.
- Nabeel Abdur Rehman, Umar Saif, and Rumi Chunara. Deep landscape features for improving vector-borne disease prediction. *arXiv preprint arXiv:1904.01994*, 2019.
- Giuseppe Scarpa, Massimiliano Gargiulo, Antonio Mazza, and Raffaele Gaetano. A cnn-based fusion method for feature extraction from sentinel data. *Remote Sensing*, 10(2):236, 2018.
- Natalia Kussul, Mykola Lavreniuk, Sergii Skakun, and Andrii Shelestov. Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, 14(5):778–782, 2017.
- Ying Li, Haokui Zhang, and Qiang Shen. Spectral–spatial classification of hyperspectral imagery with 3d convolutional neural network. *Remote Sensing*, 9(1):67, 2017.
- Amina Ben Hamida, Alexandre Benoit, Patrick Lambert, and Chokri Ben Amar. 3-d deep learning approach for remote sensing image classification. *IEEE Transactions on geoscience and remote sensing*, 56(8):4420–4434, 2018.
- Zhiguang Wang, Weizhong Yan, and Tim Oates. Time series classification from scratch with deep neural networks: A strong baseline. In *2017 International joint conference on neural networks (IJCNN)*, pages 1578–1585. IEEE, 2017.
- Andreas Kamaras and Francesc X Prenafeta-Boldú. Deep learning in agriculture: A survey. *Computers and electronics in agriculture*, 147:70–90, 2018.
- Liheng Zhong, Lina Hu, and Hang Zhou. Deep learning based multi-temporal crop classification. *Remote sensing of environment*, 221:430–443, 2019.
- Kun Yi, Qi Zhang, Shoujin Wang, Hui He, Guodong Long, and Zhendong Niu. Neural time series analysis with fourier transform: A survey. *arXiv preprint arXiv:2302.02173*, 2023.
- Hongruixuan Chen, Naoto Yokoya, and Marco Chini. Fourier domain structural relationship analysis for unsupervised multimodal change detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 198:99–114, 2023.
- Minh D Nguyen, Oscar M Baez-Villanueva, Duong D Bui, Phong T Nguyen, and Lars Ribbe. Harmonization of landsat and sentinel 2 for crop monitoring in drought prone areas: Case studies of ninh thuan (vietnam) and bekaa (lebanon). *Remote Sensing*, 12(2):281, 2020.
- Olalekan J Awujoola, PO Odion, AE Evwiekpaefe, and GN Obunadike. Multi-stream fast fourier convolutional neural network for automatic target recognition of ground military vehicle. In *Artificial Intelligence and Applications*, 2022.
- Gu Lingyun, Eugene Popov, and Dong Ge. Spectral network combining fourier transformation and deep learning for remote sensing object detection. In *2022 International Conference on Electrical Engineering and Photonics (EExPolytech)*, pages 99–102. IEEE, 2022.
- Qiaoyun Xie, Jadu Dash, Wenjiang Huang, Dailiang Peng, Qiming Qin, Hugh Mortimer, Raffaele Casa, Stefano Pignatti, Giovanni Laneve, and Simone Pascucci. Vegetation indices combining the red and red-edge spectral information for leaf area index retrieval. 2018.

- Manal Elarab, Andres M Ticlavilca, Alfonso F. Torres-Rua, Inga Maslova, and Mac McKee. Estimating chlorophyll with thermal and broadband multispectral high resolution imagery from an unmanned aerial system using relevance vector machines for precision agriculture. *International Journal of Applied Earth Observations and Geoinformation*, 43:32–42, 2015.
- Driss Haboudane, John R Miller, Elizabeth Pattey, Pablo J Zarco-Tejada, and Ian B Strachan. Hyperspectral vegetation indices and novel algorithms for predicting green lai of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, 90(3):337–352, 2004.
- H. Li, G. Liu, Q. Liu, Z. Chen, and C. Huang. Retrieval of winter wheat leaf area index from chinese gf-1 satellite data using the prosail model. *Sensors*, 18(4):1120, 2018.
- Yue Shi, Wenjiang Huang, Juhua Luo, Linsheng Huang, and Xianfeng Zhou. Detection and discrimination of pests and diseases in winter wheat based on spectral indices and kernel discriminant analysis. *Computers and Electronics in Agriculture*, 141:171–180, 2017a.
