

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/354695290>

Cotton yield estimation model based on machine learning using time series UAV remote sensing data

Article in International Journal of Applied Earth Observation and Geoinformation · December 2021

DOI: 10.1016/j.jag.2021.102511

CITATIONS
99

READS
1,067

6 authors, including:



Weicheng Xu

Guangdong Academy of Agricultural Sciences

21 PUBLICATIONS 376 CITATIONS

[SEE PROFILE](#)



Pengchao Chen

South China Agricultural University

47 PUBLICATIONS 1,242 CITATIONS

[SEE PROFILE](#)



Zhan Yilong

South China Agricultural University

10 PUBLICATIONS 382 CITATIONS

[SEE PROFILE](#)

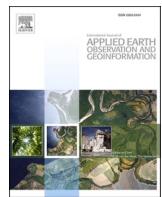
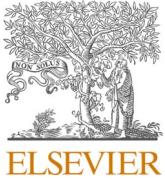


Lei Zhang

South China Agricultural University

771 PUBLICATIONS 28,535 CITATIONS

[SEE PROFILE](#)



Cotton yield estimation model based on machine learning using time series UAV remote sensing data

Weicheng Xu^{a,c}, Pengchao Chen^{a,c}, Yilong Zhan^{a,c}, Shengde Chen^{a,c,*}, Lei Zhang^{b,c},
Yubin Lan^{a,c,d,*}

^a College of Electronic Engineering, South China Agriculture University, Guangzhou 510642, China

^b College of Agricultural, South China Agriculture University, Guangzhou 510642, China

^c National Center for International Collaboration on Precision Agricultural Aviation Pesticide Spraying Technology, Guangzhou 510642, China

^d Department of Biological and Agricultural Engineering, Texas A&M University, College Station, TX 77843, USA

ARTICLE INFO

Keywords:

UAV remote sensing
Multisource data fusion
Neural networks
Cotton yield forecast

ABSTRACT

Crop yield prediction is of great practical significance for farmers to make reasonable decisions, such as decisions on crop insurance, storage demand, cash flow budget, fertilizer, water and other input factors. The traditional yield measurement method is sampling surveys, which require a large area of destructive sampling of cotton fields and consume considerable time and labor costs. This study established a cotton yield estimation model based on time series Unmanned Aerial Vehicle (UAV) remote sensing data. The U-Net semantic segmentation network is used to recognize and extract the boll opening pixels in high-resolution visible images, and the boll opening pixel percentage (BOP) is calculated according to the network extraction results. By combining the multispectral images and the pixel coverage of cotton bolls, a Bayesian regularization BP (back propagation) neural network was used to predict cotton yields. In order to simplify the input parameters of the model, the stepwise sensitivity analysis method is used to eliminate redundant variables and obtain the optimal input feature set. The experimental results show that the R^2 of the proposed model is 0.853 at the scale of 0.81 m^2 (average results of ten-fold cross validation). This study provides a method that can simultaneously meet the requirements of large-area and small-scale forecasting of cotton yields and provides a new idea for cotton yield measurement and breeding screening.

1. Introduction

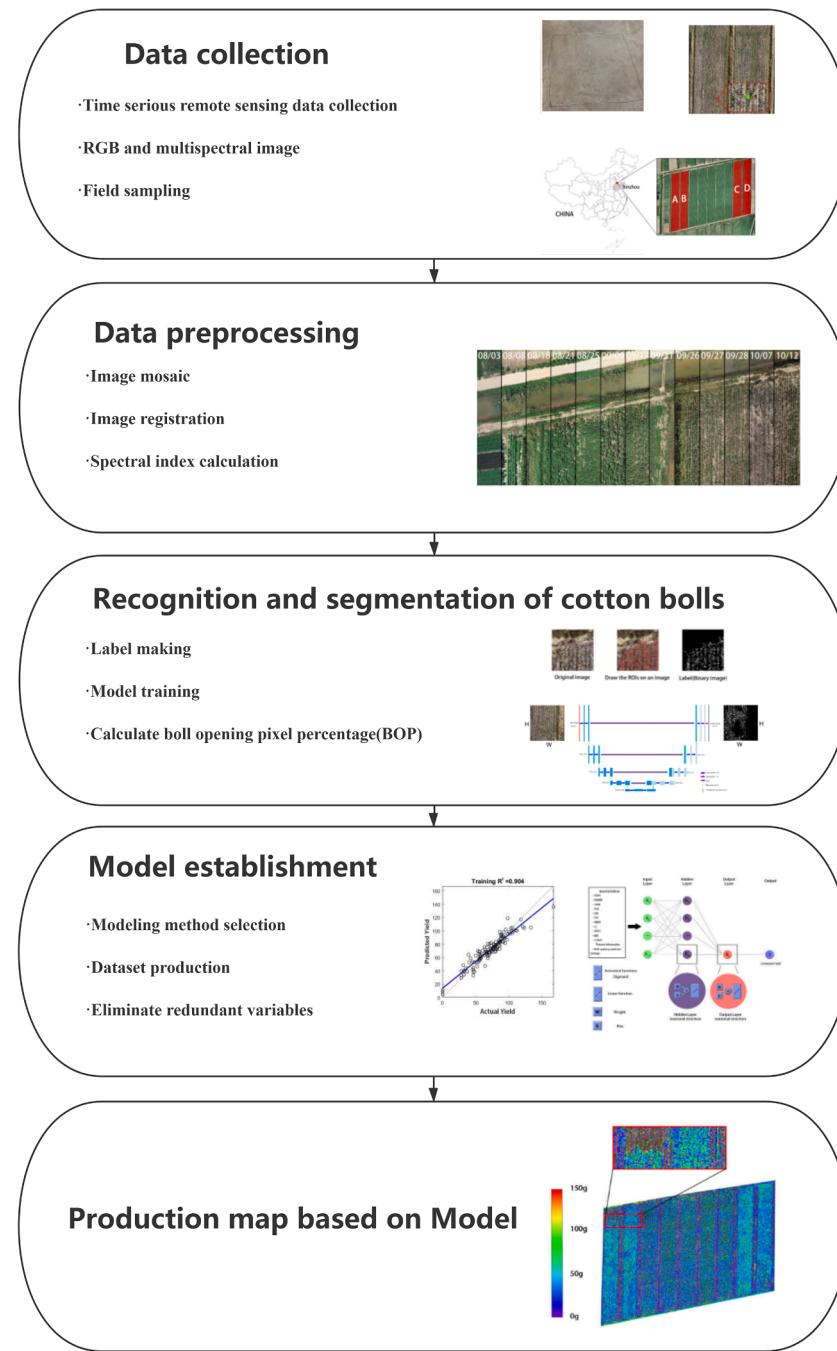
Cotton is one of the most important crops in the world. It is the main raw material of the textile industry and has important uses in national defense, medicine and the automobile industry. In agricultural production, crop yield prediction is of great practical significance for farmers to make reasonable decisions such as decisions on crop insurance, storage demand, cash flow budget, fertilizer and water input (Meina et al. 2019). In addition, yield measurement also plays a very important role in cotton breeding, and the selection of genotypes suitable for specific environments is one of the main objectives of cotton breeding research (Ashapure et al. 2020). The traditional yield measurement method is to evaluate a cotton field as a whole after a cotton harvest. Harvesting the whole cotton field or conducting destructive sampling in a large area requires considerable time and labor costs (Pantazi et al. 2016). In recent years, remote sensing technology has

developed rapidly and has been widely used in agricultural monitoring (Xu et al. 2019; Lan et al. 2010; Shrestha et al. 2017). The accuracy of yield prediction using satellite remote sensing is increasing. Many scholars use satellite data to estimate the yields of rice (Shibayama and Akiyama 1991), wheat (Iobell and Asner 2003), corn (Shrestha et al. 2017) and other crops. Although satellite remote sensing data are very valuable due to their large-scale coverage, spatial resolution is still a concern in many precision agriculture applications. Many prediction models can only provide more accurate crop yield predictions in large-scale work at the county level and above but cannot describe the detailed changes in crop yields on a smaller scale (such as a single farmland). In addition, because satellites may be blocked by clouds, their revisit time is not flexible, making it impossible to obtain timely information about crops throughout the growth cycle (Xiang and Tian 2011).

Thanks to the remarkable progress of Unmanned Aerial Vehicle

* Corresponding authors at: College of Electronic Engineering, South China Agriculture University, Guangzhou 510642, China (S. Chen and Y. Lan).

E-mail addresses: shengde-chen@scau.edu.cn (S. Chen), ylan@scau.edu.cn (Y. Lan).

