International Journal of Environment, Agriculture and Biotechnology
Vol-8, Issue-6; Nov-Dec, 2023
Peer-Reviewed International Journal
Journal Home Page Available: https://ijeab.com/
Journal DOI: 10.22161/ijeab

Analysis of Dynamic Changes of Winter Wheat in Xinye County, Henan Province Based on SVM Method

Yan Si[1], Bing-Yuh Lu[2], Yun-Shang Wang[3], Ruei-Yuan Wang*[4]

[1,4]School of Science, Guangdong University of Petrochem Technology (GDUPT), China [2]Faculty of Automation, Guangdong University of Petrochemical Technology, China

[3]Graduate Institute, Fu Jen Catholic University

*Corresponding author

Received: 13 Sep 2023; Received in revised form: 25 Oct 2023; Accepted: 05 Nov 2023; Available online: 11 Nov 2023

***Abstract*— *The guarantee of grain yield is an important issue for national security. Wheat is one of the main grain crops in China, and monitoring the spatio-temporal changes in its planting area and yield has important implications for decision-making support. With the development of remote sensing technology, estimating the long-term changes in the area of wheat planting has become a vital agricultural monitoring method. This article uses GF-1 satellite WFV sensor data to estimate the wheat planting areas in Xinye County, Henan Province in 2017, 2020, and 2023, mainly using SVM algorithm for calculation and comparison. After classification, the overall classification accuracy reaches over 95%, and the Kappa coefficient is above 0.95. The results show that the winter wheat planting area in Xinye County has shown an increasing trend over the past six years, from 34296.295 hm^2 in 2017 to 56914.662 hm^2 in 2023. By analyzing and summarizing the changes in regional crops, it has an important contribution to regional production and agricultural evaluation decision-making.*

***Keywords*— Support Vector Machine (SVM), Winter wheat; Spatial and temporal distribution, Dynamic changes, Agricultural remote sensing**

I. INTRODUCTION

Winter wheat is an important grain crop in China, and Xinye County, Henan Province, is a major grain producing county. Obtaining its agricultural spatial distribution information is of great significance for the local agricultural informatization and professional intelligent development. Traditional agricultural surveys require a significant amount of manpower, material resources, and time costs, and the final results cannot reflect spatial distribution information. The use of computer technology can further reduce costs, making survey results more intuitive and improving efficiency through timeliness and visualization. Using remote sensing satellite data to detect crops in planting areas can meet the timely and long-term estimation needs of large areas. By analyzing data from multiple periods, resolutions, and sources, data on crop distribution, area, and yield estimation can be effectively obtained.

The research on grain estimation using agricultural remote sensing is to effectively distinguish monitored crop categories from other land features through algorithms and thresholds. The commonly used method is to establish thresholds by normalizing vegetation indices to distinguish wheat from other land features. For example, Zhao et al. (2012) used the monthly normalized vegetation index to establish a linear model for winter wheat yield, which can effectively, quickly, and accurately estimate winter wheat yield [1]; Hao et al. (2017) used the normalized vegetation index as a threshold to segment and extract winter wheat, and calculated the area, achieving good results [2]; Liu et al. (2019) established a yield estimation model using normalized plant mean based on the annual average yield of rice, achieving remote sensing estimation of rice yield [3]; Ren et al. (2006) selected NDVI data ranging from 0.2 to 0.8 during the critical growth period of winter wheat and established their relationship with winter wheat yield, obtaining yield estimation data with an error of within 4%, which has high accuracy [4]. In addition, there are also methods for analyzing and identifying crops using multi-source data: for example, Feng et al. (2023) used multivariate remote sensing data to analyze Sentinel-1 SAR images and Sentinel-2 optical remote sensing images, and used SVM algorithm to achieve an overall classification accuracy of 94.3% [5].

Furthermore, different algorithms can also be used to better classify crops on the surface. For example, Sun et al. (2017) used different algorithms for supervised classification and compared the characteristics of six classification methods. The results showed that the producer and user precision of each classification method had different differences and could not be applied to all types of land objects. Each classification method had different advantages in different land object categories. Among them, the maximum likelihood classification (MLC), support vector machine (SVM) classification, and artificial neural network (ANN) classification have better overall classification accuracy and Kappa coefficient [6]. Many studies have shown that SVM has good advantages, for example, Sun et al. (2013) used SVM algorithm to monitor land use and cover changes in the Abihu Lake area in 1990, 2001, and 2011, and concluded that SVM classification method is the optimal classification method [7]; Guo et al. analyzed and evaluated six classification methods, including SVM, and concluded that the accuracy of SVM based classification methods is higher than other classification methods through comparative analysis of classification experimental data; Chen et al. (2019) used multispectral images as the basis to train and effectively recognize landslide sample points in remote sensing images using landslide area SVM detection models [8]; Zu (2018) processed the remote sensing images of Zhuhai in Phase III, used SVM algorithm for land and water separation, and conducted cross processing on the results. The extracted coastline was analyzed to obtain the coastal construction changes of Zhuhai in the past decade [9]; Zhang et al. (2016) used SVM algorithm to classify wetland water bodies, and the results showed that SVM has higher accuracy in various aspects than traditional methods, making it very suitable for wetland information extraction and monitoring in arid areas [10]. In addition, the SVM classification method has been repeatedly tested as a mature and high-precision classification algorithm.

The selection of remote sensing data is mostly based on Sentinel and Landsat series satellites. However, in recent years, with the continuous improvement of domestic satellite level, high resolution series satellites have gradually been widely used. In terms of spatial resolution, medium and low resolution satellites have certain advantages in terms of breadth and temporal resolution, which have good benefits for large research areas. At the same time, they can compare the ground conditions of different years and the same time period, which can improve the requirements for temporal resolution. Based on this, this study is using the SVM classification algorithm, mainly using data from the GF-1 satellite WFV sensor with a spatial resolution of 16 meters. Xinye County is selected as the study area, and the distribution of winter wheat crops for three years is extracted and land use change statistics are conducted to analyze the area change of local grain crop cultivation, in order to obtain visual spatial information of crops.

II. STUDY AREA AND DATA SOURCES

2.1 Study Area

Xinye County is located in the southwest of Henan Province (112° 12′ 44″~112° 35′ 42″ E, 32° 19′ 30″~32°49′ 08″ N), with a total administrative area of 1062km² (Figure 1). It belongs to the northern subtropical monsoon climate, with an average annual temperature of 16–17°C, an average annual precipitation of 800-900mm, and an average annual frost-free period of 227 days. In summer, the temperature is high, precipitation is concentrated, and drought and flood disasters are frequent. In winter, it is dry and cold, with a small amount of rain and snow. Due to its location in the transitional zone between the north and south of China, the transition zone has obvious characteristics: a warm and humid climate, distinct seasonal changes, a long frost-free period, sufficient solar radiation, and being suitable for wheat growth, making it a key area for agricultural development in China. Xinye County is located in the hinterland of the Nanyang Basin, with a flat and vast terrain and an average elevation of around 100m. It is the only county in Nanyang City that does not have a mountainous distribution. There are two main rivers within the territory: the Bai River and the Tang River, as well as numerous tributaries. The river generally passes through the county town from north to south and finally enters the Han River in Hubei, returning to the Yangtze River basin.

