

Identification and mapping of yellow-flowering rapeseed fields by combining social media data, Sentinel-2 imagery, deep learning algorithm, and Google Earth Engine

Zhenjie Liu^{a,b}, Yingyue Su^a, Xiangming Xiao^{c,*}, Yuanwei Qin^c, Jun Li^b, Luo Liu^{a,*}

^a Guangdong Province Key Laboratory for Land Use and Consolidation, South China Agricultural University, Guangzhou 510642, China

^b Hubei Key Laboratory of Intelligent Geo-Information Processing, and School of Computer Science, China University of Geosciences, Wuhan 430078, China

^c Department of Microbiology and Plant Biology, Center for Earth Observation and Modeling, University of Oklahoma, Norman, OK 73019, USA

ARTICLE INFO

Keywords:

Rapeseed mapping
Deep learning
Sentinel-2
Google Earth Engine

ABSTRACT

Rapeseed cultivation on winter-fallow fields enables a seasonal switch between agricultural and bioenergy output. In view of the spectral features during the rapeseed flowering period (RFP), numerous remote sensing studies identified and produced the maps of rapeseed fields. The RFP is frequently identified by field surveys, visual imagery interpretation, or empirical crop knowledge in a specific region, none of which is appropriate for large-scale rapeseed identification and mapping. In this research, we combine social media data on the RFP fields, Sentinel-2 imagery, and deep learning algorithm to identify and produce maps of yellow-flowering rapeseed fields. In the context of citizen science and crowdsourcing, the social media data on the fields is regarded as reference data and utilized to generate the spatial distribution of the RFP and select temporal Sentinel-2 imagery. We develop and evaluate the novel method in Anhui Province, China. The resultant rapeseed map in 2018 has a user's accuracy and a producer's accuracy of 0.93 and 0.99, respectively. To test the generalization of the knowledge-based method, we apply the deep neural network (DNN) model trained in Anhui Province to produce the maps of yellow-flowering rapeseed fields in Hubei Province and Shaanxi Province, China. The overall accuracy of the resultant rapeseed maps for Hubei Province and Shaanxi Province is 0.97 and 0.95, respectively. The proposed method provides a new way to produce rapeseed maps on a large scale, which could be used to support agricultural planning and ecological system management.

1. Introduction

As a major biofuel and cash crop, rapeseed can boost farmer incomes while lowering dependency on fossil fuels and environmental risks (Davis et al., 2009). More attractively, the seasonal rotation between rapeseed and other summer crops allows for bioenergy and agricultural production. Compared to other crop types, the bright-yellow flowers of rapeseed present a distinctive feature which is highly informative to monitor growth status and predict final yield (d'Andrimont et al., 2020, Sulik and Long, 2016). Therefore, in-season identification and mapping of yellow-flowering rapeseed fields is critical for the sustainability of agricultural systems and ecosystems, as well as food security.

For purposes of illustration, we define the abbreviation RFD as the flowering date of rapeseed, and RFP as the rapeseed flowering period determined by the start and end flowering dates. Traditional ground-

based methods for obtaining prior knowledge of RFD or RFP include field surveys, questionnaires, interviews, etc. However, these methods typically require a lot of labor and time. In comparison, remote sensing has become a feasible and efficient approach to obtain spatial and spectral information on rapeseed due to the advantages of wide coverage and synchronous observation (Liu and Zhang, 2023). Specifically, yellow-flowering rapeseed and other crop types exhibit a significant difference in the reflectance of the green, red, and near-infrared bands (Sulik and Long, 2015, 2016). Numerous studies identify and map rapeseed fields using the spectral features during the RFP (Breckling et al., 2011). For example, the research in (Wang et al., 2018) combined the colorimetric transformation and spectral features to identify the yellow-flowering rapeseed fields. Multiple yellowness indices are applied to extract rapeseed from other vegetation during the RFP, such as the ratio yellowness index (RYI) (Sulik and Long, 2015) and the

* Corresponding authors.

E-mail addresses: xiangming.xiao@ou.edu (X. Xiao), liuluo@scau.edu.cn (L. Liu).

<https://doi.org/10.1016/j.jag.2024.104047>

Received 26 March 2024; Received in revised form 19 June 2024; Accepted 16 July 2024

Available online 24 July 2024

1569-8432/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

normalized difference yellowness index (NDYI) (Sulik and Long, 2016). Generally, rapeseed and winter wheat are often cultivated from the late fall to spring seasons (Fig. S1). Many studies have also explored the classification of rapeseed and winter wheat according to the spectral features during the RFP (Rußwurm et al., 2023). In this case, prior knowledge of the RFP is essential for rapeseed identification. Since the RFP could vary over different regions (Han et al., 2022; Luo et al., 2018), these methods are not suitable for large-scale rapeseed mapping. To determine the best temporal window for rapeseed mapping, the research (Meng et al., 2020) assessed all possible combinations of temporal windows using all available cloud-free Sentinel-2 multispectral instrument (MSI) imagery. However, this method relies on a large amount of ground survey data and is computationally inefficient. In addition, the RFP may also not be homogeneous even in the same region due to variations in planting dates and topography (Wang et al., 2019). Therefore, it is still a significant and challenging task to automatically obtain the prior knowledge of the RFP on a large scale so as to determine the acquisition times of corresponding remote sensing data in different regions.

In the last decade, with social networks growing quickly, social media data in the citizen science context can be gathered and processed in real time for scientific research (Huang et al., 2024). It has shown great potential to obtain valuable geographic information in many applications, such as socioeconomic factor estimation (Zhao et al., 2020), emergency response (Liu et al., 2020), etc. Recently, social media data has also been considered and shown great potential in crop monitoring. For instance, a large and public dataset of crowdsourced road view photos from major grain-producing areas of China was established for crop type identification (Wu et al., 2021). Some studies used social media streams for regional-scale crop disease monitoring (Shankar et al., 2020). In particular, rapeseed presents a unique characteristic of social media. The public prefers to visit rapeseed fields during the flowering period and share relevant contents, photos, and videos with geo-locations and date information. In this context, social media data has great potential to capture the biological characteristics of yellow-flowering rapeseed during the RFP. In addition, similar to in-situ data from researchers' surveys, social media could also provide a way to obtain in-situ information, e.g., points of interest (POIs). However, social media data is generated in the form of text or point data with uncertainty and noise. As mentioned earlier, remote sensing is a reliable data source for rapeseed mapping, but it relies on the prior knowledge of the RFP. Therefore, the combination of remote sensing and social media data offers an unprecedented chance for large-scale rapeseed mapping.

Deep learning can automatically discover the key feature representations for classification from original imagery at multiple levels, which is essential for crop identification because of the intricate relationships between internal biochemical processes and innate relationships between external variables (Zhao et al., 2019). Deep learning has been extensively employed in related crop classification investigations (Zhong et al., 2019). For instance, a study (Kussul et al., 2017) adopted convolutional neural network (CNN) to discriminate summer crop types based on Landsat-8 operational land imager (OLI) and Sentinel-1 synthetic aperture radar (SAR) data produced during the whole vegetation season. Their results demonstrated that the CNN outperformed the traditional machine learning methods, and the resultant rapeseed map had high producer's accuracy but comparatively low user's accuracy, which could be attributed to the inclusion of the non-RFP imagery. The non-RFP imagery may not provide effective information to identify rapeseed from other crops and may also contain some noise. Despite the fact that deep learning algorithms are effective to identify rapeseed from other crops, prior knowledge of the RFP is still required, especially for large-scale rapeseed mapping.

Considering the spectral features of yellow-flowering rapeseed during the RFP, most spectral- or deep learning-based methods rely on prior knowledge to determine the RFP in a specific region and select corresponding remote sensing data. However, the RFP could be varied in

different regions, even within the same remote sensing imagery (Han et al., 2022; Wang et al., 2019). Therefore, existing methods have certain limitations for large-scale rapeseed mapping. The following are the contributions made by our work: (1) The key knowledge of RFD and RFP is obtained through the information of dates and locations from social media data; (2) the proposed knowledge-based method can automatically associate the remote sensing data produced during the RFP without considering the spatial heterogeneity of phenological periods.

2. Material and methods

2.1. Study area

Anhui Province, China (extending from 114°54'–119°37' E and 29°41'–34°38' N) is located in the transition area between warm-temperate and subtropical climate zones (Fig. 1). The average annual sunlight, average annual temperature, and average annual precipitation are 1800–2500 h, 14–17 °C, 800–1800 mm, respectively. During the rapeseed growing period from September to May, abundant precipitation and a low temperature stage around zero degrees in winter are conducive to the vernalization and overwintering of rapeseed. In late autumn and early spring, there is sufficient sunlight, and the temperature drops and rises slowly, which is suitable for the growth of rapeseed. The yield of rapeseed was 843×10^3 tons in 2018. In Anhui Province, cropping intensity is dominated by double cropping of crops in a year, mainly wheat (*Triticum aestivum*)-soybean (*Glycine max*), wheat (*Triticum aestivum*)-rice (*Oryza sativa*), rapeseed (*Brassica napus*)-rice (*Oryza sativa*), and rice (*Oryza sativa*)-rice (*Oryza sativa*).

Hubei Province, China (extending from 108°21'–116°07' E and 29°01'–33°06' N) belongs to a subtropical humid monsoon climate zone with sufficient sunlight, heat, and precipitation (Fig. 1). The average annual sunlight, average annual temperature, and average annual precipitation are 1100–2150 h, 15–17 °C, 800–1600 mm, respectively. Hubei Province and Anhui Province are both located in the middle and lower reaches of the Yangtze River and have similar climatic conditions suitable for the cultivation and growth of rapeseed. In 2018, the yield of rapeseed reached 2053×10^3 tons. In Hubei Province, cropping intensity is dominated by double cropping of crops, mainly wheat (*Triticum aestivum*)-rice (*Oryza sativa*), rapeseed (*Brassica napus*)-rice (*Oryza sativa*), and rice (*Oryza sativa*)-rice (*Oryza sativa*).