**Fig. 1.** Flow chart.

(UAV) platforms and lightweight sensor development technology, UAV remote sensing is increasingly more widely used in agricultural research (Zhang and Yemoto 2019). Using drones to fly at low altitudes to obtain high-resolution remote sensing images has become an ideal choice for agricultural applications. In recent years, UAV remote sensing has achieved high accuracy in crop biomass detection (Kanning et al. 2018), quality prediction (Yu et al. 2019; Xu et al. 2020), crop pest detection (Huang et al. 2018), field weed detection (Huang et al. 2018; Huang et al. 2020), leaf area index detection (Li et al. 2016), etc. Similar to the traditional remote sensing yield estimation method, calculating the vegetation index (VI) using UAV images is a commonly used crop yield prediction method, but the accuracy obtained by using single parameter modeling is not high. Huang (Huang et al. 2013) used NDVI to establish a cotton yield estimation model, and the R^2 was 0.47. Adding more

factors to the prediction model can improve the accuracy of the model. Zhang (Zhang Meina et al. 2019) and others installed a high-resolution digital camera on a UAV to collect RGB images; obtain the color characteristics, vegetation coverage and other parameters; and obtain higher yield prediction accuracy ($R^2 = 0.97$). In addition, using UAVs equipped with lidar to collect crop canopy surface information and height information can generate digital surface models (DSMs), which are of great help in biomass detection (Wang et al. 2019; Liu et al. 2018). The structure from motion algorithm (SFM) can be used for 3D reconstruction based on various unordered images collected. With the help of SFM, we can generate a digital surface model from the orthophoto images taken by UAVs without lidar (Lucieer et al. 2014). The fusion of time series remote sensing images to obtain the growth of crops in different time periods, which will help us quickly discover the changes in the time

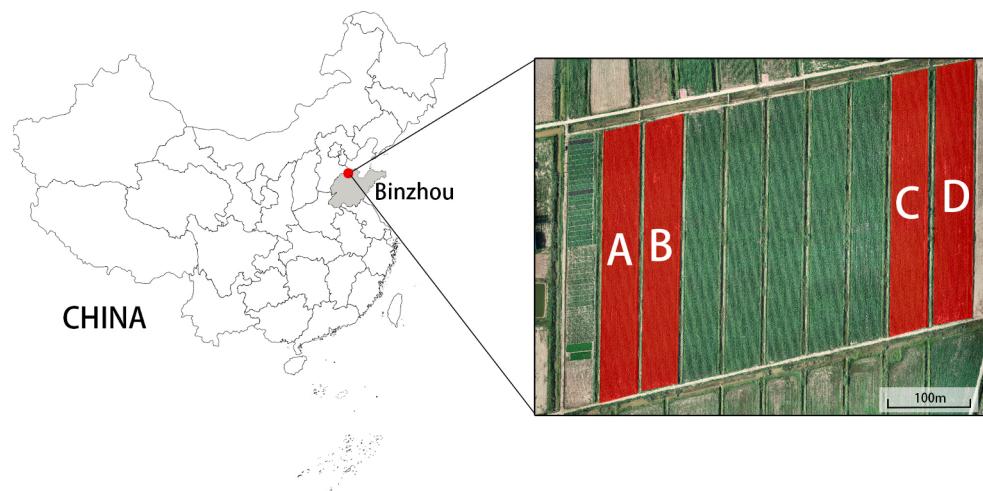


Fig. 2. Test site location. The red areas are the test fields. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

dimension of the research object. This method plays a very important role when studying the changes of ground objects in the time dimension (such as farmland classes, vegetation cover, crop growth and land desertification). (Wang et al. 2020; Huang et al. 2017). In yield prediction research, the yield measurement model based on time series data has higher accuracy and precision than the model based on single period data (Lambert et al. 2018).

Using machine vision to extract cotton bolls from high-resolution remote sensing images is very helpful to the research on the boll opening rate, defoliant effect (Yi et al. 2019) and intelligent agricultural machinery (Li et al. 2015). The remote sensing images captured by UAVs have a spatial resolution that traditional satellites cannot achieve, which makes it possible to segment cotton bolls from remote sensing images. In the past decade, some scholars have used machine vision technology to extract cotton bolls. Zhang collected near-infrared remote sensing images of cotton fields and used an object-oriented method to segment cotton bolls and background (Zhang et al. 2011). Li studied a field cotton boll segmentation algorithm based on random forest classification (Li et al. 2016). Over the last decade, a typical and state-of-the-art machine learning (ML) framework named deep learning (DL), which is developed from the traditional neural network (NN), has outperformed traditional models with considerable improvement in performance, substantial progress in developing a DL methodology for a variety of remote sensing has been observed (Yuan et al. 2020). At the same time, the rapid improvement of computer computing power has led to the rapid development of deep learning technology. Convolutional neural networks are currently the most studied and fastest-growing branch in the field of deep learning, and they have repeatedly achieved amazing results. Semantic segmentation is one of the methods used to analyze image scenes. It can complete image pixel classification and image object segmentation at the same time by labeling each pixel with semantic labels (Ma et al. 2019). The ENVI deep learning module used in this study is a tool based on the U-Net network model. It uses TensorFlowTM to train the deep learning model to recognize the features in an image according to the spatial and spectral characteristics of the image. U-net is more precise and efficient than the Fully Convolutional Neural Network (FCN), which is a traditional semantic segmentation model. U-Net is widely used in remote sensing image classification. The ENVI deep learning module used in this research is a tool developed based on the U-Net model. It uses TensorFlowTM to train the deep learning model and recognize the features in an image according to the spatial and spectral characteristics of the image. U-Net is more refined and efficient than the traditional semantic segmentation model of the FCN, and it has a wide range of applications in remote sensing image classification (Wiratama

et al. 2020; Ren et al. 2020; Xu et al. 2020).

The main purpose of this research is to establish a cotton production prediction model based on machine learning. The network model can implicitly detect the complex nonlinear relationships between the independent variables and the dependent variable in a complex system without the need for an explicit mathematical expression (Sargent. 2001). The time series of UAV multi-spectral and high-resolution visible remote sensing data are fused, combined with the yield data obtained from field sampling as the data set of this study. Comparing the performance of multiple linear regression and neural network fitting in this data set, it is found that the performance of neural network fitting is better. After all characteristic variables are sent into the neural network, delete the variables individually one by one for stepwise sensitivity analysis, find the characteristics that have a significant impact on the yield change, eliminate the redundant variables and simplify the model parameters. Fig. 1 is the flow chart of the study. The main contributions of this research are listed as follows:

- Use U-Net to establish a cotton boll segmentation recognition model based on UAV remote sensing data during the spitting period and further excavate the data information contained in high-resolution UAV images
- Remove the redundancy from the input feature set of various canopy attributes obtained by a UAV to obtain the optimal feature set
- Generate high spatial resolution field yield prediction maps, which provide the ability to automate the genotype selection process in cotton breeding experiments

2. Materials and methods

2.1. Study region

The study area of this experiment is located in Changbei village, Liubao town, Wudi County, Binzhou city, Shandong Province ($37^{\circ}56'53''N$, $117^{\circ}50'14''E$), which belongs to the Yellow River Basin cotton area. The total area of the experimental field is approximately 92 mu. In this experiment, four regions labeled A, B, C, and D were selected for the experiment, as shown in Fig. 2. The planting variety in the experimental area was Lumian 532, and the planting density was 7000 plants per mu.

2.2. Cotton boll sampling

A total of 150 sampling areas were established in the four fields of



a. A wire frame used to define a boundary

b. Sample data collection



c. Cotton ginning

d. Sample area example

Fig. 3. Sample collection and processing.

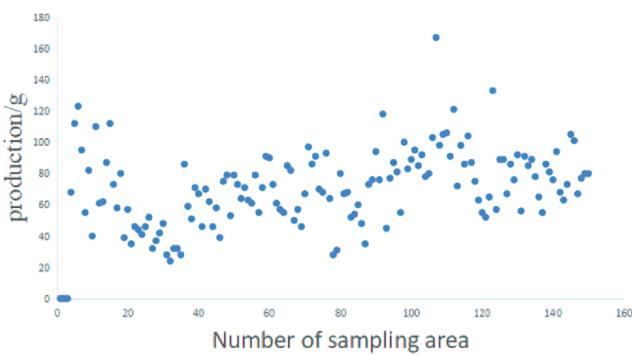


Fig. 4. Regional production chart.

A, B, C and D. Each survey point was a 90 cm × 90 cm square area. A wire frame was used to determine the area boundary (Fig. 3a), and a white cloth was used as a mark at the center of the area. In order to accurately find the sampling area in the remote sensing image, the sampling area is outlined as shown in Figure 2.3. The collection of the samples was completed manually on October 12, 2020 (Fig. 3.b), at which time the cotton had completely split. All samples were collected and ginned using a small sawtooth cotton gin (Fig. 3.c) and then dried in a ventilated and dry environment for 5 days. After ginning, all cotton bolls in the same sampling area were collected to obtain lint as a sample. Statistics were calculated after weighing each sample. Fig. 4 shows the yield of each sampling area.

2.3. Remote sensing data collection

The data collection uses two types of drones to capture visible and multispectral remote sensing data, as shown in Fig. 5. A DJI Phantom 4 RTK (Real - time kinematic) captures visible data, and a DJI Phantom 4 Multispectral captures multispectral data. The performance parameters of the two aircraft and the flight parameters when the data were collected are shown in Table 1. Both devices are equipped with centimeter-level navigation and positioning systems and high-performance imaging systems. The DJI Phantom 4 RTK integrates an RTK module, has a strong anti-magnetic interference capability and precise positioning capability, supports the connection of D-RTK 2 high-precision GNSS mobile stations, and can connect to NTRIP via a wireless network card or WiFi hotspot. The aircraft continuously records data such as the original satellite observations and camera exposure files. After the operation is completed, users can directly calculate high-precision position information through the DJI Cloud PPK (post-process kinematic) service. The mechanical shutter supports high-speed flight shooting, which can effectively avoid the reduction of the drawing accuracy caused by the jelly effect. In a sunny environment, when the wind speed is less than 4 m/s, the flying height is 100 m, the ground sampling distance (GSD) is 2.74 cm, the heading overlap rate is 80%, and the side overlap rate is 70%, the root mean squared (RMS) are calculated as follows.

$$P_1 = 1.5 + 1\text{ppm}(\text{RMS}) \quad (1)$$

$$P_2 = 1.5 + 1\text{ppm}(\text{RMS}) \quad (2)$$

Formula (1) and formula (2) respectively represent the horizontal



a. DJI Phantom 4 RTK

b. DJI Phantom 4 Multispectral

Fig. 5. Collect data with DJI Phantom 4 RTK and Phantom 4 Multispectral.

Table 1
Parameters of remote sensing data acquisition equipment.

	Phantom 4 RTK	Phantom 4 Multispectral
Resolution of CMOS sensor	20 million	2 million
Field of View (FOV)	84°	62.7°
Flight altitude	100 m	100 m
Ground sampling distance (GSD)	2.74 cm/px	5.29 cm/px
Imaging band	Visible	R (650 nm), G (560 nm), B (450 nm), RE (730 nm) and NIR (840 nm)

R-Red, G-Green, B-Blue, RE-Red Edge, and NIR-Near Infrared.

positioning accuracy and vertical positioning accuracy of a UAV during flight. ppm means that the error increases by 1 mm for every 1 km of aircraft movement.

In order to improve the shooting efficiency, the UAV will not hover when taking photos. At this time, a sufficiently high shutter speed is the key to whether an image is clear or not. During the flight, the A mode (aperture priority) was used to shoot the entire flight. We set the aperture f to the maximum value to ensure sufficient light input. In addition, we improved the shutter speed as much as possible to ensure that the images were clear enough. The three-axis stable pan tilt equipped with the camera ensured the image quality. The influence of the pitch angle change on the lens orientation when the four-axis aircraft is moving forward is also addressed. After all the data are collected, they are uploaded to the graphics workstation. Pix4dMapper is used to splice and

calculate the reflectivity to generate the orthophoto panorama and digital surface model.

The data collection started on August 8th and ended on October 12th, covering the bud period to the spitting period. A total of 13 periods of data were collected. In order to prevent cloud movement from causing inconsistencies in the exposure of each photo and thus affecting the stitched images, data collection is conducted when the sky is clear and cloudless.

2.4. Time series data fusion

There will inevitably be certain errors between the remote sensing images collected by different sensors in different periods. If two images are to be displayed in the same coordinate system, image registration is first required, which is the accuracy assurance of most remote sensing image processing and analysis (such as image mosaic, image fusion and time series analysis) (Feng et al.). Image registration is the process of geometrically aligning images with overlapping areas captured by the same or different sensors from different perspectives or at different times in the same scene (Yue et al. 2020). The calculation model used is a polynomial model, which is a common method for the geometric correction of remote sensing images. It is unrelated to the specific type of sensor and does not need to know the specific position and attitude angle change of the sensor during imaging. The polynomial equation can be used to describe any curved surface, and the coordinate conversion relationship between the corresponding pixels before and after the remote sensing image correction must also be described by a polynomial, as shown in formula (3).

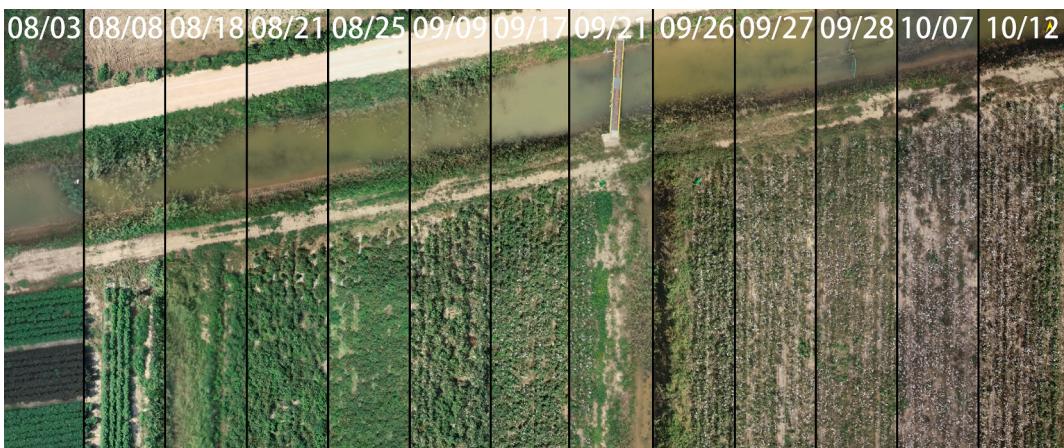


Fig. 6. Schematic diagram of the image registration effect.

Table 2
The indices used in this study.

Indices used	Formulas
Indices calculated from RGB images	
Visible-band Difference Vegetation Index (VDVI) (Wang et al. 2015)	$\frac{2G - R - B}{2G + R + B}$
Normalized Green-Red Difference Index (NGRDI) (Barrero and Perdomo 2018)	$\frac{G - R}{G + R}$
Visible Atmospherically Resistant Index (VARI) (Gitelson et al. 2002)	$\frac{G - R}{G + R - B}$
Excess Green Index (ExG) (Kim et al. 2018)	$\frac{2G - R - B}{2G + R - B}$
Indices calculated from multispectral images	
Difference Spectral Index (DSI) (Feng et al. 2020)	$NIR - R$
Ratio Spectral Index (RSI) (Zhang et al. 2019)	$\frac{NIR}{R}$
Normalized Difference Vegetation Index (NDVI) (Ni et al. 2019)	$\frac{NIR - R}{NIR + R}$
Red Edge Chlorophyll Index (CI) (Gitelson et al. 2003)	$\frac{NIR - 1}{NIR - RE}$
MERIS Terrestrial Chlorophyll Index (MTCI) (Dash and Curran 2004)	$\frac{RE + R}{RE + R}$
Enhanced Vegetation Index (EVI) (Misra et al. 2020)	$2.5(NIR - R)$
Optimized Soil-Adjusted Vegetation Index (OSAVI) (Zou and Mottus 2017)	$\frac{NIR + R + 1}{NIR + R + 0.16}$

$$\begin{cases} x = \sum_{i=0}^N \sum_{j=0}^{N-i} a_{ij} u^i v^j \\ y = \sum_{i=0}^N \sum_{j=0}^{N-i} b_{ij} u^i v^j \end{cases} \quad (3)$$

In the formula, a_{ij} and b_{ij} are polynomial coefficients, N is the number of polynomial coefficients, and the relationship between N and polynomial order n is shown in Formula (4).

$$N = \frac{(n+1)(n+2)}{2} \quad (4)$$

From Formula (4), the number of coefficients of the first-, second- and third-order two-dimensional polynomial models are 3, 6 and 10, respectively. The order of the polynomial model is generally not greater than the third order. Moreover, the higher-order polynomial model not only needs to calculate the increase of the number of coefficients, which would make the calculation amount increase rapidly, increase the correlation of model parameters and reduce the model accuracy. In addition, there is no absolute relationship between the geometric correction accuracy and the polynomial order.

We used the second-order polynomial to find the same control points with obvious features easily marked in 2 images in the experiment, such as the corners of walls, the corners of houses, telegraph poles, etc. When the number of control points reached 6, the polynomial coefficient could be calculated. When the number of control points was greater than 6, the

polynomial coefficient could be calculated using the least squares method. The root mean squared error (RMSE) can be calculated by using Formula (5). As the number of control points increased, the error gradually decreased. When the error was less than 1 px, the image registration process was completed. Due to the high precision of RTK, only a few control points were sufficient to achieve the effect.

$$RMSE_{error} = \sqrt{(v - x)^2 + (u - y)^2} \quad (5)$$

Fig. 6 shows the registration effect. The left image is at the bud stage, and the right image is at the boll-opening stage. Basically, the ground features can completely overlap.

2.5. Calculation of vegetation index

The spectral characteristics of crops are the changes in the absorption, transmission and reflection of light caused by physiological characteristics; and the physiological characteristics of crops reflect their growth status. Therefore, the spectral index has a certain correlation with crop yields. Some vegetation indices that have been widely used in yield prediction are selected in this study, as shown in Table 2.

2.6. Recognition and segmentation of cotton bolls

2.6.1. Recognition model framework

The ENVI deep learning module used in this research was developed by Esri based on the TensorFlow™ deep learning framework. The ENVINet-5 model based on the U-Net model and based on an encoder-decoder architecture is used as a tool. Pixel-level semantic segmentation can be realized, and the network structure is shown in Fig. 7. Its greatest advantage is that ENVINet-5 can be combined with the powerful remote sensing image processing function of the ENVI software and can build training samples and neural networks without using code, which greatly facilitates the use of deep learning methods for remote sensing images by researchers in nondeep learning fields.

In the U-Net model, the convolution mode is valid. This mode does not pad an image before convolution, which will cause the size of the image to be reduced after each convolution (Ronneberger et al. 2015). The traditional method is to fill a circle of pixels with a value of 0 around the image before convolution so as to ensure that the size of the image remains unchanged after convolution. However, as the number of convolutions increases, constant zero padding will increase the error of the feature and affect the recognition effect of the final model. Therefore, the author proposed a data enhancement scheme called overlap tile, which mirrors and fills pixels at the edge of an image. As shown in Fig. 8, the pixels on the right and lower sides between the yellow and blue boxes are copied by a mirror copy. We go to the left and upper side to complete the blue box to obtain the context information of the feature

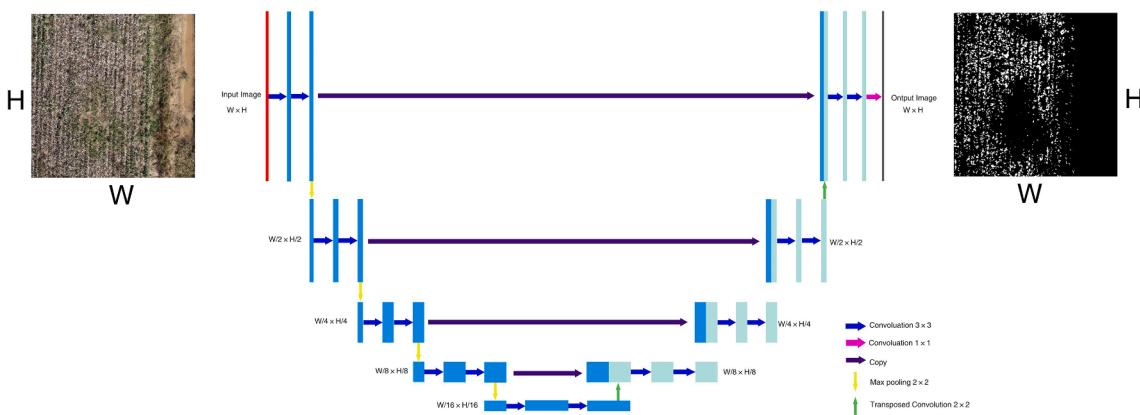


Fig. 7. Structural diagram of the U-Net model.

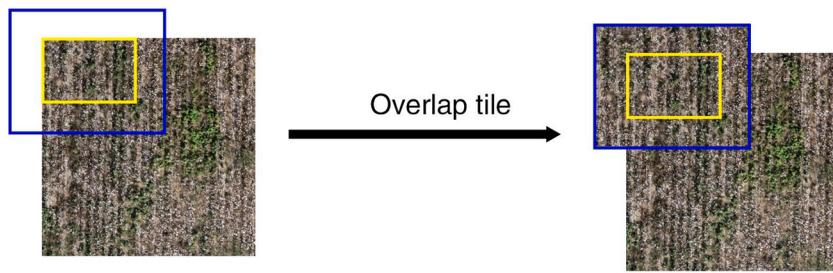


Fig. 8. Diagram of overlap tile.

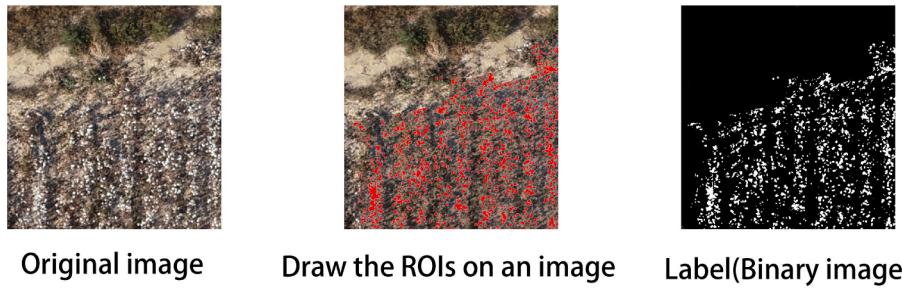


Fig. 9. Label making.

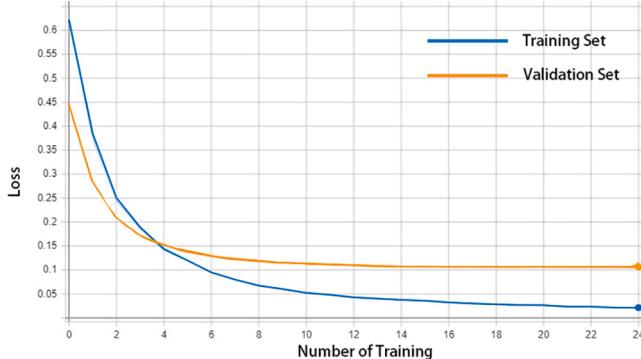


Fig. 10. Loss Function.

and at the same time keep the size of the image before and after the convolution consistent.

2.6.2. Training model

In this study, the region of interest (ROI) tool in ENVI is used to draw the ROI on the image to establish the label, which greatly facilitates the establishment of remote sensing data training set and verification set. Fig. 9 shows the production process of dataset labels, ENVI will automatically generate the label (binarized image) based on the drawn ROI and the original image for model training, each ROI is a sample. This study draws 500 ROI regions, of which 400 are used as training sets and 100 as verification sets.

The loss function is a cross entropy softmax in pixel point form, which is essentially a function evaluating the classification error of the model (Zhang et al. 2017). In the first training process, the model attempts to predict and generate a random class activation grid, which is compared with the mask band of the label grid. By fitting the loss function, the model can determine the position of the random prediction error. The internal parameters or weights of the model can be adjusted to optimize the accuracy and pass the label grid back to the model. Before training, the model extracts a square patch of a given size from the label grid rather than importing the label grid of the entire image

Table 3
Confusion matrix.

Category	Boll opening	Other objects	Total sample size	User accuracy
Boll opening	371	10	381	97.35%
Other objects	14	105	119	88.23%
Total sample size	385	115		
Producer's accuracy	96.36%	91.30%		

into the model. After setting the number of epochs to 24, the loss function of the model in the training set and verification set changes with the amount of training, as shown in Fig. 10.

2.6.3. Validation model

In order to test the extraction effect, we randomly selected 500 checkpoints to build a confusion matrix for Changji, as shown in Table 3. According to Formula (6), the overall accuracy was 92.4%. n is the number of correctly classified samples, and N is the total number of samples. In addition, we obtained a kappa coefficient of 0.8204 via Formula (7). P^o is the overall precision, and p^c is the pixel proportion with the same contingency. In conclusion, the classification results are highly consistent with the actual situation.

$$P^o = \frac{n}{N} \quad (6)$$

$$P_c = \frac{a_1 b_1 + a_2 b_2}{N \times N} \quad (7)$$

$$k = \frac{P_o - P_c}{1 - P_c} \quad (8)$$

2.7. Establishment of prediction model

2.7.1. Bayesian regularized BP neural network

The parameter calculation was completed in Ersi® ENVI 5.5®, and the values calculated for all parameters were the average values in the sampling area (circular area with a radius of 1 m).

A BP neural network usually consists of an input layer, a hidden layer

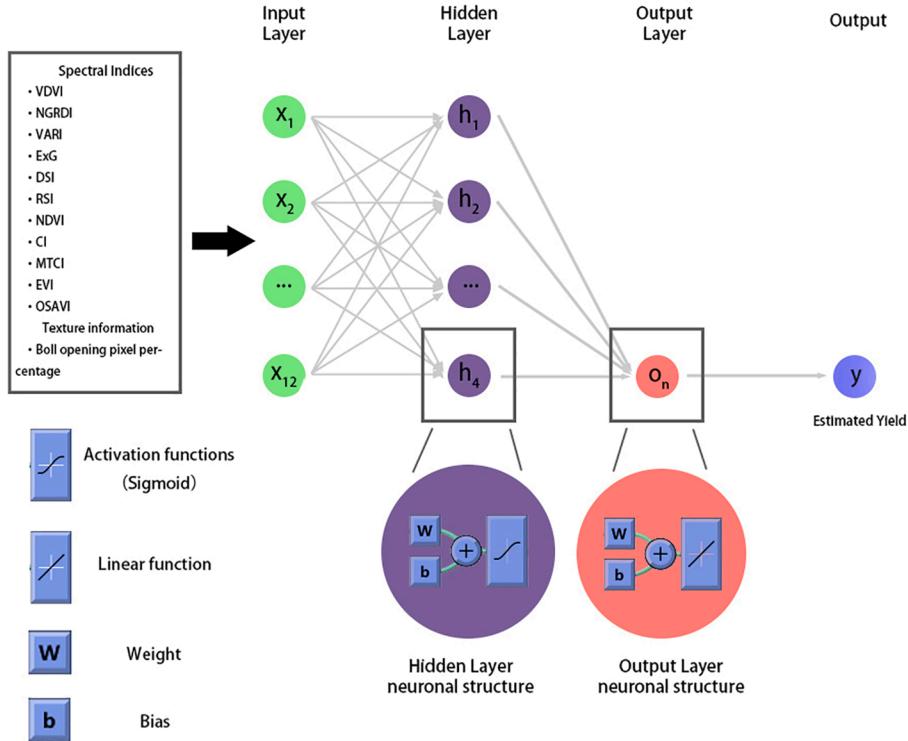


Fig. 11. BP neural network model structure.

and an output layer. The BP neural network realizes the mapping function from input to output and can approximate any nonlinear continuous function with arbitrary precision. The two-layer network (fitnet) of hidden neurons with a sigmoid activation function and linear output neurons, given consistent data and enough hidden neurons, can fit the multidimensional mapping problem well. As the numbers of hidden neurons and layers increase, the performance of the model increased on the training set but decreased on the test set, which indicates that the model is overfit due to overparameterization. The structure of the neural network is shown in Fig. 11. The Bayesian regularization training algorithm is used to train the neural network. The Bayesian regularization training algorithm introduces a correction function to the performance function based on the conventional mean squared error performance function during the neural network training process (Nielsen et al. 2011).