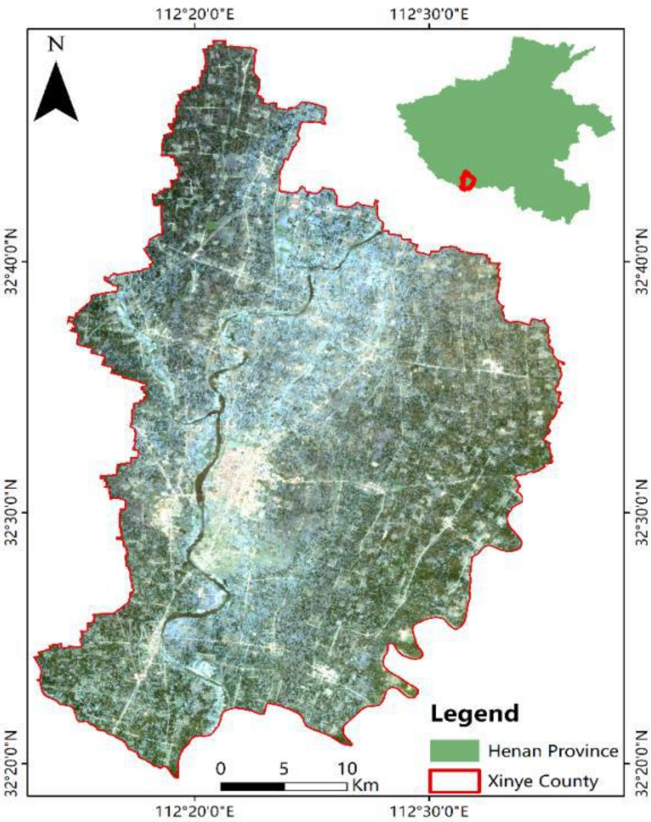


Fig.1 Regional Overview of Xinye County Prefecture

Under the influence of superior climate and hydrological resources, the perennial wheat planting area of Xinye County can reach 533000 hectares, with a total yield of 3.25 billion kilograms, accounting for 11% of the yield in Henan Province and 1% of the national yield. It is truly a major grain county.

2.2 Data Sources

The remote sensing data source of this study is the GF-1 satellite WFV sensor data from the China Resources Satellite Application Center in 2017, 2020, and 2023 (https://data.cresda.cn/#/home). In the experiment, LANDSAT-8 L2 data was also used for geographic registration of images at different time resolutions, while ASTER GDEM 30M resolution DEM was used, which was sourced from the geospatial data cloud (https://www.gscloud.cn/home).

This article is based on ENVI software, after radiation correction, and uses the Atmospheric Correction Algorithm Tool (FLAASH) to produce available GF-1 series satellite data covering the Xinye County area in 2017, 2020, and 2023.

The ROI sample data is produced by combining the sample data with human visual interpretations of Google Earth high-resolution images. The classification of this study is divided into four types: water body, wheat, buildings, and bare land. Images at different times are compared and modified to avoid changes in the selected pixels at different times, which may affect the accuracy of the classification data.

III. METHODOLOGY

The analysis image in this article is based on the GF-1/WFV satellite image with a 16m resolution spectral band as the main data source, assisted by Landsat-8 L2 level data. Sample data is constructed by identifying the spectral features of different land features, and SVM classification algorithm is used to conduct research using a human-machine interaction interpretation platform. The winter wheat planting information in the Xinye County area in 2017, 2020, and 2023 is obtained (Figure 2).

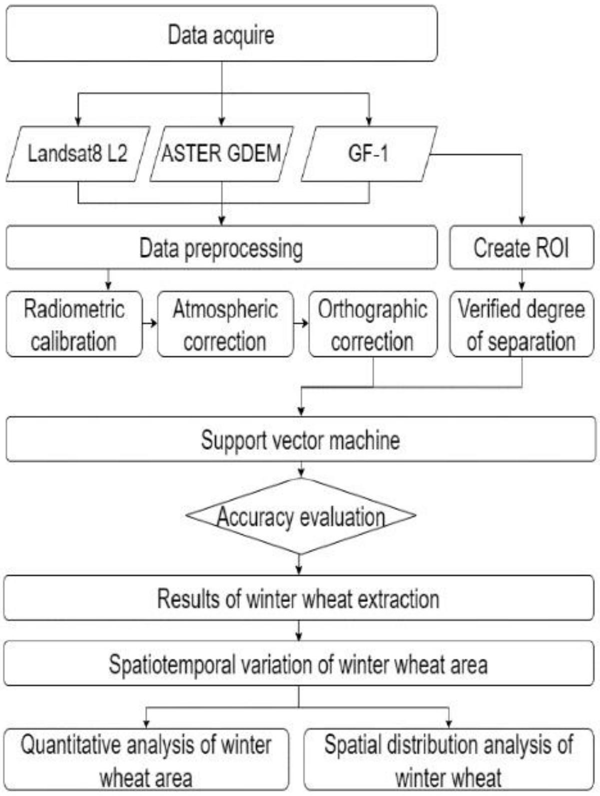
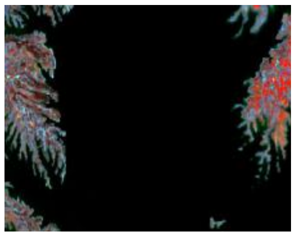
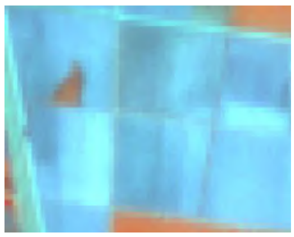


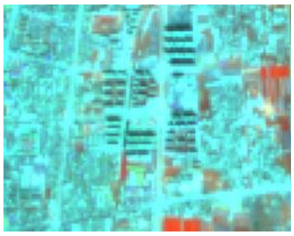
Fig.2 the Technology Roadmap of the Study

3.1 Obtaining ROI Sample Data

Due to the changes in the distribution of surface features in different years, except for rivers that have remained almost unchanged for six years, the planting of crops in different fields varies each year. The changes in bare land and wheat planting areas are significant, and manual sample data extraction is required for images from different years to establish training samples for subsequent interpretation. After repeated experiments and reference to other studies, a false color band combination of 432 was finally used for the production of ROI samples (Table 1). Under this band combination, the surface wheat showed a bright red color, while the bare ground showed a gray blue color, the buildings and roads were turquoise green, and the water body was black. Various objects can be distinguished well, which is helpful for the identification of objects and the production of sample areas, and can reduce the time cost of object identification.

Table 1 Ground Feature Characteristics of GF-1 Imagery in the False Color Band

Type	Characteristic	True color	False color
Wheat	Dark green or green appears in true colors, and red or orange red appears in false color combinations.		
Water body	The colors presented in true color and false color are similar, both being black.		
Bare land	True color appears in yellow, while false color appears in blue. There is a difference in brightness.		

Building	True colors appear in blue and white, with significant variations in brightness, while false colors appear in cyan with higher brightness cyan.		

The sample selection in the ROI area should cover the entire study area as much as possible, and the distribution of different types should also follow the principle of average distribution; Furthermore, the amount of image sample data in different time periods should be as close as possible to prevent significant differences in algorithm classification due to differences in sample data. Among them, the selection of water samples remained basically unchanged; The changes in buildings are also relatively small, with the addition of some roads and building areas in 2023; The selection of samples for bare land and wheat has relatively significant changes, and it can be observed from the images that the range of bare land has decreased, so the selection of samples has gradually decreased. The opposite is true for wheat samples (Table 2). The average separation degree of the sample pixels extracted based on the above principles is above 1.93, which meets the requirement of separation degree above 1.8.