- Jingcheng Zhang, Ruiliang Pu, Wenjiang Huang, Yuan Lin, Juhua Luo, and Jihua Wang. Using in-situ hyperspectral data for detecting and discriminating yellow rust disease from nutrient stresses. *Field Crops Research*, 134(3):165–174, 2012.
- Driss Haboudane, John R. Miller, Nicolas Tremblay, Pablo J. Zarco-Tejada, and Louise Dextraze. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment*, 81(2):416–426, 2002.
- R Darvishzadeh, AK Skidmore, Tiejun Wang, and A Vrieling. Evaluation of sentinel-2 and rapideye for retrieval of lai in a saltmarsh using radiative transfer model. In *ESA Living Planet Symposium 2019*.
- Vinay Kumar Sehgal, Debasish Chakraborty, and Rabi Narayan Sahoo. Inversion of radiative transfer model for retrieval of wheat biophysical parameters from broadband reflectance measurements. *Information Processing in Agriculture*, 3(2):107–118, 2016. ISSN 2214-3173.
- J. G. P. W. Clevers and A. A. Gitelson. Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on sentinel-2 and -3. *International Journal of Applied Earth Observations and Geoinformation*, 23 (8):344–351, 2013.
- Suvarna M Punalekar, Anne Verhoef, Tristan L Quaife, David Humphries, Louise Birmingham, and Chris K Reynolds. Application of sentinel-2a data for pasture biomass monitoring using a physically based radiative transfer model. *Remote sensing of environment*, 218:207–220, 2018. ISSN 0034-4257.
- Yongcai Zhou, Congcong Lao, Yalong Yang, Zhitao Zhang, Haiying Chen, Yinwen Chen, Junying Chen, Jifeng Ning, and Ning Yang. Diagnosis of winter-wheat water stress based on uav-borne multispectral image texture and vegetation indices. *Agricultural Water Management*, 256:107076, 2021a.
- Jae-Hyun Ryu, Hoejeong Jeong, and Jaeil Cho. Performances of vegetation indices on paddy rice at elevated air temperature, heat stress, and herbicide damage. *Remote Sensing*, 12(16):2654, 2020.
- Ye Tan, Jia-Yi Sun, Bing Zhang, Meng Chen, Yu Liu, and Xiang-Dong Liu. Sensitivity of a ratio vegetation index derived from hyperspectral remote sensing to the brown planthopper stress on rice plants. *Sensors*, 19(2):375, 2019.
- Samuel O Ihuoma and Chandra A Madramootoo. Sensitivity of spectral vegetation indices for monitoring water stress in tomato plants. *Computers and Electronics in Agriculture*, 163:104860, 2019.
- Carlos Ballester, James Brinkhoff, Wendy C Quayle, and John Hornbuckle. Monitoring the effects of water stress in cotton using the green red vegetation index and red edge ratio. *Remote Sensing*, 11(7):873, 2019.
- Tengcong Jiang, Zihe Dou, Jian Liu, Yujing Gao, Robert W Malone, Shang Chen, Hao Feng, Qiang Yu, Guining Xue, and Jianqiang He. Simulating the influences of soil water stress on leaf expansion and senescence of winter wheat. *Agricultural and Forest Meteorology*, 291:108061, 2020.
- Kangying Zhu, Zhigang Sun, Fenghua Zhao, Ting Yang, Zhenrong Tian, Jianbin Lai, Wanxue Zhu, and Buju Long. Relating hyperspectral vegetation indices with soil salinity at different depths for the diagnosis of winter wheat salt stress. *Remote Sensing*, 13(2):250, 2021.
- Max Gerhards, Martin Schlerf, Kaniska Mallick, and Thomas Udelhoven. Challenges and future perspectives of multi-/hyperspectral thermal infrared remote sensing for crop water-stress detection: A review. *Remote Sensing*, 11 (10):1240, 2019.
- Zheng Zhou, Yaqoob Majeed, Geraldine Diverres Naranjo, and Elena MT Gambacorta. Assessment for crop water stress with infrared thermal imagery in precision agriculture: A review and future prospects for deep learning applications. *Computers and Electronics in Agriculture*, 182:106019, 2021b.

