Optimal spatial resolution of remote-sensing imagery for monitoring cantaloupe greenhouses

B Cui\textsuperscript{1,2,3}, W J Huang\textsuperscript{1,2}, H C Ye\textsuperscript{1,2}, Q X Chen\textsuperscript{2}, Z C Li\textsuperscript{3}, H Y Jiang\textsuperscript{3}

\textsuperscript{1}Key Laboratory of Earth Observation Hainan Province, Hainan Institute of Aerospace Information Research Institute, Chinese Academy of Sciences, Sanya 572029, China
\textsuperscript{2}Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China
\textsuperscript{3}Sanya Institute of Remote Sensing, Sanya 572029, China

Abstract: Plastic greenhouses are vital agricultural facilities to protect cash crops from disease and insects, especially in the Hainan region of China, which has high temperature and high humidity. Remote-sensing technology is an efficient means to quickly determine the spatial distribution of plastic greenhouses on the regional scale. With the rapid development of remote-sensing technology, and especially the increasing types of high-spatial-resolution remote-sensing imagery, many studies have obtained good results by using remote-sensing technology to monitor plastic greenhouses. However, the best spatial resolution of images for monitoring plastic greenhouses has yet to be studied. To address this issue, we use cantaloupe greenhouses as the research object and GF-2 images with 1m spatial resolution as data source. We then use the re-sampling method to generate images from these data with spatial resolutions of 0.5, 2, 3, and 5 m. The details of the spatial distribution (texture features and shape features) and the spectral features of the plastic greenhouses were then extracted from images of varying spatial resolution, and a remote-sensing monitoring method for cantaloupe greenhouses was constructed based on the object-oriented random forest algorithm, which combines spectral, texture and shape features, and the monitoring results are compared. The results show that the use of 2 m spatial resolution provides the highest monitoring accuracy of cantaloupe greenhouses (overall accuracy = 94.85\% and KIA = 0.92). This study thus provides a theoretical basis for remote-sensing monitoring of greenhouse cantaloupes that satisfies the current demands of production accuracy.

Keywords: plastic greenhouse; cantaloupe; object-oriented method; random forest algorithm
1. Introduction
Plastic greenhouses are important agricultural facilities that fulfill the functions of pest control, rain control, heat preservation, and humidity control. Greenhouses reduce the number of seedlings killed by heavy rain and reduce the direct harm caused by disease and pests, thereby reducing the use of pesticides, especially in high-temperature and -humidity areas. Timely and accurate knowledge of the spatial distribution and dynamic changes of plastic greenhouses is not only necessary to ensure the steady implementation of the “vegetable basket” project but is also the basis for relevant government agencies to formulate agricultural subsidy policies. The traditional method is to collect and report statistics by administrative unit level through field investigations. However, this approach implies a long working cycle and high cost, requires tedious investigation work and is inoperable over large regions. With the development of remote-sensing technology and especially its wide application in agriculture, it is now suitable for monitoring agricultural facilities.

Remote-sensing technology has already been used to monitor plastic greenhouses in numerous studies. Lu et al. [1] constructed a threshold discrimination model based on the normalized difference vegetation index (NDVI) and cumulative days by using long-term MODIS image data and exploiting spectral changes caused by variations in crop growth in plastic greenhouses. Ou et al. [2] used Landsat images from 1990 to 2018 and used spectral reflection and vegetation index to carry out spatial-temporal dynamic monitoring based on the Google Earth Engine (GEE) platform of plastic greenhouses in protected farmland. Zhu et al. [3] introduced the texture features of agricultural greenhouses from multi-temporal Landsat images and extracted the spatial distribution of agricultural greenhouses in the Shandong province in the past 30 years by using the random forest algorithm. The average overall accuracy reached 91.63%, which improved the extraction accuracy of large-scale agricultural greenhouses by remote sensing. Wu et al. [4] used texture features and spectral features of Landsat images to construct an object-oriented plastic greenhouse extraction method based on support vector machine, and the recognition accuracy is significantly higher than that based on spectral features only.

With the development of high-resolution remote-sensing technology, more and more studies now use high-resolution remote-sensing images to extract the spatial distribution of plastic greenhouses and obtain higher recognition accuracy. Commonly used remote-sensing data sources with high spatial resolution include aerial image data [5], low-altitude unmanned aerial vehicle (UAV) image data [6], and satellite image data GF-2 [7-10] (Gao et al., 2018; Gonzalez et al., 2018; Zhao et al., 2019; Wu et al., 2019), Spot-7 [11] and worldview-2 [12]. Sun et al. [6] used UAV remote-sensing images as data source to monitor agricultural facilities based on pixel and object-oriented methods. The results show that the recognition accuracy of the latter is significantly better than that of the former. Wu et al. [10] reported that, when using a single-texture algorithm to identify plastic greenhouses, the recognition accuracy of a local binary model exceeds that of the gray-level co-incidence matrix and the pixel shape index algorithm. Shi et al. [13] proposed a three-step stratified model to extract plastic greenhouses based on GF-2 remote-sensing data; that is, they first distinguished plastic greenhouses, vegetation, and other feature types by using the double-coefficient vegetation screening index. Next, they used the high-density vegetation suppression index to screen out high-density vegetation features. Finally, the NDVI was used to distinguish low-density vegetation from plastic greenhouses.

The above studies all involve remote-sensing extraction of spatial distribution information on plastic agricultural greenhouses. With the increase of remote sensing images with high spatial resolution, many studies found that the identification accuracy is not always high when using images with higher spatial resolution, which may be because excessive spatial resolution increases the unimportant information from ground objects. Based on this, the present study selected plastic cantaloupe greenhouse as research objects and resampled GF-2 images to obtain images of varying spatial resolution, then built an object-oriented random forest algorithm with comprehensive use of the spectrum, texture, and shape features to monitor cantaloupe greenhouses. Finally, we report the best spatial resolution for monitoring cantaloupe greenhouses.
2. Materials and Method

2.1 Study region
The study area is in Ledong Li Autonomous County of Hainan Province, which is known as the “Hometown of Cantaloupe” in China. It has a tropical monsoon climate, abundant light, and warm temperatures. The annual average temperature of the coastal plain is 25.2°C, and the annual average precipitation is 1075 mm (Tan, 1981). To improve the yield and quality of the cantaloupe and reduce the damage caused by disease and insects, the cantaloupe is planted in a steel-frame plastic greenhouse. The shoulder height of the shed is 1.7 m, the top height is 3 m, and the span is 4 m. The distance between adjacent sheds is 40–60 cm. Insect-proof nets span between the sheds, and the top of each shed is covered with plastic film. The consistency of planting patterns of cantaloupe in the study area facilitates remote sensing identification. In this study, the study area is Ledong Li Autonomous County in the coastal area, which hosts densely planted cantaloupe (see Figure 1).

2.2 Data acquisition and processing
This study uses a GF-2 remote-sensing image acquired on January 16, 2020 as data source. GF-2 is the first civil optical remote sensing satellite independently developed by China with submeter spatial resolution. It is equipped with a 1 m panchromatic camera and a 4 m multispectral camera. The panchromatic band spectrum has a range of 450–900 nm and the multispectral bands include a blue band (450–520 nm), green band (520–590 nm), red band (630–690 nm), and a near-infrared band (770–890 nm). To eliminate atmospheric interference and improve the spatial resolution of the image, the GF-2 image was preprocessed by radiometric calibration, atmospheric correction, image registration, fusion, and other techniques to obtain a spatial resolution of 1 m. Images with spatial resolutions of 0.5, 2, 3, and 5 m were obtained by resampling and provided the data source for the follow-up study of optimal spatial resolution.

The field survey of cantaloupe greenhouses in Ledong Li Autonomous County was undertaken on January 14, 2020 and yielded a total of 30 cantaloupe greenhouse samples. To satisfy the classification requirements, we also comprehensively considered the Google historical high-spatial-resolution images of November 2019 and February 2020, and the visual interpretation of the study area was carried out to obtain a sample set containing the “true ground values” to verify classification and accuracy. Finally, six object categories containing a total of 1277 samples were selected including early cantaloupe greenhouses (period of low vegetation coverage), middle-late cantaloupe greenhouses (period of high vegetation coverage), vegetation, water, soil, and other (land for construction, road, beaches, and other) (Table 1). In this study, about two-thirds of the total sample was selected as the training samples to build the
classification model, and the remaining one-third of the total samples was used as the validation samples to validate the accuracy of classification model.

| Type                      | Number of training samples | Number of validated samples | Total |
|---------------------------|----------------------------|----------------------------|-------|
| Early cantaloupe greenhouses | 66                         | 33                         | 99    |
| Middle-late cantaloupe greenhouses | 61                         | 30                         | 91    |
| Vegetation                | 255                        | 128                        | 383   |
| Water                     | 60                         | 30                         | 90    |
| Bare soil                 | 249                        | 124                        | 373   |
| Other                     | 161                        | 80                         | 241   |
| **Total**                 | **852**                    | **425**                    | **1277** |

2.3. Segmentation algorithm

Image segmentation is a key step in the object-oriented method. The accurate segmentation of different ground objects directly affects the recognition accuracy. The multiresolution segmentation algorithm proposed by [15] is a bottom-up segmentation method for merging two regions. This method takes a single pixel as the starting point and iteratively merges pixels into larger units. It is an optimized process that minimizes the average heterogeneity of a given number of image objects to obtain maximum homogeneity. In the process of segmentation, the spectral, spatial and neighbourhood information of remote-sensing images are considered comprehensively, which has been widely used in remote-sensing image segmentation. Segmentation scale, shape index and compactness index are the main parameters that affect the segmentation accuracy of the multiresolution segmentation algorithm. The specific segmentation process requires many attempts to set different factor parameters before obtaining the best segmentation results.

2.4. Feature selection

The spectral reflection from plastic greenhouses is influenced both by the plastic film and by the growth state of the crops within the greenhouses. In the early growth stage of the crops, the spectral reflection from plastic greenhouses is greatly affected by soil and is like that of bare soil. With increasing crop coverage, the spectral reflectance of plastic greenhouses increases gradually under the influence of vegetation, and its reflection spectrum becomes like that of vegetation. Therefore, it is difficult to carry out remote-sensing recognition of plastic greenhouses based only on spectral characteristics. Because the spatial distribution of plastic greenhouses is regular and their size and shape are consistent, the introduction of texture and shape features in the remote-sensing recognition procedure helps to improve the recognition accuracy.

This study uses spectral features, including the brightness, the normalized difference vegetation index (NDVI) [16] and the normalized plastic greenhouse index (NDPG) [17]. We used eight texture features extracted by using the gray-level co-occurrence matrix [18], including homogeneity, contrast, dissimilarity, entropy, angular second moment, mean, variance and correlation. Shape features include asymmetry, border index, compactness, density, elliptic fit, radius of largest enclosing ellipse, radius of smallest enclosing ellipse, rectangular fit, roundness and shape index.

\[
NDVI = \frac{(NIR - RED)}{(NIR + RED)} \tag{1}
\]

\[
NDPG = \frac{(RED - GREEN)}{(RED + GREEN)} \tag{2}
\]
Table 2 The spectral features used in this study

| Type              | Feature               | Type              | Feature                          |
|-------------------|-----------------------|-------------------|----------------------------------|
| Spectral features | Brightness            | Shape features    | Asymmetry                        |
|                   | NDVI                  |                   | Border index                     |
|                   | NDPG                  |                   | Compactness                      |
| Texture features  | Homogeneity           |                   | Density                          |
|                   | Contrast              |                   | Elliptic fit                     |
|                   | Dissimilarity         |                   | Radius of largest enclosing ellipse |
|                   | Entropy               |                   | Radius of smallest enclosing ellipse |
|                   | Angular second moment |                   | Rectangular fit                  |
|                   | Mean                  |                   | Roundness                        |
|                   | Variance              |                   | Shape index                      |
|                   | Correlation           |                   |                                  |

2.5. Building the classification model

This study uses the random forest (RF) algorithm to construct a classification model for plastic greenhouses. RF was proposed by [19] and is based on the decision tree algorithm, and its integrated classification was done by combining multiple classification and regression decision trees. The RF algorithm is robust against noise and does not easily over-fit. It is also good at dealing with outliers and collinear variables and has been widely used to classify ground objects. In terms of accuracy of remote-sensing extraction of plastic greenhouses, the RF algorithm is better than other machine learning algorithms (e.g., maximum likelihood method, support vector machine, K-proximity method, decision tree) [2] [7] [11]. For these reasons, the RF algorithm is used herein to monitor cantaloupe greenhouses.

The RF algorithm requires few parameters; the key parameters are Ntree and mtry. Ntree is the number of decision trees constituting the RF, and mtry is the number of features randomly selected during the training of each decision tree. Based on results of the relevant experiments, Ntree is taken as 50 herein, and mtry is the square root of the total number of features.

2.6. Accuracy evaluation

The confusion matrix method is commonly used to verify the classification accuracy of remote-sensing images and is used herein to evaluate the classification accuracy. The confusion matrix is a comparison matrix that represents the number of samples in a certain category and the number of real samples in the category. The evaluation metrics used herein include overall accuracy (OA), producer’s accuracy (PA), user’s accuracy (UA) and the kappa coefficient (KIA).

3. Results

3.1. Results of image segmentation

The results of segmenting (be it over-segmentation or under-segmentation) remote-sensing images directly affect the classification accuracy. This study uses the multiresolution segmentation algorithm. After several tests to ensure to the extent possible the purity of greenhouse objects, the optimal segmentation parameters were finally determined for remote-sensing images with different spatial resolutions (see Table 3). Higher spatial resolution corresponds to a larger optimal segmentation scale, shape index and compactness index, mainly
because the image with high spatial resolution contains more detailed information. To avoid disturbance from unimportant information detail, a higher segmentation parameter should be used in this study to prevent over-segmentation.

| Table 3. Optimal segmentation parameters for images of different spatial resolution. |
|----------------------------------|-----------------|------------|------------|
| Spatial resolution | Segmentation scale | Shape | Compactness |
| 0.5 m | 130 | 0.9 | 0.9 |
| 1 m | 90 | 0.9 | 0.9 |
| 2 m | 80 | 0.8 | 0.9 |
| 3 m | 70 | 0.6 | 0.8 |
| 5 m | 70 | 0.6 | 0.8 |

3.2. Classification results

This study uses the object-oriented RF classification algorithm to identify plastic cantaloupe greenhouses from images of spatial resolution of 0.5, 1, 2, 3, and 5 m. The results appear in Figures 2-6. Given the reduction in spatial resolution and the influence of mixed pixels, some small objects in the study area are not identified, such as the road between the plastic greenhouses. The boundary of identified objects is also rough. However, images with higher spatial resolution correspond to more fragmented ground object, especially for areas with bare soil and/or other ground objects.

![Figure 2](image-url)  
(a) Classification results for entire region.  
(b) Classification results for key region.

**Figure 2.** Classification results for remote-sensing image with 0.5 m spatial resolution.

![Figure 3](image-url)  
(a) Classification results for entire region.  
(b) Classification results for key region.

**Figure 3.** Classification results for remote-sensing image with 1 m spatial resolution.
The validation sample data were used to evaluate the recognition accuracy of the remote-sensing images with different spatial resolutions (see results in Table 4). With decreasing spatial resolution of the images, the classification accuracy first increases and then decreases. The classification accuracy is highest (overall accuracy = 94.85%, KIA = 0.92) for a spatial resolution of 2 m, followed by a spatial resolution of 1 m (overall accuracy = 93.88%, KIA = 0.92). The images with 5 m spatial resolution produce the worst accuracy (overall accuracy = 87.57%, KIA = 0.81). When the spatial resolution of an image is too high, the classification accuracy cannot be improved, possibly because images with high spatial resolution contain too much detailed information on ground objects, which interferes with ground-object recognition.

For early cantaloupe greenhouses and middle-late cantaloupe greenhouses, the recognition accuracy for images with 1 m spatial resolution slightly exceeds that for images with 2 m spatial resolution, especially for early-cantaloupe greenhouses. During the early growth stage of cantaloupe, the low coverage means that the spectral characteristics of greenhouses are strongly affected by soil, which is like bare soil, increasing the difficulty of identifying plastic greenhouses. However, images with 1 m spatial resolution have more texture and shape features than images with 2 m spatial resolution, which improves the recognition accuracy.
Table 4. Classification results based on images of varying spatial resolution.

| Spatial resolution of images | Type                | Producer’s accuracy | User’s accuracy | Overall accuracy | KIA  |
|-----------------------------|---------------------|---------------------|-----------------|------------------|------|
| 0.5 m                       | Early cantaloupe    | 96.27%              | 90.99%          | 92.49%           | 0.90 |
|                             | Mid-later cantaloupe| 70.41%              | 100.00%         |                  |      |
| 1 m                         | Early cantaloupe    | 100.00%             | 92.06%          | 93.88%           | 0.92 |
|                             | Mid-later cantaloupe| 82.61%              | 97.05%          |                  |      |
| 2 m                         | Early cantaloupe    | 97.83%              | 85.91%          | 94.85%           | 0.92 |
|                             | Mid-later cantaloupe| 84.25%              | 97.00%          |                  |      |
| 3 m                         | Early cantaloupe    | 97.19%              | 87.04%          | 92.96%           | 0.90 |
|                             | Mid-later cantaloupe| 73.32%              | 100.00%         |                  |      |
| 5 m                         | Early cantaloupe    | 86.02%              | 78.10%          | 87.57%           | 0.81 |
|                             | Mid-later cantaloupe| 63.81%              | 71.02%          |                  |      |

4. Discussion

Remote-sensing images with high spatial resolution have gradually become the mainstream data source for mapping greenhouses. Significant research has been done to monitor greenhouses by using high-spatial-resolution images, such as GF-2 with 1 m spatial resolution, QuickBird with 2.5 m resolution, IKONOS with 4 m resolution and GeoEye-1 with 0.5 m resolution. All of these have obtained a good identification accuracy with an overall accuracy of 84.20%–97.34% [13] [20-23]. Note that remote-sensing images of higher resolution cost more than their low-resolution counterparts, and images with higher spatial resolution may include redundant detail that could hinder the effort to map greenhouses. The present work studies this question by considering images with different spatial resolutions. The results indicate that higher spatial resolution is not always for mapping greenhouses. For mapping cantaloupe greenhouses, the most accurate classification was obtained with a spatial resolution of 2 m (overall accuracy = 94.85%, KIA = 0.92). This result can provide a theoretical basis for the selection of remote sensing image to estimate greenhouse. However, images with different spatial resolutions acquired by resampling method are different from real images, more tests should be to carried out in future.

5. Conclusion

This study is based on GF-2 images constructed by using an object-oriented random forest classification method with the spectral, texture and shape features of remote-sensing images with varying spatial resolution. The images are used to monitor cantaloupe greenhouses to determine the best spatial resolution suitable for such a task. The results show that, with spatial resolution decreasing from 0.5 to 5 m, the classification accuracy first increases and then decreases. The best classification accuracy is obtained for 2 m spatial resolution (overall accuracy = 94.85%, KIA = 0.92). However, for greenhouses with crops in the early growth stage (i.e., low crop coverage), a better classification accuracy of plastic greenhouses is obtained with 1 m spatial resolution.

References

[1] Lu L, Huang D, Di L. 2015 Threshold model for detecting transparent plastic-mulched landcover using moderate-resolution imaging spectroradiometer time series data: a case study in southern Xinjiang, China J APPL REMOTE SENS. 9 p 097094
[2] Ou C, Yang J, Du Z, et al. 2020 Long-Term Mapping of a Greenhouse in a Typical
Protected Agricultural Region Using Landsat Imagery and the Google Earth Engine

Remote Sensing 12 p 55

[3] Zhu D, Liu Y, Feng Q, et al. 2020 Spatial-temporal Dynamic Changes of Agricultural Greenhouses in Shandong Province in Recent 30 Years Based on Google Earth Engine Transactions of the Chinese Society for Agricultural Machinery. 051 p 168-175.

[4] Wu C F, Deng J S, Wang K, et al. 2016 Object-based classification approach for greenhouse mapping using Landsat-8 imagery International Journal of Agricultural & Biological Engineering. 9 p 79-88.