Table 2 Sample Data Pixel Table (Unit: piece)

	2017	2020	2023
Water body	7108	7328	7012
Building	7746	5944	7852
Bare land	6230	5978	5630
Wheat	12420	13433	16513

3.2 SVM Classification Algorithm

Support Vector Machine (SVM) is a class of generalized linear classifiers that perform binary classification of data using supervised learning. Its decision boundary is the maximum-margin hyperplane that solves the learning sample. The algorithm logic is to transform the analyzed dataset into a high-dimensional new space through nonlinear algorithms, and in the new space, classify the dataset through algorithms to achieve nonlinear discrimination. The basic principle of SVM classification is to set $x_i \in R^d$ as the input mode on a sample dataset, and $y \in \{\pm 1\}$ as the output target. Let the equation for the optimal decision surface be:

$$w^t x_i + b = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,\ldots(1)$$

The weight vectors w and b offset must satisfy constraint$y_i(w^t x_i + b) \geq 1 - \xi_i$. In the formula ζ_i is the slack variable of the sample under linearly indivisible constraints, and is the degree to which the pattern deviates from the ideal situation. The essence of SVM is to fit multiple hyperplanes into one, separate two types of data, and find a decision surface that minimizes the average classification error of all data. Based on this, the optimization formula is derived as follows:

$$\phi(w,\xi) = \frac{1}{2}w^T w + c\sum_{i=1}^{n}\xi_i \ldots\ldots\ldots\ldots(2)$$

Among them, c represents a positive parameter specified by the user, which is used for SVM to support the punishment of sample correctness and error. It is a parameter that balances the complexity of the algorithm and the proportion of error samples. By balancing the complexity and learning cost of classification with limited samples, good classification results can be achieved [11-13].

3.3 Precision Evaluation

The confusion matrix is the commonly used method for evaluating the accuracy of remote sensing classification. The classification results are quantitatively rated based on classification accuracy, Kappa coefficient, and other specific values. The calculation formula for Kappa coefficient is as follows:

$$kappa = \frac{p_o - p_e}{1 - p_e} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

Where $p_o = \frac{\text{sum of diagonal elements}}{\text{sum of matrix elements}}$, which is the consistency observed between the two sample data;

$$p_e = \frac{\sum_i \ Sum\ of\ elements\ in\ the\ i{-}th\ row * Sum\ of\ elements\ in\ the\ i{-}th\ column}{(\sum All\ elements\ of\ the\ matrix)^2}$$

, which is the consistency of opportunities between two samples.

The Kappa coefficient is obtained by combining different precision parameters, with values between [-1, 1]. The closer it is to 1, the higher the consistency of classification and the better the classification effect.

IV. ANALYSIS AND RESULT

4.1 Winter Wheat Extracted by SVM Algorithm

Based on the requirements of average distribution in the study area, average number of sampling points, and average distance of sampling points, the sample data was created and classified using the SVM algorithm to obtain the classification of ground objects at different time resolutions (Figure 3). The water body has more obvious features in the false color image, and the overall separation degree is good, which is the blue part in the figure. However, the variation between buildings and bare land in the frequency band is relatively small, and there are bright pixels in both the bare land and buildings in the false color, which affects the classification results of the algorithm, as are the white and orange parts in the image. Finally, winter wheat performs particularly well in the frequency band, making it relatively distinguishable from bare land and buildings, as shown in the green part of the image.

The accuracy verification part involves manually selecting test sample points and evaluating classification accuracy. Ultimately, the overall classification accuracy of images in 2017, 2020, and 2023 can reach over 95%, and the Kappa coefficient remains above 0.95. According to the quantitative rating system, it is closer to "1", indicating good overall classification performance. However, due to the optical resolution of the image, the classification results of smaller-scale features are not obvious, but their errors are still within an acceptable range, which has little impact on the overall research data.

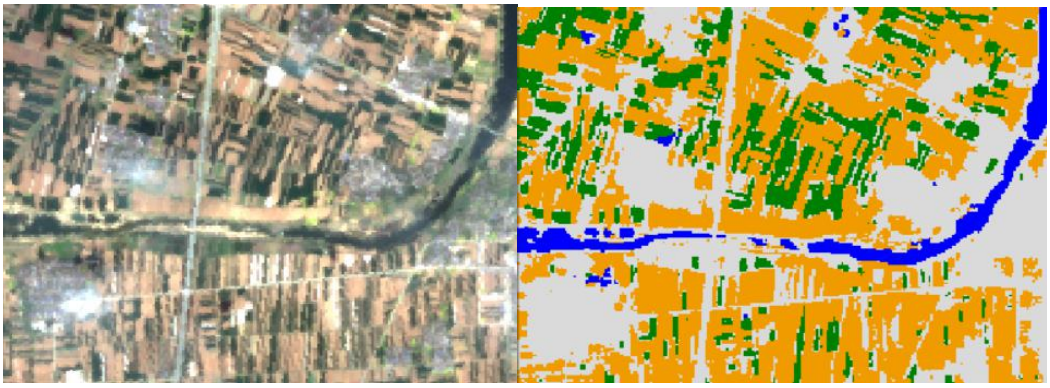


Fig.3 Comparison of Remote Sensing Images and Classification Results

4.2 Analysis of Dynamic Changes in Winter Wheat

According to the classification results of the SVM algorithm, the distribution map of winter wheat planting areas in the study area was obtained, and it can be clearly observed that the winter wheat planting areas in Xinye County increased year by year from 2017 to 2023 (Figure 4). In 2017, there were also significant scattered patches in the wheat planting area, without a large-scale contiguous planting area. By 2023, it has been connected as a large-scale planting area. According to satellite image analysis, the added area is the original bare land area. The central part of Xinye County is the county seat, and the overall area has not changed much. In 2017, there was a large amount of bare land in the north and south of the county town, and some wheat fields were interspersed among them, resulting in a large number of scattered wheat field areas. By 2023, a large amount of wasteland in the south and north will be reused, resulting in a significant increase in wheat cultivation areas.

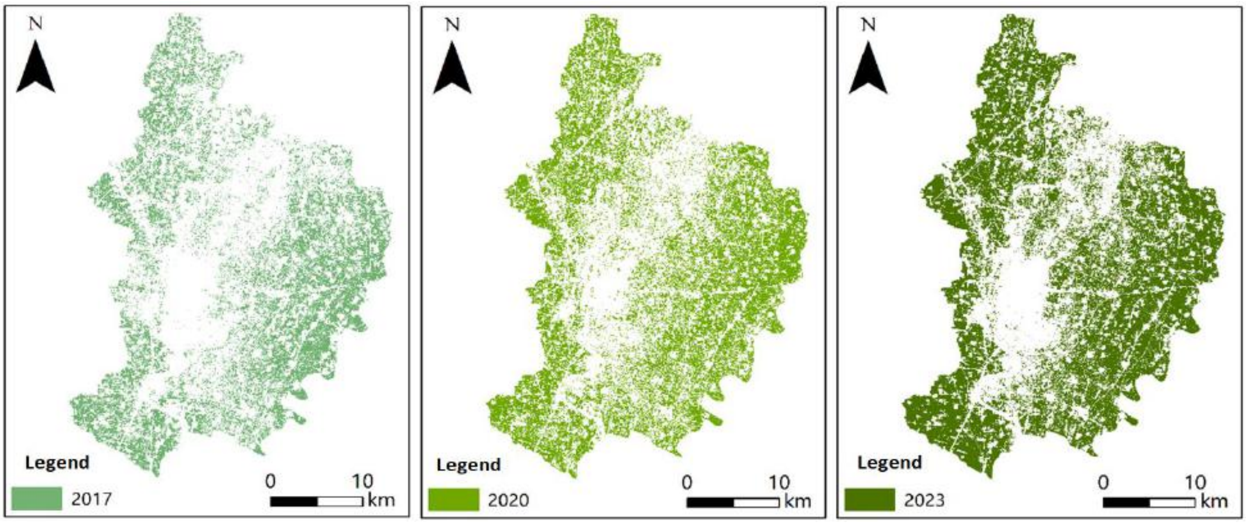


Fig.4 Temporal and Spatial Changes of Wheat Planting Areas in Xinye County

Using ArcGIS tool to conduct area statistics on the winter wheat planting areas in Xinye County for three years (Figure 5), the analysis shows that the total area of Xinye County is 106200 hm^2, and the winter wheat area in April 2017 was 34296.295 hm^2, accounting for 32% of the total area of the county; In April 2020, the winter wheat area was 46113.737 hm^2, accounting for 43% of the county's total area; The winter wheat area in April 2023 was 56914.662 hm^2, accounting for 54% of the entire county. The area of wheat planting areas has been increasing year by year, and the upward trend is obvious. At present, a circular surrounding planting area centered around the county town and river has been formed. The expansion of winter wheat cultivation in the future can develop into a large number of bare land near the river north of the central county town.

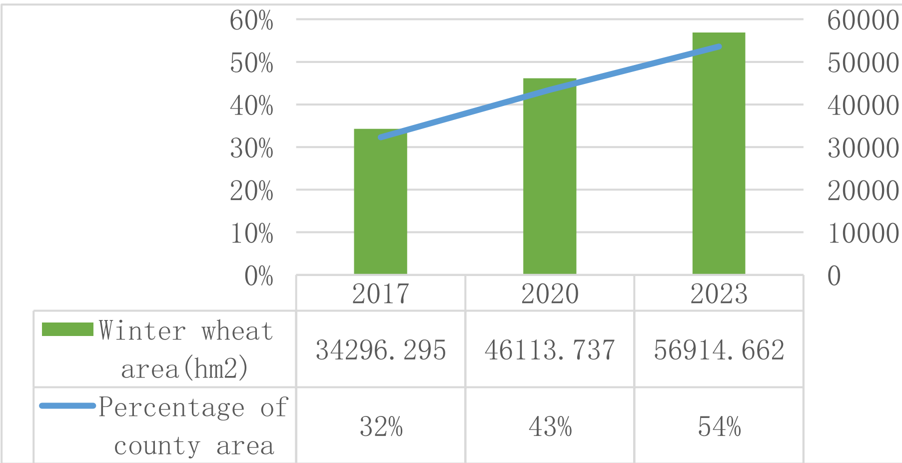


Fig.5 Changes in Winter Wheat Planting Area in Xinye County

V. CONCLUSION

This article uses the GF-1/WFV sensor image to extract the distribution information of winter wheat crops in Xinye County, Nanyang City, Henan Province in 2017, 2020, and 2023. High resolution satellite images are used as reference values to establish sample area (ROI), and accuracy verification is conducted to assist in dividing four different types of feature pixels within the research area. SVM algorithm is used to classify and extract the distribution information of winter wheat crops from the WFV sensor, and the area is calculated. The final extraction effect meets the accuracy requirements. The main focus of this study is to verify the effectiveness of GF-1 imaging combined with SVM algorithm.

The overall classification accuracy of the SVM algorithm in this study is 95%. In the face of a small number of categories, moderate area, small terrain fluctuations, and relatively significant spectral differences of the research object, it has produced good results, and the visual effect is also relatively good. From the research process, the SVM algorithm has relatively high requirements for selecting research samples, and under better sample conditions, the classification effect will be better. In the study, the spectral differentiation between bare land and buildings was relatively low. As a result, the classification of the two types of features was unclear, but it did not affect the classification and extraction of winter wheat.

In terms of agricultural monitoring and analysis, the winter wheat planting area in Xinye County has been continuously increasing in the past six years, reaching 54% of the total county area in 2023. By utilizing a large amount of potential land, the planting area of food crops has been increased, ensuring food production.

This article obtained better classification results in a small scale, and in future research, it can be attempted to use SVM algorithm to study winter wheat in large areas and verify its effectiveness in large-scale space. Overall, this study utilizes remote sensing data sources and technology to monitor agricultural crop yields, which will help China achieve the goal of digitalization and intelligence in agriculture and move towards intelligent agriculture.

ACKNOWLEDGEMENTS

The author is grateful for the research grants given to Ruei-Yuan Wang from GDUPT Talents Recruitment (No.2019rc098), in Guangdong Province, China, and Academic Affairs in GDUPT for Goal Problem-Oriented Teaching Innovation and Practice Project Grant No.701-234660.

REFERENCES

[1] Zhao, W., He, Z., He, J., and Zhu, L. Remote sensing estimation for winter wheat yield in Henan based on the MODIS-NDVI data. Geographical Research, 2012, 31(12): 2310-2320. DOI:10.11821/yj2012120018.

[2] Hao, Z., Zhao, H., and Jiang, Y. Extraction of winter wheat area information based on the improved NDVI density slicing method . South to North Water Transfers and Water Science & Technology, 2017, 15(3) : 67272, 93.

[3] Liu, S., Niu, C., and Bian, L., Remote sensing estimation of rice yield based on NDVI．Jiangsu Agricultural Sciences，2019，47 (3): 193-198. doi:10.15889 / j.issn.1002－1302.2019.03.047

[4] Ren, J., Chen, Z., and Tang, H. Regional scale remote sensing-based yield estimation of winter wheat by using MODIS-NDVI data：A case study of Jining City in Shandong Province. Chinese Journal of Applied Ecology, 2006, 17(12): 2371-2375.

[5] Feng, Q., Ren, Y., Yao, X., Niu, B., Chen, B., and Zhao, Y. Identification of Winter Wheat in Huang-Huai-Hai Plain Based on Multi-source Optical Radar Data Fusion. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(2):160-168.

[6] Sun, K., and Lu, T. Comparison of Supervised Classification Methods in Remote Sensing Image Classification. JIANGXI KEXUE, 2017, 35(3):6.DOI:10.13990/j.issn1001-3679.2017.03.009.

[7] Sun, C., Li, X., and Meng, X. Analysis on Land Use and Land Cover Change and Driving Force of Ebinur Lake Region during the Past 20 Years Based on Support Vector Machine Classification. Xinjiang Agricultural Sciences, 2013, 50(7):1322-1329. DOI:10.6048/j.issn.1001-4330.2013.07.021

[8] Chen, S., Kang, Q., Shen, Z., and Zhou, R. Landslide Detection Based on Color Feature Model and SVM in Remote Sensing Imagery. Spacecraft Recovery & Remote Sensing, 2019, 40(6):10. DOI：10.3969/j.issn.1009-8518.2019.06.011

[9] Zu, J. Coastal Line Extraction in Zhuhai City Based on Supervised Classification and Multi source Remote Sensing Data. China Science and Technology Horizon, 2018(22): DOI：10.3969/j.issn.1671-2064.2018.22.087

[10] Zhang, X., and An, F. Research on Remote Sensing Classification for Ulungur Lake Wetlands Based on SVM. HUBEI AGRICULTURAL SCIENCES, 2016, 55(16):5. DOI:10.14088/j.cnki.issn0439-8114.2016.16.018.

[11] Zhang, F., Xue, Y., Li, Y., and Ding, X. Object-oriented building extraction of multi-source remote sensing imagery based on SVM. Remote Sensing for Land & Resources, 2008(2):27-29+47. DOI:10.6046/gtzyyg.2008.02.07.

[12] He, D., Xiao, Y., Xiao, X., Huang, Y., and Zhou, Q. Application of the Support Vector Machine in Remote Sensed Image Processing. Urban Geotechnical Investigation & Surveying, 2006(3):4. DOI:10.3969/j.issn.1672-8262.2006.03.008.

[13] Zhu, H., and Jia, Y. Remote Sensing Image Classification Based on Support Vector Machines. Science Technology and Engineering, 2010(15): 3659-3663. DOI:10.3969/j.issn.1671-1815.2010.15.023.