The network training error function is

$$E_D = \sum_{i=1}^n (t_i - t_x)^2 \quad (9)$$

t_i is the actual output, and t_x is the expected output.

The mean squared sum of all network weights is

$$E_W = \frac{1}{m} \sum_{i=1}^m w_i^2 \quad (10)$$

The performance function of the network is

$$F(w) = \alpha E_W + \beta E_D \quad (11)$$

where α and β denote the regularization coefficients that affect the model training. The parameters α and β affect the complexity and smoothness of the network, respectively. If α and β are extremely small, the network will be overfit and underfit, respectively.

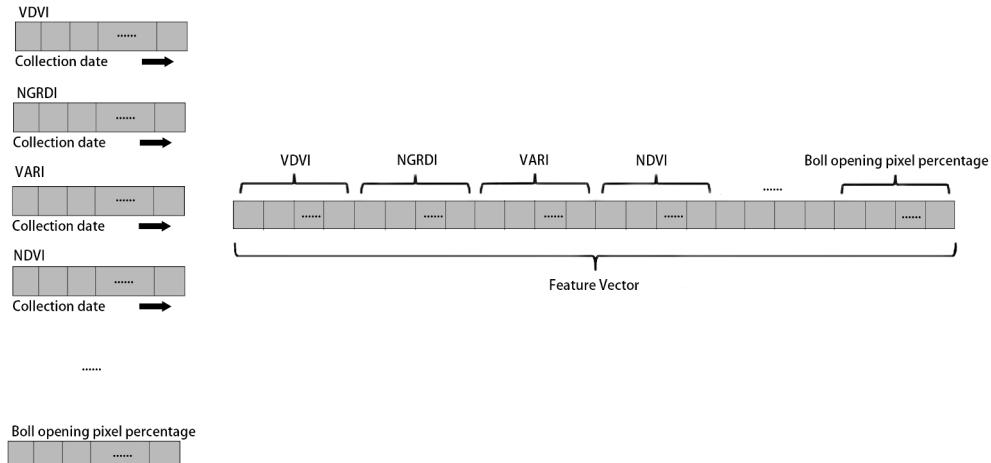


Fig. 12. Forming the feature vector by concatenating all the attributes.

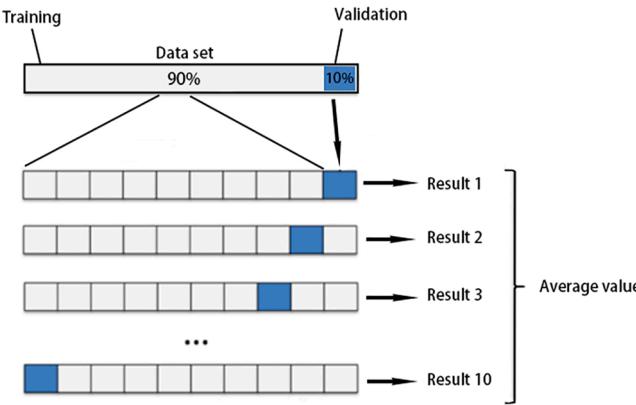


Fig. 13. K-fold cross validation.

In the network training, the Bayesian regularization algorithm adjusts α and β to achieve the best results. The network weight is regarded as a random variable, and the posterior probability density function of the weight after the training set is given by the Bayesian rule.

$$P(w|D, \alpha, \beta, M) = \frac{P(D|w, \beta, M)P(w|\alpha, M)}{P(D|\alpha, \beta, M)} \quad (12)$$

where D is the data of the training set, and M is the neural network model. $P(D|\alpha, \beta, M)$ is a standardized factor to ensure that the overall probability is 1, $P(w|\alpha, M)$ is the prior probability density function of the weight vector, and $P(D|w, \beta, M)$ is the probability density function when the weight is given. Assume that the noise and weight vectors present in the sample data follow a Gaussian distribution.

$$\begin{cases} P(D|w, \beta, M) = \frac{\exp(-\beta E_D)}{Z_n(\beta)} \\ P(w|\alpha, M) = \frac{\exp(-\alpha E_w)}{Z_m(\alpha)} \end{cases} \quad (13)$$

In formula (12), $Z_n(\beta) = (\pi/\beta)^{n/2}$ and $Z_m(\alpha) = (\pi/\alpha)^{m/2}$.

Bring formula (13) into (12). The optimal weight vector should have the maximum posterior probability $P(w|D, \alpha, \beta, M)$, which is equivalent to the minimum regularization performance function $F(w)$. The optimal solutions of α and β at the minimum point w_0 are

$$\begin{cases} \alpha = \frac{\gamma}{2E_w(w_0)} \\ \beta = \frac{m - \gamma}{2E_w(w_0)} \end{cases} \quad (14)$$

where γ is the effective number of network parameters determined by the training set.

2.7.2. Data set construction

In order to prepare the input data set using the time series spectral index and crop canopy attributes, the extracted eigenvalues are connected to form feature vectors (Fig. 12), and the yield of each sample collected on the ground is taken as the target feature. The 150 feature vectors obtained can be directly sent to the neural network for training.

2.8. K-fold cross validation

Cross-validation is performed in the case of insufficient data. Since all the samples were collected and measured manually, the total number of samples was not sufficiently large. Therefore, K-fold cross-validation (Jung et al., 2020) was performed as shown in Fig. 13. K was equal to 10. We selected 90% of the samples for modeling and repeated the process 10 times. The average R^2 was used as the evaluation standard.

3. Results

3.1. Calculation of cotton boll coverage percentage

Through the U-Net model, a class activation grid with the same size as the input image can be obtained. After the activation grid image is binarized, the average value in the sample square can be calculated to obtain the percentage of boll cotton pixels in the sample square, as shown in Fig. 14.

3.2. Use all input variables to build a production forecast model

3.2.1. Linear regression modeling

The least square multiple linear regression method is used to establish the regression model, and the 10 fold cross validation method is used to verify the effect. The results of 10 times are shown in Fig. 15. The abscissa in the figure is the predicted value of the model output and the ordinate is the real value.

It can be seen from Fig. 14 that the production forecast results fluctuate greatly. The maximum value of R^2 is 0.6937, the minimum value is 0.6244, the average is 0.6648, and the average MSE is 212.2910. This shows that the linear regression method has poor generalization ability in this data set, R^2 is generally at a low level, and the prediction accuracy is not high.

3.2.2. BP neural networks modeling

In order to determine the number of neurons in the hidden layer, the initial number of hidden neurons is set as 1. Each time the number of neurons is increased, the model is cross validated, and the average value of R^2 is taken as the standard to evaluate the prediction performance. The results are shown in Fig. 16. The figure shows that when the number of neurons in the hidden layer is 5, the overfitting phenomenon begins to appear, which leads to the R^2 of the model in the training set being 1. When calculating the average, the abnormal data are eliminated and do not participate in the calculation of the average. When the number of hidden layer neurons is more than 5, the overfitting phenomenon is very serious. In ten experiments, the R^2 is 1 more than 5 times, which indicates that the model is in the state of fitting the noise or nonrepresentative features in the data set due to overparameterization. Therefore, the case in which the number of hidden layer neurons is more than 5 is not considered.

K-fold cross validation was used to evaluate the performance of the R^2 and MSE on the training set and validation set (Table 4). The model with the best performance in the 10-fold cross validation will be displayed. Fig. 17 shows the change in the MSE with the training epochs (left), the fitting effect of the model (middle) and the error histogram (right).

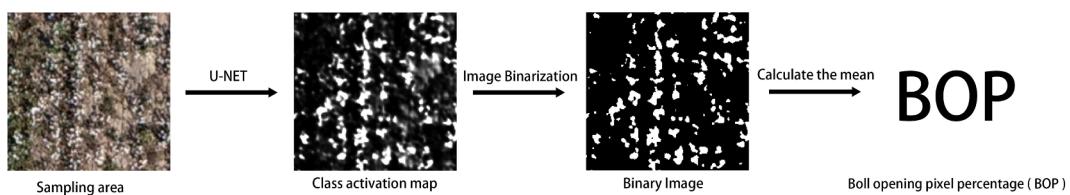


Fig. 14. BOP calculation process.

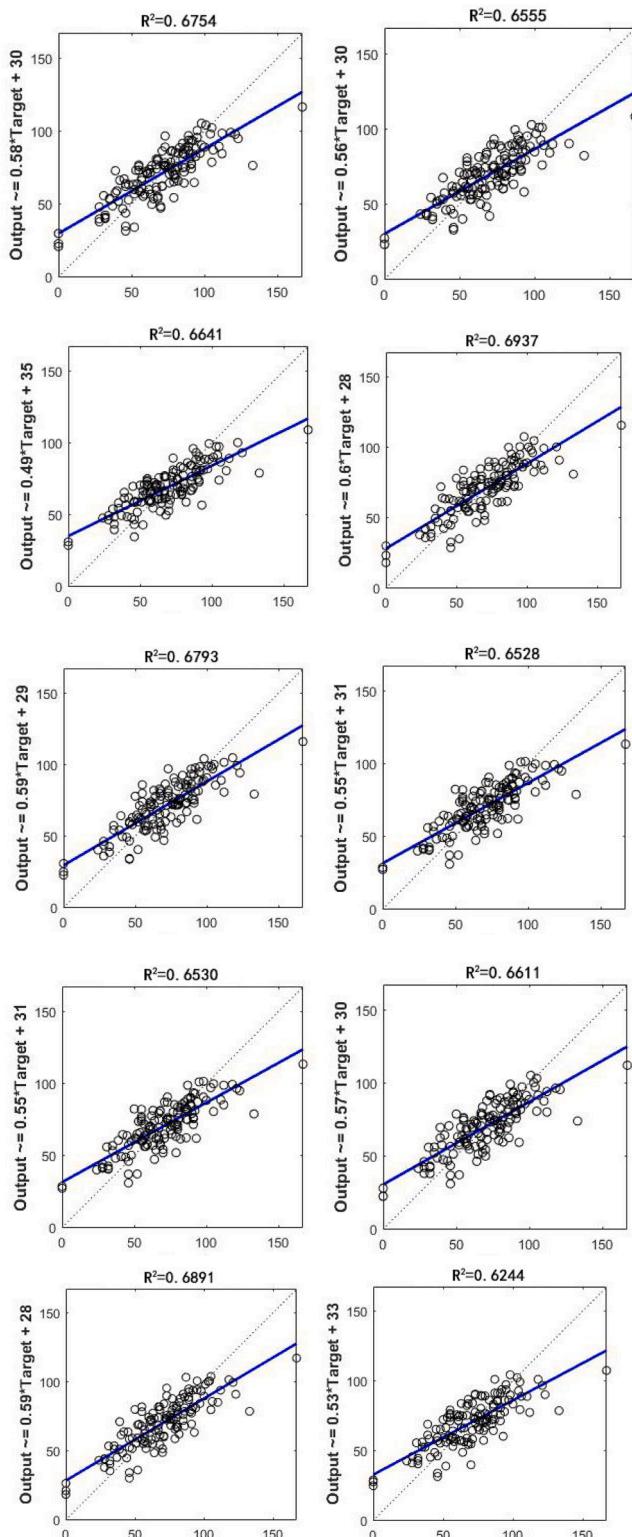


Fig. 15. Ten fold cross validation results - linear regression.

3.3. Stepwise sensitivity analysis

Removing any variables from the best input variable set will result in a significant reduction in the R^2 of the test set. In the stepwise sensitivity analysis, one variable was eliminated at a time. The artificial neural network model was run 10 times to calculate the average value of the correlation coefficient and mean squared error (Table 5). The one-sided

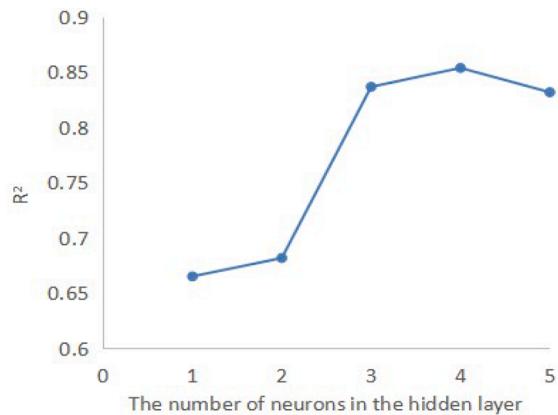
Fig. 16. The relationship between the number of neurons and R^2 .

Table 4
Model fitting effect parameters.

	Best of the 10 folds	Average
Training R^2	0.904	0.854
Training MSE	68.313	96.062
Testing R^2	0.835	0.806
Testing MSE	102.466	205.049

test hypothesis is used to test whether removing any variable significantly reduces the average R^2 of the test set. The Z test was used to test whether the 95% confidence interval was significant. The results show that the removal of the NGRDI, ExG, DSI, CI and EVI will not significantly reduce the prediction effect of the model. According to the results of the step-by-step sensitivity analysis, the input variables of the model are optimized, and the model is simplified by effectively reducing the input variables while controlling the accuracy of the model. K-fold cross validation was used to evaluate the performance of the simplified model, and the results are shown in Table 6. The simplified input variable set is used to establish the model again and conduct stepwise sensitivity analysis. The results are shown in Table 7. The table shows that reducing any parameter will cause a significant reduction in the prediction accuracy.

3.4. Visualization of yield estimation

The image is input as a matrix in MATLAB, and the output value is mapped to the same position after each pixel value is calculated by the model. Finally, the output matrix is visualized as an image to obtain the yield map (Fig. 18). Using the simplified model to predict the cotton yield in the experimental area, the yield map can be more intuitive to see the yield difference distribution in the field, which provides convenience for future research work. The resolution of the yield map is 90 cm, and the corresponding actual area of each pixel is 0.81 m^2 . The base map is the visible image of the region, and the transparent part of the yield map indicates that the yield of the region is 0.

4. Discussion

This research applies neural network methods to time series UAV visible and multispectral remote sensing images to predict cotton yields. The yield prediction model based on all input variables has higher accuracy and a lower mean squared error ($R^2 = 0.854$, $\text{MSE} = 96.062$). The result of stepwise sensitivity analysis shows that there is some redundancy in the initial input variable set, which means that removing redundant variables from the input set will not have a significant impact on the accuracy of the prediction model. The experimental results in 3.3

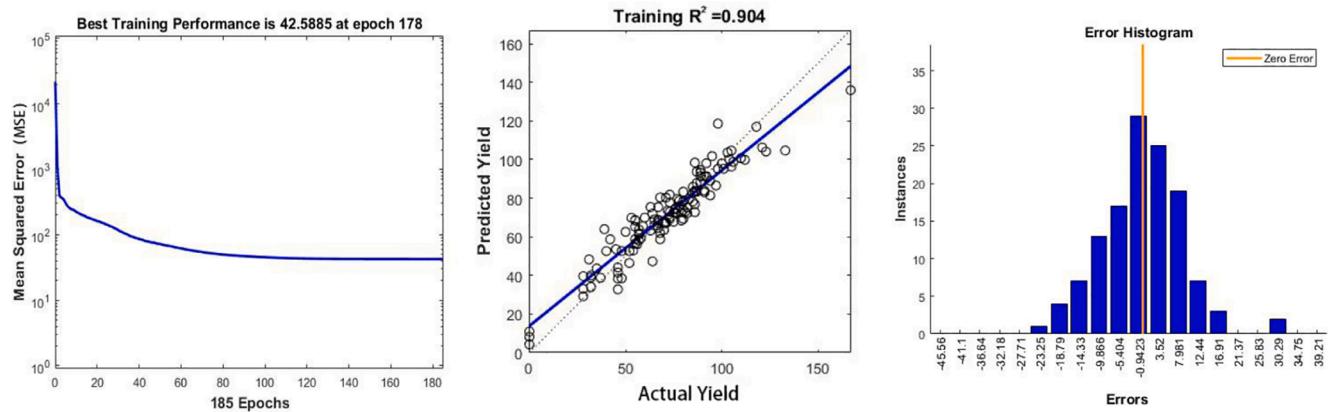


Fig. 17. Model fitting effect.

Table 5

One variable was removed at a time; and the R^2 , mean squared error (MSE), and Z statistics are shown (averages of the 10-fold cross validation).

Parameter Removed	Training R^2	Training MSE	Detection result
VDVI	0.789	141.602	Significant difference
NGRDI	0.850	94.854	No significant difference
VARI	0.805	124.611	Significant difference
ExG	0.854	99.064	No significant difference
DSI	0.848	103.244	No significant difference
RSI	0.795	136.768	Significant difference
NDVI	0.823	120.655	Significant difference
CI	0.837	109.104	No significant difference
MTCI	0.806	127.052	Significant difference
EVI	0.850	102.112	No significant difference
OSAVI	0.824	114.473	Significant difference
BOP	0.811	122.995	Significant difference

Table 6

Model fitting effect parameters (simplified).

	Best of the 10 folds	Average
Training R^2	0.907	0.853
Training MSE	45.363	98.0655
Testing R^2	0.826	0.801
Testing MSE	99.084	198.13

Table 7

One variable was removed at a time; and the R^2 , mean squared error (MSE), and Z statistics are shown (average of the 10-fold cross validation).

Parameter Removed	Training R^2	Training MSE	Detection result
VDVI	0.746	166.805	Significant difference
VARI	0.769	148.270	Significant difference
RSI	0.783	164.285	Significant difference
NDVI	0.829	120.017	Significant difference
MTCI	0.743	174.758	Significant difference
OSAVI	0.810	121.858	Significant difference
BOP	0.771	149.995	Significant difference

prove this point. Eliminating redundant variables can reduce the input variables and simplify the model. Another stepwise sensitivity analysis (Table 6) shows that when the reduced set of input variables is the optimal subset, deleting any of these inputs will affect the prediction accuracy of the network. In this study, six spectral indices and one parameter (BOP) obtained from deep learning segmentation of cotton bolls in the images were used as the input variables of the prediction model. A BP neural network was used to establish the prediction model, and a Bayesian regularization algorithm was used to train the model.

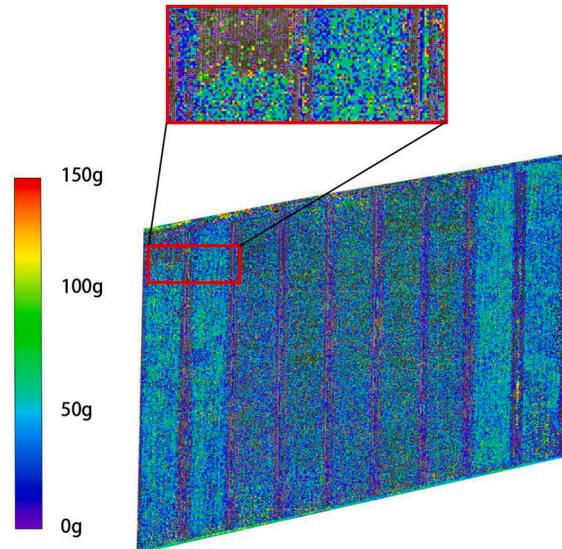


Fig. 18. Yield map.

In agriculture, it is important to identify the variables that contribute to crop yields because these variables can help farmers and agricultural scientists reduce environmental impacts while improving agricultural sustainability (Khoshroo et al. 2018). The results of the stepwise sensitivity analysis showed that the VDVI, VARI, RSI, NDVI, MTCI, OSAVI and BOP had greater impacts on cotton yields. The first five features were spectral features and the BOP was a texture feature.

Using UAV remote sensing technology to extract canopy attributes over the entire growing season for yield prediction has the characteristics of high accuracy, easy use and high cost performance. It is very important to collect high-quality UAV data to calculate the BOP and spectral index, which may be affected by the UAV flight height, longitudinal and lateral overlap ratio and spatial resolution. In addition, the quality of the sensor will also affect the quality of the calculated vegetation index (such as the NDVI), but in the process of collecting experimental data, we use the same multispectral sensor, and the neural network can better fit the nonlinear relationship; therefore, the reliability of the model can be guaranteed.

In this study, the geographical location is a constraint. In the future, more experiments will be conducted to test the robustness of the current model across geographical locations and time. In addition, multisite trials may require additional analysis to include more input characteristics such as the soil type, weather information, etc. This study only focused on the estimation of cotton yields. The portability of this method

to other crops will require the replacement of cotton crop-specific attributes, such as the number of bolls, boll size and boll volume, as well as new crop-specific crop attributes. In general, this study proposes a machine learning framework based on UAV remote sensing data instead of traditional manual yield measurements for cotton yield prediction, which can accurately predict cotton yields and obtain a high-resolution yield map. It can not only help improve the economics of agricultural research but also help to make informed crop management decisions based on yield prediction, especially in the future in the case of breeding trials involving thousands of plots.

5. Conclusions

In this study, a Bayesian regularized BP neural network and an ENVINet-5 semantic segmentation model were used to establish a cotton yield prediction model based on multi-time high-resolution visible and multispectral UAV remote sensing images. The established model predicts cotton yields on the 0.81 m^2 scale, the R² is 0.904, and the average R² of ten cross-validations is 0.854. Through sensitivity analysis, redundant input variables are eliminated, and the optimal subset of input variables is obtained to simplify the model. By using deep learning to extract cotton bolls from high-resolution RGB images, the texture information in the data are further mined. The results of this study can provide valuable tools for cotton breeding research. If breeders can incorporate this technology, it is possible to increase the efficiency of genotype selection. Using this technology, agricultural scientists, crop managers and farmers can obtain more accurate crop information and make appropriate crop management decisions to improve the efficiency and sustainability of agricultural operations.

CRediT authorship contribution statement

Weicheng Xu: Conceptualization, Validation, Formal analysis, Investigation, Software, Writing – original draft. **Pengchao Chen:** Investigation, Resources, Data curation. **Yilong Zhan:** Investigation, Resources, Data curation. **Shengde Chen:** Writing – review & editing, Funding acquisition. **Lei Zhang:** Investigation, Resources. **Yubin Lan:** Conceptualization, Investigation, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

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

Acknowledgments

This research was supported by The Leading Talents of Guangdong Province Program (2016LJ06G689), the Science and Technology Planning Project of Guangdong Province (2017B010117010), the China Agriculture Research System (CARS-15-22), The 111 Project (D18019) and the Science and Technology Planning Project of Guangzhou (201807010039).

References

- Ashapure, A., Jung, J., Chang, A., Oh, S., Yeom, J., Maeda, M., Maeda, A., Dube, N., Landivar, J., Hague, S., Smith, W., 2020. Developing a machine learning based cotton yield estimation framework using multi-temporal UAS data. *ISPRS J. Photogramm. Remote Sens.* 169, 180–194.
- Barrera, O., Perdomo, S.A., 2018. RGB and multispectral UAV image fusion for Gramineae weed detection in rice fields. *Precis. Agric.* 19 (5), 809–822.
- Dash, J., Curran, P.J., 2004. The MERIS terrestrial chlorophyll index. *Int. J. Remote Sens.* 25 (23), 5403–5413.
- Feng, R., Shen, H., Bai, J., Li, X. Advances and opportunities in remote sensing image geometric registration: a systematic review of state-of-the-art approaches and future research directions. *IEEE Geosci. Remote Sens. Mag.* 2–25.
- Feng, S., Cao, Y., Xu, T., Yu, F., Chen, C., Zhao, D., Jin, Y., 2020. Inversion based on high spectrum and NSGA2-ELM algorithm for the nitrogen content of japonica rice leaves. *Spectrosc. Spectral Anal.* 40, 2584–2591.
- Gitelson, A.A., Gritz, J.Y., Merzlyak, M.N., 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* 160 (3), 271–282.
- Gitelson, A.A., Kaufman, Y.J., Stark, R., Rundquist, D., 2002. Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* 80 (1), 76–87.
- Huang, C., Yang, H., Li, Y., Zhang, M., Lv, H., Zhu, A.-X., Yu, Y., Luo, Y.u., Huang, T., 2017. Quantification effect of reforestation to soil erosion in subtropical monsoon regions with acid red soil by sediment fingerprinting. *Environ. Earth Sci.* 76 (1) <https://doi.org/10.1007/s12665-016-6349-z>.
- Huang, H., Deng, J., Lan, Y., Yang, A., Deng, X., Zhang, L., Wen, S., Jiang, Y., Suo, G., Chen, P., 2018a. A two-stage classification approach for the detection of spider mite-infested cotton using UAV multispectral imagery. *Remote Sens. Lett.* 9 (10), 933–941.
- Huang, H., Lan, Y., Deng, J., Yang, A., Deng, X., Zhang, L., Wen, S., 2018b. A semantic labeling approach for accurate weed mapping of high resolution UAV imagery. *Sensors* 18 (7), 2113. <https://doi.org/10.3390/s18072113>.
- Huang, H., Lan, Y., Yang, A., Zhang, Y., Wen, S., Deng, J., 2020. Deep learning versus Object-based Image Analysis (OBIA) in weed mapping of UAV imagery. *Int. J. Remote Sens.* 41 (9), 3446–3479.
- Huang, Y., Sui, R., Thomson, S.J., Fisher, D.K., 2013. Estimation of cotton yield with varied irrigation and nitrogen treatments using aerial multispectral imagery. *Int. J. Agric. Biol. Eng.* 6, 37–41.
- Jung, K., Bae, D.-H., Um, M.-J., Kim, S., Jeon, S., Park, D., 2020. Evaluation of nitrate load estimations using neural networks and canonical correlation analysis with K-fold cross-validation. *Sustainability* 12 (1), 400. <https://doi.org/10.3390/rs12010400>.
- Kanning, M., Kühlung, I., Trautz, D., Jarmer, T., 2018. High-resolution UAV-based hyperspectral imagery for LAI and chlorophyll estimations from wheat for yield prediction. *Remote Sens.* 10 (12), 2000. <https://doi.org/10.3390/rs10122000>.
- Khoshroo, A., Emrouznejad, A., Ghaffarizadeh, A., Kasraei, M., Omid, M., 2018. Sensitivity analysis of energy inputs in crop production using artificial neural networks. *J. Cleaner Prod.* 197, 992–998.
- Kim, D.-W., Yun, H., Jeong, S.-J., Kwon, Y.-S., Kim, S.-G., Lee, W., Kim, H.-J., 2018. Modeling and testing of growth status for Chinese cabbage and white radish with UAV-Based RGB imagery. *Remote Sensing* 10 (4), 563. <https://doi.org/10.3390/rs10040563>.
- Lambert, M.-J., Traoré, P.C.S., Blaes, X., Baret, P., Defourny, P., 2018. Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's cotton belt. *Remote Sens. Environ.* 216, 647–657.
- Lan, Y., Thomson, S.J., Huang, Y., Hoffmann, W.C., Zhang, H., 2010. Current status and future directions of precision aerial application for site-specific crop management in the USA. *Comput. Electron. Agric.* 74 (1), 34–38.
- Li, Y., Cao, Z., Lu, H., Xiao, Y., Zhu, Y., Cremers, A.B., 2016. In-field cotton detection via region-based semantic image segmentation. *Comput. Electron. Agric.* 127, 475–486.
- Li, Y.A., Cao, Z.G., Xiao, Y., Lu, H., Zhu, Y.J., 2015. A Novel Denoising Autoencoder Assisted Segmentation Algorithm for Cotton Field. *IEEE, New York*, pp. 588–593.
- Liu, K., Shen, X., Cao, L., Wang, G., Cao, F., 2018. Estimating forest structural attributes using UAV-LiDAR data in Ginkgo plantations. *ISPRS J. Photogramm. Remote Sens.* 146, 465–482.
- Lobell, D.B., Asner, G.P., 2003. Comparison of Earth Observing-1 ALI and Landsat ETM+ for crop identification and yield prediction in Mexico. *IEEE Trans. Geosci. Remote Sens.* 41 (6), 1277–1282.
- Lucieer, A., Jong, S.M.d., Turner, D., 2014. Mapping landslide displacements using Structure from Motion (SfM) and image correlation of multi-temporal UAV photography. *Prog. Phys. Geogr.-Earth Environ.* 38 (1), 97–116.
- Ma, L., Liu, Y.u., Zhang, X., Ye, Y., Yin, G., Johnson, B.A., 2019. Deep learning in remote sensing applications: a meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* 152, 166–177.
- Meina, Z., Aijing, F., Jianfeng, Z., Xiaolan, L., 2019. Cotton yield prediction using remote visual and spectral images captured by UAV system. *Trans. Chinese Soc. Agric. Eng.* 35, 91–98.
- Misra, G., Cawkwell, F., Wingler, A., 2020. Status of phenological research using sentinel-2 data: a review. *Remote Sensing* 12 (17), 2760. <https://doi.org/10.3390/rs12172760>.
- Ni, Z., Yang, Z., Li, W., Zhao, Y., He, Z., 2019. Decreasing trend of geohazards induced by the 2008 Wenchuan earthquake inferred from time series NDVI data. *Remote Sensing* 11.
- Nielsen, J.K., Christensen, M.G., Cemgil, A.T., Godsill, S.J., Jensen, S.H., 2011. Bayesian interpolation and parameter estimation in a dynamic sinusoidal model. *IEEE Trans. Audio Speech Lang. Process.* 19 (7), 1986–1998.
- Pantazi, X.E., Moshou, D., Alexandridis, T., Whetton, R.L., Mouazen, A.M., 2016. Wheat yield prediction using machine learning and advanced sensing techniques. *Comput. Electron. Agric.* 121, 57–65.
- Ren, Y., Yu, Y., Guan, H., 2020. DA-CapsUNet: a dual-attention capsule U-net for road extraction from remote sensing imagery. *Remote Sensing* 12.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (Eds.), *Lecture Notes in Computer Science*. Springer International Publishing Ag, Cham, pp. 234–241.
- Sargent, D.J., 2001. Comparison of artificial neural networks with other statistical approaches - Results from medical data sets. *Canacer* 91 (S8), 1636–1642.

- Shibayama, M., Akiyama, T., 1991. Estimating grain yield of maturing rice canopies using high spectral resolution reflectance measurements. *Remote Sens. Environ.* 36 (1), 45–53.
- Shrestha, R., Di, L., Yu, E.G., Kang, L., SHAO, Y.-Z., BAI, Y.-q., 2017. Regression model to estimate flood impact on corn yield using MODIS NDVI and USDA cropland data layer. *J. Integr. Agric.* 16 (2), 398–407.
- Wang, D., Wan, B., Qiu, P., Zuo, Z., Wang, R., Wu, X., 2019. Mapping height and aboveground biomass of mangrove forests on hainan island using UAV-LiDAR sampling. *Remote Sensing* 11 (18), 2156. <https://doi.org/10.3390/rs11182156>.
- Wang, M., He, G., Ishwaran, N., Hong, T., Bell, A., Zhang, Z., Wang, G., Wang, M., 2020. Monitoring vegetation dynamics in East Rennell Island World Heritage Site using multi-sensor and multi-temporal remote sensing data. *Int. J. Digital Earth* 13 (3), 393–409.
- Wiratama, W., Lee, J., Sim, D., 2020. Change detection on multi-spectral images based on feature-level U-net. *IEEE Access* 8, 12279–12289.
- Xiang, H., Tian, L., 2011. Development of a low-cost agricultural remote sensing system based on an autonomous unmanned aerial vehicle (UAV). *Biosyst. Eng.* 108, 174–190.
- Xu, W., Deng, X., Guo, S., Chen, J., Sun, L., Zheng, X., Xiong, Y., Shen, Y., Wang, X., 2020a. High-resolution U-Net: preserving image details for cultivated land extraction. *Sensors* 20.
- Xu, W., Lan, Y., Li, Y., Luo, Y., He, Z., 2019. Classification method of cultivated land based on UAV visible light remote sensing. *Int. J. Agric. Biol. Eng.* 12, 103–109.
- Xu, W., Yang, W., Chen, S., Wu, C., Chen, P., Lan, Y., 2020b. Establishing a model to predict the single boll weight of cotton in northern Xinjiang by using high resolution UAV remote sensing data. *Comput. Electron. Agric.* 179, 105762. <https://doi.org/10.1016/j.compag.2020.105762>.
- Yi, L., Lan, Y., Kong, H., Kong, F., Huang, H., Han, X., 2019. Exploring the potential of UAV imagery for variable rate spraying in cotton defoliation application. *Int. J. Precis. Agric. Aviat.* 1.
- Yu, F., Du, W., Guo, Z., Zhou, C., Wang, D., Xu, T., 2019. UAV hyperspectral inversion modeling of rice nitrogen content based on WOA-ELM. *Int. J. Precis. Agric. Aviat.* 2, 43–48.
- Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., Xu, H., Tan, W., Yang, Q., Wang, J., Gao, J., Zhang, L., 2020. Deep learning in environmental remote sensing: Achievements and challenges. *Remote Sens. Environ.* 241, 111716. <https://doi.org/10.1016/j.rse.2020.111716>.
- Yue, W., Jun-Wei, L., Chen-Zhuo, Z., Zhuang-Fei, B., Qi-Guang, M., Wen-Ping, M., Mao-Guo, G., 2020. Computational intelligence in remote sensing image registration: a survey. *Int. J. Autom. Comput.*
- Zhang, H., Yemoto, K., 2019. UAS-based remote sensing applications on the Northern Colorado Limited Irrigation Research Farm. *Int. J. Precis. Agric. Aviat.* 1, 1–10.
- Zhang, K., Zhang, D.P., Jing, C.X., Li, J.C. (2017). Scalable Softmax Loss for Face Verification. In: International Conference on Systems and Informatics. IEEE, New York, pp. 491–496.
- Zhang, X., Li, D., Yang, W., Wang, J., Liu, S., 2011. A fast segmentation method for high-resolution color images of foreign fibers in cotton. *Comput. Electron. Agric.* 78 (1), 71–79.
- Zhang, Y., Wang, L., Bai, Y., Yang, L., Lu, Y., Zhang, J., Li, G., 2019a. Nitrogen nutrition diagnostic based on hyperspectral analysis about different layers leaves in maize. *Spectrosc. Spectral Anal.* 39, 2829–2835.
- Zou, X., Möttus, M., 2017. Sensitivity of common vegetation indices to the canopy structure of field crops. *Remote Sensing* 9 (10), 994. <https://doi.org/10.3390/rs9100994>.
- Wang, X., Wang, M., Wang, S., Wu, Y., 2015. Extraction of vegetation information from visible unmanned aerial vehicle images. *Trans. Chinese Soc. Agric. Eng.* 31, 152–157 (In Chinese).
- Zhang, M., Ai, F., Zhou, J., Lv, X., 2019b. Cotton yield prediction using remote visual and spectral images captured by UAV system. *Trans. Chinese Soc. Agric. Eng.* 35, 91–98.