Article

High-Resolution Mapping of Maize in Mountainous Terrain Using Machine Learning and Multi-Source Remote Sensing Data

Luying Liu, Jingyi Yang, Fang Yin * and Linsen He

Shaanxi Key Laboratory of Land Consolidation, School of Land Engineering, Chang'an University, Xi'an 710054, China; 2023235005@chd.edu.cn (L.L.); jingyiyang@chd.edu.cn (J.Y.); 2023135011@chd.edu.cn (L.H.)
* Correspondence: yinf@chd.edu.cn; Tel.: +86-29-8233-9936

Abstract: In recent years, machine learning methods have garnered significant attention in the field of crop recognition, playing a crucial role in obtaining spatial distribution information and understanding dynamic changes in planting areas. However, research in smaller plots within mountainous regions remains relatively limited. This study focuses on Shangzhou District in Shangluo City, Shaanxi Province, utilizing a dataset of high-resolution remote sensing images (GF-1, ZY1-02D, ZY-3) collected over seven months in 2021 to calculate the normalized difference vegetation index (NDVI) and construct a time series. By integrating field survey results with time series images and Google Earth for visual interpretation, the NDVI time series curve for maize was analyzed. The Random Forest (RF) classification algorithm was employed for maize recognition, and comparative analyses of classification accuracy were conducted using Support Vector Machine (SVM), Gaussian Naive Bayes (GNB), and Artificial Neural Network (ANN). The results demonstrate that the random forest algorithm achieved the highest accuracy, with an overall accuracy of 94.88% and a Kappa coefficient of 0.94, both surpassing those of the other classification methods and yielding satisfactory overall results. This study confirms the feasibility of using time series high-resolution remote sensing images for precise crop extraction in the southern mountainous regions of China, providing valuable scientific support for optimizing land resource use and enhancing agricultural productivity.

Keywords: machine learning; NDVI; time series; classification; remote sensing; *Zea mays*

1. Introduction

Maize (*Zea mays* L.) is one of China's three major food crops, known for its high production potential and significant economic benefits. It serves multiple purposes, including as food and animal feed, and various industrial uses, making it strategically important for ensuring food security [1]. Maize is widely cultivated across China due to its high yield, strong drought resistance, cold tolerance, adaptability to poor soils, and overall environmental resilience. As a typical dryland crop, maize has distinct water requirements at different growth stages. According to relevant studies, under high-yield conditions, the total water demand for summer maize throughout its growth period ranges from 417.30 mm to 507.45 mm. Specifically, the water requirement during the seedling stage is 16.80–33.75 mm; during the jointing stage, it is 94.35–130.8 mm; during the tasseling stage, it is 92.85–108.15 mm; and during the grain-filling stage, it is 181.05–267.0 mm [2]. Crop coefficient (Kc) refers to the ratio of the water requirement of a crop during different growth stages to the reference crop evapotranspiration [3]. It is a key parameter in calculating evapotranspiration and water demand and plays an important guiding role in precision

irrigation and water conservation in agriculture. According to FAO56, the crop coefficients for maize during the early, mid, and late growth stages are 0.3, 1.2, and 0.6, respectively [4]. The crop coefficient is influenced by various factors, such as crop type, soil properties, climatic conditions, and irrigation methods. The crop coefficient varies for different crops. Even for the same crop, the coefficient fluctuates as vegetation grows and as surface characteristics and environmental conditions change [5]. Therefore, in practical applications, it is necessary to adjust the crop coefficient appropriately based on local factors such as climate, soil, irrigation methods, and crop varieties.

In recent years, both the area under maize cultivation and its yield have shown stable growth. Timely and accurate acquisition of spatial distribution information for maize planting can assist agricultural departments in optimizing resource allocation, rationally planning maize farmland, and providing data for the formulation of local agricultural subsidy policies. This is crucial for improving agricultural production efficiency [6–10].

With the rapid development of remote sensing technology, it has been widely applied in areas such as vegetation classification, environmental pollution monitoring, earthquake monitoring, land-use planning, crop pest and disease monitoring, and crop yield surveys [11]. By acquiring high-resolution images and multispectral data through satellites, drones, and sensors, a wide range of surface information can be captured, accurately identifying crop types, boundaries, shapes, and environmental changes. Some scholars have combined these data with machine learning methods, providing an important tool for crop monitoring [12]. This approach enables the automatic identification, classification, and yield prediction of different crop types and has been widely applied in the field of crop area identification. For example, Chen Yuehao et al. [13] used GF-2 satellite data and two different classifiers—Maximum Likelihood and Support Vector Machine—to identify and extract tomatoes in the Yuanmou hot zone. Similarly, Yang Yanjun [14] constructed an NDVI time series covering the full growth cycle of crops using GF-1 satellite images and employed various classification methods such as Maximum Likelihood, Minimum Distance, Mahalanobis Distance, Support Vector Machine, and Artificial Neural Network to classify crops in the southern region of Tangshan City, Hebei Province. Wei Pengfei et al. [15] used multi-temporal GF-1 satellite remote sensing images and combined typical vegetation indices of major crops in the study area, employing classification methods such as Maximum Likelihood, Support Vector Machine, and Decision Trees to classify the crops. Their results indicated that the Decision Tree method was the best, successfully extracting the spatial distribution of soybean, rice, maize, and sweet potato planting areas in the study region. Qiao Shuting et al. [16] used time-series Sentinel-2 satellite remote sensing data, combined with field survey data of typical ground features, and applied the Random Forest classification algorithm to successfully extract a remote sensing dataset of the main crop planting distribution in the Sanjiang Plain for the years 2020–2022.

Previous research has demonstrated that the integration of remote sensing technology with machine learning significantly enhances the ability to process and analyze remote sensing data, particularly achieving remarkable results in crop information extraction. However, most studies have focused on large, relatively flat plain areas, with relatively few studies on small, irregular plots in southern mountainous regions where the terrain is more complex. In small, irregular plots, the spatial resolution of medium- to low-resolution images (such as Sentinel-2) is insufficient to capture subtle changes, making it difficult to reveal the complexity of mountainous terrain and plot diversity. Additionally, due to the low temporal resolution of high-resolution images, relying solely on it makes it challenging to provide continuous time-series data, complicating the monitoring of crop growth and land-use changes in mountainous areas.

To address this challenge, this study encompasses the entire crop growth cycle by utilizing multi-source satellite remote sensing images from GF-1, ZY-3, and ZY1-02D to construct an NDVI time series for seven months in 2021 (January, February, March, May, August, November, and December). Drawing on the rich texture and spectral characteristics of the data, along with field survey results, visual interpretation was conducted to select training samples. Four classification algorithms—Gaussian Naive Bayes, Artificial Neural Network, Support Vector Machine, and Random Forest—were employed to classify the study area, accompanied by a comparative accuracy analysis of the results. The Random Forest algorithm, which demonstrated the highest accuracy, was chosen to identify and extract maize planting areas in Shangzhou District. This study serves as a reference for crop classification and precision agricultural management using remote sensing technology in mountainous and hilly regions with complex terrain.

2. Materials and Methods

2.1. Study Area

The study area is located in Shangzhou District, Shangluo City, Shaanxi Province (Figure 1). The main crop in this area, maize (*Zea mays* L.), typically enters the sowing and seedling stage from mid-April to early June; the jointing stage occurs from mid-June to early July; the tasseling stage occurs from mid-July to early August; the grain-filling and maturity stage occurs from mid-August to early September; and the harvesting stage occurs from mid-September to early October, as shown in Figure 2. The NDVI curve changes during the phenological periods of maize are shown in Figure 3. Shangzhou District lies in the southeastern part of the Shaanxi province, on the southern slopes of the eastern section of the Qinling Mountains and the upper reaches of the Danjiang River. It borders Danfeng County to the east, Shanyang County to the south, and connects to Lantian and Zhashui counties to the west via the Qinling mountain range. To the north, it adjoins Luonan County and is situated between latitudes 33°38′–34°11′ N and longitudes 109°30′–110°14′ E. The district extends 67.5 km from east to west and 65 km from north to south, covering a total area of 2672 km^2 [17].

Shangzhou, positioned in the mid-latitudes, benefits from the natural barrier of the Qinling Mountains to the northwest, which protects the region from cold air intrusions. The southeast-facing valleys promote the influx of warm temperate air, resulting in a monsoon, semi-humid mountainous climate typical of the transitional zone at the southern edge of the warm temperate belt. The area experiences four distinct seasons, characterized by mild winters and cool summers. Winters and springs are prolonged, while summers and autumns are brief, with the region enjoying a balance of water and heat resources. However, there are significant interannual variations in temperature and precipitation, along with frequent natural disasters such as droughts, floods, and hailstorms. The annual average temperature is 12.8 °C, with July being the hottest month (averaging 24.8 °C) and January the coldest (averaging 0.3 °C).

The combined effects of climate, terrain, and soil conditions have established maize as one of the main crops in the region. Irrigation methods primarily rely on traditional channel irrigation as well as modern sprinkler and drip irrigation systems. Proper irrigation practices can effectively reduce water resource waste and ensure the normal growth of crops during dry seasons. In 2022, the gross domestic product (GDP) of Shangzhou District reached 16.168 billion yuan. By industry, the added value of the primary industry (agriculture, forestry, animal husbandry, and fisheries) was 2.030 billion yuan, the secondary industry (mining; manufacturing; production and supply of electricity, heat, gas, and water; and construction) was 4.3403 billion yuan, and the tertiary industry (services) was for 9.798 billion yuan.
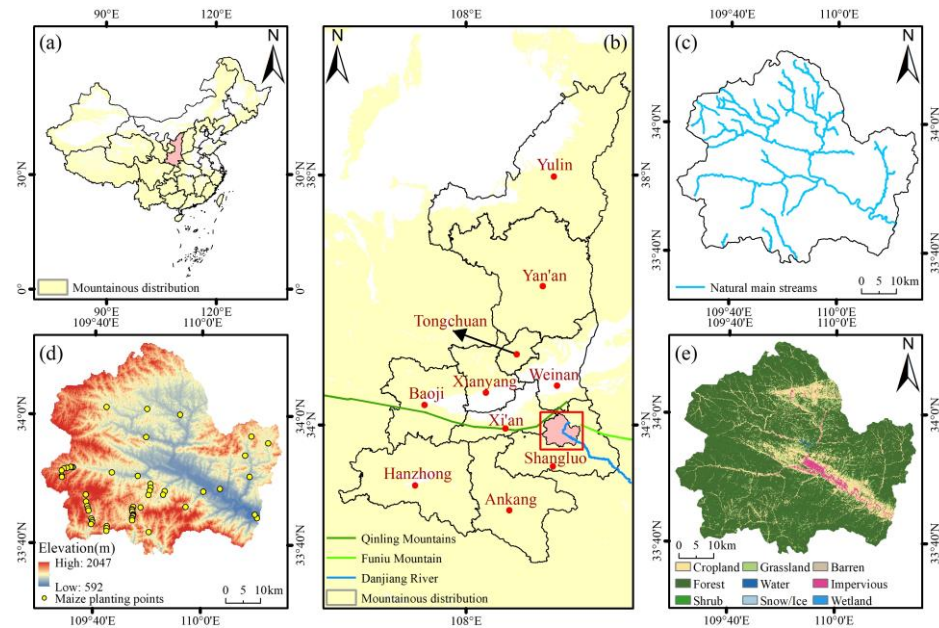
Figure 1. Location of the study area and land-use features: (**a**) administrative divisions of China; (**b**) administrative divisions of Shaanxi Province; (**c**) main natural rivers in Shangzhou District; (**d**) elevation map and distribution of maize planting points; (**e**) land-use status map of the study area in 2023.
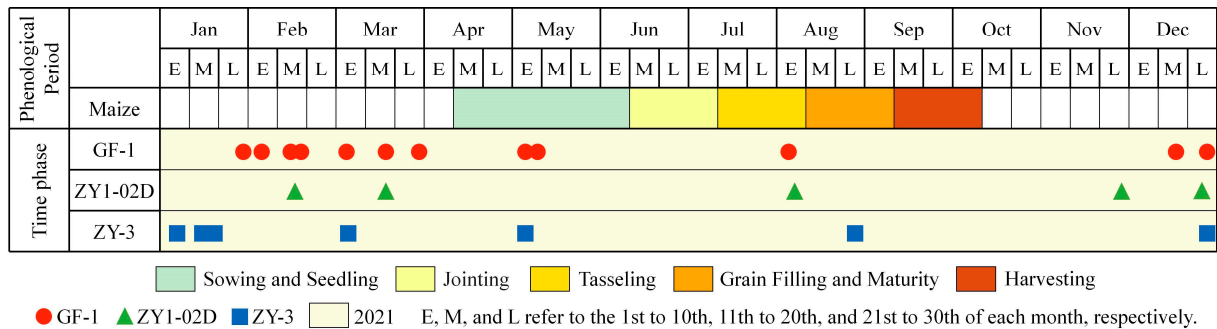

Figure 2. Maize phenological period and image acquisition dates.
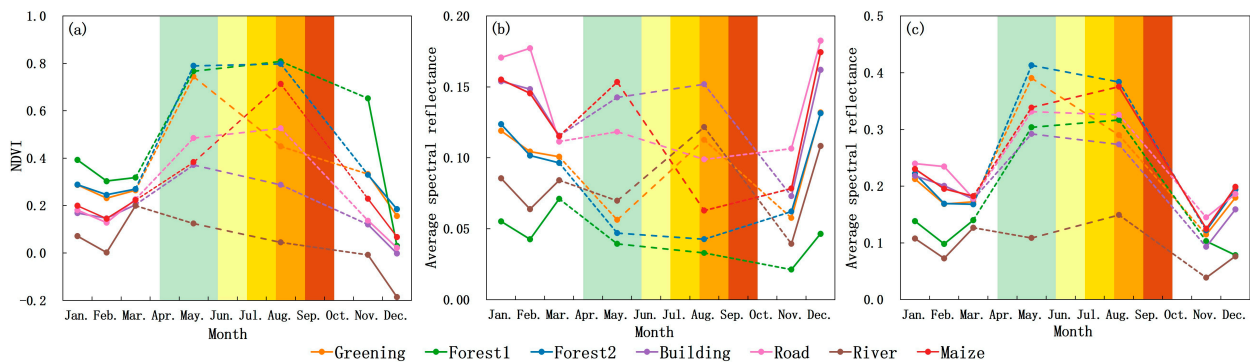
Figure 3. NDVI time series curves and the arithmetic mean spectral reflectance for different land cover types: (**a**) NDVI time series for major land cover types; (**b**) average spectral reflectance in the red band for major land cover types; (**c**) average spectral reflectance in the near-infrared band for major land cover types; the solid line is used to represent periods with continuous data, while the dashed line is used to connect periods with missing data.

2.2. Data and Preprocessing

2.2.1. Remote Sensing Images

Based on the administrative boundaries of Shangzhou District and the phenological period of maize, remote sensing images from 2020 to 2024 were selected. However, each year exhibited varying degrees of data gaps or incomplete coverage of the study area. Although the 2021 data also had missing months, they were still the best available data compared to other years. Therefore, this study utilized three high-resolution satellite images covering the area in 2021, including images from Gaofen-1 (GF-1), Ziyuan-1 02D (ZY1-02D), and Ziyuan-3 (ZY-3). A total of 42 scenes were obtained in 2021. Among them, six scenes were obtained during the maize sowing and seedling stage from mid-April to early June, three were obtained scenes during the tasseling stage from mid-July to early August, and two scenes were obtained during the grain filling and maturity stage from mid-August to early September. No images were acquired for the jointing and harvesting stages, mainly due to heavy cloud cover during these periods, which prevented effective acquisition. Data acquisition for other time periods is as follows: 10 scenes in January, 6 in February, 8 in March, 2 in November, and 5 in December. All data were downloaded from the China Resource Satellite Data and Application Center (http://www.cresda.com.cn (accessed on 9 October 2023)). However, the high-resolution data provided by this platform are not open to international researchers. If international researchers need to access remote sensing data, they can browse through the Natural Resources Satellite Remote Sensing Cloud Service Platform (https://www.sasclouds.com/ (accessed on 21 September 2023)) and obtain data from the SPACE WILL platform (http://en.spacewillinfo.com/ (accessed on 13 January 2025)). Other relevant information is provided in Table 1, and the acquisition times of the different images are shown in Figure 2.

Table 1. Summary of data information.

Data Type	Resolution	Time Range	Data Source	Data Characteristics
Remote Sensing Data (GF-1)	Panchromatic: 2 m Multispectral: 8 m	2021	http://www.cresda.com.cn (accessed on 9 October 2023)	High-resolution optical remote sensing data, useful for land feature identification and monitoring
Remote Sensing Data (ZY1-02D)	Panchromatic: 2.5 m Multispectral: 10 m	2021	http://www.cresda.com.cn (accessed on 9 October 2023)	Used for obtaining surface information, supporting resource survey applications
Remote Sensing Data (ZY-3)	Nadir panchromatic: 2.1 m Nadir multispectral: 5.8 m	2021	http://www.cresda.com.cn (accessed on 9 October 2023)	Features multiple imaging modes suitable for surveying and mapping
Google Earth High-resolution Images	-	2020	https://earth.google.com/ (accessed on 25 October 2023)	Google Earth allows free browsing of high-resolution satellite imagery from around the world
Digital Elevation Model (DEM)	30 m	-	http://www.gscloud.cn (accessed on 31 October 2024)	Represents terrain elevation data and supports terrain analysis
Evapotranspiration	1 km	2019–2023	https://data.tpdc.ac.cn (accessed on 15 December 2024)	Reflects surface water evaporation and plant transpiration conditions
Rainfall	1 km	2019–2023	https://data.tpdc.ac.cn (accessed on 15 December 2024)	Reflects precipitation data, important for agriculture and other sectors
Temperature	-	2019–2023	https://www.ncei.noaa.gov (accessed on 15 December 2024)	Reflects atmospheric temperature conditions, fundamental for climate analysis
Land Use Data	30 m	2023	https://zenodo.org (accessed on 15 December 2024)	Represents land use types and their distribution

GF-1 Data [18,19]: The Gaofen-1 (GF-1) satellite was successfully launched on 26 April 2013. The satellite is equipped with two high-resolution cameras, which provide 2 m panchromatic and 8 m multispectral imaging, as well as four wide-swath multispectral cameras with 16 m resolution. The acquired data include panchromatic images with a spatial resolution of 2 m and multispectral images with a spatial resolution of 8 m, the latter featuring four bands: blue, green, red, and near-infrared.

ZY1-02D Data [20]: The Ziyuan-1 02D (ZY1-02D) satellite was launched on 12 September 2019. It is equipped with both a visible-near infrared camera and a hyperspectral camera. The acquired data include panchromatic images with a 2.5 m spatial resolution and multispectral images with a 10 m spatial resolution.

ZY-3 Data [21]: The ZY-3 satellite was successfully launched on 9 January 2012. It is equipped with four optical cameras, including a 2.1 m resolution nadir-viewing panchromatic Time Delayed and Integration–Charge Coupled Devices (TDI-CCD) camera, two 3.6 m resolution forward and backward-viewing panchromatic TDI-CCD cameras, and a 5.8 m resolution nadir-viewing multispectral camera. The acquired data include panchromatic images with a 2.1 m spatial resolution and multispectral images with a 5.8 m spatial resolution, the latter featuring four bands: blue, green, red, and near-infrared.

Due to the fragmented terrain of the study area, the resolution of remote sensing data is crucial for the accuracy of crop classification. Therefore, this study compares three Chinese remote sensing satellites with the widely used Landsat 8 and Sentinel-2 satellites, as shown in Table 2, further highlighting the advantages of Chinese high-resolution satellite data for crop classification in mountainous and fragmented terrains.

Table 2. Comparison of parameters for GF-1, ZY1-02D, ZY-3, Landsat 8, and Sentinel-2 satellites.

Satellite	GF-1	ZY1-02D	ZY-3	Landsat 8	Sentinel-2
Sensor	PMS	VNIC	-	OLI	MSI
Number of bands	5	9	5	9	13
Spatial resolution (m)	Panchromatic: 2 Multispectral: 8	Panchromatic: 2.5 Multispectral: 10	Nadir panchromatic: 2.1 Nadir multispectral: 5.8	Panchromatic: 15 Multispectral: 30	10, 20, 60
Revisit period (days)	41	55	59	16	5
Swath width (km)	60	115	Nadir panchromatic: 50 Nadir multispectral: 52	185	290
Application	Provide high-quality, high-resolution data for land-use planning, environmental monitoring, resource management, disaster response, and other applications.	Applied to land resource surveys, urban and rural construction, statistical surveys, environmental monitoring, precision agriculture, disaster monitoring, and other areas.	Applied to land resource surveying and monitoring, disaster prevention and mitigation, agriculture, forestry, water conservancy, ecological environment, urban planning and construction, transportation, major national projects, and other areas.	Applied to watershed and regional ecological environment monitoring, land use type extraction, biomass estimation, crop growth monitoring in reclamation areas, vegetation coverage inversion, crop planting area estimation, and other areas.	Mainly used for global high-resolution and high-revisit land observation, biophysical change mapping, monitoring of coastal and inland water areas, as well as disaster mapping, and other applications.

2.2.2. Preprocessing

For the selected remote sensing images, preprocessing was conducted using ENVI 5.3 software to ensure data accuracy and lay a solid foundation for subsequent analysis [22–24]. First, radiometric calibration, atmospheric correction, and orthorectification were applied to the multispectral data for each acquisition phase, while radiometric calibration and orthorectification were also performed on the corresponding panchromatic data. Radiometric calibration converts the digital values of the image into physical quantities such as radiance, reflectance, or surface temperature, providing a reliable data foundation for subsequent analysis. This step was completed using the Radiometric Calibration tool. After radiometric calibration, atmospheric correction was performed to eliminate the effects of atmospheric transmission on the image, making it more realistic and reliable. Atmospheric correction was performed using the FLAASH Atmospheric Correction tool, with parameters including sensor altitude, ground elevation, atmospheric model, aerosol model, aerosol retrieval, initial visibility, and spectral files, with adjustments made based on the characteristics of the image data. Orthorectification was used to eliminate geometric distortions from the image and precisely align it with the geographic coordinate system, ensuring spatial accuracy. This process was implemented using the RPC Orthorectification Workflow tool, with bilinear resampling chosen and appropriate output pixel size set according to the resolution of the image data. Finally, the NNDiffuse Pan Sharpening tool was used to perform image fusion between the multispectral and panchromatic data.

After preprocessing, the resolution of the GF-1 images was 2 m, the ZY-02D images had a resolution of 2.5 m, and the ZY-3 images had a resolution of 2.1 m. To facilitate

subsequent data operations and change analysis, the resolutions of the ZY1-02D and ZY-3 image data were resampled to a uniform 2 m. By mosaicking and cropping, a dataset of remote sensing images for Shangzhou District covering seven months was created, followed by geometric correction to eliminate or correct any geometric errors in the images.

The study area is extensive and encompasses various land-use types. Each high-resolution remote sensing image contains rich geographic information, but its large data volume can make processing cumbersome. After mosaicking and cropping the multi-source remote sensing images to create the NDVI time series, the data volume increased further, complicating rapid processing in a single batch. Due to the limited hardware resources of the computer, such as the CPU and memory, directly processing large-scale datasets may lead to excessive CPU load, high memory usage, and even system crashes, resulting in slower processing speeds. To improve data processing efficiency, a chunking method was employed [25,26]. In ENVI 5.3 software, the Simple Frame Subset tool was used to divide the entire study area image into smaller blocks, with both the row and column numbers set to 6. Edge blank blocks were removed, resulting in 28 smaller, more manageable sub-blocks. The specific chunking criteria were based on the image size and the computer's memory capacity, ensuring that the size of each sub-block was suitable for memory processing and preventing memory overflow due to excessive data loading. The entire process was carried out on a computer with 16 GB of RAM and a 6-core, 12-thread processor, and the chunking operation took approximately 7 min. This approach effectively reduces memory requirements and enhances data processing efficiency, facilitating the successful handling and classification of large-scale remote sensing images.

2.2.3. Sample Data

Field surveys were conducted in Shangzhou District, Shangluo City, from 13 June to 15 June 2023 and from 30 June to 1 July 2024. During these surveys, handheld GPS (Garmin eTrex309X, Manufacturer: Garmin Ltd., Olathe, KS, USA) devices were used to collect location information for maize planting areas within the study region. A total of 74 maize planting points were recorded.

In addition, based on the NDVI time series curves of different land cover types, detailed observations and analysis were conducted using high-resolution images from Google Earth (May 2020) and the images acquired during the maize tasseling period in August 2021. This approach was used to identify and distinguish various land cover types within the study area. Through visual interpretation, random and evenly distributed training samples were selected, resulting in distribution data for seven land cover types, totaling 11,223 samples. This included 166 river samples, 757 road samples, 641 building samples, 2545 forest type 1 (shady slopes) samples, 2504 forest type 2 (sunny slopes) samples, 2178 maize samples, and 2432 (planted protective forests, as well as various plants in residential areas and parks) greening samples. These samples provide critical data for distinguishing different land cover types and establish a solid foundation for subsequent classification model training and validation.

2.2.4. Other Data

To ensure the reliability and representativeness of the research results, this study selected 2021 as the primary year for analysis. Meanwhile, temperature, precipitation, and evapotranspiration data from 2019 to 2023 for the study area were collected, with detailed information provided in Table 1. The annual variation trends of the three types of data are shown in Figure 4.
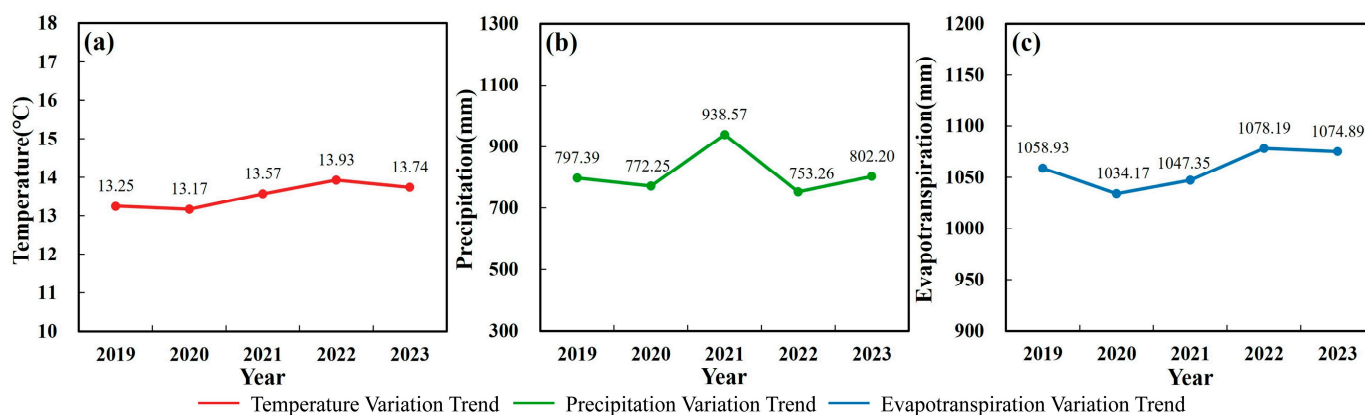
Figure 4. Annual variation trends of temperature, precipitation, and evapotranspiration in the study area from 2019 to 2023: (**a**) annual variation trend of temperature; (**b**) annual variation trend of precipitation; (**c**) annual variation trend of evapotranspiration.

As shown in Figure 4, the precipitation in 2021 (938.57 mm) was relatively high, while both the temperature (13.57 °C) and evapotranspiration (1047.35 mm) were close to the multi-year average, indicating that the climate conditions in 2021 were generally representative. Furthermore, considering that maize cultivation in the study area exhibits good adaptability to changes in temperature and precipitation, the 2021 data can objectively reflect the spatial distribution characteristics of maize cultivation in the region.

2.3. NDVI Timeseries Construction

2.3.1. NDVI Calculation

NDVI (normalized difference vegetation index) is a widely used remote sensing metric for assessing and monitoring plant health, vegetation cover, and biomass. The calculation formula is as follows:

$$NDVI = (NIR - Red)/(NIR + Red) \tag{1}$$

where NIR (near infrared) refers to the reflectance value in the near-infrared band; Red refers to the reflectance value in the red band.

NDVI changes over time, reflecting the growth stages of crops [27,28]. The differences in NDVI time series curves are also quite pronounced among different land cover types or crop species [29]. Therefore, ENVI 5.3 software was used to calculate the NDVI for images from each month and to synthesize these into a 2021 NDVI time series dataset, yielding NDVI time series curves for typical land cover types such as maize, forest, road, building, and river. The maize time series curve clearly shows the entire growth process from seedling stage to ear and grain formation stages.

2.3.2. NDVI Time Series Curve and Spectral Feature Analysis of Maize

Different crops exhibit significant variations in their growth cycles, which are reflected in the morphology of their NDVI time series curves. The peak values and the timing of these peaks are distinctive for each crop and follow certain patterns. Within the same area, despite being influenced by factors such as climate, soil, and management practices, the growth processes of the same crop tend to follow relatively consistent trends. This regularity provides a theoretical basis for land cover identification based on time series remote sensing data. By analyzing these time series, it becomes possible to effectively identify crop types and monitor their growth conditions, offering valuable insights for precision agriculture and resource management.

By combining sample data obtained from field surveys and visual interpretation, the average NDVI values corresponding to maize sample points in the NDVI images for

each month were calculated, resulting in the NDVI time series curve for maize over seven months. Using the same approach, NDVI time series curves for other typical land cover types can also be derived. Additionally, the arithmetic mean spectral reflectance of the red band (B3) and the near-infrared band (B4) for typical land cover types was compared with the NDVI curve, as shown in Figure 3. It is worth noting that the NDVI values of rivers in Figure 3a are negative, which is a normal phenomenon. This is because the spectral reflectance characteristics of water bodies typically result in a higher reflectance in the red band than in the near-infrared band, leading to a negative NDVI value. This phenomenon is expected and reflects the unique spectral reflectance behavior of water bodies.