- Zhongsheng Cao, Xia Yao, Hongyan Liu, Bing Liu, Tao Cheng, Yongchao Tian, Weixing Cao, and Yan Zhu. Comparison of the abilities of vegetation indices and photosynthetic parameters to detect heat stress in wheat. *Agricultural and Forest Meteorology*, 265:121–136, 2019.
- Sarah W Shivers, Dar A Roberts, and Joseph P McFadden. Using paired thermal and hyperspectral aerial imagery to quantify land surface temperature variability and assess crop stress within California orchards. *Remote Sensing of Environment*, 222:215–231, 2019.
- Yukimasa Kaneda, Shun Shibata, and Hiroshi Mineno. Multi-modal sliding window-based support vector regression for predicting plant water stress. *Knowledge-Based Systems*, 134:135–148, 2017.
- Peter Harrington. *Machine learning in action*. Simon and Schuster, 2012.
- Alvaro Fuentes, Sook Yoon, Sang Cheol Kim, and Dong Sun Park. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors*, 17(9):2022, 2017.
- Aravind Krishnaswamy Rangarajan and Raja Purushothaman. Disease classification in eggplant using pre-trained vgg16 and msvm. *Scientific reports*, 10(1):1–11, 2020.
- Xiu Jin, Lu Jie, Shuai Wang, Hai Jun Qi, and Shao Wen Li. Classifying wheat hyperspectral pixels of healthy heads and fusarium head blight disease using a deep neural network in the wild field. *Remote Sensing*, 10(3):395, 2018.
- Jiangyong An, Wanyi Li, Maosong Li, Sanrong Cui, and Huanran Yue. Identification and classification of maize drought stress using deep convolutional neural network. *Symmetry*, 11(2):256, 2019.
- Albert Cruz, Yiannis Ampatzidis, Roberto Pierro, Alberto Materazzi, Alessandra Panattoni, Luigi De Bellis, and Andrea Luvisi. Detection of grapevine yellows symptoms in *Vitis vinifera* L. with artificial intelligence. *Computers and electronics in agriculture*, 157:63–76, 2019.
- Jayme Garcia Arnal Barbedo. Plant disease identification from individual lesions and spots using deep learning. *Biosystems Engineering*, 180:96–107, 2019.
- Ke Lin, Liang Gong, Yixiang Huang, Chengliang Liu, and Junsong Pan. Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network. *Frontiers in plant science*, 10:155, 2019.
- Yue Shi, Liangxiu Han, Wenjiang Huang, Sheng Chang, Yingying Dong, Darren Dancey, and Lianghao Han. A biologically interpretable two-stage deep neural network (bit-dnn) for hyperspectral imagery classification. *arXiv preprint arXiv:2004.08886*, 2020.
- Edna Chebet Too, Li Yujian, Sam Njuki, and Liu Yingchun. A comparative study of fine-tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture*, 161:272–279, 2019.
- Mohammed Brahimi, Saïd Mahmoudi, Kamel Boukhalfa, and Abdelouhab Moussaoui. Deep interpretable architecture for plant diseases classification. In *2019 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, pages 111–116. IEEE, 2019.
- Thomas Lillesand, Ralph W Kiefer, and Jonathan Chipman. *Remote sensing and image interpretation*. John Wiley & Sons, 2015.
- Jan Behmann, Jörg Steinrücken, and Lutz Plümer. Detection of early plant stress responses in hyperspectral images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93:98–111, 2014a.
- Oren Rippel, Jasper Snoek, and Ryan P Adams. Spectral representations for convolutional neural networks. *Advances in neural information processing systems*, 28, 2015.
- Lu Chi, Guiyu Tian, Yadong Mu, Lingxi Xie, and Qi Tian. Fast non-local neural networks with spectral residual learning. In *Proceedings of the 27th ACM International Conference on Multimedia*, pages 2142–2151, 2019.
- Davoud Ashourloo, Ali Akbar Matkan, Alfredo Huete, Hossein Aghighi, and Mohammad Reza Mobasher. Developing an index for detection and identification of disease stages. *IEEE Geoscience and Remote Sensing Letters*, 13(6):851–855, 2016.
- Jan Behmann, Jörg Steinrücken, and Lutz Plümer. Detection of early plant stress responses in hyperspectral images. *Isprs Journal of Photogrammetry and Remote Sensing*, 93(7):98–111, 2014b.
- M. E. Jakubauskas, D. R. Legates, and J. H. Kastens. Crop identification using harmonic analysis of time-series avhrr ndvi data. *Computers and Electronics in Agriculture*, 37(1):127–139, 2002.
- A. K. Mahlein, M. T. Kuska, S. Thomas, D. Bohnenkamp, E. Alisaac, J. Behmann, M. Wahabzada, and K. Kersting. Plant disease detection by hyperspectral imaging: from the lab to the field. 8(2):238–243, 2017.
- D. P. Roy and L. Yan. Robust landsat-based crop time series modelling. *Remote Sensing of Environment*, 2018.

- Bethany A. Bradley, Robert W. Jacob, John F. Hermance, and John F. Mustard. A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite ndvi data. *Remote Sensing of Environment*, 106(2):137–145, 2007.
- Yang Shao, Ross S. Lunetta, Brandon Wheeler, John S. Iiames, and James B. Campbell. An evaluation of time-series smoothing algorithms for land-cover classifications using modis-ndvi multi-temporal data. *Remote Sensing of Environment*, 174:258–265, 2016.
- M. E. Jakubauskas. Time series remote sensing of landscape-vegetation interactions in the southern great plains. *Sensing*, 68(10):1021–1030, 2002.
- Rogier De Jong, Sytze De Bruin, Allard De Wit, Michael E. Schaepman, and David L. Dent. Analysis of monotonic greening and browning trends from global ndvi time-series. *Remote Sensing of Environment*, 115(2):692–702, 2011.
- Toshihiro Sakamoto, Masayuki Yokozawa, Hitoshi Toritani, Michio Shibayama, Naoki Ishitsuka, and Hiroyuki Ohno. A crop phenology detection method using time-series modis data. *Remote Sensing of Environment*, 96(3):366–374, 2005.
- Wenjiang Huang, Junjing Lu, Huichun Ye, Weiping Kong, A Hugh Mortimer, and Yue Shi. Quantitative identification of crop disease and nitrogen-water stress in winter wheat using continuous wavelet analysis. *International Journal of Agricultural and Biological Engineering*, 11(2):145–152, 2018. ISSN 1934-6352.
- P Shanmugapriya, S Rathika, T Ramesh, and P Janaki. Applications of remote sensing in agriculture—a review. *Int. J. Curr. Microbiol. Appl. Sci*, 8:2270–2283, 2019.
- Moussa El Jarroudi, Louis Kouadio, Mustapha El Jarroudi, Jürgen Junk, Clive Bock, Abdoul Aziz Diouf, and Philippe Delfosse. Improving fungal disease forecasts in winter wheat: A critical role of intra-day variations of meteorological conditions in the development of septoria leaf blotch. *Field Crops Research*, 213:12–20, 2017. ISSN 0378-4290.
- Yue Shi, Wenjiang Huang, and Xianfeng Zhou. Evaluation of wavelet spectral features in pathological detection and discrimination of yellow rust and powdery mildew in winter wheat with hyperspectral reflectance data. *Journal of Applied Remote Sensing*, 11(2):026025, 2017b.
- T. Rumpf, A. K. Mahlein, U. Steiner, E. C. Oerke, H. W. Dehne, and L. Plümer. Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. *Computers and Electronics in Agriculture*, 74(1):91–99, 2010.
- Yueru Chen, Yijing Yang, Wei Wang, and C-C Jay Kuo. Ensembles of feedforward-designed convolutional neural networks. In *2019 IEEE International Conference on Image Processing (ICIP)*, pages 3796–3800. IEEE. ISBN 1538662493.
- Yue Shi, Wenjiang Huang, Pablo González-Moreno, Belinda Luke, Yingying Dong, Qiong Zheng, Huiqin Ma, and Linyi Liu. Wavelet-based rust spectral feature set (wrsfs): A novel spectral feature set based on continuous wavelet transformation for tracking progressive host-pathogen interaction of yellow rust on wheat. *Remote Sensing*, 10(4):525, 2018.