Cotton Fiber Quality Estimation Based on Machine Learning Using Time Series UAV Remote Sensing Data

Weicheng Xu,1,2,3,†, Weiguang Yang,1,2,3,†, Pengchao Chen,1,2,3, Yilong Zhan,1,2,3, Lei Zhang,3,4, and Yubin Lan,1,2,3,5,*

1 College of Electronic Engineering, South China Agricultural University, Guangzhou 510642, China
2 Guangdong Laboratory for Lingnan Modern Agriculture, Guangzhou 510642, China
3 National Center for International Collaboration on Precision Agricultural Aviation Pesticide Spraying Technology, Guangzhou 510642, China
4 College of Agriculture, South China Agricultural University, Guangzhou 510642, China
5 Department of Biological and Agricultural Engineering, Texas A&M University, College Station, TX 77843, USA

*Correspondence: ylan@scau.edu.cn
† These authors contributed equally to this work.

Abstract: As an important factor determining the competitiveness of raw cotton, cotton fiber quality has received more and more attention. The results of traditional detection methods are accurate, but the sampling cost is high and has a hysteresis, which makes it difficult to measure cotton fiber quality parameters in real time and at a large scale. The purpose of this study is to use time-series UAV (Unmanned Aerial Vehicle) multispectral and RGB remote sensing images combined with machine learning to model four main quality indicators of cotton fibers. A deep learning algorithm is used to identify and extract cotton boll pixels in remote sensing images and improve the accuracy of quantitative extraction of spectral features. 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 results of this study show that the $R^2$ of the prediction model established by a neural network is improved by 29.67% compared with the model established by linear regression. When the spectral index is calculated after removing the soil pixels used for prediction, $R^2$ is improved by 4.01% compared with the ordinary method. The prediction model can well predict the average length, uniformity index, and micronaire value of the upper half. $R^2$ is 0.8250, 0.8014, and 0.7722, respectively. This study provides a method to predict the cotton fiber quality in a large area without manual sampling, which provides a new idea for variety breeding and commercial decision-making in the cotton industry.

Keywords: UAV remote sensing; cotton fiber quality; inversion; semantic segmentation

1. Introduction

Nondestructive and low-cost large-scale prediction of crop quality parameters is of great help to the government and farmers in making correct agronomic decisions [1]. Getting crop character index data early is helpful to effectively screen high-quality varieties from a large number of varieties, improve crop quality, and reap subsequent economic benefits [2]. The quality of cotton fiber directly determines the quality of cotton yarn, and it is also a decisive factor in the value of cotton [3]. Although the traditional indoor detection method has high precision, it is time-consuming, labor-consuming, and expensive; therefore, it is necessary to explore new methods to predict cotton quality quickly and on a large scale [4]. At present, the indexes widely used to evaluate the quality of cotton fiber mainly include the upper half mean length, uniformity, breaking strength, maturity, micronaire value, etc. Taking fiber length as an example, the current measurement methods can be divided into manual measurement and machine measurement [5,6]. The hand-pull
length test is convenient and does not need special test conditions; the test results are highly representative but inefficient. The machine measurement mainly uses a HVI (High Volume Instrument) high-capacity fiber tester to measure the average length of the upper half [7]. The results obtained by the above two measurement methods are different in sample sampling position, self-error, and sample uniformity, so there are also different evaluation standards.

The chemical bonds in the chemical molecular structure of plant biochemical components vibrate under the irradiation of a certain radiation level, resulting in a difference in spectral reflection and absorption at some characteristic wavelengths, which leads to different spectral reflectance, and the change of spectral reflectance at this wavelength is very sensitive to the change of chemical component content. This is the principle behind using remote sensing technology to determine crop chemical parameters. At present, there are two mainstream ways to use remote sensing technology to detect crop quality. For tea [8], tobacco [9], and other crops, the content of biochemical components in leaves and stems (such as nitrogen) is an important index to evaluate quality. The correlation between remote sensing data and biochemical components in leaves or stems can be established directly to evaluate their quality. For rice [10], wheat [11], cotton [12], corn [13], and other crops, the grain and fiber are the harvest objects of economic yield, and the biochemical components of leaves or stems cannot be directly used as indicators to evaluate the quality. Therefore, these crops first establish the correlation between remote sensing parameters and biochemical components in leaves, stems, or bolls, and then take the non-remote sensing model between biochemical components in leaves, stems, or bolls and crop quality indicators as the link. At the same time, according to the remote sensing (image) data obtained at different growth and development stages of crops, the main factors of crop quality formation are determined through analysis, and the comprehensive evaluation model of crop quality is established according to multiple factors. The factors retrieved from remote sensing images and background GIS data (Geographic Information System) are evaluated, and finally, crop quality grade distribution information is obtained.

With the rapid development of remote sensing technology, it is becoming more and more widely used in agricultural management and agricultural condition monitoring [14]. As a low-altitude remote sensing system, a UAV has the advantages of convenience, flexibility, low cost, and high resolution; it has been used to estimate various crop parameters in many instances [15,16]. For example, the UAV is equipped with visible light, multispectral, hyperspectral, LiDAR, and other sensors to collect spectral information of the crop canopy for estimating chlorophyll content [17], nitrogen content [18], disaster extent [19], yield [20], LAI (leaf area index) [21], etc. In addition, the use of a UAV to evaluate the efficacy of cotton defoliants has also achieved good results, which is conducive to achieving precise spraying of the defoliant [22,23]. The formation of cotton fiber quality is the result of multiple factors, such as genetic characteristics [24], environmental conditions [25], and cultivation technical measures [26,27]. If all factors are considered comprehensively, it will be difficult to carry out the design and implementation of the experiment. It is a reasonable scheme to use a remote sensing mechanism and composite model to indirectly monitor cotton fiber quality. Previous studies have shown that cotton fiber has the highest correlation with spectral reflectance and the widest range of sensitive bands in the boll opening period. The sensitive bands for predicting fiber length, strength, and maturity are 350–920 nm and 1400–2500 nm. The sensitive bands for predicting micronaire values are 350–950 nm in the visible range and 1400–2500 nm in the shortwave infrared range, which shows that it is feasible to use spectral reflectance in cotton fiber quality inversion [28].

The remote sensing image taken by an UAV has a centimeter-level ultra-high resolution, which can be used to identify and segment the cotton bolls from the image with the help of computer vision [29,30]. Extracting cotton boll pixels from remote sensing images can eliminate the influence of soil pixels when establishing the quality prediction model and make the prediction result of the model more accurate. At present, there is some research on extracting cotton bolls using machine learning, mainly based on traditional
classification methods such as object-oriented [31] and random forest [32]. The appearance of a Full Convolution Neural Network (FCN) realizes the application of a deep learning model in the field of semantic segmentation and constantly derives various optimization models. U-Net adopts a completely different feature fusion method [33]. Compared with the operation of summing the corresponding points of the feature map during FCN fusion, it splices features together in the channel dimension to form a feature map with higher dimensions, which can extract image features at more scales [34]. At present, U-Net has been widely used in the field of remote sensing; the ENVINet-5 used in this study is a semantic segmentation network with U-Net as the core.

The purpose of this study is to (1) reveal the response of cotton fiber quality differences in remote sensing data, (2) achieve efficient acquisition of cotton fiber quality parameters at field scale, and (3) generate a visual heat map of the field distribution of key quality parameters of cotton fiber.

2. Materials and Methods

2.1. Research Area

Figure 1 shows the geographical location of the experimental site, which is located in Binzhou City, Shandong Province, China. These fields belong to Shandong Lvfeng Agriculture Group Co. Ltd., with a total area of about 61,000 square meters and a planting density of 11 plants per square meter.

Figure 1. Test site location.

A flowchart outlining the overall methodology is presented in Figure 2. Subsequent sections describe all the steps in detail.

2.2. Field Sample Collection

Four fields were selected and marked A, B, C, and D, and 150 square sampling points were set with a side length of 0.9 m and marked with a wire frame and white cloth (Figure 3a) for the identification of the ROI (region of interest). After collection, all samples were put into air-permeable gauze bags and labeled (Figure 3b). A small sawtooth gin was used (Figure 3c) to remove the cotton seeds to obtain the lint, which was air-dried in a ventilated and dry environment for 5 days and sent for inspection. The quality parameter testing was completed in a professional organization, and the testing organization received a qualification certificate. The test method is the physical property test method of HVI cotton fiber. Taking GB/T 20392-2006 as the test standard, five indexes of average length, uniformity index, breaking specific strength, micronaire value, and upper half mean length are measured. Table 1 shows the maximum, minimum, mean, and coefficient of variation (CV) for the five parameters.
Figure 2. Overall flow chart.

Figure 3. Sample collection and processing. (a) A wire frame used to define a boundary; (b) sample data; (c) cotton ginning; and (d) a sample area example.
Table 1. Quality parameter statistics.

<table>
<thead>
<tr>
<th>Upper Half Mean</th>
<th>Uniformity</th>
<th>Breaking Tenacity/cN•tex-1</th>
<th>Micronaire Value</th>
<th>Elongation/%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length/mm</td>
<td>Index/%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max</td>
<td>29.6</td>
<td>85.1</td>
<td>30.4</td>
<td>5.4</td>
</tr>
<tr>
<td>Min</td>
<td>26</td>
<td>80.4</td>
<td>25.5</td>
<td>4</td>
</tr>
<tr>
<td>Mean</td>
<td>27.85</td>
<td>82.98</td>
<td>27.68</td>
<td>4.76</td>
</tr>
<tr>
<td>CV</td>
<td>0.024</td>
<td>0.009</td>
<td>0.036</td>
<td>0.058</td>
</tr>
</tbody>
</table>

It can be seen from the data in the table that elongation has no significant difference in the data collected and the data distribution is relatively concentrated; therefore, only the average length of the upper half, the uniformity index, the fracture specific strength, and the micronaire value are predicted in this study.

2.3. Data Acquisition Equipment

Data collection uses two types of drones to capture visible light and multispectral remote sensing data, as shown in Figure 4a. A DJI Phantom 4 RTK captures visible light data, and a DJI Phantom 4 Multispectral captures multispectral data. The parameters of the two aircraft and the flight parameters when the data were collected are shown in Table 2. In order to obtain centimeter-level accurate positioning, the time synchronization system is used to reduce the error between the flight control system, camera, and the RTK’s clock system to a microsecond level. Combined with the equipment altitude information, the center point positions of the camera lens and the antenna are compensated in real time so that the image can obtain accurate geographical position information. The data collection time was selected at 2 p.m. on a cloudless day, and the calibration plate was photographed to obtain the reflectance (Figure 4b).

Table 2. Equipment parameter table.

<table>
<thead>
<tr>
<th></th>
<th>Phantom 4 RTK</th>
<th>Phantom 4 Multispectral</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMOS pixel</td>
<td>5472 × 3648</td>
<td>1600 × 1300</td>
</tr>
<tr>
<td>Field of View (FOV)</td>
<td>84°</td>
<td>62.7°</td>
</tr>
<tr>
<td>Flight altitude</td>
<td>100 m</td>
<td>100 m</td>
</tr>
<tr>
<td>Ground sampling distance</td>
<td>2.74 cm/px</td>
<td>5.29 cm/px</td>
</tr>
<tr>
<td>Overlap rate</td>
<td>70%</td>
<td>70%</td>
</tr>
<tr>
<td>Imaging band</td>
<td>Visible light</td>
<td>Red (650 nm ± 16 nm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Green (560 nm ± 16 nm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Blue (450 nm ± 16 nm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Red Edge (730 nm ± 16 nm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Near Infrared (840 nm ± 26 nm)</td>
</tr>
</tbody>
</table>
The preprocessing of remote sensing data is completed on the Power Leader™ PR490P computing server. DJI Terra is used to splice the images taken by the UAV. The software will automatically read the EXIF (exchangeable image file format) information of the images and find the same name points in a single image to generate an orthophoto of the research area.

2.4. Pixel-Level Fusion of Time Series RGB and Multispectral Data

In this study, 13 periods of remote sensing image data of cotton from the bud stage to the complete boll opening were collected and superimposed with pixels in the direction of the channel to realize the data fusion of the time series (Figure 5). The most important step in the fusion process is to reduce the error of image registration to less than 1 pixel, which is the premise of accurate feature extraction. A polynomial correction model was used for image registration, and the coefficients \((a_{ij} \text{ and } b_{ij})\) in the polynomial were calculated by the least squares method using the image coordinates of the ground control point and the ground coordinates of its same name point. Because the data has centimeter-level positioning accuracy, it does not need a lot of control points to meet the accuracy requirements. The update process of image coordinates is shown in Formula (1), where \((x, y)\) is the original coordinate and \((u, v)\) is the coordinate of the same name point.

\[
\begin{align*}
    x &= \sum_{i=0}^{N} \sum_{j=0}^{N-i} a_{ij}u^iv^j \\
    y &= \sum_{i=0}^{N} \sum_{j=0}^{N-i} b_{ij}u^iv^j
\end{align*}
\] (1)

![Figure 5](image-url)

**Figure 5.** Schematic diagram of the data fusion effect. (a) Time-series data fusion and (b) RGB and multispectral data fusion.
Formula (2) shows the relationship between the number \( N \) of polynomial coefficients and the polynomial order \( n \). In general, the second-order polynomial model can handle most image registrations.

\[
N = \frac{(n + 1)(n + 2)}{2}
\]  

(2)

the second-order polynomial model needs at least 6 control points. When the number of control points exceeds 6, the error is calculated according to Formula (3). The error is reduced by increasing the control points to make them smaller than one pixel.

\[
RMSE_{error} = \sqrt{(v - x)^2 + (u - y)^2}
\]  

(3)

2.5. Cotton Boll Pixel Recognition Based on Deep Learning

2.5.1. Network Structure

In this study, an ENVI deep learning module is used to extract cotton peach pixels from remote sensing images, with the feature extraction part based on a U-Net network model (Figure 6a). Before convolution, pixels with a value of 0 are usually filled around the image to ensure the same size before and after convolution; however, as the number of model layers increases, the degree of feature map abstraction increases, and the error will continue to accumulate. A U-Net model will not perform this operation, which will make it difficult to restore the original resolution of the image. The “overlap tile” scheme is used to reflect and fill pixels at the edge of the image to solve this problem (Figure 6b). This method preserves the context information of edge features while maintaining the original resolution of the image.

![Network Structure](image)

**Figure 6.** Network structure. (a) U-Net model structure and (b) overlap tile.
2.5.2. Training Model

Once a set of ROIs is defined for each training raster (Figure 7a), the Deep Learning Labeling Tool automatically creates a label raster in the form of a binarized image (Figure 7b). In this study, 500 ROI regions were extracted, and the training set and the test set were divided according to a ratio of 4 to 1.

![Figure 7](image)

**Figure 7.** Label making. (a) A drawn ROI in the image and (b) a labeled-binary image.

In this study, the cross-entropy loss function is used to evaluate the classification error of the model. In the semantic segmentation task, it will calculate each pixel one by one. Initially, the model will generate a random class activation grid, calculate the loss function between the grid and the label, and then update the model parameters through the gradient descent method to converge the model. The image will be divided into a given size in advance and then sent to the network to avoid excessive calculation pressure.

2.5.3. Cotton Boll Opening Pixel Percentage

The extraction of cotton boll pixels is a binary task. The proportion of cotton boll opening pixel percentage (BOP) can be obtained by calculating the average value in the sampling area after the output results are binarized, as shown in Figure 8.

![Figure 8](image)

**Figure 8.** BOP calculation process.

2.6. Establishment of a Prediction Model

2.6.1. Bayesian Regularized BP Neural Network

In this study, a two-layer BP neural network was built to explore the mapping relationship between cotton fiber quality and spectral index (Figure 9). The Bayesian regularization training algorithm introduces the regularization coefficient into the traditional loss function and uses the Gaussian distribution as the prior probability density distribution, which can improve the overfitting problem in the model training.
2.6. Establishment of a Prediction Model

2.6.1. Bayesian Regularized BP Neural Network

In this study, a two-layer BP neural network was built to explore the mapping relationship between cotton fiber quality and spectral index (Figure 9). The Bayesian regularization training algorithm introduces the regularization coefficient into the traditional loss function and uses the Gaussian distribution as the prior probability density distribution, which can improve the overfitting problem in the model training.

Figure 9. BP neural network model structure.

The traditional loss function is the sum of the squared errors, calculated according to Formula (4) where \( t_i \) is the true value and \( x_i \) is the predicted value of the model.

\[
E_D = \sum_{i=1}^{n} (t_i - x_i)^2
\]

\[
E_W = \frac{1}{m} \sum_{i=1}^{m} w_i^2
\]

\[
F(w) = \alpha E_W + \beta E_D
\]

Formula (6) is the regularization performance function of the network, where \( E_{\text{reg}} \) is the mean square sum of all weighted parameters in the network and \( \alpha \) and \( \beta \) affect the complexity and smoothness of the network, respectively; they can be calculated by Formula (9) when the effective network parameter is \( \gamma \). As the number of network training increases, they are constantly updated. All weights and biases in the network will be given random values in the first training. Suppose that the noise in \( D \) (datasets) and the weight vector in \( M \) (neural network model) obey Gaussian distributions in probability density; a posteriori probability density function of parameters in the network after an iteration can be calculated using the Bayesian formula.

\[
P(w|D, \alpha, \beta, M) = \frac{P(D|w, \beta, M)P(w|\alpha, M)}{P(D|\alpha, \beta, M)}
\]

\[
\begin{align*}
P(D|w, \beta, M) &= \exp\left(-\beta E_D\right) \sqrt{\left(\frac{\beta}{\pi}\right)^n} \\
P(w|\alpha, M) &= \exp\left(-\alpha E_W\right) \sqrt{\left(\frac{\alpha}{\pi}\right)^n}
\end{align*}
\]
\begin{equation}
\begin{cases}
\alpha = \frac{\gamma}{2E_W(w_0)} \\
\beta = \frac{m - \gamma}{2E_W(w_0)}
\end{cases}
\end{equation}

2.6.2. Training Model

Figure 10 shows some spectral indicators used in this study and the band range required for calculation. These indicators have been widely used in the study of vegetation physiological and biochemical monitoring indicators.

The spectral data of each sampling point is linked in the time direction to form a feature vector (Figure 11) and matched with the measured values one by one to form a dataset. Finally, there are 150 sets of eigenvectors in the dataset for network training and testing.

**Figure 10.** Spectral index and its required band range.

**Figure 11.** Connect all-time series eigenvectors.
3. Results
3.1. Cotton Boll Extraction

Figure 12 shows the convergence of the loss function on the training set and the verification set. When the epoch reaches 24, the model converges almost completely.

A total of 500 points are randomly selected in the image as the test set, and the confusion matrix method is used to test the accuracy of the model. The results are shown in Table 3. Formula (10) is the calculation method for overall accuracy, where \(n\) is the number of correctly classified samples and \(N\) is the total number of samples. Formula (12) is the calculation method for the Kappa coefficient, where \(P_o\) is the overall accuracy and \(P_c\) is the accidental consistent pixel scale. The overall accuracy and Kappa coefficient are 92.4% and 0.8204, respectively, which show that the classification results of the model are highly consistent with the actual situation.

\[
p_o = \frac{n}{N} \quad (10)
\]

\[
p_c = \frac{a_1b_1 + a_2b_2}{N \times N} \quad (11)
\]

\[
k = \frac{p_o - p_c}{1 - p_c} \quad (12)
\]

Table 3. Confusion matrix.

<table>
<thead>
<tr>
<th>Identification Type</th>
<th>Cotton Boll</th>
<th>Other Objects</th>
<th>Total</th>
<th>User’s Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cotton boll</td>
<td>371</td>
<td>10</td>
<td>381</td>
<td>97.35%</td>
</tr>
<tr>
<td>Other objects</td>
<td>14</td>
<td>105</td>
<td>119</td>
<td>88.23%</td>
</tr>
<tr>
<td>Total</td>
<td>385</td>
<td>115</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Producer’s accuracy</td>
<td>96.36%</td>
<td>91.30%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.2. All Parameters Participate in Modeling

Using the upper half mean length prediction as an example, this paper explains how to build the prediction model. The modeling and prediction methods of uniformity index, fracture specific strength, and micronaire value are consistent with the method steps shown below. Two methods were used to calculate the spectral index in the sampling area. The first is to calculate the average value in the region. The second is to use the segmentation results of U-Net to calculate the average value after removing the non-cotton peach pixels.
in the region. The dataset is divided into a training set and a verification set according to the ratio of 4:1, and the model is evaluated by ten-fold cross-validation.

3.2.1. Least Squares Modeling

A preliminary correlation analysis on the upper half mean length is performed using a multivariate linear equation. The dataset is divided into 10 equal parts and verified with a ten-fold cross-validation method. Figure 13 shows the ten-fold cross-validation results of the correlation between the estimated results of the linear regression model and the measured results of the sampling.

![Figure 13. Ten-fold cross-validation results—least squares modeling.](image)

According to the results of the cross validation, $R^2$ is generally low (average 0.56) and fluctuates greatly (MSE is 0.21). This may be due to the complex relationship between cotton fiber quality parameters, spectrum, and texture features and the traditional linear regression model, which makes it difficult to express the mapping relationship between them.

3.2.2. BP Neural Networks Modeling

The hyperparameters of the neural network (such as the number of hidden layers and the number of neurons contained in each hidden layer) need to be set before training. In general, the two-layer network and linear output neurons with sigmoid activation functions can fit well the multi-dimensional mapping problem given a certain amount of data. For more hidden neurons and layers, the model’s performance has increased in the training set but decreased in the test set. This study uses a trial-and-error method to determine the number of neurons in the hidden layer. Figure 14 shows the $R^2$ for the different network structures.

![Figure 14. The relationship between the number of hidden layer neurons and network performance.](image)
As can be seen from Figure 14, with the increase in the number of neurons, there is an overall fluctuation trend, in which four and nine are the positions of the two peaks in the broken line diagram. From the principal analysis of the neural network, the more neurons, the more complex the data features the model can fit. At the same time, the model may learn some noise or irrelevant features in the dataset, thus affecting the performance of the model in the test set. Comparing the accuracy of the model on the test set when the number of hidden neurons is four and nine, the results show that the latter has higher accuracy.

There are twelve nodes in the input layer, which input various spectral indexes of time series and BOP, respectively. The hidden layer contains nine neurons, which are connected with the twelve nodes of the input layer, and then the average length of the upper half is output through the output layer, which contains a linear function. Figure 15 shows the ten-fold cross-validation results of the correlation between the estimated results of the BP neural networks model and the measured results of sampling.

![Figure 15. Ten-fold cross-validation results—BP neural network.](image)

3.2.2. BP Neural Networks Modeling

The hyperparameters of the neural network (such as the number of hidden layers and the number of neurons contained in each hidden layer) need to be set before training. In general, the two-layer network and linear output neurons with sigmoid activation functions can fit well the multi-dimensional mapping problem given a certain amount of data. For more hidden neurons and layers, the model's performance has increased in the training set but decreased in the test set. This study uses a trial-and-error method to determine the number of neurons in the hidden layer. Figure 14 shows the R² for the different network structures.

![Figure 14. The relationship between the number of hidden layer neurons and network performance.](image)

3.2.3. Comparison of Results

It can be seen from the results that the prediction effect of neural network modeling is much better than that of linear regression, so a BP neural network is selected as the framework for a prediction model. In addition, in order to explore the impact of soil pixels on fiber quality prediction performance, 10 repeated modeling experiments were carried out. When calculating the spectral index, the average value of boll cottonseed pixels segmented by U-Net was used instead of the average value in the region. Table 4 shows the accuracy of the estimation model in three cases; the estimation accuracy is higher when using the spectral index calculated by pixel points, excluding soil, for modeling, and R² has a 4.01% increase.

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Max R²</th>
<th>Average R²</th>
<th>Minimum MSE</th>
<th>Average MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear regression</td>
<td>0.6404</td>
<td>0.5593</td>
<td>0.4012</td>
<td>0.212</td>
</tr>
<tr>
<td>BP neural network</td>
<td>0.9084</td>
<td>0.7952</td>
<td>0.0676</td>
<td>0.1094</td>
</tr>
<tr>
<td>BP neural network (remove soil pixels)</td>
<td>0.9098</td>
<td>0.8271</td>
<td>0.075</td>
<td>0.104</td>
</tr>
</tbody>
</table>
3.3. Remove Redundant Variables

In order to eliminate redundant variables, a sensitivity analysis was conducted for all input parameters of the model (using the one-by-one elimination method). After each parameter is proposed, the model effect will be cross-verified, and $R^2$ and MSE will be counted (Table 5). One-sided test (95% confidence interval) is used to evaluate whether the influence of the removed variables on the model is significant.

Table 5. Results of stepwise sensitivity analysis.

<table>
<thead>
<tr>
<th>Parameter Removed</th>
<th>Training $R^2$</th>
<th>Training MSE</th>
<th>Whether Significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>VDVI</td>
<td>0.811</td>
<td>129.686</td>
<td>×</td>
</tr>
<tr>
<td>NGRDI</td>
<td>0.8221</td>
<td>92.938</td>
<td>×</td>
</tr>
<tr>
<td>VARI</td>
<td>0.8193</td>
<td>122.695</td>
<td>×</td>
</tr>
<tr>
<td>ExG</td>
<td>0.8268</td>
<td>97.148</td>
<td>×</td>
</tr>
<tr>
<td>DSI</td>
<td>0.821</td>
<td>101.328</td>
<td>×</td>
</tr>
<tr>
<td>RSI</td>
<td>0.7678</td>
<td>134.852</td>
<td>✓</td>
</tr>
<tr>
<td>NDVI</td>
<td>0.796</td>
<td>118.739</td>
<td>✓</td>
</tr>
<tr>
<td>CI</td>
<td>0.8098</td>
<td>107.188</td>
<td>×</td>
</tr>
<tr>
<td>MTCI</td>
<td>0.7788</td>
<td>125.136</td>
<td>✓</td>
</tr>
<tr>
<td>EVI</td>
<td>0.823</td>
<td>100.196</td>
<td>×</td>
</tr>
<tr>
<td>OSAVI</td>
<td>0.7969</td>
<td>112.557</td>
<td>✓</td>
</tr>
<tr>
<td>BOP</td>
<td>0.6381</td>
<td>121.079</td>
<td>✓</td>
</tr>
</tbody>
</table>

Not significant model input parameters were deleted according to the results of the step sensitivity analysis. After removing redundant variables, the model performance parameters are shown in Table 6.

Table 6. Comparison of model performance parameters.

<table>
<thead>
<tr>
<th></th>
<th>Cross Validation Maximum (Minimum) Value</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training $R^2$</td>
<td>0.8702</td>
<td>0.825</td>
</tr>
<tr>
<td>Training MSE</td>
<td>0.0735</td>
<td>0.1053</td>
</tr>
<tr>
<td>Validation $R^2$</td>
<td>0.8061</td>
<td>0.7826</td>
</tr>
<tr>
<td>Validation MSE</td>
<td>0.1025</td>
<td>0.1986</td>
</tr>
</tbody>
</table>

When using the simplified input variable set for prediction, the accuracy of the model remains basically unchanged in the training and test sets. In addition, the accuracy level remains at the same level. The result of the stepwise sensitivity analysis shows that each of the remaining input parameters has a significant impact on the performance of the estimation model.

3.4. Summary of Cotton Fiber Quality Parameter Prediction Models

Using the same method, the prediction models of uniformity index, fracture specific strength, and micronaire value are established. It is found that the optimal input variables of the four-index prediction model are consistent, which shows that the fiber quality parameters of cotton are consistent in the formation of physiological characteristics. RSI, NDVI, MTCI, OSAVI, and BOP can be used to predict cotton fiber quality parameters. The model performance parameters are shown in Table 7.
Table 7. Comparison of prediction model construction results.

<table>
<thead>
<tr>
<th></th>
<th>Training R²</th>
<th>Validation R²</th>
<th>Training MSE</th>
<th>Validation MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper half mean length/mm</td>
<td>0.825</td>
<td>0.7826</td>
<td>0.1053</td>
<td>0.1986</td>
</tr>
<tr>
<td></td>
<td>0.8702</td>
<td>0.8061</td>
<td>0.0735</td>
<td>0.1025</td>
</tr>
<tr>
<td>Uniformity Index/±%</td>
<td>0.8014</td>
<td>0.7523</td>
<td>0.1903</td>
<td>0.288</td>
</tr>
<tr>
<td></td>
<td>0.8591</td>
<td>0.7966</td>
<td>0.117</td>
<td>0.2195</td>
</tr>
<tr>
<td>Breaking tenacity/cN tex⁻¹</td>
<td>0.7264</td>
<td>0.6536</td>
<td>0.3132</td>
<td>0.2112</td>
</tr>
<tr>
<td></td>
<td>0.8212</td>
<td>0.7884</td>
<td>0.1591</td>
<td>0.2965</td>
</tr>
<tr>
<td>Micronaire value</td>
<td>0.7722</td>
<td>0.7655</td>
<td>0.012</td>
<td>0.0167</td>
</tr>
<tr>
<td></td>
<td>0.8267</td>
<td>0.8142</td>
<td>0.0151</td>
<td>0.0126</td>
</tr>
</tbody>
</table>

It can be seen from the prediction results that the prediction model has good prediction performance for the average length, uniformity index, and micronaire value of the upper half. $R^2$ is around 0.8, but the prediction ability for fracture-specific strength is poor, and the average $R^2$ is only 0.7264. It is speculated that the reason for this result is that the characteristics sensitive to the change of fracture-specific strength have not been added to the input variables of the model. The average errors of the four indexes in the verification within one square meter are 0.5502, 0.6625, 0.5674, and 0.1595, respectively.

3.5. Visualization of Quality Prediction

An image is inputted into MATLAB in the form of a matrix, the output matrix is visualized in the form of an image, and the visible light in the orthophoto image is taken as the base map to obtain the distribution map of cotton fiber quality parameters. From Figure 16, we can more intuitively see the distribution of cotton fiber quality parameters in the field. In practice, it is meaningless to predict the fiber quality at the pixel level with a BOP of 0, so the model eliminates the pixel with a BOP of 0.

![Figure 16](image_url)
4. Discussion

This study used a time series of UAV remote sensing images to estimate the cotton fiber quality index, and good accuracy was achieved. High-resolution images in visible light can make up for the deficiencies of traditional multispectral sensors in spatial resolution to a certain extent. Input of time series fusion features can improve the accuracy of the model, and machine learning models further improves the accuracy. Combined with machine learning and multi-source data fusion, it realizes efficient and accurate monitoring of cotton fiber quality parameters.

4.1. Correlation between Reflectivity and Quality Parameters of Cotton Fiber

The spectral and texture information of UAV images has been widely used in crop phenotypic research. The absorption of substances in the near-infrared spectral region mainly includes the frequency combination and frequency doubling vibration absorption of chemical bond groups with higher energy (C-H, N-H, O-H, S-H, C=O, and C=C). It contains almost all the information of hydrogen-containing groups in organic matter, which is rich in information, so a multispectral vegetation index has a strong correlation with phenotypic traits [35–37]. Compared with multispectral images and RGB images, they usually have a higher spatial resolution, which helps to extract more abundant texture information and achieve a fixed point, as well as a quantitative and accurate collection of spectral features. However, a single data source has great limitations [38]. RGB images cannot obtain spectral information in the NIR band, and multispectral images will lose a lot of texture features, such as BOP and soil pixel masks, in this research; these indicators both have a greater contribution to improving the monitoring accuracy of cotton fiber quality (Tables 5 and 6). Moreover, Thompson et al. showed that the vegetation index based on the red edge band can be a good grading score of cotton fiber maturity, and the correlation was not breed-specific, suggesting that spectral monitoring of cotton fiber quality may be used for multiple cotton genotypes [39].

NIR spectroscopy is generally considered a cost-effective alternative to traditional laboratory methods and systems [40]. Because the infrared reflectance spectrum is sensitive to cellulose content, there is a strong correlation between fiber quality and the reflectance spectrum. