

Article

A Study of Apple Orchards Extraction in the Zhaotong Region Based on Sentinel Images and Improved Spectral Angle Features

Jingming Lu , Weiwei Song *, Xiaoqing Zuo, Daming Zhu and Qunlan Wei

Faculty and Land Engineering, Kunming University of Technology, Kunming 650093, China; 20212101049@stu.kust.edu.cn (J.L.); zxq@kust.edu.cn (X.Z.); dmzhu@kust.edu.cn (D.Z.); weiqunlan@stu.kust.edu.cn (Q.W.)

* Correspondence: weiweisong@kust.edu.cn

Abstract: Zhaotong City in Yunnan Province is one of the largest apple growing bases in China. However, the terrain of Zhaotong City is complicated, and the rainy weather is more frequent, which brings difficulties to the identification of apple orchards by remote sensing. In this paper, an improved spectral angle feature is proposed by combining the Spectral Angle Mapper and Sentinel-1 data. Based on the Google Earth Engine and Sentinel image, a random forest classifier was used to extract apple orchards in the Ganhe Reservoir area, Zhaoyang District, Zhaotong City, which provides a theoretical basis for extracting the spatial distribution and sustainable development of the local apple industry. The classification results show that the improved spectral angle characteristics can improve the overall accuracy and F1 score of apple orchards. The RGB band combined with NDVI, GLCM, and improved spectral angle features obtained the most favorable results, and the F1 score and overall accuracy were 88.89% and 84.44%, respectively, which proved the reliability of the method in identifying apple orchards in Zhaotong City.

Keywords: crop classification; Google Earth Engine; random forests; spectral angle mapper; time-series analysis



Citation: Lu, J.; Song, W.; Zuo, X.; Zhu, D.; Wei, Q. A Study of Apple Orchards Extraction in the Zhaotong Region Based on Sentinel Images and Improved Spectral Angle Features. *Appl. Sci.* **2023**, *13*, 11194. <https://doi.org/10.3390/app132011194>

Academic Editor: Jesús Montero Martínez

Received: 7 September 2023

Revised: 3 October 2023

Accepted: 8 October 2023

Published: 11 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Remote sensing technology can provide important data for crop monitoring. This can help guide agricultural production and management [1]. Zhaotong City, located in southwestern China, boasts the largest apple production area. By 2022, its apple cultivation area had exceeded 800,000 acres, establishing it as a vital component of the local apple industry [2]. However, because of the complex topography of the Zhaotong City area, apple orchards are scattered, and field surveys are difficult to conduct. Simultaneously, with the implementation of the western rural revitalization strategy, sustainable development of the apple industry in the Zhaotong City area needs to be implemented. Therefore, it is very necessary to study the remote sensing identification method of apple orchards in the Zhaotong area. This can provide a theoretical basis for the sustainable development of the Zhaotong apple industry, which is of great significance to local and national economic development.

Scholars worldwide have mostly researched crops, such as wheat, corn, and rice, all of which have achieved good results [3–6]. However, little research and corresponding progress have been made in apple orchard extraction [7–9]. While monitoring crops via remote sensing in southwestern China, it was observed that the critical growth period is frequently obstructed by clouds and rain, significantly impacting the accuracy and timeliness of monitoring and hindering the acquisition of optimal temporal optical crop images [10]. Furthermore, the complex and fragmented terrain in the southwest of China is often affected by information, such as brightness and shadows, which exacerbates the

“pepper and salt” phenomenon and leads to misclassification. Therefore, it is necessary to build a remote sensing method to identify apple orchards to solve these problems.

During the classification of apple orchards, the precision of the classification is greatly influenced by the samples selected and the input features. The spatial homogeneity of the sample selection allows it to be representative. The selection of input features that enable the classification algorithm to learn characteristics specific to apple orchards is the key to improving the accuracy of apple orchard recognition.

Numerous studies have shown that time-series data, as an input feature, can effectively deal with the “same spectrum and different objects” problems and positively impact the final classification results. For example, based on the GaoFen-1 image data of time series, Li et al. adopted the time-series remote sensing image classification method of DTW-K Means to classify tree species in the Harbin experimental forest farm, and achieved good results [11]. Belgiu et al. evaluates how a time-weighted dynamic time warping (TWDTW) method that uses Sentinel-2 time series performs when applied to pixel-based and object-based classifications of various crop types in three different study areas [12].

With the continuous development of radar technology, Synthetic Aperture Radar (SAR) technology has been developed and has wide applications in crop monitoring due to its all-weather, all-day observation and independence from clouds and rain. Erasmi et al. demonstrate the synergy of optical and radar satellite data for land cover mapping in tropical regions [13]. Gebhardt et al. analyzed the usefulness and potential of SAR data for vegetation characterization and plant physiological parameter estimation [14].

Meanwhile, the spectral angle mapper classifies features based on their spectral properties [15], focusing on the differences among bands in the image, and does not rely on prior knowledge. To a certain extent, this method reduces the influence of factors, such as the “same spectrum and different objects.” Grzegorzewski et al. used MODIS-EVI data to identify soybean and corn using the Spectral Angle Mapper [16]. Li et al. used multi-temporal Landsat 8 OLI images, combined with spectral Angle mapping and decision tree classification, to extract the distribution of main crops in the study area [17]. These studies show that the spectral angle Mapper is less affected by terrain in the classification process, which reduces the influence of factors, such as “foreign objects in the same spectrum”, to a certain extent. However, image classification tasks with numerous samples and input features incur significant labor and time costs. Google Earth Engine (GEE), with its powerful data processing and analysis capabilities, can directly invoke preprocessed mass images and multiple algorithms and effectively address these problems [18,19]. Several studies have recently been conducted on image classification using GEE [20–23]. These studies show that GEE can provide massive data and cloud-computing support for remote sensing research.

To effectively identify apple orchards in Zhaotong City, an improved spectral Angle mapper method is proposed in this paper. Based on the GEE platform and Sentinel-1 and Sentinel-2 images, apple orchards, cultivated land, artificial surface, water body, and other land types were selected, and the improved spectral angle features and other features were integrated to achieve accurate identification of apple orchards in the Ganhe Reservoir area of Zhaoyang District, Zhaotong City, by the random forest (RF) classification method.

The structure of this article is as follows: Section 2 details the study area and data sources, including remote sensing and sample data. Section 3 details the research methods for apple orchard identification, including the sample selection methods, principles of the spectral angle mapper, improved spectral angle features, input feature selection, classification experimental design, RF classification methods, and accuracy evaluation. Section 4 describes the apple orchard extraction results and accuracy evaluation. Section 5 summarizes the discussion, and Section 6 summarizes the conclusions.

2. Study Area and Data Sources

2.1. Study Area

Zhaotong City is located in the hinterland of the Wumeng Mountains in the north-eastern Yunnan province, at the junction of Yunnan, Sichuan, and Guizhou provinces, and has a complex topography with undulating terrain [24], and it has many local climate types and significant vertical variation, with an average annual temperature of 11.6 °C, an average annual minimum temperature of 7.3 °C, an average annual maximum temperature of 18.3 °C, and an annual frost-free period of about 220 days. Additionally, the sum of daily temperatures greater than or equal to 10 °C during the year is 3,217 °C, the average annual sunshine hours is 1902 h, and the average annual rainfall is 738.2 mm. It is a great national base for apple cultivation in the south.

The study area (Figure 1) is located next to Ganhe Reservoir, southeastern Zhaoyang District, Zhaotong City (103°43'55" E–103°45'17" E, 27°16'50" N–27°17'41" N), which is one of the best apple growing areas in Zhaoyang District. The study area had numerous crop species and was large and highly representative; therefore, it was chosen as the study area. The study area was divided into five categories, according to land use characteristics: apple orchards, croplands, artificial surfaces, water bodies, and other land. Most of the apple orchards in the study area were in full bloom, with apple trees starting to bud from March to April and the stable growth period from June to August, with high vegetation cover in the orchards and fruit ripening in September to October.

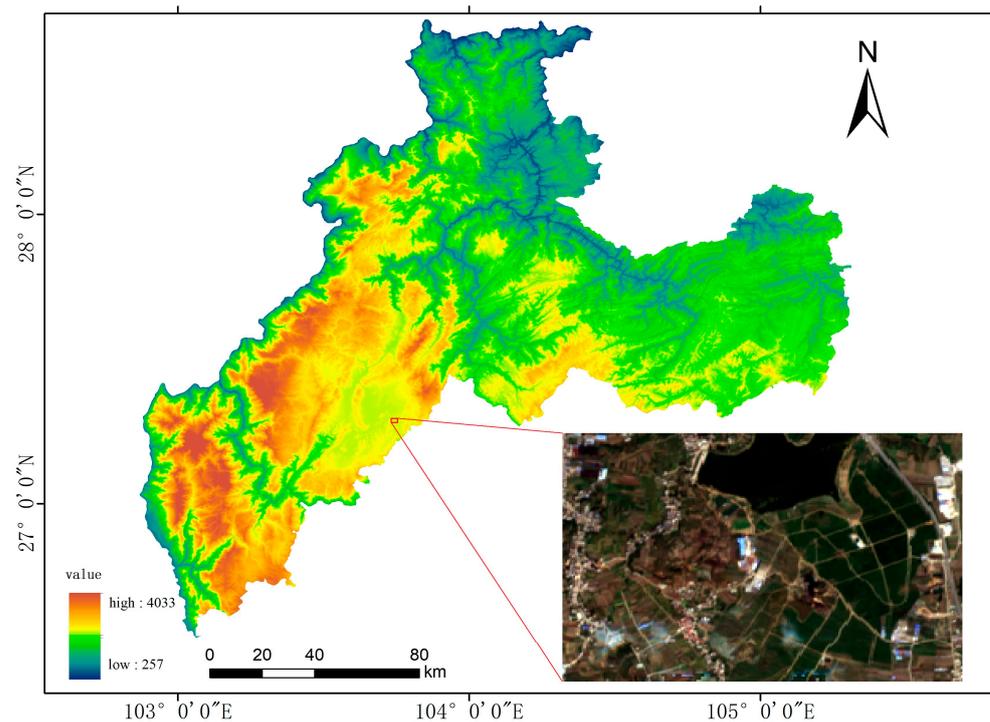


Figure 1. Sentinel-2 images of the study area.

The image size of the study area was 251 rows and 157 columns, with a total area of 3.94 km². According to the visual interpretation, the proportions of apple orchards, croplands, artificial surfaces, water bodies, and other lands were 9.77%, 13.56%, 8.46%, 62.11%, and 6.1%, respectively. The presence of thin clouds in the October Sentinel-2 images and the complex topography of the study area made it challenging to distinguish apple orchards using RGB Bands alone.

2.2. Data Sources

2.2.1. Remote Sensing Data

The Sentinel-1 (5–20 m) and Sentinel-2A (10~60 m) data used in this study were obtained from GEE.

For the GRD data of Sentinel-1, backscattering coefficients of vertical–horizontal (VH) polarization and vertical–vertical (VV) polarization are used. Average images for each month in 2021 were collected, totaling 12 images. Average processing can reduce the pollution caused by noise.

For the surface reflectance data of Sentinel-2, there are 13 spectral bands. And only B2, B3, B4, B7, and B8 are selected in this paper. September to October is the fruiting period of the apple orchard, and the characteristics of the apple orchard are more prominent [25]. Therefore, the monthly average Sentinel-2 image for October 2022 was selected. Although it is affected by the number of images and cloud cover, it still covered most of the study area.

The geographic coordinate system of all the data used in this study was GCS_WGS_1984. The band information of the remote sensing image data is described in Table 1.

Table 1. Remote sensing image waveform information.

Sensor Used	Bands	Descriptions	Resolution
Sentinel-1	VV	5.405 GHz	10 m
	VH	5.405 GHz	10 m
Sentinel-2	Blue	496.6 nm (S2A)/492.1 nm (S2B)	10 m
	Green	560 nm (S2A)/559 nm (S2B)	10 m
	Red	664.5 nm (S2A)/665 nm (S2B)	10 m
	NIR	835.1 nm (S2A)/833 nm (S2B)	10 m
	Red Edge 3	782.5 nm (S2A)/779.7 nm (S2B)	20 m

2.2.2. Samples

Since thin clouds partially obscured the October 2022 Sentinel-2 image, this study used GEE as the basis for visual interpretation based on the October 2022 Sentinel-2 image, with the March 2022 Sentinel-2 and Google Earth images as secondary references, to obtain five types of sample areas: apple orchards, croplands, artificial surfaces, water bodies, and other lands. An unbalanced number of samples will affect the accuracy of the classifier. Therefore, this paper uses the random point tool in ArcGIS 10.6 software; random samples were generated in the sample area, according to the area share of each site obtained from visual interpretation. A total of 1098 samples were generated, of which 770 were training samples and 328 were validation samples, as listed in Table 2.

Table 2. Training and validation samples.

No.	Land Type	Training Samples/Pixels	Validation Samples/Pixels
1	Apple Orchards	137	58
2	Croplands	435	186
3	Artificial Surfaces	95	41
4	Water Bodies	60	25
5	Other Lands	43	18
Total		770	328

3. Methods

The overall workflow of the apple orchards in the study area is shown in Figure 2. The specific steps include (1) sample selection; (2) improved spectral angle features; (3) input features; (4) experimental design; and (5) classification and accuracy evaluation.

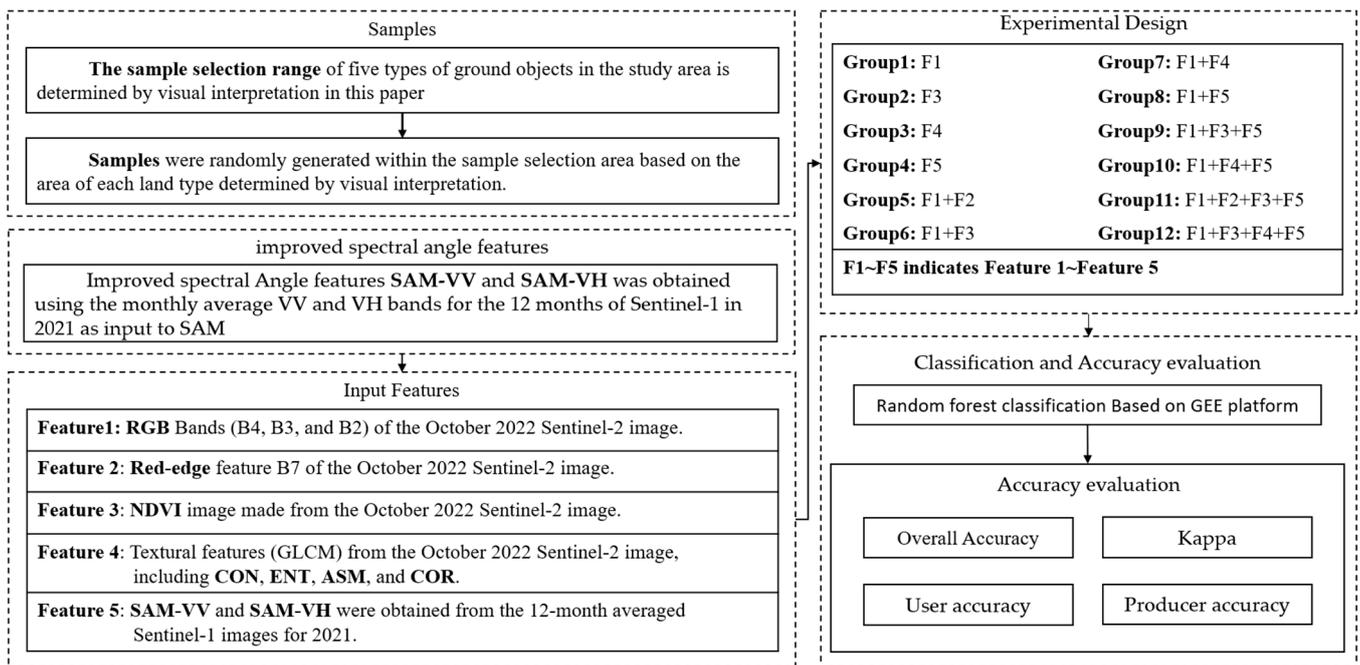


Figure 2. Flow chart of remote sensing identification of apple orchards.

Based on GEE, this study first conducted a visual interpretation of high-resolution images; determined five types of sample selection ranges, including apple orchards; and randomly generated 1098 sample points within the sample selection range, according to the area share of each location obtained from visual interpretation. The VV and VH bands in the monthly average Sentinel-1 images over 12 months were then calculated using a spectral angle mapper to generate the corresponding spectral angle features, SAM_VV and SAM_VH.

To analyze the applicability of different features for apple orchard extraction in the Zhaotong City area, we designed 12 feature combinations based on the RGB band, red-edge feature, NDVI, texture feature, and improved spectral Angle feature. The reasons for selecting these features are that the RGB band can provide basic spectral information for classification, NDVI and red-edge features can better reflect the characteristics of vegetation, and texture features [26] can provide spatial information.

Among them, NDVI is calculated directly from the Sentinel-2 image of October 2022, while texture feature needs to calculate the gray image of the Sentinel-2 image of October 2022. Then, the gray-level co-occurrence matrix (GLCM, window size: 3 × 3 pixels) was used to obtain four texture features: contrast (CON), entropy (ENT), angular second moment (ASM), and correlation (COR). These features are computed on the GEE platform.

Finally, an RF classifier in GEE was used to extract the apple orchard distribution information, and the apple orchard extraction and overall accuracy of the different methods were evaluated.

3.1. Sample Point Selection

The following samples selection process was used to identify apple orchards in the study area:

3.1.1. Visual Interpretation to Obtain the Samples Selection Range

The study area mainly includes five land types: apple orchard, cultivated land, artificial surface, water body, and other lands. Due to the difficulty of field sampling, visual interpretation was used to identify features in the study area. As the apple orchards from the October 2022 Sentinel-2 imagery were difficult to distinguish from croplands, and there was a thin cloud contamination, we used the March 2022 Sentinel-2 imagery and Google

Earth imagery as secondary references for apple orchards, cultivated land, and water bodies. Based on the October 2022 Sentinel-2 imagery of apple orchards, vectorization operations were carried out on apple orchards, croplands, artificial surfaces, and water bodies, and the land types that were still difficult to distinguish were classified as other land types. The final vectorized images of apple orchards, croplands, artificial surfaces, water bodies, and other lands were identified as sample candidates areas, as shown in Figure 3.

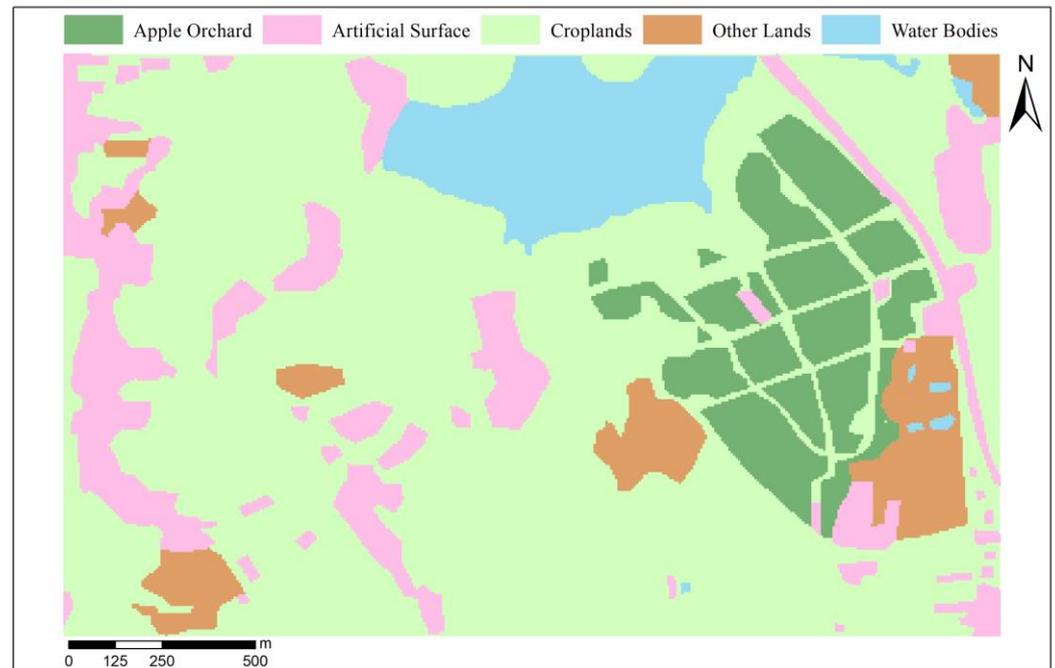


Figure 3. Candidate areas for sample points based on visual interpretation.

3.1.2. Random Generation of Samples

To make the samples more representative, reasons for the differences should be considered when selecting the samples to obtain sample diversity and comprehensiveness and improve the quality of the training dataset, which can effectively ensure the accuracy of algorithm learning and prediction [27]. The sampling points were selected randomly. This study used the random point creation tool in ArcGIS 10.6 to generate random points based on the sample candidate areas obtained from visual interpretation to obtain a high-quality sample.

To randomly place a specified number of points in each polygon, the Create Random Points tool uses a standard polygon-partitioning algorithm to partition the polygon into multiple triangles of varying sizes. A triangle is randomly selected for this polygon. The two legs of the triangle become the two axes on which a random point is placed. A value is selected randomly along the axis. The same procedure is then performed for the other axis or leg of the triangle. These two random values are used to place the points. This point lies within the parallelogram created by the two axes of the triangle. If a point fell outside the triangle, it was mirrored in the defined triangle. This process was repeated until the specified number of points was placed in the polygon and was repeated for each polygon [28].

The distribution of the random sample points generated based on the sample point candidate area is shown in Figure 4.

Finally, each category of randomly generated samples is assigned the attribute “label” and the corresponding label value (0 Apple Orchards; 1 Croplands; 2 Artificial Surfaces; 3 Water Bodies; and 4 Other Lands).

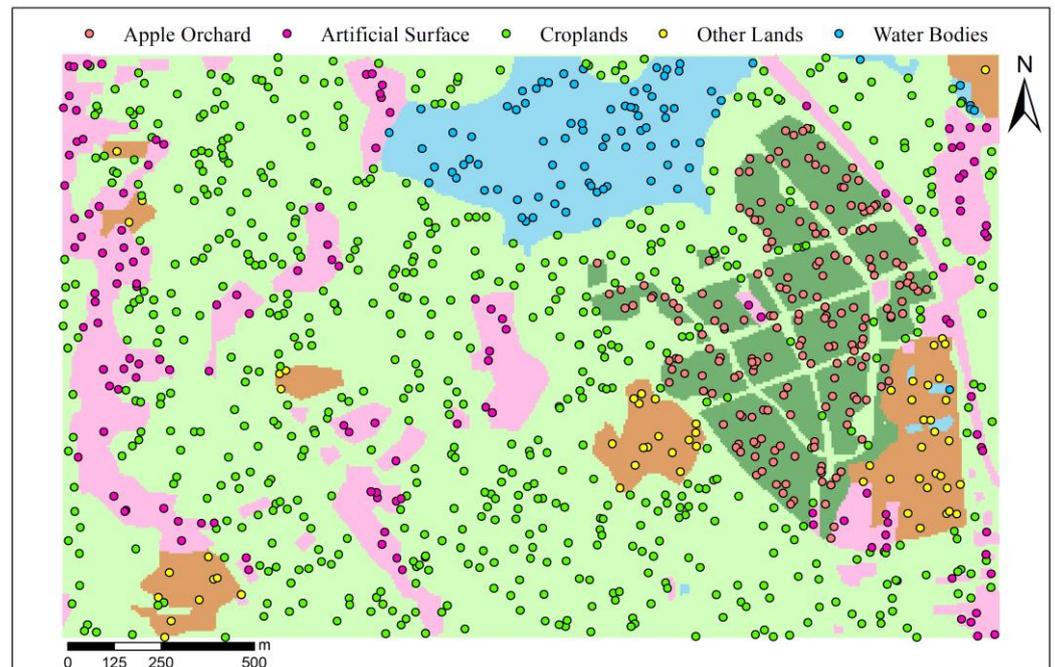


Figure 4. Sample selection results.

3.2. Spectral Angle Mapper

3.2.1. Principle of the Spectral Angle Mapper

The spectral angle mapper (SAM) treats the n bands in an image as an n -dimensional space, where each band is treated as a vector, and the value of each band is simply the length of the vector corresponding to that band. The SAM first generates a standard reference spectrum using the mean value of k known feature samples as a standard, which is then matched to the original spectrum of the image element. The image is classified by calculating the generalized angle θ between the standard reference spectrum and the original spectrum of the image element (Figure 5). The smaller the angle θ , the more likely the image element belongs to a known feature class. Since the angle between two vectors is independent of the vector length and is unaffected by factors such as solar illumination [29], SAM can suppress shadowing effects and highlight the reflective properties of the target; it is a simple, fast, and effective classification method.

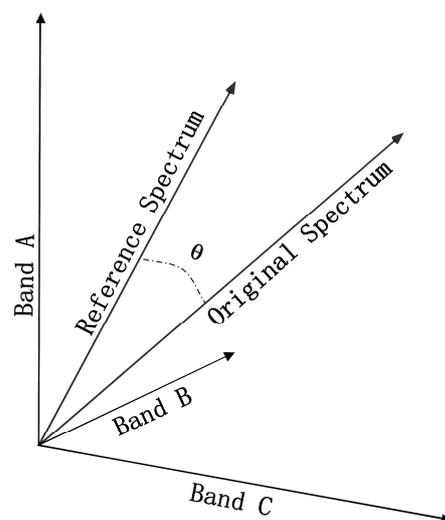


Figure 5. Principle of the spectral angle mapper.

3.2.2. Improved Spectral Angle Features

The traditional SAM uses different bands in a single time-phase image as inputs to construct a spectral angle feature. Although this type of spectral angle feature can suppress the shadow effect to a certain extent and highlight the target reflectance, it has limitations in the Zhaotong City area, where the terrain is complex and cloudy. Therefore, we propose an improved spectral angle feature based on the SAM. Compared to the traditional SAM, this method uses the VV and VH bands of Sentinel-1 images of multiple time phases as inputs to obtain two spectral angle features with time-series characteristics that are unaffected by clouds and rain.

As the growth cycle of apple trees is different from that of other crops in a year, the VV and VH bands of the monthly average Sentinel-1 images from January to December 2021 were used as the basis for this study to obtain two characteristic images, SAM-VV and SAM-VH, through the SAM.

Using SAM-VV as an example, the steps are as follows:

- (1) The VV bands of monthly averaged Sentinel-1 images from January to December 2021 were obtained.
- (2) Ten randomly selected image elements were used as standard image elements.
- (3) In the same month, the VV values of the ten standard image elements were averaged to obtain the reference mean values of the VV bands from January to December 2021.
- (4) Using the SAM, the VV band from January to December 2021 was used as the input to calculate the original spectrum $\vec{\alpha}$ of the individual image elements.
- (5) Using the SAM, we constructed a reference spectrum $\vec{\beta}$ of the VV using the VV averages from January to December 2021 as input.
- (6) The angle between the original spectrum $\vec{\alpha}$ and reference spectrum $\vec{\beta}$ of a single image element to obtain the spectral angle coefficient θ of a single image element was calculated using the following formula:

$$\theta = \cos^{-1}\left(\frac{\vec{\alpha} \cdot \vec{\beta}}{\|\vec{\alpha}\| \cdot \|\vec{\beta}\|}\right), \theta \in \left(0, \frac{\pi}{2}\right) \quad (1)$$

Finally, the operation in Step (6) was performed for each image element, and the calculated spectral angle coefficients constituted the SAM-VV features.

The improved spectral angle features of SAM-VV and SAM-VH are shown in Figure 6.

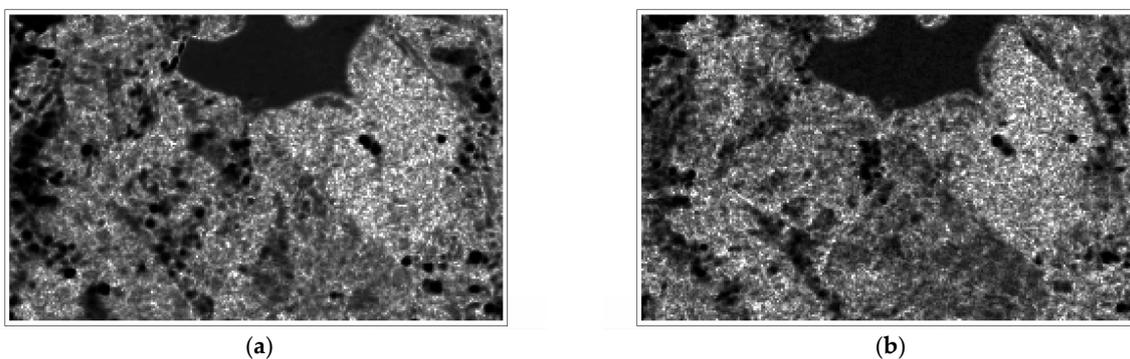


Figure 6. Improved spectral angle features: (a) SAM-VV and (b) SAM-VH.

3.3. Selection of Input Features

Five types of features were selected for this study to compare the impacts of different features on the classification results: RGB Bands, Red-edge, NDVI, texture, and improved spectral angle features.

- (1) Feature 1: RGB Bands (B4, B3, and B2) of Sentinel-2 images from October 2022.
- (2) Feature 2: Red-edge feature B7 of the October 2022 Sentinel-2 image.

- (3) Feature 3: NDVI image made from the October 2022 Sentinel-2 image.
- (4) Feature 4: Textural features (GLCM) from the October 2022 Sentinel-2 image, including CON, ENT, ASM, and COR.
- (5) Feature 5: SAM-VV and SAM-VH were obtained from the 12-month averaged Sentinel-1 images for 2021.

The red-edge feature B7 of the October 2022 Sentinel-2 image, the NDVI image obtained from the October 2022 Sentinel-2 image, and the GLCM textural features are shown in Figure 7.

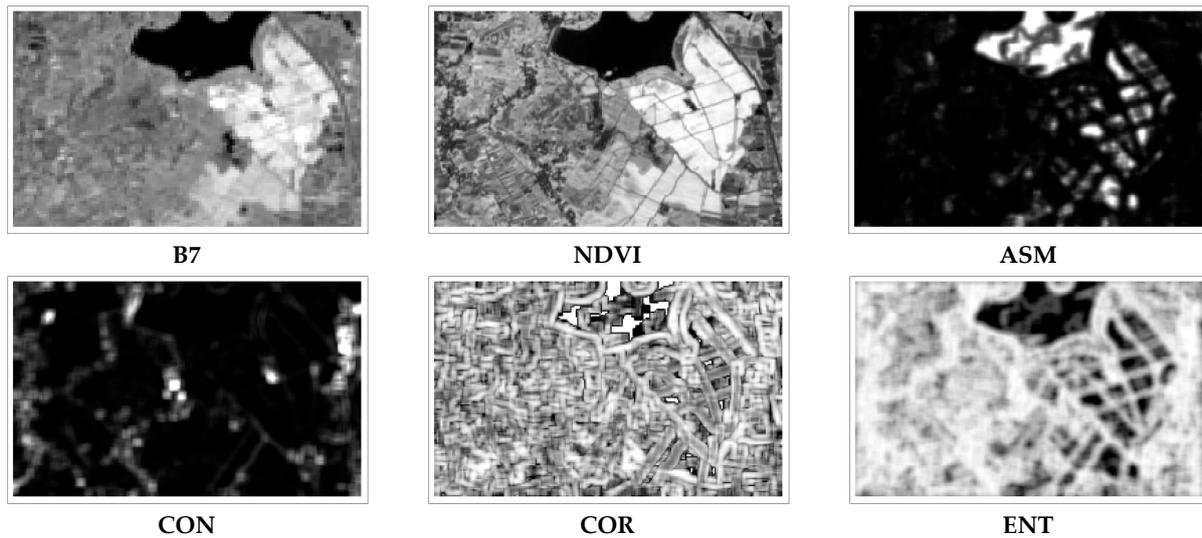


Figure 7. Red-edge features (B7), NDVI, and textural features (ASM, CON, COR, ENT).

3.4. Experiment Design

To better compare the effects of different input feature combinations on the classification effect, 12 feature combinations were designed to participate in the classification experiments for features Feature 1–Feature 5, as described in Section 3.3. The 12 feature combinations are listed in Table 3. (1) G1–G4 were used to explore the recognition ability to compare RGB Bands, NDVI, textural features, and improved spectral angle features of apple orchards during the hanging stage. (2) G5–G8 were designed to compare the improved spectral angle features with the red-edge band B7, NDVI, and textural features to recognize apple orchards at the hanging stage. (3) G9–G12 aims to explore the ability of different feature combinations to improve the distinguishability of apple orchards from other land types.

Table 3. Input feature combination design table.

	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5
G1	✓				
G2			✓		
G3				✓	
G4					✓
G5	✓	✓			
G6	✓		✓		
G7	✓			✓	
G8	✓				
G9	✓		✓		✓
G10	✓			✓	✓
G11	✓	✓	✓		✓
G12	✓		✓	✓	✓

3.5. Random Forest Classification

3.5.1. Principle of Random Forest Classifier

Random Forests (RF) [30] is a machine learning method for classification and regression. It is an integrated model consisting of multiple independent decision trees. Each decision tree predicts the classification result, and the classification result of the random forest is determined by these decision trees together.

The RF algorithm has superior accuracy, can process large batches of data quickly and efficiently, and has strong noise immunity to noisy datasets, and has the ability to process high-dimensional features while ranking variables by importance [31–33], which is a significant advantage over other algorithms. Therefore, this study used an RF model to classify the input features.

3.5.2. Random Forest Classifier Based on GEE

Due to the large number of input features and samples, this study classifies the input features by calling the random forest classifier *ee.Classifier.smileRandomForest* on a high-performance GEE, where the *ee.Classifier.smileRandomForest* classifier contains the following parameters:

- (1) **NumberOfTrees**: The number of decision trees to create.
- (2) **VariablesPerSplit**: The number of variables per split. Unless otherwise specified, the square root of the number of variables was used.
- (3) **MinLeafPopulation**: Only creates nodes whose training set contains at least this many points; default:1.
- (4) **BagFraction**: The fraction of input to the bag per tree.
- (5) **MaxNodes**: The maximum number of leaf nodes in each tree. If unspecified, it defaults to no limit.
- (6) **Seed**: The randomization seed.

Algorithm 1 describes a specific algorithm for RF classification based on GEE.

Algorithm 1 Random forest classification based on GEE

Input: The randomly generated **samples** in Section 3.1; The combination of features **InputBands**;

The number of decision trees **Number of Trees**. Original **image** composed of InputBands

Output: Classification results; four types of accuracy evaluation indicators for OA, kappa, UA and PA

```

1  Train_points = 70% samples;           //  Division of the training set
2  Test_points = 30% samples;           //  Division of the test set
3  Train_label = [Train_points.label]; //  Get the vector of label values for the training set
4  Test_label = [Test_points.label];     //  Get the vector of label values for the test set
5  for each InputBands do:
6    Train RFclassifier = RandomForest(Number of Tree, Train_label, InputBands); //  Training RFclassifier
7    Result = RFclassifier(image, InputBands); //  Classify with classifier and get results
8    RFclassifier (Test_label).errorMatrix; //  Get Confusion Matrix
9    Calculation Accuracy with errorMatrix; //  Calculate OA, kappa, UA and PA
10 end for

```

To consider both the accuracy and computational efficiency of the random forest algorithm and to prevent the algorithm from overfitting, this study conducted a control experiment based on RGB Bands, and the number of decision trees with high overall accuracy was determined as the number of optimal decision trees between 10 and 200. The experimental results showed that when the number of decision trees was 50 or 70, the overall accuracy (OA) obtained by RF classification was higher. The experimental results are described in Figure 8. The number of decision trees in the classifier was set to 70, and the other parameters were set to their default values.

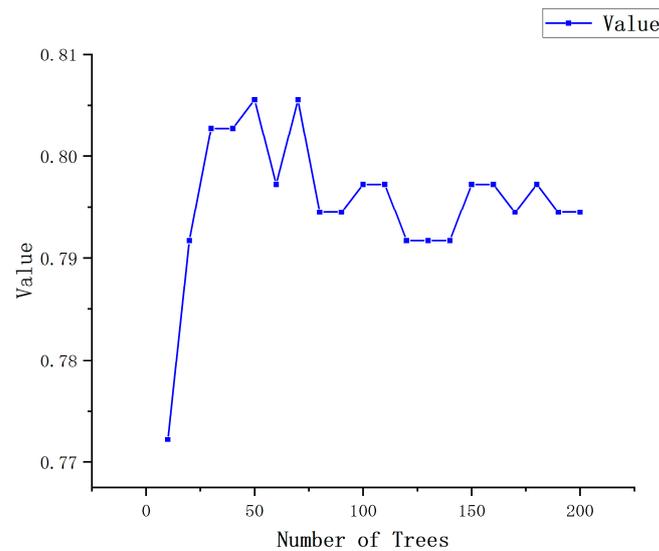


Figure 8. Optimal number selection for the number of trees based on RGB bands.

3.6. Accuracy Evaluation

To evaluate the accuracy of remote sensing in identifying apple orchard distribution, this study calculated the *OA*, kappa coefficient, user accuracy (*UA*), and producer accuracy (*PA*) of the classification results based on the confusion matrix and conducted an accuracy evaluation. The overall accuracy and kappa coefficient were used to compare the overall accuracy, and the *UA* and *PA* were used to evaluate the classification accuracy of a particular target within a multi-classification target, as shown in the following equations:

$$\text{Overall Accuracy} = \frac{TP + TN}{FN + TP + FP + TN} \quad (2)$$

$$\text{kappa} = \frac{(TP + FP) \times (TP + FN) \times (FP + TN) \times (FN + TN)}{(FN + TP + FP + TN)^2} \quad (3)$$

$$UA = \frac{TP}{TP + FP} \quad (4)$$

$$PA = \frac{TP}{TP + FN} \quad (5)$$

where *TP* is the number of pixels for which positive categories were judged as positive, *FP* is the number of pixels for which negative categories were judged as positive, *FN* is the number of pixels for which positive categories were judged as negative, and *TN* is the number of pixels for which negative categories were judged as negative.

To evaluate the recognition accuracy of apple orchards individually, this study used the combined *PA* and *UA* [34,35] to evaluate the apple orchard categories in the classification results, which are defined by the following formula:

$$F1 \text{ Score} = \frac{(\beta^2 + 1) \times PA \times UA}{\beta^2 \times PA + UA} \quad (6)$$

where β is the weight of *PA* and *UA*. When $\beta > 1$, *UA* is more important than *PA*. Conversely, when $\beta < 1$, *PA* is more important than *UA*. In this study, we used $\beta = 1$ to equalize the weights of *UA* and *PA*.

4. Results and Analysis

4.1. Comparison of Apple Orchard Extraction Results for Different Feature Combinations

To compare the classification results of different feature combinations, a total of 12 groups of feature combinations designed in Section 3.4 were used to conduct RF classification experiments, compare the classification results of these 12 feature combinations, and produce a graph of the classification results in the study area, as shown in Figure 9.

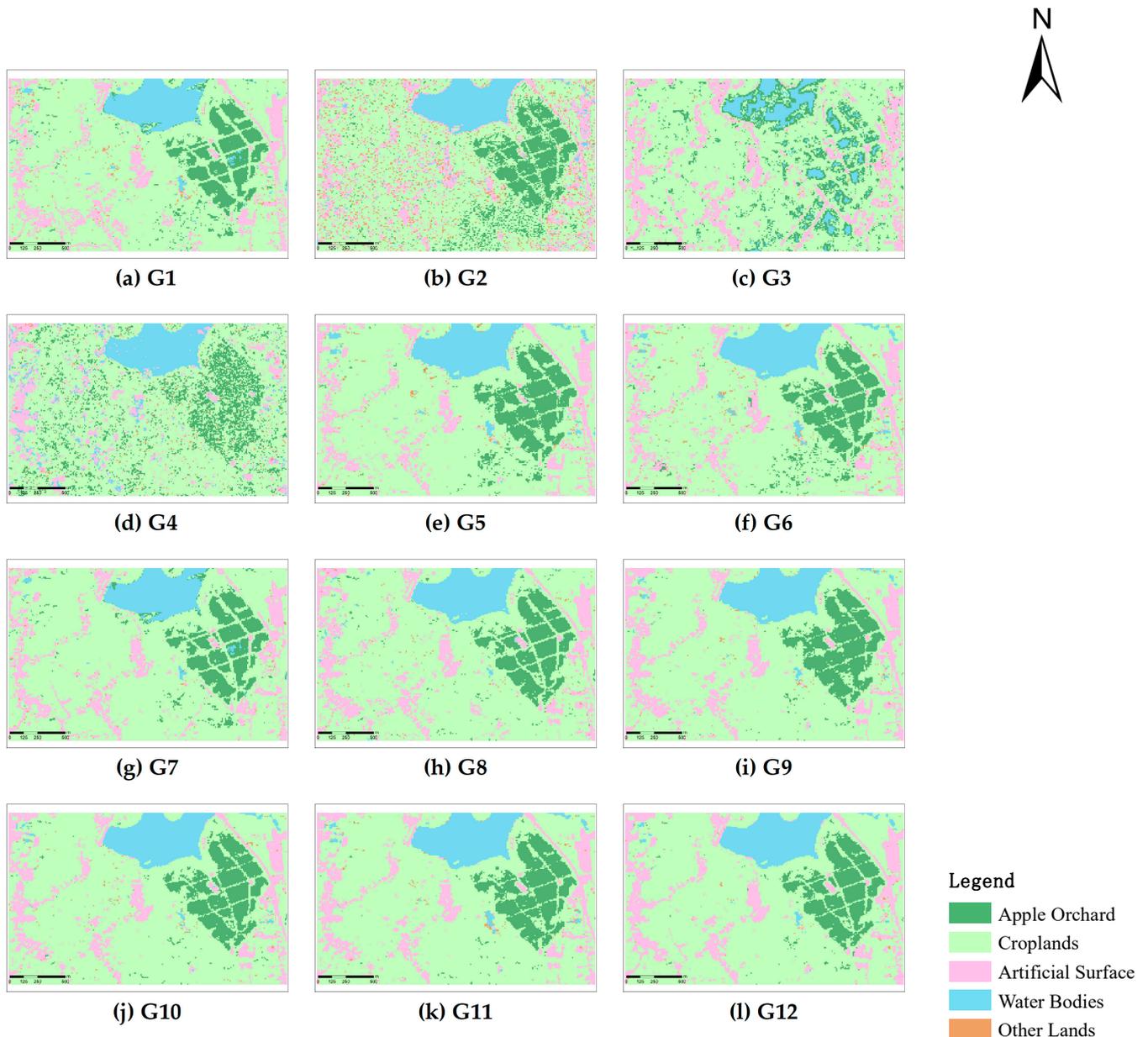


Figure 9. Comparison of classification results for different combinations of features. (a–l) are the results of G1–G12 in Table 3.

Comparing the classification results of G1 to G12, the G1 classification results (Figure 9a) could better identify apple orchards. However, identifying artificial surfaces was poor, and there was slight confusion between water bodies and apple orchards. The classification results for G2 (Figure 9b) identified apple orchards. However, there was a clear “Salt and pepper phenomenon” when the other land types were identified. G3 (Figure 9c) did not distinguish apple orchards from water bodies. G4 (Figure 9d) showed a significant im-

provement in the confusion between apple orchards and water body boundaries compared to G1 and improved the recognition of artificial surfaces. However, apple orchards and other cultivated land could not be distinguished.

G5 and G6 (Figure 9e,f) could identify apple orchards and still exhibited the phenomenon of mixing apple orchards and cultivated land. However, there was still a lack of accuracy in identifying apple orchards and croplands, and the identification of artificial surfaces was insufficient. G7 (Figure 9g) identified the types of features contained in the thin cloud-covered areas of the study area, although there was still a clear mixing of apple orchards and water bodies. Compared to G5, G6, and G7, G8 (Figure 9h) was better at identifying the features contained in the apple orchards and the thin cloud-covered area in the study area and better recognized other features. Nevertheless, its recognition was still more fragmented.

The classification of G9 to G12 (Figure 9i–l) still contained fragmented image elements. However, the number of fragmented image elements in G12 was lower and more consistent with the range of features obtained by visual interpretation in Section 3.1. The above analysis shows that the extraction of apple orchards by combining RGB Bands, NDVI, textural features, GLCM, and improved spectral angle features is optimal and can be applied to the extraction of apple orchards in the Zhaoyang District, Zhaotong City.

4.2. Analysis of Apple Orchard Extraction Accuracy for Different Feature Combinations

To further verify the apple orchard extraction accuracy of different feature combinations, the RF classification accuracy of different classification feature combinations was analyzed, and the accuracy of different classification feature combinations is shown in Table 4.

Table 4. Results of random forest classification using different features.

Combination of Features	OA/%	Kappa	UA/%	PA/%	F1 Score/%
G1	80.56	66.95	93.88	75.41	83.64
G2	70.83	52.65	76.27	73.77	75.00
G3	56.67	29.18	45.83	18.03	25.88
G4	67.78	45.29	60.71	55.74	58.12
G5	80.56	66.41	88.68	77.05	82.46
G6	80.56	66.63	88.89	78.69	83.48
G7	78.06	63.61	91.84	73.77	81.82
G8	80.83	67.17	92.16	77.05	83.93
G9	79.44	66.33	88.68	77.05	82.46
G10	83.06	70.70	92.45	80.33	85.96
G11	82.78	70.55	92.73	83.61	87.93
G12	84.72	74.69	92.86	85.25	88.89

Analysis of Table 4 shows that:

Comparing the accuracy of feature combinations G1 to G4, G1 had an OA of 80.56%, an F1 score of 83.64%, and a kappa coefficient of 66.95%. G2 had an OA of 70.83%, an F1 score of 75%, and a kappa coefficient of 52.65%. G3 had an OA of 56.67%, an F1 score of 25.88%, and a kappa coefficient of 29.18%. G4 had an OA of 67.78%, an F1 score of 58.12%, and a kappa coefficient of 45.29%. All classification accuracies for G1 were higher than those for G2, G3, and G4, indicating that the overall classification of RGB Bands in the study area in October 2022 was better, whereas using only the improved spectral angle features or texture features distinguishing all land use types in the study area was less effective and had lower OA and F1 scores.

Comparing the classification accuracies of feature combinations G5 to G8, the OA and F1 scores of G5, G6, and G8 were greater than 80% when the RGB Bands were combined with other features, achieving a relatively good classification result. Although the results of G4 (improved spectral angle features) were not good when classified alone, the OA, kappa coefficient, and F1 score of G8 (RGB Bands + improved spectral angle features) were the best

for the combination of features G5–G8, reaching 80.83%, 67.17%, and 83.93%, respectively, indicating that the combination of RGB Bands with improved spectral angle features can effectively improve the separability of the feature space, which in turn improves the classification accuracy and is more suitable for apple orchard extraction in the Zhaotong City area.

Comparing the classification accuracies of feature combinations G9 to G12, we observed that G10 outperformed G9 in terms of OA, kappa coefficient, and F1 score, indicating that the improved spectral angle feature was more effective than NDVI in identifying apple orchards. The OA, kappa coefficient, and F1 score of G12's feature combinations were the highest at 84.72%, 74.69%, and 88.89%, respectively. Among the 12 feature combinations, G11 had a slightly lower OA, kappa coefficient, and F1 score, indicating that the addition of texture features could improve the separability of features in the study area to a certain extent and that the combination of RGB Bands, NDVI, texture features, GLCM, and improved spectral angle features could effectively identify apple orchards.

5. Discussion

The complex topography and cloudy and rainy weather in Zhaotong City are important reasons for the difficulty in accurately identifying apple orchards. Selecting samples and features are two important factors affecting the classification results during the classification process. However, during sample selection, the influence of human and objective factors prevents effective manual sampling. Therefore, based on visual interpretation, to obtain the sample selection area of the study space, we used the random generation of samples through the Create Random Points Tool in ArcGIS 10.6, which attempts to solve these problems regarding feature selection.

This study proposes an improved spectral angle feature using the VV and VH bands of 12 months of SAR images as the input of the Spectral Angle Mapper and obtains a spectral angle feature with time-series characteristics that is unaffected by clouds [36,37]. Based on the GEE platform, a series of features were constructed, and the Ganhe Reservoir area in Zhaoyang District of Zhaotong City was classified by random forest classification. From the results and accuracy of apple orchard identification, it is poor at distinguishing all land use types in the study area only by using improved spectral angle features or texture features. The feature combination of the RGB band and improved spectral angle feature can effectively improve the separability of the feature space, and compared with the feature combination of RGB and NDVI, the combination of RGB and the improved spectral angle feature has a better recognition effect for apple orchards. The experimental results show that the method of combining texture features, such as Song et al. [7], has poor accuracy in the study area. The combination of spectral angle features and spectral features has obvious advantages in identifying different ground features, and can identify the types of ground features contained in the thin cloud coverage area to a certain extent.

However, there are a few 'broken pixels' in the classification result. The reason for this may be that in October, some of the cultivated land pixels in the study area are highly similar in spectrum and space. In addition, the improved spectral angle features of this paper are calculated from 12 months of monthly average SAR images. This will cause these pixels to have a certain similarity to the apple orchard pixels in the improved spectral angle features, resulting in a "broken pixels". Therefore, in the case of similar spectral and spatial information, using the monthly average SAR image with a large pixel value difference between apple orchards and other cultivated land to calculate the improved spectral angle features can alleviate the problem of "broken pixels". Most of the classification results for other lands were classified as croplands, which differs significantly from the results obtained from visual interpretation; this is an important factor affecting the OA of the RF classification.

In subsequent work on identifying apple orchards in the entire Zhaotong City area, it was not easy to interpret visually because of the large scope of the images. Therefore, we could use other data products with higher accuracy to complete the selection of samples

without manual intervention, which could effectively avoid the misclassification of land classes due to visual interpretation. Meanwhile, there are numerous mixed pixels present within the study area. Therefore, it is possible to use the hybrid pixel decomposition technique [38] for the decomposition of pixels within the study area, which allows for an increase in the accuracy of the classification. In addition, when producing such special spectral angle features, the effects of SAR images on apple orchards and other land classes in particular months can be considered to obtain better quality spectral angle features. In addition, as apple orchards require special light intensity, topographic factors can be considered as one of the input features when performing apple orchard identification for the entire Zhaotong City area, which is conducive to fully learning the topographic features and thus reducing the error and omission rates of apple orchards and other vegetation types.

6. Conclusions

In this study, an improved spectral angle feature based on SAM was proposed. The apple orchards in the study area were classified using GEE, which is computationally powerful and has strong online visualization and computational analysis capabilities. The classification results showed that this feature could effectively improve the identification accuracy of apple orchards in the study area. In summary, the improved spectral angle feature proposed in this study can effectively identify apple orchards in the Zhaotong City area under conditions of complex terrain topography with cloudy and rainy terrain by combining RGB Bands, NDVI, and texture features (GLCM). The specific findings are as follows:

(1) In this study, using Sentinel-1 and Sentinel-2 images as data sources, the RGB Bands, red-edge features, NDVI, texture features, and improved spectral angle features were used in the GEE platform to identify apple orchards in the study area using the RF algorithm. By comparing the classification accuracy of different classification features, this study observed that the texture features facilitated the identification of apple orchards and could not effectively identify other land classes. The combination of the RGB band and spectral angle features can effectively improve the extraction accuracy and overall accuracy of apple orchards, which are 80.3% and 67.17%, respectively, compared to the single RGB band classification. When the RGB Bands were combined with NDVI, GLCM, and spectral angle features, the apple orchard extraction accuracy and overall accuracy reached their best values of 88.89% and 84.44%, respectively.

(2) The research method used in this study can provide a reference for recognition tasks in complex terrains and has practical value. However, when constructing the spectral angle features, the VV and VH bands of the 12 months of SAR images have a certain redundancy, and the study area contains many mixed image elements, resulting in the “pretzel phenomenon” in the study results, which has a certain impact on the extraction accuracy of apple orchards. Therefore, the VV and VH bands, which have more obvious differences from other land types, can be considered in constructing spectral angle features to reduce redundancy, reduce computational effort, and improve the separability of spectral angle features. Concurrently, in subsequent work on identifying apple orchards in the entire Zhaotong City area, other data products with higher accuracy can be used to complete the selection of samples without manual intervention to avoid misclassification of land types due to visual interpretation.

Author Contributions: Conceptualization, J.L., W.S. and X.Z.; software, X.Z. and D.Z.; validation, Q.W.; resources, W.S. and J.L.; writing—original draft preparation, J.L.; writing—review and editing, J.L. and W.S.; supervision, X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Yunnan Province Key Research and Development Program (No. 202202AD080010).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the Copernicus project for providing Sentinel-1 and Sentinel-2 images for free and the GEE platform for image preprocessing of Sentinel-1 and Sentinel-2 images.