[5] Tarantino E and Figorito B. 2012 Mapping Rural Areas with Widespread Plastic Covered Vineyards Using True Color Aerial Data Remote Sensing 4 p 1913-1928.

[6] Sun Y, Han J, Chen Z, et al. 2018 Monitoring Method for UAV Image of Greenhouse and Plastic-mulched Landcover Based on Deep Learning Transactions of the Chinese Society for Agricultural Machinery. 49 p 133-140.

[7] Gao M, Jiang Q, Zhao Y, et al. 2018 Comparison of plastic greenhouse extraction method based on GF-2 remote-sensing imagery Journal of China Agricultural University. (in Chinese) 22 p 125-134.

[8] González-Yebra Óscar, Aguilar M, Abderrahim N, et al. 2018 Methodological proposal to assess plastic greenhouses land cover change from the combination of archival aerial orthomages and Landsat data Biosystems Engineering 175 p 36-51.

[9] Zhao L, Ren H, Yang L. 2019 Retrieval of agriculture greenhouse based on GF-2 remote sensing images Remote sensing technology and application 34 p 677-684.

[10] Wu J, Liu X, Bo Y, et al. 2019 Plastic greenhouse recognition based on GF-2 data and multi-texture features Transactions of the Chinese Society of Agricultural Engineering 35 p 173-183.

[11] Balcić F, Senel G, and Goksel C. 2020 Object-Based Classification of Greenhouses Using Sentinel-2 MSI and SPOT-7 Images: A Case Study from Anamur (Mersin), Turkey IEEE J-STARS. 13 p 2769-2777.

[12] Aguilar M, Nemmouche A, Novelli A, et al. 2016 Object-Based Greenhouse Mapping Using Very High Resolution Satellite Data and Landsat 8 Time Series Remote Sensing 8 p 513.

[13] Shi L, Huang X, Zhong T, et al. 2020 Mapping Plastic Greenhouses Using Spectral Metrics Derived From GaoFen-2 Satellite Data IEEE J-STATS. 13 p 49-59.

[14] Tan W. 1981 Climate and agriculture in Ledong County, Hainan Island Tropical geography 1 p 33-38.

[15] Baätz M, Schäpe A, Strobl J, et al. 2000 Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation Angewandte Geographische Informationen-Verarbeitung XII p 12-23.

[16] Rouse J, Haas R, Schell J, et al. 1974 Monitoring the Vernal Advancement of Retrogradation of Natural Vegetation NASA/GSFC Type II, Final Report. Greenbelt, MD, USA p 1-371.

[17] Cui B, Huang W J, Ye H C, Li Z C and Chen Q X. 2021 Object-oriented greenhouse cantaloupe identification by remote sensing technology 9th International Conference on Agro-Geoinformatics p 1-6.

[18] Haralick R. 1973 Texture features for image classification Studies in Media and Communication 3 p 610-621.

[19] Breiman L. 2001 Random Forests MACH LEARN. 45 p 5-32.

[20] Agüera F, Aguilar F J, and Aguilar M A. 2008 Using texture analysis to improve per-pixel classification of very high resolution images for mapping plastic greenhouses,” ISPRS J. Photogrammetry Remote Sens. 63(6) p 635-646.

[21] Mura M D, Benediktsson J A, Waske B, and Bruzzone L. 2010 Morphological attribute profiles for the analysis of very high resolution images IEEE Trans. Geosci. Remote Sens. 48(10) p 3747-3762.

[22] Agüera F. and Liu J G. 2009 Automatic greenhouse delineation from QuickBird and Ikonos satellite images Computers and Electronics in Agriculture 66(2) 191-200.

[23] Aguilar M A, Bianconi F, Aguilar F J, Fernández I. 2014 Object-based greenhouse classification from GeoEye-1 and WorldView-2 stereo imagery Remote Sens. 6 3554-3582.
Acknowledgments
This work was supported by Hainan Provincial Key R&D Program of China (ZDYF2019084), Hainan Provincial Major Science and Technology Program of China (ZDKJ2019006), National special support program for high-level personnel recruitment (Wenjiang Huang) and Ten-Thousand Talents Program (Wenjiang Huang).