Due to the missing data for some months, this study does not constitute a complete time series dataset. Therefore, based on previous research findings [1,14,30–32] and field survey data, the following conclusions can be drawn. In Shangzhou District, the period from mid-April to early June is typically the sowing and seedling stage for maize. During this early stage, the surface is mainly bare soil with low vegetation cover and weak absorption capacity. Compared to other land covers, the red band reflectance of maize during the sowing period is relatively high, while the near-infrared band reflectance is slightly lower. As maize begins to emerge and grow, the NDVI value rises from low levels.

From mid-June to early July, maize enters the jointing stage, during which rapid growth leads to a significant increase in NDVI values. From mid-July to early August, maize reaches the tasseling phase, where vegetation cover is high, and NDVI values continue to rise, reaching a peak. During this phase, near-infrared band reflectance also reaches its highest value, significantly exceeding that of other land cover types, while the red band reflectance remains low.

As maize progresses into the grain-filling and maturity stage, both NDVI values and near-infrared band reflectance begin to decrease, while red band reflectance gradually increases. By mid-September to early October, the harvest period for maize, vegetation cover significantly decreases, returning to levels similar to those observed at sowing. By analyzing the changes in these three reflectance curves (red, near-infrared, and NDVI), maize can be effectively distinguished from other land cover types, providing a solid foundation for subsequent classification and identification.

2.4. Classification Methods

This study uses the Python programming language, with Python 3.9 and PyCharm 2024 as the development environment to build classification models. All data are divided into training, validation, and test sets in a ratio of 8:1:1 and standardized to improve the model's training performance. Subsequently, four classification models—Gaussian Naive Bayes, Artificial Neural Network, Support Vector Machine, and Random Forest—are constructed. Hyperparameter tuning is performed using a combination of grid search and cross-validation to select the optimal model parameters. Finally, the optimized model is validated using the test set to evaluate its accuracy.

2.4.1. Gaussian Naive Bayes

Gaussian Naive Bayes (GNB) is a machine learning classification algorithm based on probabilistic models and Gaussian distributions. It assumes that the conditional probability of each feature follows a Gaussian distribution and applies Bayes' theorem to calculate the posterior probabilities of a sample belonging to each class based on the given feature distribution. The class of the sample is determined by maximizing the posterior probability [33,34].

GNB is especially effective in handling continuous features and typically provides good classification performance, especially when the features exhibit a Gaussian distribu-

tion. Additionally, GNB has a relatively low computational complexity, making it suitable for large-scale datasets. It can quickly train and classify with a large sample size, and the training time is relatively short. The var_smoothing parameter for the GNB is set to 0.001. This parameter is used to smooth the variance and prevent the occurrence of zero variance in the training data for certain classes.

2.4.2. Support Vector Machine

Support Vector Machine (SVM) is a general linear classifier introduced by Vapnik et al. in 1995, which performs binary classification based on supervised learning. The fundamental idea of SVM is to map data points in a high-dimensional space to a lower-dimensional space and find an optimal hyperplane that separates the data points into two classes [35,36]. Common kernels include linear, polynomial, and radial basis function kernels [37]. Choosing the appropriate kernel function is crucial for SVM performance. By using kernel functions, SVM can handle both linearly separable and complex non-linear classification problems. SVM is widely favored in practical applications for its strong classification performance and robustness. The SVM utilizes a radial basis function (RBF) kernel, with the penalty parameter C set to 100.

2.4.3. Random Forest Classification Algorithm

Random Forest (RF) is a powerful ensemble learning method that constructs multiple decision trees and combines their results through voting or averaging to obtain the final prediction [38–40]. It can effectively handle high-dimensional data and large-scale samples. During training, RF builds each decision tree by randomly sampling subsets from the original dataset and selecting a subset of features at each node split. RF excels in prediction and classification performance compared to single decision trees. By integrating multiple models, it effectively handles overfitting, improves generalization [41], and is robust to missing values and imbalanced data. The RF model consists of 300 trees. For each tree, the maximum number of features is set to the square root of the total number of input features. A minimum of two samples is required to split a node, and each tree's leaf node must contain at least one sample.

2.4.4. Artificial Neural Network

Artificial Neural Network (ANN) are computational models inspired by biological neural systems designed to simulate and process complex information tasks. The core concept of ANN is to mimic the connections between neurons in the human brain, enabling input-to-output mapping through layers of processing. In this study, Multilayer Perceptron (MLP) is used as the implementation of ANN. MLP is a feedforward neural network with an input layer, multiple hidden layers, and an output layer [42]. Each layer comprises multiple neurons (or nodes), where the output of one layer becomes the input for the next layer [43,44]. The strength of MLP lies in its use of multiple nonlinear processing layers to extract and transform feature information from the input data. This hierarchical structure allows the MLP to learn and approximate complex nonlinear functions, exhibiting strong generalization ability, especially when handling large-scale and high-dimensional datasets. The ANN consists of two hidden layers, with 128 and 64 neurons in the first and second layers, respectively.

2.5. Evaluation Methods

The accuracy of the classification results is assessed using a confusion matrix, which is generated by comparing the location and classification of each reference pixel with those in the classified image. Key evaluation metrics include the Kappa coefficient, Overall Accuracy (OA), User's Accuracy (UA), and Producer's Accuracy (PA) [45,46]. These metrics provide

a comprehensive evaluation of image classification accuracy from various perspectives. The accuracy standard deviation is further calculated based on the following formula for standard deviation:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \tag{2}$$

where σ represents the standard deviation; x_i represents the i-th data point; μ represents the mean of the data; N represents the total number of data points.

3. Results

3.1. Classification Results Analysis

Based on Google Earth images, field survey data, and NDVI time series curves for different land cover types, visual interpretation was conducted in the study area to select sample points for various land cover types, resulting in data for seven categories. The study area was classified using four classification algorithms: Gaussian Naive Bayes, Artificial Neural Network, Support Vector Machine, and Random Forest. This successfully extracted spatial distribution information for typical land cover types in Shangzhou District, including maize, road, building, and river.

To enhance the reliability of the decision-making process, bootstrap simulation was used to assess the predictive confidence of the remote sensing classification model. By performing resampling with a replacement on the original dataset, multiple simulated datasets were generated, allowing the evaluation of the model's performance across different datasets and providing reliable confidence intervals and error estimates for the classification model. The confusion matrix for the classification results was calculated, and four metrics were used to evaluate the classification accuracy of each machine learning method: user's accuracy, overall accuracy, producer's accuracy, and the Kappa coefficient. The user's accuracy represents the proportion of samples that truly belong to a specific category and are correctly classified as such. The producer's accuracy indicates the probability that the classification results at a given location on the map match the corresponding sample in the validation data. The overall accuracy refers to the proportion of correctly classified pixels in the classification results relative to the total number of pixels. While the user's accuracy and producer's accuracy provide insights into the performance of individual categories, overall accuracy and the Kappa coefficient assess the overall classification performance. The classification accuracies for each method are summarized in Table 3.

Table 3. Classification accuracy of various machine learning methods.

Method	User's Accuracy (%)	Producer's Accuracy (%)	OA (%)	Kappa	Mean Accuracy (%)	Standard Deviation of Accuracy
	Maize	Maize				
GNB	90.51%	78.71%	78.74%	0.74	78.56%	0.0027
ANN	91.67%	94.47%	91.49%	0.90	91.14%	0.0028
SVM	90.80%	93.68%	90.41%	0.88	90.29%	0.0021
RF	97.05%	92.80%	94.88%	0.94	94.01%	0.0017

All methods except for Gaussian Naive Bayes achieved high accuracy in identifying maize planting areas, with overall and average accuracies above 90% and Kappa coefficients greater than 0.85. The Random Forest algorithm performed the best, achieving an overall accuracy of 94.88% and a Kappa coefficient of 0.94, significantly outperforming the other algorithms. Additionally, Random Forest's average accuracy was 94.01%, with a very low standard deviation of 0.0017, indicating high stability and consistency. Compared to

Gaussian Naive Bayes, Artificial Neural Network, and Support Vector Machine, Random Forest's overall accuracy was higher by 16.14%, 3.39%, and 4.47%, respectively, with Kappa coefficients exceeding these by 0.2, 0.04, and 0.06, respectively. In user and producer accuracy, Random Forest achieved the highest user accuracy (97.05%), while Artificial Neural Network achieved the highest producer accuracy (94.47%).

In conclusion, random forest demonstrated the best performance in terms of accuracy, stability, and consistency, followed by Artificial Neural Network and Support Vector Machine, with Gaussian Naive Bayes showing the lowest accuracy. This indicates that Random Forest is particularly effective for accurately identifying maize locations in Shangzhou District's complex terrain.

Figure 5 illustrates the distribution of classification accuracy across 100 iterations for the four algorithms. Both Random Forest and Support Vector Machine exhibit high stability, with minimal fluctuations across iterations, though Random Forest shows smaller fluctuations than Support Vector Machine, indicating a more robust model and lower error. In contrast, Artificial Neural Network and Gaussian Naive Bayes displayed larger fluctuations across iterations, indicating lower stability. Overall, the Random Forest classification algorithm demonstrates the highest reliability for classification tasks.
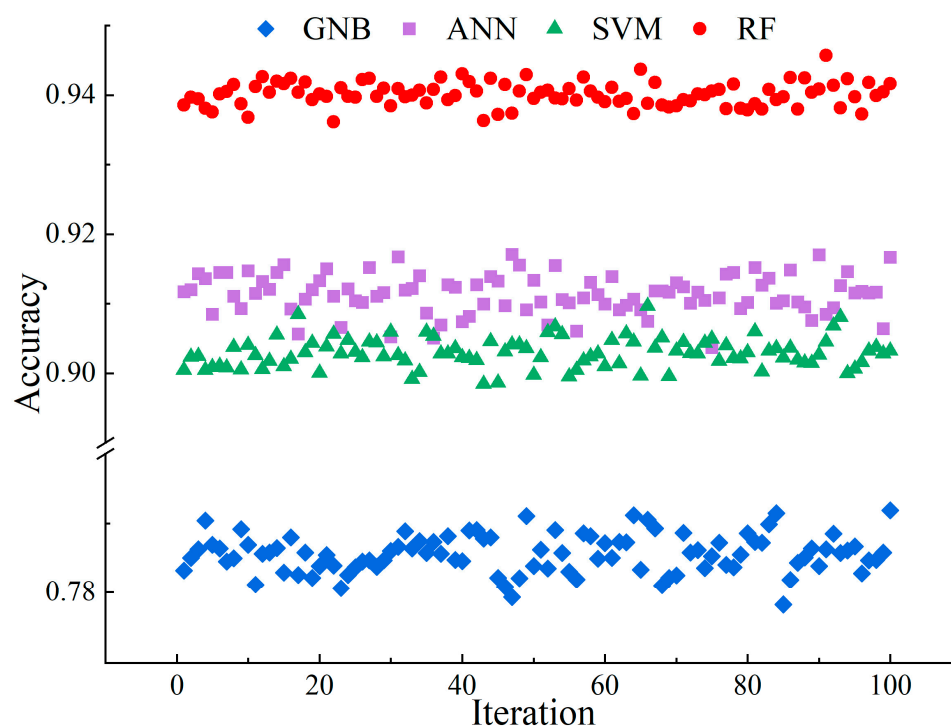

Figure 5. Distribution of accuracy for each machine learning method.

In Shangzhou District, the substantial terrain variations lead to different levels of solar radiation across surface positions, creating shaded and sunlit slopes in the images, which in turn affects the spectral reflectance of surface vegetation. Forest 1 and Forest 2 correspond to the shaded and sunlit slopes of the forest, respectively. For clarity, these are combined in the resulting map, as shown in Figures 6 and 7.

Field survey results reveal that the study area is characterized by interwoven mountain ranges and extensive forest coverage. The dense vegetation and complex terrain limit the expansion of arable land, resulting in the significant fragmentation of farmland. Although maize is the primary crop, its distribution is scattered, with few large, contiguous planting areas.
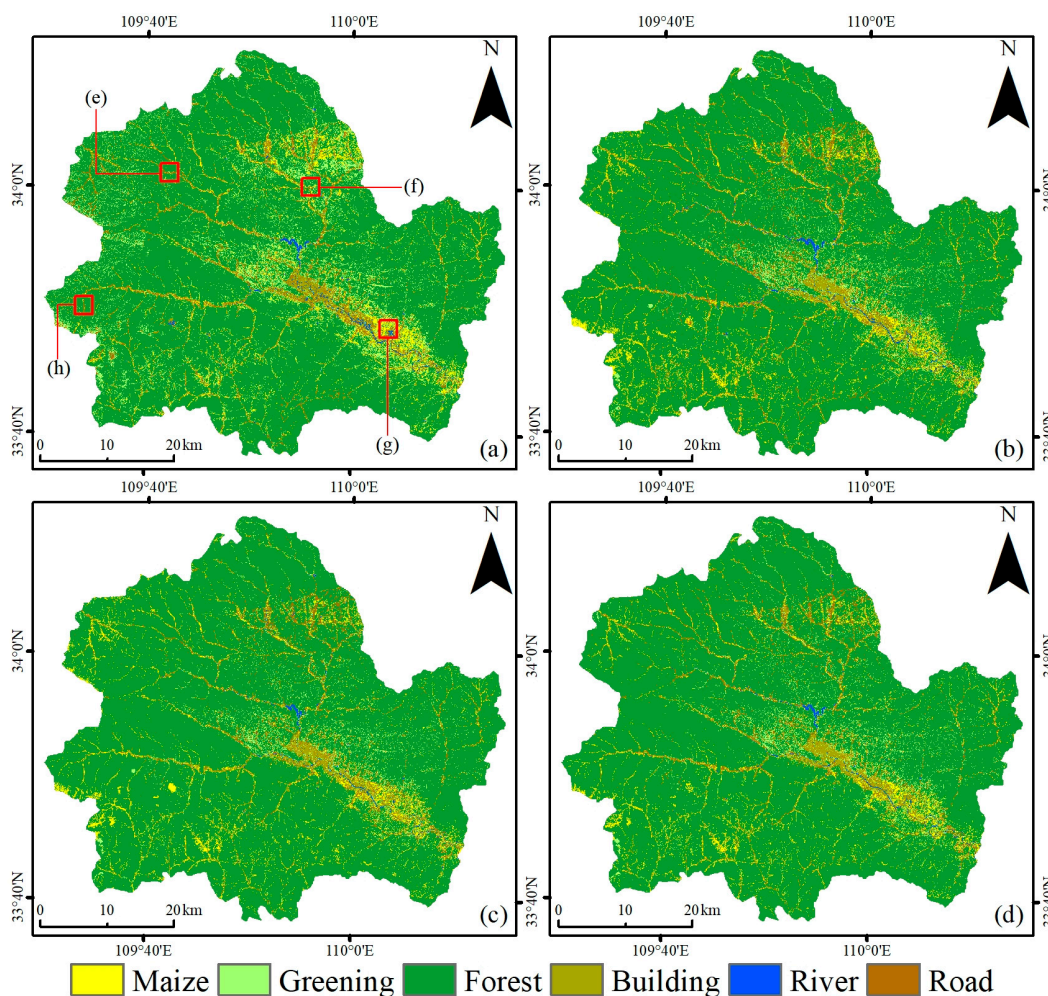
(e) is a typical area in the mid-altitude region, (f) and (g) are typical areas in the
low-altitude region, and (h) is a typical area in the high-altitude region.

Figure 6. Classification results of various machine learning methods: e, f, g, and h represent the four
typical regions of the study area; (**a**) Gaussian Naive Bayes classification results; (**b**) Artificial Neural
Network classification results; (**c**) Support Vector Machine classification results; (**d**) Random Forest
classification results.

When comparing the four classification results to remote sensing images, most land
cover types were accurately identified, yielding generally satisfactory classification out-
comes. As shown in Figure 6, forests cover approximately 86.65% of the total area, with
clear delineation of their location and boundaries. Developed areas are concentrated in the
southeastern valley, making up 2.02% of the total area. Despite their smaller size, linear
features like rivers and roads are distinctly visible in the classification results. Maize fields
are mostly located in valleys and along roads where the terrain is relatively flat, supporting
cultivation and the transport of agricultural products.

The four classification algorithms performed well for identifying linear features and
larger land areas. However, for maize, which is grown in smaller, fragmented plots, the
results showed varying degrees of misclassification and omissions. Of the methods used,
the Random Forest algorithm demonstrated better performance in maintaining the integrity
of crop plot boundaries and reducing the misclassification of small patches.
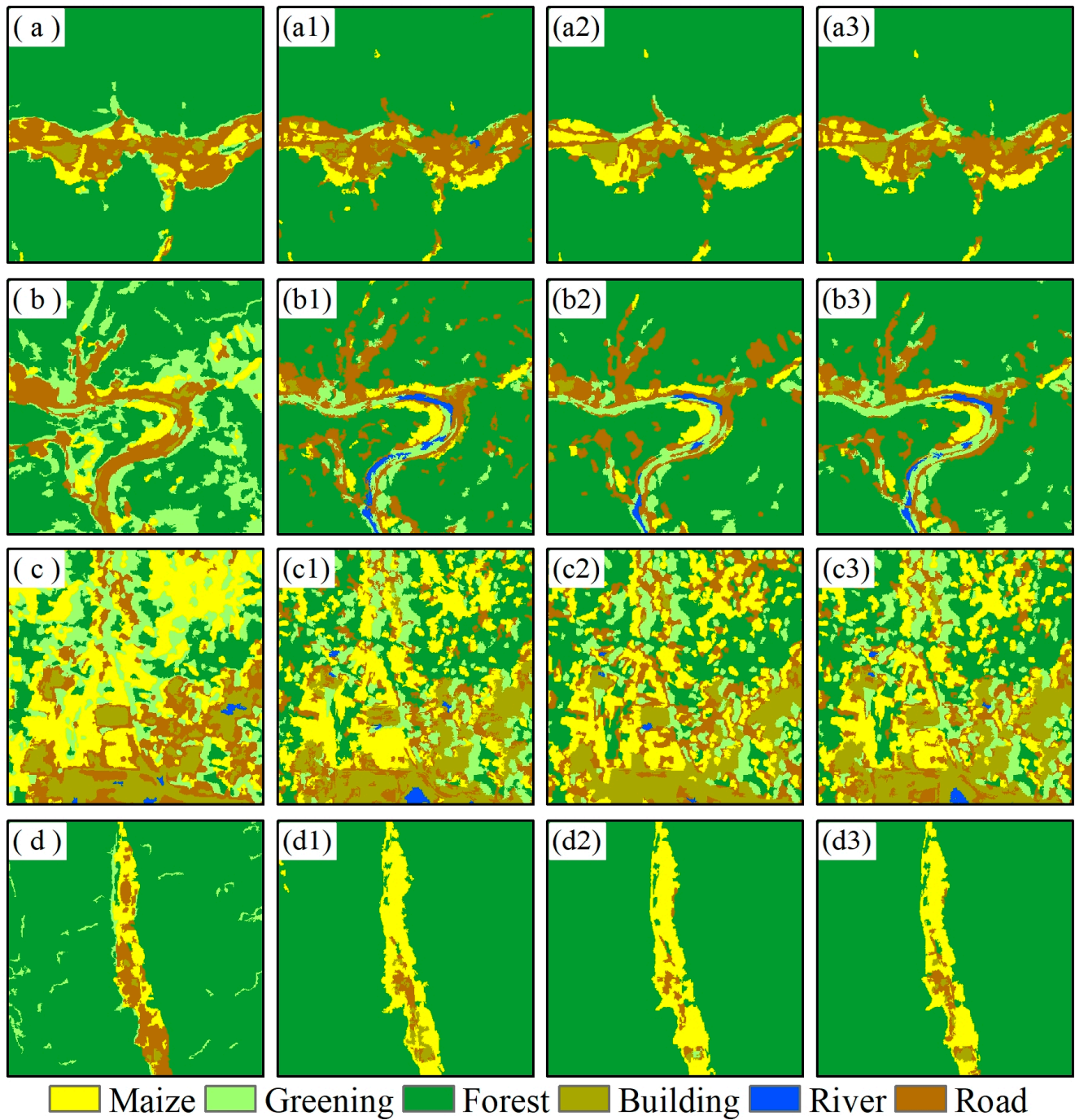
Figure 7. Typical area classification results of various machine learning methods: (**a**–**d**) represent the classification results using Gaussian Naive Bayes for four typical regions of the study area; (**a1,b1,c1,d1**) represent the classification results using Artificial Neural Network; (**a2,b2,c2,d2**) represent the classification results using Support Vector Machine; (**a3,b3,c3,d3**) represent the classification results using Random Forest.

3.2. Extraction of Maize Planting Areas

Using the Random Forest classification algorithm, the spatial distribution of maize planting areas in Shangzhou District was extracted, revealing a relatively uniform overall distribution. However, due to variations in topography, external environmental conditions, the application of agricultural practices, and local policy differences, the area and distribution of maize cultivation significantly differ among townships. Overall, the southern areas have a slightly higher distribution than the northern ones, particularly in the southeastern towns of Shahezi and Yecun, as well as in the southwestern regions of Yanchihe Township, Heishan Town, and Yangyuhe Town, as shown in Table 4.

Table 4. Maize cultivation area in each township.

Town Name	Planting Area (km^2)	Town Name	Planting Area (km^2)
Yangyuhe Town	6.73	Banqiao Town	5.61
Sanshilipu Township	3.61	Heilongkou Town	5.31
Beikuanping Town	3.75	Yangxie Town	6.17
Muhuguan Town	3.24	Heishan Town	7.77
Dajing Town	2.98	Yancun Township	3.26
Urban areas	1.81	Sanchahe Township	1.72
Jinlingsi Town	1.17	Shangguanfang Township	2.29
Chenyuan Street Office	1.00	Xijing Township	1.73
Yecun Town	15.76	Yaoshi Town	6.01
Yanchihe Township	9.17	Shahezi Town	11.20
Machihe Township	3.87	Majie Town	1.36

Four typical areas were selected for result presentation and analysis based on different elevation levels. Area a is located in the mid-elevation region, areas b and c are in the low-elevation region, and area d is situated in the high-elevation region. As shown in Figure 8, cultivated land in areas a, b, and d is mainly situated near residential areas on mid-slopes and distributed along roads, with relatively small maize plots surrounded by extensive forested regions. In contrast, area c is located in a flatter region near an industrial building, where farmland is more concentrated, though fragmentation is still present. Overall, the maize planting area in region c is larger than in regions a, b, and d.

The NDVI curve variations for maize across the four areas are generally consistent and align with typical maize growth patterns, showing peak NDVI values of 0.79, 0.71, 0.69, and 0.82, though with some numerical differences. These variations may be influenced by factors like elevation and soil fertility specific to each area. Climatic elements, including temperature, precipitation, and sunlight, significantly affect crop growth and health; drought or excessive rainfall, for instance, can inhibit growth and reduce NDVI values. Soil type also plays a role, as more fertile soils better support healthy crop development. Furthermore, agricultural management practices, such as fertilization, irrigation, and pest control, impact crop conditions, leading to NDVI differences for the same crop across regions or over different years. The severe fragmentation of farmland in this mountainous study area aligns with observations from field surveys.
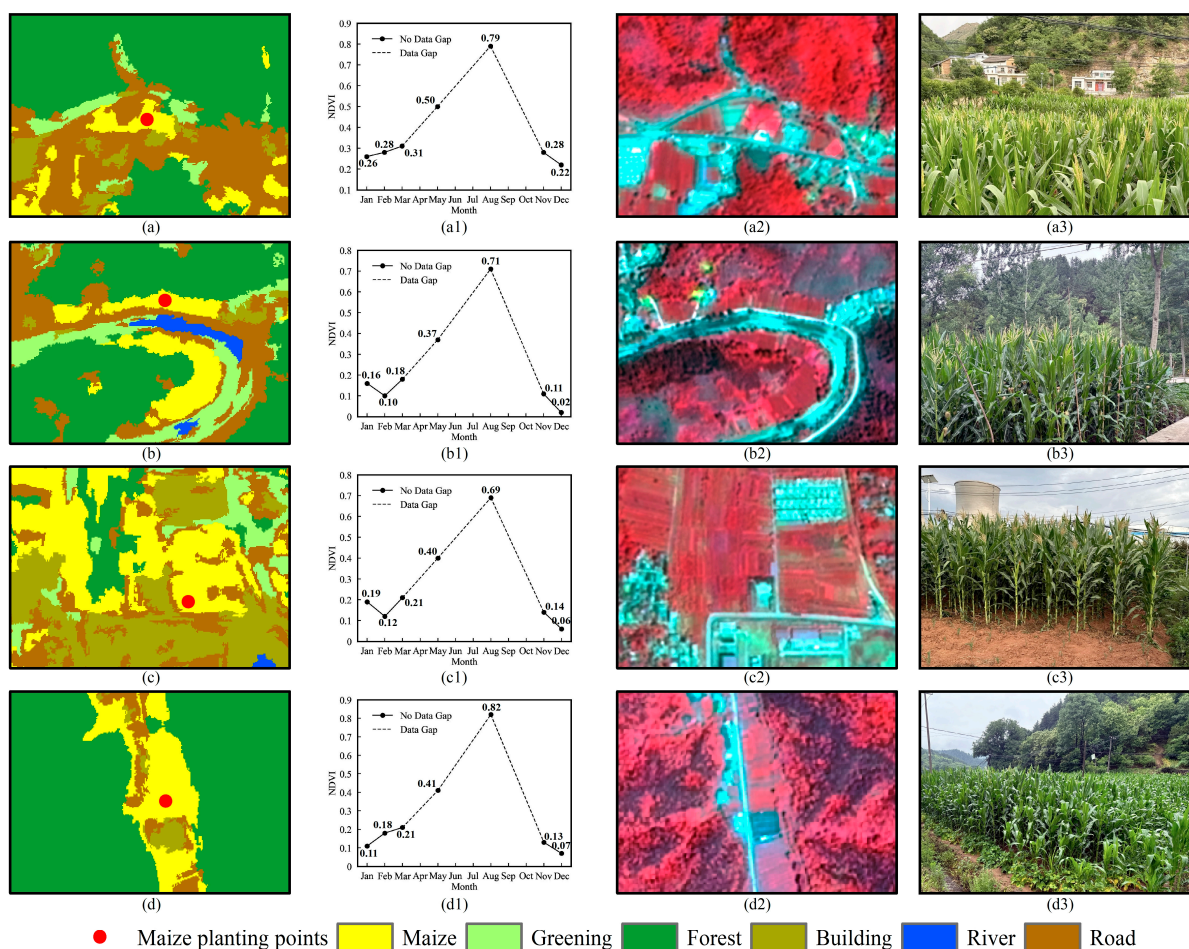
Figure 8. Distribution of main crop maize in typical areas: (**a**–**d**) are the classification results of typical regions; (**a1**,**b1**,**c1**,**d1**) are the NDVI curves for maize in these regions; (**a2**,**b2**,**c2**,**d2**) are the images of these regions; (**a3**,**b3**,**c3**,**d3**) are field photographs of these regions.

4. Discussion

This study utilizes high-resolution satellite images from GF-1, ZY1-02D, and ZY-3 satellites, combined with NDVI time series from multiple months, to identify and extract maize—the primary crop in Shangzhou District, Shaanxi Province—and distinguish other land cover types, such as forested and green areas within the study area.

4.1. The Impact of Topographical Characteristics and Temporal Differences in Image Data Selection on Classification Results

The terrain in the southern Shaanxi hilly region is complex, with the distribution of mountains and hills leading to significant spatial heterogeneity in agricultural land. In the hilly areas, maize is typically concentrated in valleys and flat slopes, whereas its distribution is limited by the terrain in steep and high-altitude areas. Most studies have focused on plains or areas with large plots, often relying on medium- to low-resolution remote sensing data. Compared to plain areas, land use in hilly regions is more fragmented, with maize cultivation areas being more dispersed and individual cultivation plots being smaller. When using medium- to low-resolution remote sensing data for research, it is often difficult to accurately capture the subtle terrain variations of mountains and the complex boundaries of small plots, which can result in mixed vegetation types within a single pixel, affecting the accuracy of classification. Therefore, conducting research in areas with complex terrain requires higher spatial resolution of remote sensing data. To address this challenge, this study selected three high-resolution remote sensing data from

GF-1, ZY1-02D, and ZY-3 in 2021 for crop classification, fully utilizing the advantages of high-resolution remote sensing data. In the process of visual interpretation, the detailed surface features and textures were successfully captured, allowing for clear identification and differentiation of the boundaries of different plots.

Due to the temporal difference between the 2021 remote sensing data and the field survey data, land use changes (such as the conversion of cultivated land to built-up land or other land types) in the study area may have some impact on the accuracy of the analysis and the validity of the results. However, according to the actual field survey results, the main crops in the study area are maize and wheat. The maize planting patterns and plot distributions have remained relatively stable over the past few years, and the planting locations and areas have not undergone significant changes due to the temporal discrepancy. Additionally, most of the selected sample points are located in stable farming areas in valleys or along roads, where cultivation has a long history and land use changes occur gradually. These areas are highly representative and consistent. Therefore, the 2021 remote sensing data can accurately reflect the current maize planting areas, and the impact of the temporal difference on the results is limited. However, for future large-scale areas with more frequent land use changes, the temporal difference may have a greater impact on the analysis results. In such cases, it is necessary to strengthen monitoring and analysis of land use changes to ensure temporal consistency between field surveys and the data used.

4.2. Maize Growth Dynamics Analysis and Comparison of Remote Sensing Classification Algorithm Accuracy

The dynamic changes in maize throughout its growth cycle were analyzed using the NDVI time series curve, which clearly reveals the different characteristics of each key growth stage, including sowing, seedling emergence, jointing, tasseling, and harvesting. The NDVI values exhibit a clear pattern of variation at different growth stages: during the sowing and early seedling stages, the NDVI values are relatively low, indicating that the crop is still in the early growth phase with low vegetation cover. As the growth progresses, the plants become increasingly vigorous, and the NDVI values begin to rise, reaching their peak during the tasseling stage, marking the peak of maize growth. After entering the maturation and harvesting stages, the NDVI values significantly decline as the plants deteriorate and water content decreases. This dynamic change not only helps distinguish the growth characteristics of maize from other crops but also provides a reliable basis for crop recognition using remote sensing data. In agricultural monitoring and crop classification, the NDVI change characteristics exhibit significant differences between crops, effectively differentiating crop types and improving the accuracy of crop recognition. This provides high-quality training samples for subsequent classification models, helping to reduce misclassification and omission errors, thereby enhancing classification accuracy, especially in complex terrains or regions with alternating crop distributions. Moreover, these analysis results offer strong support for crop growth monitoring and precision agriculture management, enabling the accurate capture of crop growth dynamics during different growth stages. This facilitates precise agricultural decision-making and real-time scheduling, providing data support for formulating scientific agricultural policies and optimizing planting management.

In remote sensing image classification, the choice of classification classifier plays a crucial role in determining both classification accuracy and efficiency. Zhang Peng et al. [47] conducted a study on crop classification at the plot scale in complex planting areas using WorldView-2 images. The results showed that the RF algorithm achieved an overall accuracy of 79.07%, outperforming ANN and K-Nearest Neighbors (KNNs). Similarly, Zheng et al. [48] found that in the fusion of visible light and multispectral data acquired by drones for crop classification, the RF algorithm achieved the highest overall accuracy of

97.77%, significantly surpassing other machine learning models such as SVM, ANN, KNN, and Classification and Regression Trees (CARTs). These findings are consistent with the results of this study, where the RF algorithm effectively distinguishes maize from other land cover types, achieving an overall accuracy of 94.88%, which is higher than that of GNB, ANN, and SVM. Analysis of the classification results from the four methods reveals the presence of significant "salt-and-pepper" noise. This phenomenon may be attributed to the inherent complexity of the land cover types, leading to the occurrence of "same object different spectra" and "different objects same spectra" within the image, as well as mixed pixels at the boundaries of land cover types containing multiple categories. These mixed radiometric values can cause misclassification. As a result, there is some degree of mixed classification between maize and other vegetation types, which makes the boundary between crop planting areas and other vegetation types less distinct. Compared to the other three classification algorithms, the "salt-and-pepper" noise in the RF classification results was generally reduced. This can be attributed to its robustness in handling data noise and outliers. Even in cases with incomplete or anomalous data, the RF algorithm can provide reliable classification results and effectively address the complex nonlinear relationships within the feature space. It demonstrates superior performance in crop classification tasks involving high-dimensional feature spaces and is well-suited for large-scale datasets.

4.3. Limitations of the Study and Future Directions

Although this study achieved high-precision extraction of maize in hilly areas with significant topographic variations and small-scale plots, there are still some limitations. First, due to the long revisit period of high-resolution satellite images and the variable climatic conditions of hilly areas, especially the frequent rainfall and humid weather during summer and autumn, cloud cover often obstructs satellite passes, which affects the continuity of the data. At the same time, the complex terrain of mountainous areas often interferes with the satellite's view, causing certain regions to be obscured by the terrain, which prevents effective observation of the target areas. Therefore, the remote sensing images obtained in this study do not represent complete data for all 12 months but rather a time series constructed from data of only 7 months. This limitation prevents a comprehensive coverage of the entire crop growth cycle and leads to some degree of "same spectrum, different objects " phenomena. Additionally, the sample data used in this study were acquired through manual visual interpretation based on prior experience, which may introduce some errors and pose challenges for extracting other crops planted in the study area. For crop classification in larger study areas, the strategy for selecting training samples should be adjusted. During the selection process, overly dense training samples should be avoided, ensuring the representativeness of the samples, reducing "same spectrum, different objects" phenomena, and improving classification accuracy.

Future research should consider integrating other types of image data, such as Sentinel-2 images. The revisit period of Sentinel-2 is 5 days (near the equator), allowing for rapid acquisition of new images even in the presence of cloud cover, providing finer temporal resolution. Combining it with existing high-resolution images can help fill the current data gaps, and in terms of both temporal and spatial resolution, it will contribute to constructing a more complete NDVI time series. With the continuous development of remote sensing technology and the successive launches of various ground monitoring satellites, China's remote sensing image data will become increasingly abundant, offering a more diversified and enriched data source for the remote sensing extraction of crop planting areas. Furthermore, while multi-source high-resolution image fusion provides rich spatial and spectral detail for precise crop classification, it presents challenges such as high data dimensionality, redundancy, inter-data correlation, and heavy processing demands.

Traditional machine learning methods often lack the feature-learning capabilities required to address these issues effectively. As deep learning continues to advance in land cover classification, future research should explore suitable deep learning models to streamline multi-source image fusion processing, reduce classification errors, and enhance accuracy. Such advancements are essential for detailed crop recognition in mountainous regions and hold significant promise for large-scale crop classification.

5. Conclusions

This study focuses on Shangzhou District and examines the capability to identify and extract maize planting areas using multi-source remote sensing images combined with NDVI time series and various machine learning methods. First, the acquired remote sensing images are preprocessed, and NDVI calculations are performed over seven months to construct a time series. Visual interpretation of the study area is conducted using sample data and high-resolution Google Earth images, resulting in the classification of seven land cover types. Four classification algorithms—Gaussian Naive Bayes, Artificial Neural Network, Support Vector Machine, and Random Forest—are applied to identify maize planting areas and assess classification accuracy. Comparative analysis highlights the strengths and potential improvements of machine learning methods for maize extraction. Key findings include the following:

(1) Utilizing multi-source high-resolution remote sensing images and NDVI time series effectively distinguishes various land features, accurately capturing the growth patterns of major crops in Shangzhou. Calculating average NDVI values for maize samples over the months enhances understanding of growth patterns, improving sample selection and reducing classification errors.

(2) All classification algorithms, except Gaussian Naive Bayes, achieved good accuracy, with overall accuracies exceeding 90% and Kappa coefficients above 0.85. Among them, the random forest algorithm performed best in identifying maize planting areas, with an overall accuracy of 94.88% and a Kappa coefficient of 0.94, indicating its suitability for classifying hilly and mountainous regions in southern Shaanxi.

(3) The combination of multi-source remote sensing images, NDVI time series, and machine learning methods show significant potential for crop identification in hilly and mountainous areas. This approach provides valuable insights for local farmers, helping them understand maize growth, optimize planting strategies, and support rational land planning and utilization.

Author Contributions: Conceptualization, F.Y. and L.L.; methodology, L.L.; software, J.Y.; validation, L.L.; formal analysis, L.H.; investigation, L.H.; resources, F.Y.; data curation, J.Y.; writing—original draft preparation, F.Y.; writing—review and editing, F.Y.; visualization, F.Y.; supervision, J.Y.; project administration, F.Y.; funding acquisition, F.Y. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The original data in this study cannot be publicly disclosed. For any other generated data, please contact the author for access.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Lv, X.; Zhang, X.; Yu, H.; Lu, X.; Zhou, J.; Feng, J.; Su, H. Extraction of Maize Distribution Information Based on Critical Fertility Periods and Active–Passive Remote Sensing. *Sustainability* **2024**, *16*, 8373. [CrossRef]

2. Liu, Z.; Xiao, J.; Liu, Z.; Nan, J.; Chang, J. Study on Water Requirement and Consumption Rules of Summer Maize with High-yield. *Water Sav. Irrig.* **2011**, *32*, 4–6.

3. Feng, L.; Sun, Z.; Xiao, J.; Liu, Y.; Hou, Z.; Tian, J.; Yin, X. Effects on Soil Water Consumption of Maize Field in Differente Soil Micro-catchment Pattern. *Res. Soil Water Conserv.* **2011**, *18*, 213–216.

4. Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. *Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements*; FAO Irrigation and Drainage Paper 56; Food and Agriculture Organization of the United Nations: Rome, Italy, 1998; pp. 109–111.

5. Liu, Z.; Li, L.; Li, H.; Liu, N.; Wang, H.; Shao, L. Changes and Influencing Factors of Crop Coefficient of Summer Maize During the Past 40 Years in the North China Plain. *Chin. J. Eco-Agric.* **2023**, *31*, 1355–1367.

6. Meng, H.; Li, C.; Liu, Y.; Gong, Y.; He, W.; Zou, M. Corn Land Extraction Based on Integrating Optical and SAR Remote Sensing Images. *Land* **2023**, *12*, 398. [CrossRef]

7. Chen, R.; Sun, L.; Chen, Z.; Wuyun, D.; Sun, Z. Early Identification of Corn and Soybean Using Crop Growth Curve Matching Method. *Agronomy* **2024**, *14*, 146. [CrossRef]

8. Wei, P.; Ye, H.; Qiao, S.; Liu, R.; Nie, C.; Zhang, B.; Song, L.; Huang, S. Early Crop Mapping Based on Sentinel-2 Time-Series Data and the Random Forest Algorithm. *Remote Sens.* **2023**, *15*, 3212. [CrossRef]

9. Meng, M.; Zhang, K.; Huang, Y.; Li, N.; Guo, Z.; Zhou, Z. Crop Classification Based on G-CNN Using Multi-Scale Remote Sensing Images. *Remote Sens. Lett.* **2024**, *15*, 941–950. [CrossRef]

10. Yang, B.; Zhu, W.; Rezaei, E.E.; Li, J.; Sun, Z.; Zhang, J. The Optimal Phenological Phase of Maize for Yield Prediction with High-Frequency UAV Remote Sensing. *Remote Sens.* **2022**, *14*, 1559. [CrossRef]

11. Tan, S.; Liu, J.; Lu, H.; Lan, M.; Yu, J.; Liao, G.; Wang, Y.; Li, Z.; Qi, L.; Ma, X. Machine Learning Approaches for Rice Seedling Growth Stages Detection. *Front. Plant Sci.* **2022**, *13*, 914771. [CrossRef] [PubMed]

12. Campi, P.; Modugno, A.F.; De Carolis, G.; Pedrero Salcedo, F.; Lorente, B.; Garofalo, S.P. A Machine Learning Approach to Monitor the Physiological and Water Status of an Irrigated Peach Orchard under Semi-Arid Conditions by Using Multispectral Satellite Data. *Water* **2024**, *16*, 2224. [CrossRef]

13. Chen, Y.; He, G.; Li, J.; Shi, L.; Fang, H.; Shi, Z. Tomato Recognition in Yuanmou Hot Area Based on Object-Oriented GF-2 Remote Sensing Data. *J. Henan Agric. Sci.* **2021**, *50*, 170–180.

14. Yang, Y.; Zhan, Y.; Tian, Q.; Gu, X.; Yu, T.; Wang, L. Crop Classification Based on GF-1/WFV NDVI Time Series. *Trans. Chin. Soc. Agric. Eng.* **2015**, *31*, 155–161.

15. Wei, P.; Xu, X.; Yang, G.; Li, Z.; Wang, J.; Chen, G. Remote Sensing Classification of Crops Based on the Change Characteristics of Multi-phase Vegetation Index. *J. Agric. Sci. Technol.* **2019**, *21*, 54–61.

16. Qiao, S.; Ye, H.; Liu, R.; Guo, A.; Zhang, B.; Qian, B.; Wei, P.; Huang, W. A Dataset of Remote Sensing Monitoring of Planting Distribution for Major Crops in Sanjiang Plain from 2020 to 2022. *China Sci. Data* **2023**, *8*, 336–346. [CrossRef]

17. Chen, W.; Ding, X.; Zhao, R.; Shi, S. Application of Frequency Ratio and Weights of Evidence Models in Landslide Susceptibility Mapping for the Shangzhou District of Shangluo City, China. *Environ. Earth Sci* **2016**, *75*, 64. [CrossRef]

18. Su, Q.; Lv, J.; Fan, J.; Zeng, W.; Pan, R.; Liao, Y.; Song, Y.; Zhao, C.; Qin, Z.; Defourny, P. Remote Sensing-Based Classification of Winter Irrigation Fields Using the Random Forest Algorithm and GF-1 Data: A Case Study of Jinzhong Basin, North China. *Remote Sens.* **2023**, *15*, 4599. [CrossRef]

19. Wang, Q.; Li, J.; Jin, T.; Chang, X.; Zhu, Y.; Li, Y.; Sun, J.; Li, D. Comparative Analysis of Landsat-8, Sentinel-2, and GF-1 Data for Retrieving Soil Moisture over Wheat Farmlands. *Remote Sens.* **2020**, *12*, 2708. [CrossRef]

20. Wei, D.; Liu, K.; Xiao, C.; Sun, W.; Liu, W.; Liu, L.; Huang, X.; Feng, C. A Systematic Classification Method for Grassland Community Division Using China's ZY1-02D Hyperspectral Observations. *Remote Sens.* **2022**, *14*, 3751. [CrossRef]

21. Luo, H.; Li, L.; Zhu, H.; Kuai, X.; Zhang, Z.; Liu, Y. Land Cover Extraction from High Resolution ZY-3 Satellite Imagery Using Ontology-Based Method. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 31. [CrossRef]

22. Kuang, X.; Guo, J.; Bai, J.; Geng, H.; Wang, H. Crop-Planting Area Prediction from Multi-Source Gaofen Satellite Images Using a Novel Deep Learning Model: A Case Study of Yangling District. *Remote Sens.* **2023**, *15*, 3792. [CrossRef]

23. Zha, D.; Cai, H.; Zhang, X.; He, Q.; Chen, L.; Qiu, C.; Xia, S. Extracting Lotus Fields Using the Spectral Characteristics of GF-1 Satellite Data. *Phyton* **2022**, *91*, 2297–2311. [CrossRef]

24. Shao, C.; Yang, G.; Sun, W.; Zuo, Y.; Ge, W.; Yang, S. Construction Method of a Spartina Alterniflora Index Based on Hyperspectral Satellite Images. *Natl. Remote Sens. Bull.* **2024**, *28*, 635–648. [CrossRef]

25. Li, F.; Chen, K.; Yi, D.; Yang, Y. Application Research of 3D Laser Scanning Technology in 1:500 Topographic Map Fine Mapping. *J. Geol. Hazards Environ. Preserv.* **2022**, *33*, 92–97.

26. Ma, F.; Bao, J.; Gao, Q.; Zhang, F.; Sun, X. Digital Twin Modeling and Processing for On-Orbit Interpretation of Satellite-Based SAR. *J. Beijing Univ. Chem. Technol. (Nat. Sci. Ed.)* **2024**, *51*, 77–88.

27. Zhao, J.; Wang, L.; Yang, H.; Wu, P.; Wang, B.; Pan, C.; Wu, Y. A Land Cover Classification Method for High-Resolution Remote Sensing Images Based on NDVI Deep Learning Fusion Network. *Remote Sens.* **2022**, *14*, 5455. [CrossRef]

28. Yao, J.; Wu, J.; Xiao, C.; Zhang, Z.; Li, J. The Classification Method Study of Crops Remote Sensing with Deep Learning, Machine Learning, and Google Earth Engine. *Remote Sens.* **2022**, *14*, 2758. [CrossRef]

29. Kong, F.; Li, X.; Wang, H.; Xie, D.; Li, X.; Bai, Y. Land Cover Classification Based on Fused Data from GF-1 and MODIS NDVI Time Series. *Remote Sens.* **2016**, *8*, 741. [CrossRef]

30. Wu, B.; Zhang, F.; Liu, C.; Zhang, L.; Luo, Z. An Integrated Method for Crop Condition Monitoring. *Natl. Remote Sens. Bull.* **2004**, *10*, 498–514.

31. Shi, J.; Wu, T.; Huang, Q.; Luo, J.; Ren, Y.; Xu, X. Geo-Parcel Crop Remote Sensing Classification via Coupling with Time Series Features of NDVI and Texture. *J. South. Agric.* **2024**, 1–15.

32. Ding, Y.; Zhao, K.; Zheng, X.; Jiang, T. Temporal Dynamics of Spatial Heterogeneity over Cropland Quantified by Time-Series NDVI, near Infrared and Red Reflectance of Landsat 8 OLI Imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *30*, 139–145. [CrossRef]

33. Ontivero-Ortega, M.; Lage-Castellanos, A.; Valente, G.; Goebel, R.; Valdes-Sosa, M. Fast Gaussian Naive Bayes for Searchlight Classification Analysis. *Neuroimage* **2017**, *163*, 471–479. [CrossRef] [PubMed]

34. Jiang, W.; Zhang, M.; Long, J.; Pan, Y.; Ma, Y.; Lin, H. HLEL: A Wetland Classification Algorithm with Self-Learning Capability, Taking the Sanjiang Nature Reserve I as an example. *J. Hydrol.* **2023**, *627*, 130446. [CrossRef]

35. Mashaba-Munghemezulu, Z.; Chirima, G.J.; Munghemezulu, C. Delineating Smallholder Maize Farms from Sentinel-1 Coupled with Sentinel-2 Data Using Machine Learning. *Sustainability* **2021**, *13*, 4728. [CrossRef]

36. Kumar, P.; Gupta, D.K.; Mishra, V.N.; Prasad, R. Comparison of Support Vector Machine, Artificial Neural Network, and Spectral Angle Mapper Algorithms for Crop Classification Using LISS IV Data. *Int. J. Remote Sens.* **2015**, *36*, 1604–1617. [CrossRef]

37. Liu, X.; Wang, L.; Liu, X.; Li, L.; Zhu, X.; Chang, C.; Lan, H. Multispectral versus Texture Features from ZiYuan-3 for Recognizing on Deciduous Tree Species with Cloud and SVM Models. *Sci. Rep.* **2023**, *13*, 7369. [CrossRef]

38. Immitzer, M.; Vuolo, F.; Atzberger, C. First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. *Remote Sens.* **2016**, *8*, 166. [CrossRef]

39. Belgiu, M.; Drăguţ, L. Random Forest in Remote Sensing: A Review of Applications and Future Directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]

40. Wei, P.; Zhu, W.; Zhao, Y.; Fang, P.; Zhang, X.; Yan, N.; Zhao, H. Extraction of Kenyan Grassland Information Using PROBA-V Based on RFE-RF Algorithm. *Remote Sens.* **2021**, *13*, 4762. [CrossRef]

41. Zhang, L.; Liu, Z.; Liu, D.; Xiong, Q.; Yang, N.; Ren, T.; Zhang, C.; Zhang, X.; Li, S. Crop Mapping Based on Historical Samples and New Training Samples Generation in Heilongjiang Province, China. *Sustainability* **2019**, *11*, 5052. [CrossRef]

42. Chowdhury, M.S. Comparison of Accuracy and Reliability of Random Forest, Support Vector Machine, Artificial Neural Network and Maximum Likelihood Method in Land Use/Cover Classification of Urban Setting. *Environ. Chall.* **2024**, *14*, 100800. [CrossRef]

43. Khalifani, S.; Darvishzadeh, R.; Azad, N.; Seyed Rahmani, R. Prediction of Sunflower Grain Yield under Normal and Salinity Stress by RBF, MLP and, CNN Models. *Ind. Crops Prod.* **2022**, *189*, 115762. [CrossRef]

44. Maier, H.R.; Galelli, S.; Razavi, S.; Castelletti, A.; Rizzoli, A.; Athanasiadis, I.N.; Sànchez-Marrè, M.; Acutis, M.; Wu, W.; Humphrey, G.B. Exploding the Myths: An Introduction to Artificial Neural Networks for Prediction and Forecasting. *Environ. Model. Softw.* **2023**, *167*, 105776. [CrossRef]

45. Fan, J.; Zhang, X.; Zhao, C.; Qin, Z.; De Vroey, M.; Defourny, P. Evaluation of Crop Type Classification with Different High Resolution Satellite Data Sources. *Remote Sens.* **2021**, *13*, 911. [CrossRef]

46. Fan, J.; Defourny, P.; Zhang, X.; Dong, Q.; Wang, L.; Qin, Z.; De Vroey, M.; Zhao, C. Crop Mapping with Combined Use of European and Chinese Satellite Data. *Remote Sens.* **2021**, *13*, 4641. [CrossRef]

47. Zhang, P.; Hu, S. Fine Crop Classification by Remote Sensing in Complex Planting Areas Based on Field Parcel. *Trans. Chin. Soc. Agric. Eng.* **2019**, *35*, 125–134.

48. Zheng, Z.; Yuan, J.; Yao, W.; Kwan, P.; Yao, H.; Liu, Q.; Guo, L. Fusion of UAV-Acquired Visible Images and Multispectral Data by Applying Machine-Learning Methods in Crop Classification. *Agronomy* **2024**, *14*, 2670. [CrossRef]