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

References

1. Wen, B. A Preliminary Study on Building an Ecological Apple Orchard in Zhaotong, Yunnan. *Appl. Technol. Inf. Fruit Tree* **2011**, *10*, 41–42.
2. Cheng, Z.; Ren, J.; Tang, H.; Shi, Y.; Liu, J. Progress and perspectives on agricultural remote sensing research and applications in China. *J. Remote Sens.* **2016**, *20*, 748–767.
3. Huang, J.; Hou, Y.; Su, W.; Liu, J.; Zhu, D. Mapping corn and soybean cropped area with GF-1 WFV data. *Trans. Chin. Soc. Agric. Eng.* **2017**, *33*, 164–170.
4. Wang, D.; Zhao, A.; Li, J.; Zhang, A. Extraction model of winter wheat planting information based on GF-1 data and unsupervised classification. *Sci. Technol. Eng.* **2019**, *19*, 95–100.
5. Zhang, C.; Han, Y.; Li, F.; Gao, S.; Song, D.; Zhao, H.; Fan, K.; Zhang, Y.N. A New CNN-Bayesian Model for Extracting Improved Winter Wheat Spatial Distribution from GF-2 imagery. *Remote Sens.* **2019**, *11*, 619. [[CrossRef](#)]
6. Zhang, H.; Tian, T.; Zhang, Q.; Lu, Z.; Shi, C.; Tan, C.; Luo, M.; Qian, C. Study on Extraction of Paddy Rice Planting Area in Low Fragmented Regions based on GF-1 WFV Images. *Remote Sens. Technol. Appl.* **2019**, *34*, 785–792.
7. Song, R.; Ning, J.; Liu, X.; Chang, Q. Apple Orchard Extraction with QuickBird Imagery Based on Texture Features and Support Vector Machine. *Trans. Chin. Soc. Agric. Mach.* **2017**, *48*, 188–197.
8. Dai, J. Apple Orchard Extraction Based on High Resolution Images and Multi-temporal Mid-resolution Images. *Chin. J. Agric. Resour. Reg. Plan.* **2022**, *43*, 140–148.
9. Dong, F.; Zhao, G.; Wang, L.; Zhu, X.; Chang, C. Remote sensing techniques of apple orchard information extraction based on linear spectral unmixing with measured data. *Chin. J. Appl. Ecol.* **2012**, *23*, 3361–3368.
10. Xie, L.; Zhang, H.; Li, H.; Wang, C. A unified framework for crop classification in southern China using fully polarimetric, dual polarimetric, and compact polarimetric SAR data. *Int. J. Remote Sens.* **2015**, *36*, 3798–3818. [[CrossRef](#)]
11. Li, J.; Xu, J. Tree Species Classification of Time Series Remote Sensing Images by Dynamic Time Warping. *J. Northeast. For. Univ.* **2017**, *45*, 56–61.
12. Belgiu, M.; Csillik, O. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sens. Environ.* **2018**, *204*, 509–523. [[CrossRef](#)]
13. Erasmi, S.; Twele, A. Regional land cover mapping in the humid tropics using combined optical and SAR satellite data—A case study from Central Sulawesi, Indonesia. *Int. J. Remote Sens.* **2009**, *30*, 2465–2478. [[CrossRef](#)]
14. Gebhardt, S.; Huth, J.; Nguyen, L.D.; Roth, A.; Kuenzer, C. A comparison of TerraSAR-X Quadpol backscattering with RapidEye multispectral vegetation indices over rice fields in the Mekong Delta, Vietnam. *Int. J. Remote Sens.* **2012**, *33*, 7644–7661. [[CrossRef](#)]
15. Kruse, F.A.; Lefkoff, A.B.; Boardman, J.W.; Heidebrecht, K.B.; Shapiro, A.T.; Barloon, P.J.; Goetz, A.F.H. The spectral image processing system (SIPS)-interactive visualization and analysis of imaging spectrometer data. *Earth Space Sci. Inf. Syst. (ESSIS)* **1993**, *283*, 192–201.
16. Grzegozewski, D.M.; Johann, J.A.; Uribe-Opazo, M.A.; Mercante, E.; Coutinho, A.C. Mapping soya bean and corn crops in the State of Paraná, Brazil, using EVI images from the MODIS sensor. *Int. J. Remote Sens.* **2016**, *37*, 1257–1275. [[CrossRef](#)]
17. Li, X.; Wang, H.; Li, X.; Chi, D.; Tang, Z.; Han, C. Study on Crops Remote Sensing Classification based on Multi-temporal Landsat 8 OLI Images. *Remote Sens. Technol. Appl.* **2019**, *34*, 389–397.
18. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [[CrossRef](#)]
19. Kumar, L.; Mutanga, O. Google Earth Engine Applications Since Inception: Usage, Trends, and Potential. *Remote Sens.* **2018**, *10*, 1509. [[CrossRef](#)]
20. Luo, C.; Qi, B.; Liu, H.; Guo, D.; Shao, Y. Using Time Series Sentinel-1 Images for Object-Oriented Crop Classification in Google Earth Engine. *Remote Sens.* **2021**, *13*, 561. [[CrossRef](#)]
21. Inoue, S.; Ito, A.; Yonezawa, C. Mapping Paddy Fields in Japan by Using a Sentinel-1 SAR Time Series Supplemented by Sentinel-2 Images on Google Earth Engine. *Remote Sens.* **2020**, *12*, 1622. [[CrossRef](#)]
22. Xie, S.; Liu, L.; Zhang, X.; Yang, J.; Chen, X.; Gao, Y. Automatic Land-Cover Mapping using Landsat Time-Series Data based on Google Earth Engine. *Remote Sens.* **2019**, *11*, 3023. [[CrossRef](#)]
23. Pan, L.; Xia, H.; Zhao, X.; Guo, Y.; Qin, Y. Mapping Winter Crops Using a Phenology Algorithm, Time-Series Sentinel-2 and Landsat-7/8 Images, and Google Earth Engine. *Remote Sens.* **2021**, *13*, 2510. [[CrossRef](#)]
24. Yuan, Y.; Wen, Q.; Zhao, X.; Liu, S.; Zhu, K.; Hu, B. Identifying Grassland Distribution in a Mountainous Region in Southwest China Using Multi-Source Remote Sensing Images. *Remote Sens.* **2022**, *14*, 1472. [[CrossRef](#)]

25. Yan, Y.; Tang, X.; Zhu, X.; Yu, X. Optimal Time Phase Identification for Apple Orchard Land Recognition and Spatial Analysis Using Multitemporal Sentinel-2 Images and Random Forest Classification. *Sustainability* **2023**, *15*, 4695. [[CrossRef](#)]
26. Maillard, P. Comparing Texture Analysis Methods through Classification. *Photogramm. Eng. Remote Sens.* **2003**, *69*, 357–367. [[CrossRef](#)]
27. Li, X.; Liu, S.; Li, L.; Jing, Y.; Fan, W.; Wu, L. Automatic Interpretation of Spatial Distribution of Winter Wheat Based on Random Forest Algorithm to Optimize Multi-temporal Features. *Trans. Chin. Soc. Agric. Mach.* **2019**, *50*, 218–225.
28. ESRI. How Create Random Points Works. Available online: <https://desktop.arcgis.com/en/arcmap/10.6/tools/data-management-toolbox/how-create-random-points-works.htm> (accessed on 27 March 2023).
29. Petropoulos, G.P.; Vadrevu, K.P.; Gavriil, X.; George, K.; Marko, S. A Comparison of Spectral Angle Mapper and Artificial Neural Network Classifiers Combined with Landsat TM Imagery Analysis for Obtaining Burnt Area Mapping. *Sensors* **2010**, *10*, 1967–1985. [[CrossRef](#)]
30. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
31. He, Z.; Hu, J.; Cai, Z. Remote sensing identification for *Artemisia argyi* integrating multi-temporal GF-1 and GF-6 images. *Trans. Chin. Soc. Agric. Eng.* **2022**, *38*, 186–195.
32. Tao, L.; Hu, Z. Crop planting structure identification based on Sentinel-2A data in hilly region of middle and lower reaches of Yangtze River. *Bull. Surv. Mapp.* **2021**, *7*, 39–43.
33. Van Beijma, S.; Comber, A.; Lamb, A. Random forest classification of salt marsh vegetation habitats using quad-polarimetric airborne SAR, elevation and optical RS data. *Remote Sens. Environ.* **2014**, *149*, 118–129. [[CrossRef](#)]
34. Reis, S.; Tasdemir, K. Identification of hazelnut fields using spectral and Gabor textural features. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 652–661. [[CrossRef](#)]
35. Aksoy, S.; Akcay, H.G.; Wassenaar, T. Automatic Mapping of Linear Woody Vegetation Features in Agricultural Landscapes Using Very High Resolution Imagery. *IEEE Trans. Geosci. Remote Sens.* **2009**, *48*, 511–522. [[CrossRef](#)]
36. Jing, X.; Wang, J.; Wang, J.; Huang, W.; Liu, L. Classifying forest vegetation using sub-region classification based on multi-temporal remote sensing images. *Remote Sens. Technol. Appl.* **2008**, *23*, 394–397.
37. Xu, J.; An, Y.; Liu, S.; Han, K. Discussion on classification for Sentinel-1A SAR data in mountainous plateau based on backscatter features—A case study in Anshun city. *J. Guizhou Norm. Univ. (Nat. Sci.)* **2016**, *34*, 15–19+38.
38. Li, M.; Guo, S.; Chen, J.; Chang, Y.; Sun, L.; Zhao, L.; Li, X.; Yao, H. Stability Analysis of Unmixing-Based Spatiotemporal Fusion Model: A Case of Land Surface Temperature Product Downscaling. *Remote Sens.* **2023**, *15*, 901. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.