Shaanxi Province is located in Loess Plateau, China, extending from 105°29'–111°15' E and 31°42'–39°35' N (Fig. 1). Due to its large span in latitude, the northern, central, and southern regions respectively belong to moderate-temperate, warm-temperate, and subtropical climate zones. The average annual sunlight, average annual temperature, and average annual precipitation are 1300–2900 h, 9–16 °C, 400–900 mm, respectively. The precipitation and temperature in central and southern regions can meet the requirements for rapeseed growth. Rapeseed experiences typical budding and flowering times since the winter is humid and the temperature rises early in the spring. The yield of rapeseed was 369×10^3 tons in 2018. In Shaanxi Province, cropping intensity is dominated by double cropping of crops, mainly wheat (*Triticum aestivum*)-corn (*Zea mays*) and rapeseed (*Brassica napus*)-rice (*Oryza sativa*).

2.2. Data

2.2.1. Social media data and preprocessing for rapeseed locations and rapeseed flowering dates

Sina Weibo, or simply Weibo, is the largest social media platform in China, with over 462 million active users in 2018. We collect geo-tagged Weibo data with the specific keyword “rapeseed” (in Chinese) through an open platform application programming interface (API). Specifically, the geo-tagged Weibo data covering the study areas in 2018 is obtained. In total, there are 593, 493, and 720 geo-tagged Weibo data in Anhui

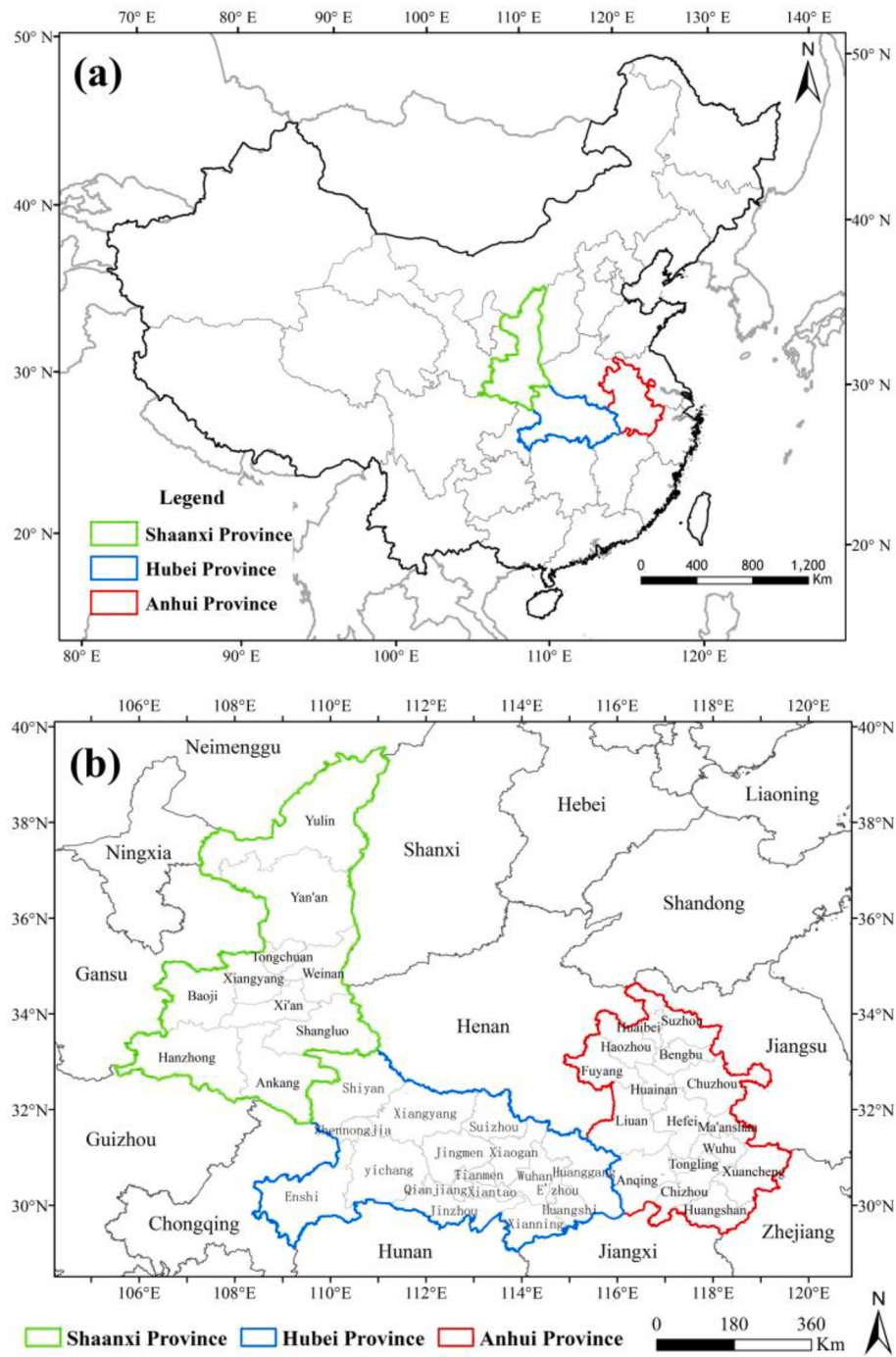


Fig. 1. (a) Location of study areas and (b) zoom-in view of Anhui Province, Hubei Province, and Shaanxi Province, China.

Province, Hubei Province, and Shaanxi Province, respectively. Each Weibo data contains information about time (date) and location (longitude and latitude).

Since the raw Weibo data can be very noisy, we develop a procedure to clean the data that is irrelevant to the rapeseed flowering. First, we adopt a Python module, namely Jieba word segmentation, to extract semantic unit information from Weibo messages. Then, we remove the Weibo data containing specific words irrelevant to our research, such as “Miss the rapeseed,” “Flowering season is over,” “Come back when the rapeseed is blooming,” “Not yet flowering,” etc. Even though these data are related to rapeseed, they may not be right to describe the RFD. After data cleaning, there are a total of 472, 406, and 577 geo-tagged Weibo data covering Anhui Province, Hubei Province, and Shaanxi Province,

respectively. The preprocessed data in Anhui Province are shown in Fig. 2, overlaid with a rapeseed extent map (Liu and Zhang, 2023).

2.2.2. Sentinel-2 data and preprocessing

We collect Sentinel-2 imagery covering the study areas through the Google Earth Engine (GEE). Considering the time required for Sentinel-2 surface reflectance (SR) data production, Sentinel-2 top of atmosphere (TOA) reflectance data are used for in-season mapping of yellow-flowering rapeseed fields. Concerning the Sentinel-2 data, bad observations are first identified by the FMask algorithm (Zhu et al., 2015), including cloud (shadow), cirrus, and snow/ice. Specifically, the parameters of cloud probability, cloud displacement index (CDI), cirrus band reflectance, maximum distance (km) to search for cloud shadows,

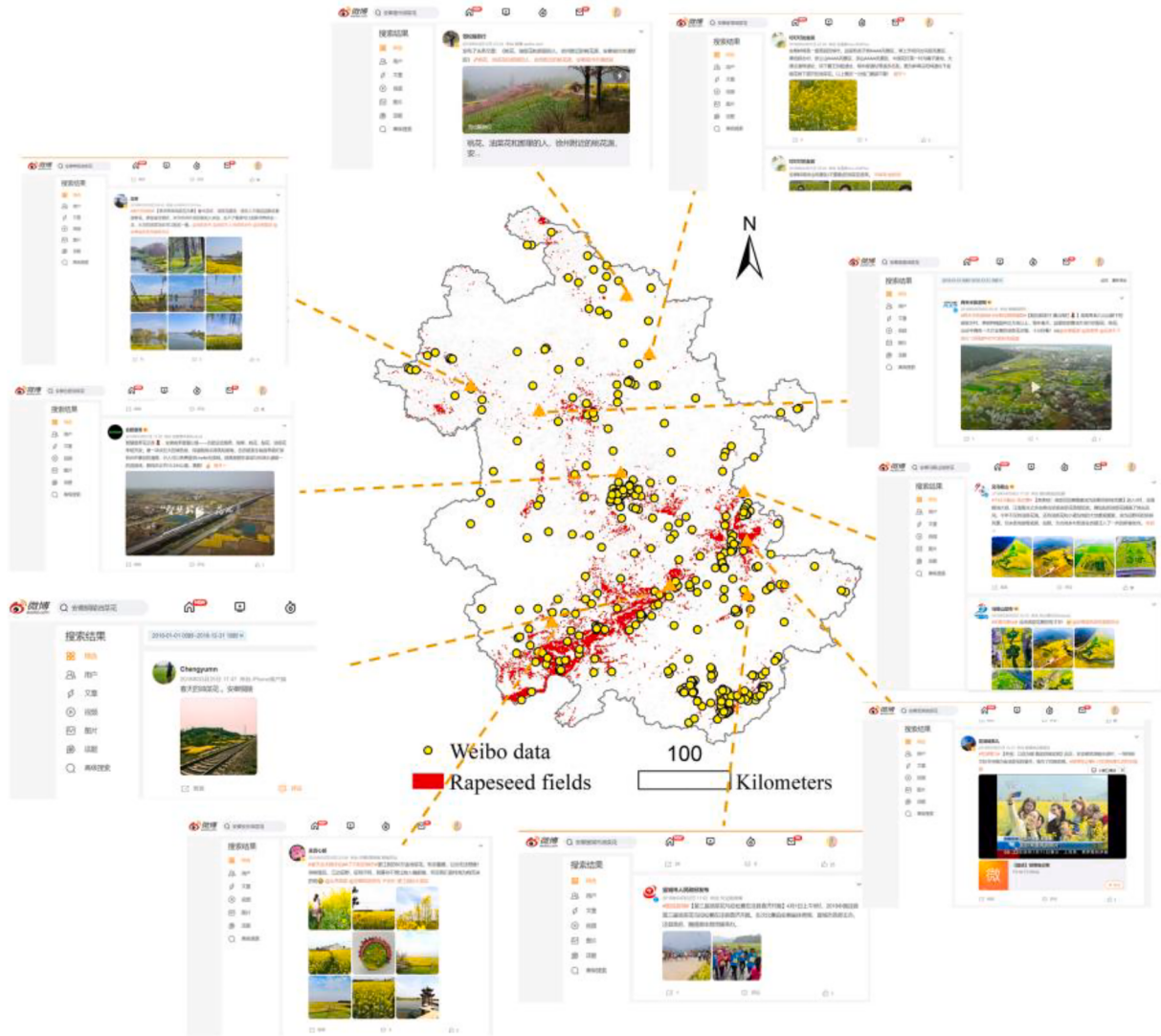


Fig. 2. Sina Weibo messages related to the flowering period of rapeseed in Anhui Province.

and distance (m) to dilate the edge of cloud-identified objects are respectively set as 0.4, -0.5 , 0.01, 10, 50. The number of pixel-based valid observations for different latitude ranges and months in 2018 is presented in Fig. S2.

Spectral bands and spectral indices have been widely used to identify rapeseed fields by detecting specific signals during the RFP (Fig. 3). Here, we calculate the following spectral indices: (1) NDVI, (2) enhanced vegetation index (EVI) (Huete et al., 1997) (Tucker, 1979), (3) land surface water index (LSWI) (Xiao et al., 2005), (4) normalized difference soil index (NDSI) (Deng et al., 2015), (5) normalized difference vegetation index (NDVI) (Tucker, 1979), and (6) normalized difference water index (NDWI) (McFeeters, 1996) from the Sentinel-2 TOA reflectance data.

2.2.3. DEM and air temperature data

The National Aeronautics and Space Administration (NASA) provides a publicly- and freely-available digital elevation model (DEM), namely NASADEM (Crippen et al. 2016). The NASADEM data with a 30-m spatial resolution are collected through GEE, which is used as a covariate for the interpolation of air temperature data. Topographic slope data are also calculated from the NASADEM data.

The average air temperature data in 2018 are obtained from the Chinese Meteorological Administration, which is interpolated to a

spatial resolution of 1 km through ANUSPLIN v.4.4 software. ANUSPLIN is provided for interpolation of noisy multi-variate data using thin plate smoothing splines (Hutchinson and Xu, 2013). Latitude and longitude are the independent variables, and elevation is adopted as a covariate to establish the change in air temperature along the altitudinal gradient. The SPLINE program is used to fit partial thin plate smoothing spline functions of independent variables. The LAPGRD program is used to calculate values and Bayesian standard error estimates of partial thin plate smoothing spline surfaces.

2.3. Methods

Fig. 4 presents the workflow of the proposed knowledge-based method for the identification and mapping of yellow-flowering rapeseed fields. First, we delineate regions of interest (ROIs) for rapeseed and non-rapeseed as ground reference data based on the Weibo POIs data, temporal spectral features, and visual interpretation of Sentinel-2 and Google Earth imagery. The locations and dates of the Weibo POIs data are then used to generate maps of RFD and RFP through spatial interpolation. Spectral bands and spectral indices extracted from 6 valid Sentinel-2 TOA reflectance data closest to the RFD are input into a deep neural network (DNN) model to identify yellow-flowering rapeseed fields. The ROIs, stratified random samples, and statistical data are used

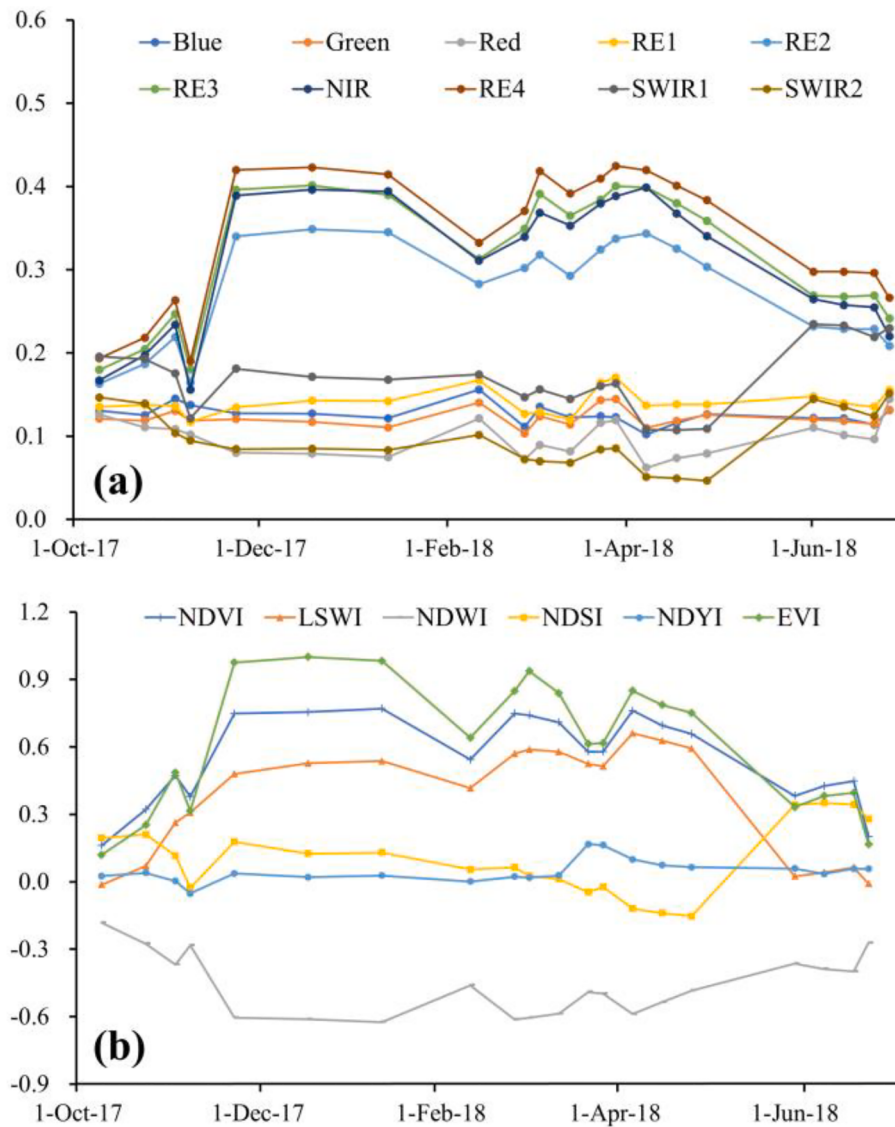


Fig. 3. Temporal profiles of Sentinel-2 (a) spectral bands of blue, green, red, red edge 1 (RE1), red edge 2 (RE2), red edge 3 (RE3), near infrared (NIR), red edge 4 (RE4), short-wave infrared 1 (SWIR1) and short-wave infrared 2 (SWIR2), and (b) spectral indices of normalized difference vegetation index (NDVI), land surface water index (LSWI), normalized difference water index (NDWI), normalized difference soil index (NDSI), normalized difference yellowness index (NDYI), and enhanced vegetation index (EVI) for rapeseed. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

for accuracy assessment. In the following paragraphs, we describe the above steps in detail.

2.3.1. Development of rapeseed ROIs data

To develop rapeseed ROIs, the recorded locations and dates of rapeseed are first extracted from preprocessed geo-tagged Weibo data in Anhui Province. The Weibo POIs data are overlaid with the closest Sentinel-2 imagery and Google Earth imagery (Fig. S3). Due to the influence of cloud cover, it may not be possible to obtain cloudless imagery near Weibo POIs data. Moreover, the noisy nature of social media is another limitation for rapeseed ROIs development within a certain geographical range from the Weibo POIs data. In this context, 160 rapeseed ROIs (16,008 pixels) within 1 km from the Weibo POIs data are first delineated by visually interpreting Google Earth imagery and temporal Sentinel-2 spectral reflectance (e.g., red and SWIR bands) (Shen et al., 2009, Wilson et al., 2015) and spectral indices (e.g., NDYI) (d'Andrimont et al., 2020, Sulik and Long, 2016). Then, another 1545 rapeseed ROIs (233,344 pixels) are evenly delineated within Anhui

Province. In total, 1705 ROIs of rapeseed are developed (Fig. 5a-d). The majority of rapeseed ROIs are distributed in southern Anhui province, where the cropping system is dominated by rapeseed-rice. Since the predominant cropping system in northern Anhui province is wheat-rice, there are significantly fewer rapeseed ROIs (Fig. 5a). The shape of manually developed rapeseed ROIs is irregular, which mainly takes into account the fragmentation of many rapeseed fields in the study areas, especially in the mountainous regions (e.g., Fig. 5c).

637 ROIs of non-rapeseed (1,088,233 pixels) are delineated in a stratified and even distribution sampling strategy from the available Sentinel-2 and Google Earth imagery. The non-rapeseed ROIs are widely distributed in Anhui Province, which include bare soil (14,308 pixels), construction land (101,882 pixels), forest (510,103 pixels), water (134,741 pixels), winter wheat (90,760 pixels) and other green vegetations (236,439 pixels). Some examples of ROIs of non-rapeseed are shown in Fig. 5e-g.

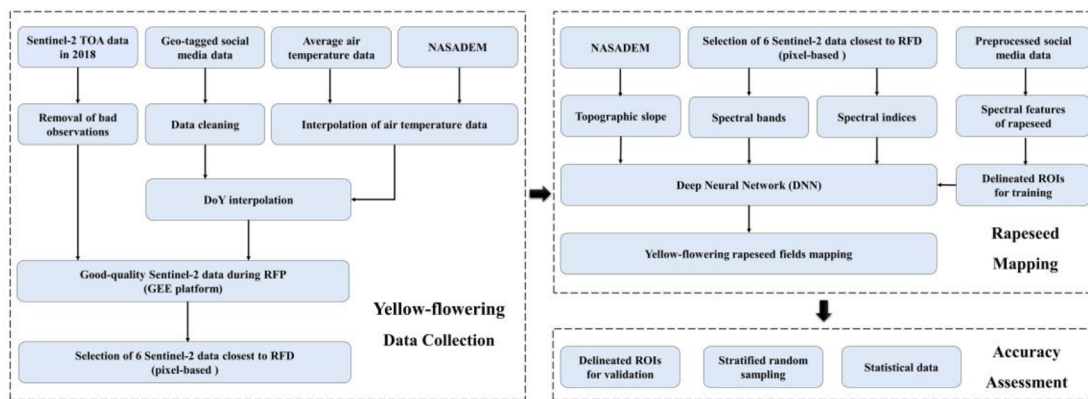


Fig. 4. The workflow for yellow-flowering rapeseed fields mapping by combining social media data, Sentinel-2 imagery, deep learning, and Google Earth Engine (GEE). It includes the processes of data collection, rapeseed mapping and accuracy assessment. The acronyms are defined as: top of atmosphere (TOA), National Aeronautics and Space Administration digital elevation model (NASADEM), day of year (DoY), rapeseed flowering period (RFP), rapeseed flowering date (RFD), and regions of interest (ROIs).

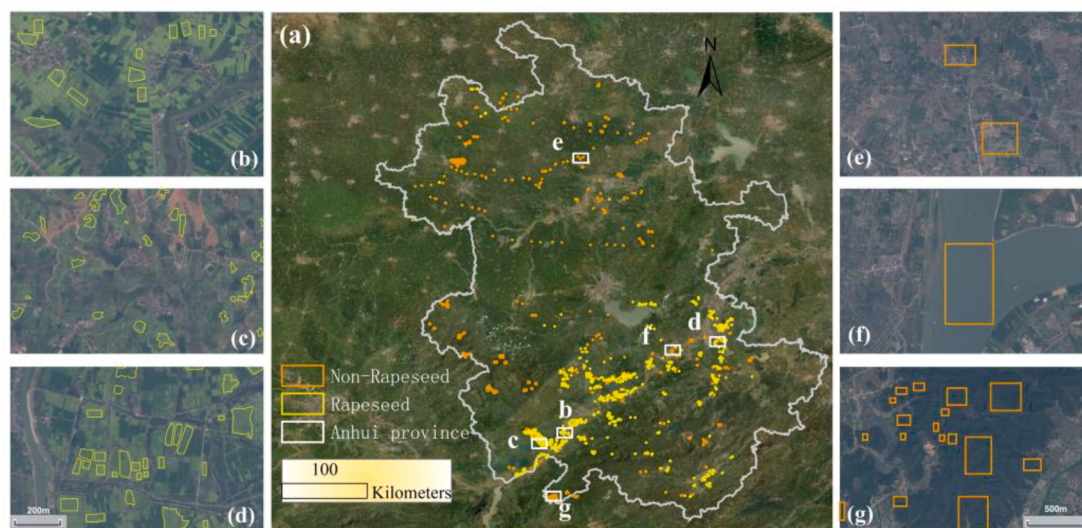


Fig. 5. (a) Ground reference data for model training and validation. (b-d) The zoom-in views for the regions of interest (ROIs) of rapeseed type. (e-g) The zoom-in views for the ROIs of non-rapeseed landcover.

2.3.2. Maps of rapeseed flowering date (RFD) and rapeseed flowering period (RFP)

Based on the recorded dates and locations of the Weibo POIs data, the day of year (DoY) is spatially interpolated into the map of RFD using thin plate smoothing splines of ANUSPLIN v.4.4 software. Since the RFD is sensitive to air temperature (Luo et al., 2018, Marjanović-Jeromela et al., 2019), the average air temperature data in 2018 is used as a covariate for the spatial interpolation of DoY. The default parameter settings of the SPLINE and LAPGRD programs are used.

Since the Weibo POIs data may be generated during early or late RFP, and the RFP generally lasts approximately 30 days (Wang et al., 2016), a window of 30 days as a tolerance is used for the RFD map to generate the RFP maps. Specifically, the period for 30 days before and after RFD (60 days in total) is identified as the RFP in this study. For each pixel in the RFD map, the start date and end date of the RFP are calculated by subtracting 30 days and adding 30 days from the RFD, respectively. After that, the RFP maps can be obtained from the RFD map, including the starting RFP map and ending RFP map.

2.3.3. Selection of the Sentinel-2 data during RFP

After excluding the bad observations of Sentinel-2 TOA reflectance data described in Section 2.2.2, the number of valid observations for

Anhui Province in 2018 is presented in Fig. S2. Based on the starting RFP map and ending RFP map generated in Section 2.3.2, the number of valid observations during the RFP in 2018 is counted (Fig. S4). There is at least one valid observation for 99.82 % of the pixels. Considering the spatial heterogeneity of valid observations and flowering periods, we construct pixel-based connections between the RFP maps and 6 valid Sentinel-2 TOA reflectance data. For the pixels with more than 6 valid data, the 6 valid data that are closest to the RFD are selected for subsequent identification of yellow-flowering rapeseed fields. For the pixels with less than 6 valid data, the valid data closest to the RFD is duplicated for supplementation, so that each pixel has 6 valid data. For the 3,056,085 pixels without any valid data, they are masked out. After that, the spectral bands and spectral indices (i.e., NDVI, EVI, NDWI, NDSI, NDYI and LSWI) of the 6 valid data can be calculated and obtained.

2.3.4. Rapeseed mapping algorithm – deep neural network (DNN)

In this study, a straightforward five-layer DNN model architecture is used, including one input layer, three hidden layers, and one output layer. The rectified linear unit (ReLU) activation function (Glorot et al., 2011) is used for the hidden layers, which have 32, 64, and 128 neuron nodes, respectively. Batch normalization technique and dropout regularization with a rate of 0.3 are adopted to prevent overfitting and

enhance the generalization ability of the DNN model (Ioffe and Szegedy, 2015, Srivastava et al., 2014). Since the flowering period usually only lasts for approximately 30 days (Wang et al., 2016), the collected 6 valid data are likely to contain the Sentinel-2 imagery that are produced during the non-RFP. To avoid the interference of these data, we develop a new DNN model to predict rapeseed and non-rapeseed. The pipeline of training and prediction for the proposed DNN model is given (Fig. 6). During the training process, the 6 valid data are input in a parallel manner, instead of directly using all bands of the 6 valid data for model training. Then, only the highest probability among the 6 predicted probabilities is used to calculate the training loss. The network parameters are adjusted by using the gradient descent optimization algorithm. Shared multilayer perception (MLP) (Qi et al., 2017) is adopted to realize weight sharing and reduce network training parameters. Similarly, the 6 valid data are parallelly input into the trained DNN model, and the highest probability is retained as the final result of the prediction.

Spectral bands and spectral indices can provide information for deep learning-based remote sensing classification (Wang et al., 2022, Yao et al., 2022). To determine the input variables of the DNN model, we perform a variable importance analysis (Gregorutti et al., 2017) using 225,670 and 797,945 pixels of random samples from the 1705 rapeseed ROIs and the 637 non-rapeseed ROIs, respectively, and exclude the variables with an importance score less than 0.01. According to Table S1, 9 spectral bands and 6 spectral indices are selected. Topographic slope data is also used as an input band for the DNN model. In total, there are 91 bands for each sample. Then, we respectively slice the 15 Sentinel-2 bands and topographic slope band for each of the 6 valid data from the 91 bands of each sample. Eventually, 6 valid data (each containing 16 bands) are input into the DNN model (Fig. 6).

The 1705 rapeseed ROIs and the 637 non-rapeseed ROIs delineated in Section 2.3.1 are split into a training dataset and a validation dataset in accordance with the 70:30 ratio. The overall accuracy (OA), Kappa, and F1 score of the validation dataset are reported by averaging ten independent Monte Carlo runs. We train the DNN model with a batch size of 2048 and use Adam Optimizer with an initial learning rate of $1e^{-4}$. The DNN model is trained for 100 epochs, and early stopping is adopted to avoid overfitting.

2.3.5. Application of the trained DNN model in Hubei Province and Shaanxi Province

To test the generalization of the knowledge-based method, the

trained DNN model is directly applied for rapeseed mapping in Hubei Province and Shaanxi Province, which have distinctly different environments and climates. The same steps for processing data collected in Anhui Province are applied to the data collected in Hubei Province and Shaanxi Province. Specifically, we spatially interpolate the DoY to the RFD map with a spatial resolution of 10 m based on the geo-tagged Weibo POIs data. The RFP maps are obtained by subtracting 30 days and adding 30 days from the RFD map, respectively. Then, the 6 valid Sentinel-2 TOA reflectance data closest to the RFD during RFP are selected through GEE and used for the extraction of spectral bands and spectral indices. After that, the DNN model trained in Anhui Province is directly applied to the yellow-flowering rapeseed mapping in these two provinces.

2.3.6. Accuracy assessment

The RFD map derived from Weibo POIs data is validated based on the 160 rapeseed ROIs delineated in Section 2.3.1. Specifically, one ground reference sample is randomly generated in each ROI. In total, there are 160 ground reference samples for validation. The 160 ground reference samples are overlaid with the closest cloudless Sentinel-2 imagery according to the DoY of these samples on the RFD map. The accuracy of the RFD map is assessed by the proportion of 160 ground reference samples that appear in distinctive yellowness in the closest cloudless Sentinel-2 imagery.

To validate the resultant maps of yellow-flowering rapeseed fields, stratified random sampling is adopted to assess the resultant yellow-flowering rapeseed maps (Olofsson et al., 2014). There are 389 samples and 1168 samples of rapeseed and non-rapeseed in Anhui Province, respectively. There are 378 samples and 1133 samples of rapeseed and non-rapeseed in Hubei Province, respectively. There are 377 samples and 1131 samples of rapeseed and non-rapeseed in Shaanxi Province, respectively. According to the results of sample allocation to strata, we generate stratified random samples on the rapeseed and non-rapeseed maps (Fig. 7). The stratified random samples are visually interpreted based on the Sentinel-2 spectral profiles and available Google Earth imagery during the RFP. These stratified random samples are then used to generate confusion matrices, user accuracy (UA), producer accuracy (PA), overall accuracy (OA), and Kappa coefficient for accuracy assessment.

In addition, we calculate the area of yellow-flowering rapeseed fields from the 2018 resultant rapeseed maps for the study areas, which are contrasted with the statistics on the rapeseed sown area released by the

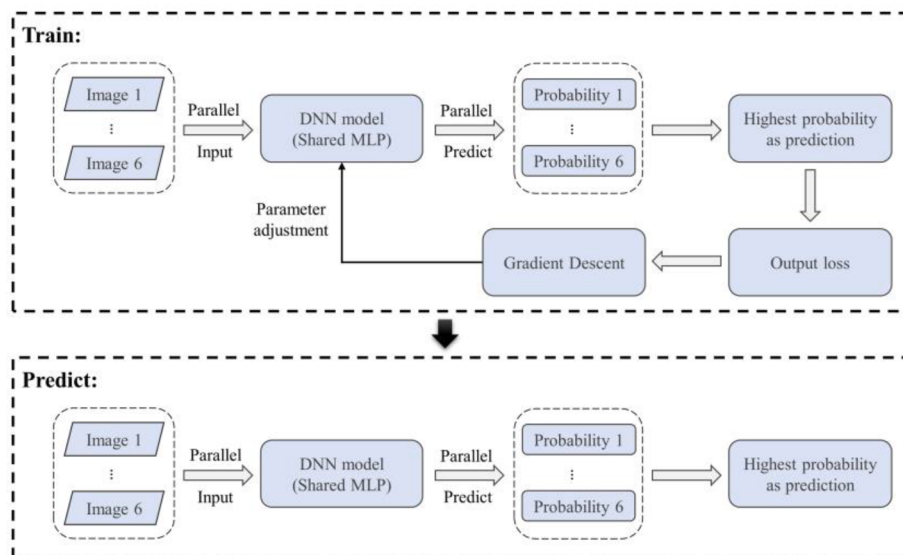


Fig. 6. The pipeline of training and prediction of the proposed DNN model. The acronyms are defined as: deep neural network (DNN), multilayer perception (MLP).

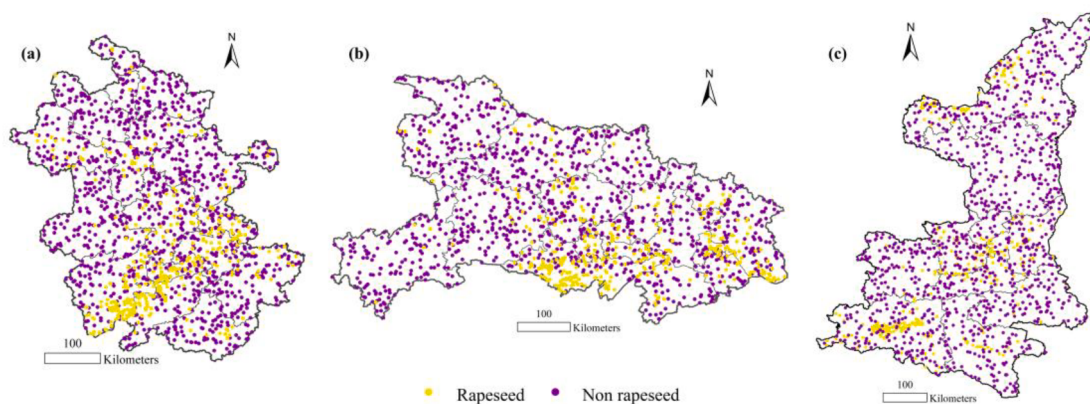


Fig. 7. Distribution of stratified random reference samples for (a) Anhui Province, (b) Hubei Province and (c) Shaanxi Province.

Bureau of Statistics of Anhui Province, Hubei Province, and Shaanxi Province.

3. Results

3.1. Maps of rapeseed flowering period in 2018 for Anhui Province

The RFD and RFP maps are generated by the spatial interpolation of geo-tagged Weibo DoY, considering the average air temperature data as a covariate (Fig. 8). It can be observed that earlier flowering dates are distributed in the southern regions, which may be due to their warmer climate. The spatial distribution of the RFP is presented in the form of the flowering starting date (Fig. 8c) and ending date (Fig. 8d) by subtracting 30 days and adding 30 days from the RFD map (Fig. 8b), respectively. The RFP in most areas of Anhui Province is between the 68th day (i.e., March 9th) and the 133rd day (i.e., May 13th), which is consistent with the flowering period of rapeseed in Anhui Province

(generally from mid-March to the end of April).

The 160 ground reference samples described in Section 2.3.6 are used for validating the RFD map. 95 % of the ground reference samples appear distinctive yellowness in the closest cloudless Sentinel-2 imagery, which indicates that the obtained RFD map can well capture the flowering time information of rapeseed. In addition, we also compare the RFD map with that generated from the Sentinel-2-based maximum NDVI for the 160 rapeseed ROIs (Fig. 9a). Here we empirically set the time range from January to May 2018 for extracting the maximum NDVI in order to exclude other high NDVI values that occur in non-RFP situations. It can be found that taking a window of 30 days as the tolerance for the RFD from the Weibo POIs data to generate the RFP maps can overcome the uncertainty of the social media data. A random cloudless rapeseed site (Fig. 9b-d) is taken as an example for illustration, the RFP of this site is between the 68th day (i.e., March 9th) and the 128th day (i.e., May 8th). Concerning the rapeseed fields in yellow (Fig. 9a), the DoY (Fig. 9c) corresponding to maximum NDVI (Fig. 9b) mainly includes 83

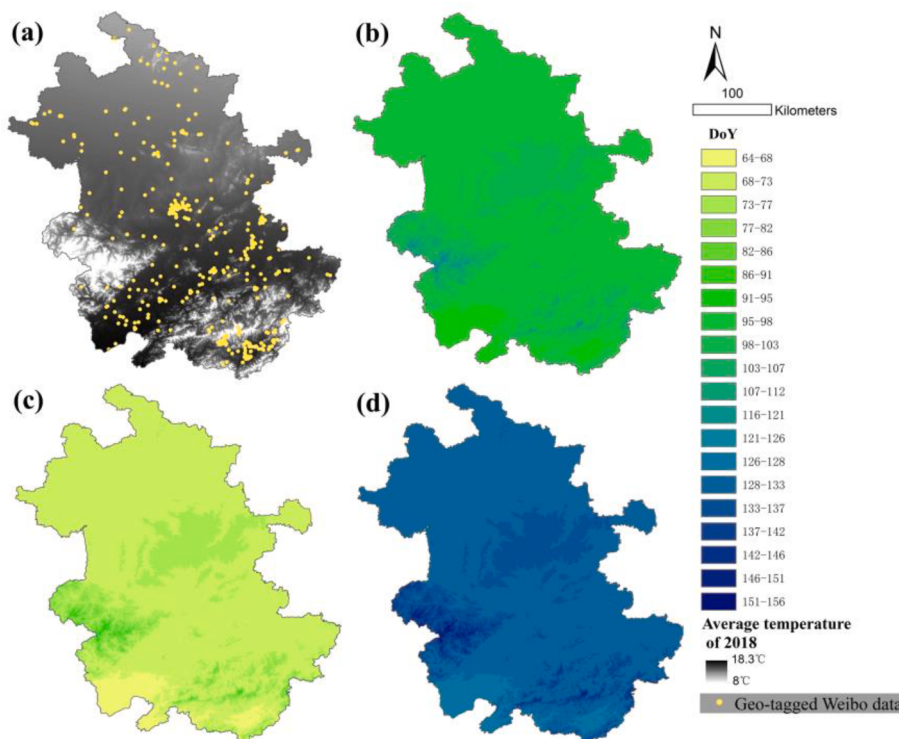


Fig. 8. Spatial distribution of (a) preprocessed geo-tagged Weibo data, (b) RFD, (c) starting RFP and (d) ending RFP of Anhui Province in 2018. The acronyms are defined as: flowering date (RFD), rapeseed flowering period (RFP).

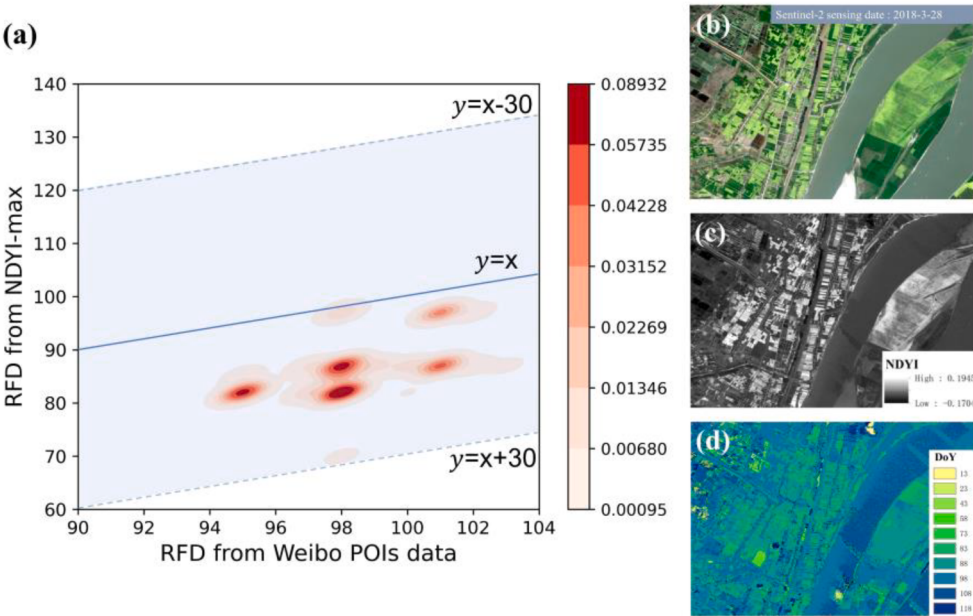


Fig. 9. (a) Comparison of the RFD derived from geo-tagged Weibo POIs data and the RFD derived from Sentinel-2-based maximum NDYI for 160 rapeseed ROIs in Anhui Province. A random rapeseed site (116.8563, 30.2255) and its (b) Sentinel-2 imagery collected during RFP, (c) map of Sentinel-2-based maximum NDYI from January to May 2018, and (d) map of DoY corresponding to Figure c are presented. The acronyms are defined as: flowering date (RFD), rapeseed flowering period (RFP), points of interest (POIs), regions of interest (ROIs), day of year (DoY), normalized difference yellowness index (NDYI).

and 88, which are within the RFP.

3.2. A map of rapeseed in 2018 for Anhui Province and its accuracy assessment

Based on the trained DNN model described in Section 2.3.4, we identify and map rapeseed in Anhui Province. A map of yellow-flowering rapeseed fields at 10-m spatial resolution is produced (Fig. 10). First, the accuracy assessment of the validation dataset is given

with a 95 % confidence interval (Table 1). The F1 score, OA, and Kappa are up to 0.97, 0.99, and 0.97, respectively. Then, the accuracy of the

Table 1
Accuracy assessment of the validation dataset with a 95% confidence interval.

F1 score	OA	Kappa
0.974 ± 0.0025	0.9938 ± 0.0006	0.9705 ± 0.0028

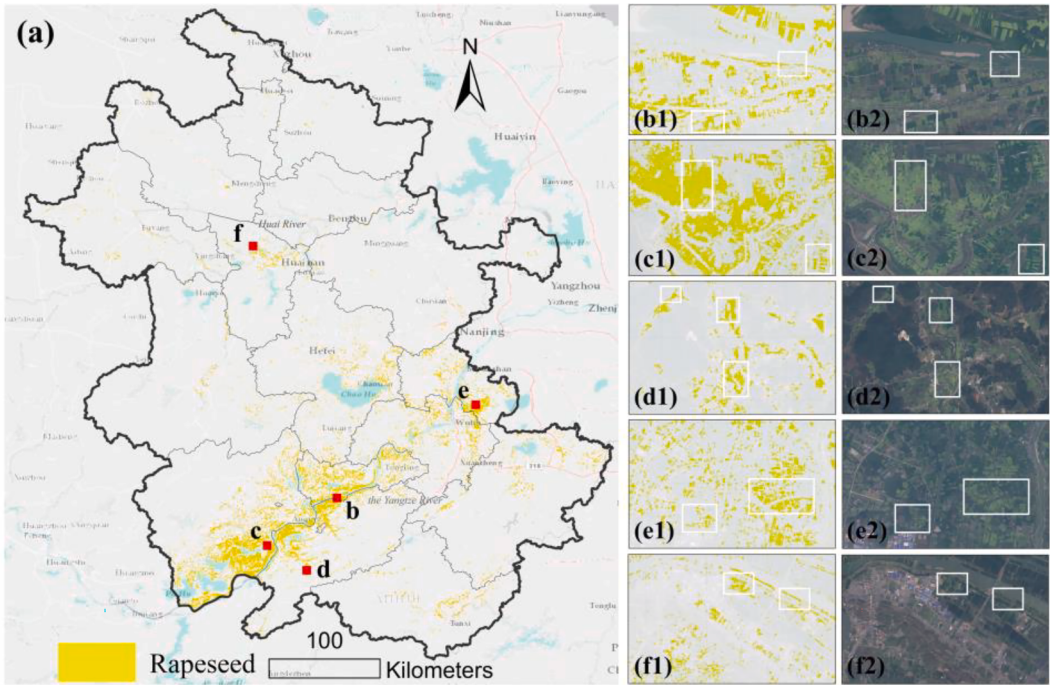


Fig. 10. (a) Rapeseed mapping for Anhui Province in 2018. Five regions in Figure a are randomly selected, which are denoted as b, c, d, e, f, and the corresponding zoom-in views are displayed in Fig. b1, c1, d1, e1, f1, and the views of Sentinel-2 imagery are displayed in b2, c2, d2, e2, f2.

resultant rapeseed map for Anhui Province is evaluated by the stratified random samples. The confusion matrix is calculated to assess the yellow-flowering rapeseed map (Table 2). Specifically, the UA and PA of rapeseed are 0.93 and 0.99, respectively. In addition, we count the number of stratified random reference samples with different landcover types identified as rapeseed by visual interpretation (Table S2). It can be observed that rapeseed is primarily confused with bare soil, construction land, water, and other green vegetations in Anhui Province.

It can be observed from the resultant rapeseed map (Fig. 10a) that a large area of yellow-flowering rapeseed fields is distributed in the southern Anhui Province. To further illustrate the rapeseed map, five random regions are presented in zoom-in views (Fig. 10b1-f1). Comparing with the Sentinel-2 imagery collected during the RFP (Fig. 10b2-f2), the rapeseed field distribution is highly consistent with the ground truth distribution. For instance, as marked by the white rectangles, the yellow-flowering rapeseed fields are well distinguished from other kinds of land cover.

Here, the area of yellow-flowering rapeseed fields in Anhui Province is presented across various scales (Fig. S5). The counties with relatively large areas of yellow-flowering rapeseed fields are concentrated in the southwestern region of Anhui Province, 5 of which have an area over 10,000 ha. Then, we sum the area of yellow-flowering rapeseed fields at the prefecture/city scale from the resultant rapeseed map. The largest area of rapeseed fields occurs in Anqing (56.62×10^3 ha). Totally, the area of the yellow-flowering rapeseed fields in Anhui Province is 212.87×10^3 ha.

3.3. Generalization of the knowledge-based method for rapeseed mapping in Hubei Province and Shaanxi Province

Fig. S6 and Fig. S7 present the resultant rapeseed maps of Hubei Province and Shaanxi Province, respectively. As can be observed, the majority of rapeseed is found throughout Hubei Province’s middle-southern and eastern regions, as well as Shaanxi Province’s southern, central, and northern regions. And the rapeseed fields are easily identified from other adjacent land cover types, including winter wheat fields, fallow fields, impervious surfaces, and rivers.

The UA and PA of the rapeseed map for Hubei Province are 0.88 and 0.99, respectively (Table 3). As for Shaanxi Province, the UA and PA of the rapeseed map are 0.83 and 0.99, respectively (Table 4). The proposed knowledge-based method for yellow-flowering rapeseed mapping can be easily transferred to other large-scale areas with quite robust performance. According to Table S2, rapeseed is mainly confused with bare soil, forest, and other green vegetations in Hubei Province and Shaanxi Province.

Here, we also estimate the area of yellow-flowering rapeseed fields across the scales from county to prefecture/city and province for Hubei Province and Shaanxi Province. The area of yellow-flowering rapeseed fields at the county scale and prefecture/city scale is presented in Figs. S8 and S9. Totally, the area of yellow-flowering rapeseed fields in Hubei Province and Shaanxi Province was 253.09×10^3 ha and 78.84×10^3 ha in 2018, respectively.

3.4. Accuracy comparison with machine learning methods

The mapping results obtained from the proposed knowledge-based method are compared with those from traditional machine learning

methods, including extreme gradient boosting (XGBOOST) and random forest (RF) (Table 5). First of all, the accuracy of our proposed method is better than traditional methods, which could be explained by avoiding the interference of the data collected during the non-RFP period among the 6 valid data when using the newly proposed DNN model. In addition, since the proposed method can automatically select the Sentinel-2 data produced during the RFP without considering the spatial heterogeneity of phenological periods, the DNN model trained in Anhui Province can be easily transferred to Hubei Province and Shaanxi Province with quite robust performance. This is the main advantage of our method over data-driven methods. In contrast, the accuracies of these two provinces as acquired from XGBOOST and RF show a substantial decline compared to the results in Anhui Province.

3.5. Comparison with the statistical data

The sown area of rapeseed in 2018 for Anhui Province, Hubei Province, and Shaanxi Province reported from the agricultural statistical data was 357.02×10^3 ha, 932.97×10^3 ha, and 176.70×10^3 ha, respectively. In comparison, the area of yellow-flowering rapeseed fields derived from the resultant rapeseed maps is generally smaller than the statistical sown area. The difference in rapeseed area derived from the satellite-based maps and the agricultural statistical dataset needs to be investigated.

One factor is that the agricultural statistical dataset reports rapeseed sown area, while our satellite-based maps report the rapeseed field area at the yellow-flowering stage, which is close to harvest area. Climate and other factors can also affect the harvested area. On the one hand, the waterlogging damage caused by the continuous rain in September and October 2017 and the freezing damage caused by the continuous low temperature and heavy snow in January 2018 severely affected rapeseed growth in the study areas. On the other hand, agricultural statistical data could be skewed because of policy factors (Seto et al., 2000). For example, rapeseed growers tend to over-report the sown area of rapeseed or just sow without subsequent management to obtain more subsidies (Khiri et al., 2017).

4. Discussion

4.1. Reliability of the proposed knowledge-based method

The collection of remote sensing data for identification and mapping of yellow-flowering rapeseed fields relies on the RFP, which is region-specific and even field-specific (Wang et al., 2019; Zhang et al., 2022). We explore the potential of social media as a new data source to obtain prior knowledge of the RFP on a large scale. To mitigate the potential uncertainty and time deviation of the RFD map derived from social media data, the RFP maps are generated by adding and subtracting 30 days from the RFD map for the inclusion of Sentinel-2 imagery during the RFP (Figs. 8 and 9). Furthermore, we develop a new DNN model for rapeseed identification, in which the 6 valid data are input in a parallel manner (Fig. 6). Combining the RFP maps and the new DNN model, the final imagery used for training and prediction is very likely to be produced during the real flowering period. Therefore, the critical issue of automatic selection of the RFP-related remote sensing imagery corresponding to different regions can be addressed, thus allowing large-scale rapeseed mapping.

Table 2
Accuracy assessment of the yellow-flowering rapeseed map for Anhui Province.

Land cover types		Reference			UA	PA	OA	Kappa
		Rapeseed	Non-rapeseed	Total				
Map	Rapeseed	362	27	389	0.9306	0.9891	0.9801	0.9458
	Non-rapeseed	4	1165	1169	0.9966	0.9773		
—	Total	9891	1192	1558				

Table 3

Accuracy assessment of the rapeseed map for Hubei Province.

Land cover types		Reference			UA	PA	OA	Kappa
		Rapeseed	Non-rapeseed	Total				
Map	Rapeseed	332	46	378	0.8783	0.9852	0.9662	0.9067
	Non-rapeseed	5	1128	1133	0.9956	0.9608		
	Total	337	1174	1511				

Table 4

Accuracy assessment of the rapeseed map for Shaanxi Province.

Land cover types		Reference			UA	PA	OA	Kappa
		Rapeseed	Non-rapeseed	Total				
Map	Rapeseed	313	64	377	0.8302	0.9874	0.9549	0.8730
	Non-rapeseed	4	1127	1131	0.9965	0.9463		
	Total	317	1191	1508				

Table 5

Comparison of the mapping accuracy obtained from different methods for the study areas. The acronyms are defined as: extreme gradient boosting (XGBOOST), random forest (RF).

Study area	Methods	UA	PA	OA	Kappa
Anhui Province	Proposed	0.93	0.99	0.98	0.95
	XGBOOST	0.65	0.97	0.92	0.73
	RF	0.60	0.99	0.91	0.69
Hubei Province	Proposed	0.88	0.99	0.97	0.91
	XGBOOST	0.43	0.95	0.87	0.52
	RF	0.44	0.98	0.88	0.55
Shaanxi Province	Proposed	0.83	0.99	0.95	0.87
	XGBOOST	0.44	0.80	0.86	0.49
	RF	0.29	0.90	0.85	0.37

Some studies extracted peak flowering period based on the variations during the RFP (Han et al., 2022, Sulik and Long, 2015, 2016). For instance, SWIR bands can observe the decline in water contents in soil and rapeseed during the RFP (Wilson et al., 2015). The reduction of chlorophyll in rapeseed petals causes an increase in red band reflectance (Shen et al., 2009). In general, temporal profiles of spectral bands or spectral indices are built to capture these changes during the RFP (d'Andrimont et al., 2020). Fig. S10 presents temporal spectral profiles of rapeseed and non-rapeseed landcover, including winter wheat, paddy

rice, and forest. Different from non-rapeseed landcover, the NDVI profile has an obvious peak, while the NDVI profile shows a downward trend during the RFP for rapeseed. Taking the temporal profile of NDVI as an example, the RFP can be detected by identifying the peaks (d'Andrimont et al., 2020). As can be observed from Fig. 11, the NDVI value of the peak is relatively high (>0.15) in 2017. However, a similar peak cannot be found during the same periods in 2018, due to the absence of cloudless Sentinel-2 TOA data in the peak flowering period. Even if there is valid data within the RFP, it may be difficult to capture the signals for the RFP extraction without valid data during the peak flowering period.

4.2. Potential sources of uncertainty and future work

It is challenging to identify the RFP across different regions and produce an accurate rapeseed map over a large spatial domain. In this study, the accuracy of the resultant rapeseed maps could be affected by some factors. First, due to the notoriously noisy nature of social media, some data irrelevant to the subject may also be included, despite data cleaning. These incorrect data will affect the spatial interpolation of RFD, especially in those areas with limited social media data. In view of the potential uncertainty caused by social media data, the following plausible future research lines are considered. The location information from text contents can enrich the geo-tagged social media data via natural language processing (Cervone et al., 2016). In addition,

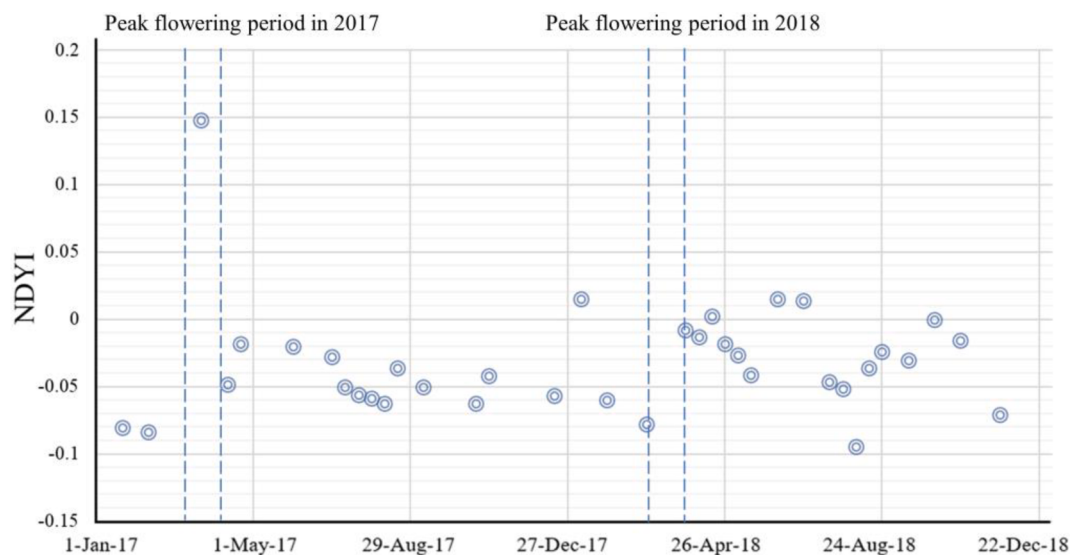


Fig. 11. Temporal profile of normalized difference yellowness index (NDYI) for the rapeseed without cloudless Sentinel-2 data during the peak flowering period in 2018 (location: [112.4729, 31.3682]).

although there are some limitations in monitoring the RFP from temporal profiles of spectral indices (e.g., NDYI), it has the potential to enhance the RFP maps generated from social media data, especially for areas with insufficient social media data.

Another potential source of uncertainty comes from the availability of Sentinel-2 data. Due to the susceptibility of optical imagery to atmospheric opacity and satellite revisit time limitations, Sentinel-2 TOA data may not be available during the RFP. Some rapeseed fields may be missed because of the inconspicuous yellowness of rapeseed at the very beginning or end of the RFP. In future research, we can adopt a few measures to alleviate this uncertainty. First, some approaches for the reconstruction of temporal Sentinel-2 data can be considered, such as the Gaussian mixture model (Mouret et al., 2022). In addition, Sentinel-1 SAR data can be included in our proposed method to increase the number of valid data. In addition, the balance between Sentinel-2 spectral redundancy and information gain should be further explored to realize more accurate model prediction based on deep learning (Inglada et al., 2022).

5. Conclusion

There is a need to identify rapeseed flowering period and produce accurate yellow-flowering rapeseed maps. In this article, a knowledge-based method is proposed for yellow-flowering rapeseed identification and mapping at a large scale by combining social media data, Sentinel-2 imagery, deep learning algorithm, and GEE. For Anhui Province, the UA and PA of the resultant rapeseed map in 2018 are 0.93 and 0.99, respectively. Moreover, our knowledge-based method can also be directly applied for rapeseed mapping in Hubei Province and Shaanxi Province with robust performance. It can be concluded that the proposed method has great potential to identify and map yellow-flowering rapeseed fields at the national scale and provides essential information for agricultural management.

CRedit authorship contribution statement

Zhenjie Liu: Writing – original draft, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Yingyue Su:** Software, Methodology, Data curation. **Xiangming Xiao:** Supervision, Funding acquisition, Conceptualization. **Yuanwei Qin:** Validation, Investigation. **Jun Li:** Supervision, Data curation. **Luo Liu:** Visualization, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This study was supported in part by research grants from the U.S. National Science Foundation (1911955).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jag.2024.104047>.

References

- Breckling, B., Laue, H., Pehlke, H., 2011. Remote sensing as a data source to analyse regional implications of genetically modified plants in agriculture—oilseed rape (*brassica napus*) in northern germany. *Ecol. Indic.* 11 (4), 942–950.
- Cervone, G., Sava, E., Huang, Q., Schnebele, E., Harrison, J., Waters, N., 2016. Using twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study. *Int. J. Remote Sens.* 37 (1), 100–124.
- Crippen, R., Buckley, S., Agram, P., Belz, E., Gurrola, E., Hensley, S., Kobrick, M., Laval, M., Martin, J., Neumann, M., Nguyen, Q., Rosen, P., Shimada, J., Simard, M., Tung, W., 2016. NASADEM global elevation model: Methods and progress. *Int. Arch. Photogramm. Remote Sens. Spat. Inform. Sci.* 41, 125–128.
- d'Andrimont, R., Taymans, M., Lemoine, G., et al., 2020. Detecting flowering phenology in oil seed rape parcels with Sentinel-1 and-2 time series. *Remote Sens. Environ.* 239, 111660.
- Davis, S.C., Anderson-Teixeira, K.J., DeLucia, E.H., 2009. Life-cycle analysis and the ecology of biofuels. *Trends. Plant Sci.* 14 (3), 140–146.
- Deng, Y., Wu, C., Li, M., Chen, R., 2015. RNDSI: A ratio normalized difference soil index for remote sensing of urban/suburban environments. *Int. J. Appl. Earth Obs. Geoinf.* 39, 40–48.
- Glorot, X., Bordes, A., Bengio, Y., 2011. Deep sparse rectifier neural networks. *Proceedings of the International Conference on Artificial Intelligence and Statistics*, pp. 315–323.
- Gregorutti, B., Michel, B., Saint-Pierre, P., 2017. Correlation and variable importance in random forests. *Stat. Comput.* 27 (3), 659–678.
- Han, J., Zhang, Z., Cao, J., Luo, Y., 2022. Mapping rapeseed planting areas using an automatic phenology-and pixel-based algorithm (APPA) in Google Earth Engine. *Crop J.* 10 (5), 1483–1495.
- Huang, X., Wang, S., Lu, T., Liu, Y., Serrano-Estrada, L., 2024. Crowdsourced geospatial data is reshaping urban sciences. *Int. J. Appl. Earth Obs. Geoinf.* 103687.
- Huete, A., Liu, H., Batchily, K., Van Leeuwen, W., 1997. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sens. Environ.* 59 (3), 440–451.
- Hutchinson, M. F., Xu, T. B., 2013. ANUSPLIN Version 4.4 User Guide. Canberra: Fenner School of Environment and Society, the Australian National University.
- Inglada, J., Michel, J., Hagolle, O., 2022. Assessment of the usefulness of spectral bands for the next generation of Sentinel-2 satellites by reconstruction of missing bands. *Remote Sens.* 14 (10), 2503.
- Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *International Conference on Machine Learning* 448–456.
- Khair, K.A., Ismail, N., Balu, N., 2017. Competitiveness of the rapeseed industry in China. *Oil Palm Indust. Econ. J.* 17 (2), 8–24.
- Kussul, N., Lavreniuk, M., Skakun, S., Shelestov, A., 2017. Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geosci. Remote Sens. Lett.* 14 (5), 778–782.
- Liu, Z., Qiu, Q., Li, J., Wang, L., Plaza, A., 2020. Geographic optimal transport for heterogeneous data: Fusing remote sensing and social media. *IEEE Geosci. Remote Sens.* 59 (8), 6935–6945.
- Liu, W., Zhang, H., 2023. Mapping annual 10 m rapeseed extent using multisource data in the Yangtze River Economic Belt of China (2017–2021) on Google Earth Engine. *Int. J. Appl. Earth Obs. Geoinf.* 117, 103198.
- Luo, T., Zhang, J., Khan, M.N., Liu, J., Xu, Z., Hu, L., 2018. Temperature variation caused by sowing dates significantly affects floral initiation and floral bud differentiation processes in rapeseed (*Brassica napus* L.). *Plant Sci.* 271, 40–51.
- Marjanović-Jeromela, A., Terzić, S., Jankulovska, M., Zorić, M., Kondić-Špika, A., Jocković, M., Hristov, N., Crnobarac, J., Nagl, N., 2019. Dissection of year related climatic variables and their effect on winter rapeseed (*Brassica napus* L.) development and yield. *Agronomy* 9 (9), 517.
- McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* 17 (7), 1425–1432.
- Meng, S., Zhong, Y., Luo, C., Hu, X., Wang, X., Huang, S., 2020. Optimal temporal window selection for winter wheat and rapeseed mapping with Sentinel-2 images: A case study of Zhongxiang in China. *Remote Sens.* 12 (2), 226.
- Mouret, F., Albughdadi, M., Duthoit, S., Kouamé, D., Rieu, G., Tournet, J.Y., 2022. Reconstruction of Sentinel-2 derived time series using robust Gaussian mixture models—Application to the detection of anomalous crop development. *Comput. Electron. Agric.* 198, 106983.
- Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* 148, 42–57.
- Qi, C. R., Su, H., Mo, K., Guibas, L. J., 2017. Pointnet: Deep learning on point sets for 3D classification and segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 652–660.
- Rußwurm, M., Courty, N., Emonet, R., Lefevre, S., Tuia, D., Tavenard, R., 2023. End-to-end learned early classification of time series for in-season crop type mapping. *ISPRS J. Photogramm. Remote Sens.* 196, 445–456.
- Seto, K.C., Kaufmann, R.K., Woodcock, C.E., 2000. Landsat reveals China's farmland reserves, but they're vanishing fast. *Nature* 406 (6792), 121.
- Shankar, P., Bitter, C., Liwicki, M., 2020. Digital crop health monitoring by analyzing social media streams. *IEEE/TTU International Conference on Artificial Intelligence for Good* 87–94.
- Shen, M., Chen, J., Zhu, X., Tang, Y., 2009. Yellow flowers can decrease NDVI and EVI values: Evidence from a field experiment in an alpine meadow. *Can. J. Remote Sens.* 35 (2), 99–106.

- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15 (1), 1929.
- Sulik, J.J., Long, D.S., 2015. Spectral indices for yellow canola flowers. *Int. J. Remote Sens.* 36 (10), 2751–2765.
- Sulik, J.J., Long, D.S., 2016. Spectral considerations for modeling yield of canola. *Remote Sens. Environ.* 184, 161–174.
- Wang, D., Fang, S., Yang, Z., Wang, L., Tang, W., Li, Y., Tong, C., 2018. A regional mapping method for oilseed rape based on HSV transformation and spectral features. *ISPRS Int. J. Geo-Inf.* 7 (6), 224.
- Wang, C., Hai, J., Yang, J., Tian, J., Chen, W., Chen, T., Luo, H., Wang, H., 2016. Influence of leaf and silique photosynthesis on seeds yield and seeds oil quality of oilseed rape (*Brassica napus* L.). *Eur. J. Agron.* 74, 112–118.
- Wang, L., Wu, B., Zhang, M., Xing, Q., 2019. Winter wheat and rapeseed classification during key growth period by integrating multi-source remote sensing data. *J. Geo-Inf. Sci.* 21 (7), 1121–1131.
- Wang, L., Wang, J., Liu, Z., Zhu, J., Qin, F., 2022. Evaluation of a deep-learning model for multispectral remote sensing of land use and crop classification. *Crop J.* 10 (5), 1435–1451.
- Wilson, R.H., Nadeau, K.P., Jaworski, F.B., Tromberg, B.J., Durkin, A.J., 2015. Review of short-wave infrared spectroscopy and imaging methods for biological tissue characterization. *J. Biomed. Opt.* 20 (3), 030901.
- Wu, F., Wu, B., Zhang, M., Zeng, H., Tian, F., 2021. Identification of crop type in crowdsourced road view photos with deep convolutional neural network. *Sensors* 21 (4), 1165.
- Xiao, X., Boles, S., Liu, J., Zhuang, D., Frolking, S., Li, C., Salas, W., Moore III, B., 2005. Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. *Remote Sens. Environ.* 95 (4), 480–492.
- Yao, J., Wu, J., Xiao, C., Zhang, Z., Li, J., 2022. The classification method study of crops remote sensing with deep learning, machine learning, and Google Earth engine. *Remote Sens.* 14 (12), 2758.
- Zhao, N., Cao, G., Zhang, W., Samson, E.L., Chen, Y., 2020. Remote sensing and social sensing for socioeconomic systems: A comparison study between nighttime lights and location-based social media at the 500 m spatial resolution. *Int. J. Appl. Earth Obs. Geoinf.* 87, 102058.
- Zhao, H., Chen, Z., Jiang, H., Jing, W., Sun, L., Feng, M., 2019. Evaluation of three deep learning models for early crop classification using sentinel-1A imagery time series-A case study in Zhanjiang, China. *Remote Sens.* 11 (22), 2673.
- Zhong, L., Hu, L., Zhou, H., 2019. Deep learning based multi-temporal crop classification. *Remote Sens. Environ.* 221, 430–443.
- Zhu, Z., Wang, S., Woodcock, C.E., 2015. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sens. Environ.* 159, 269–277.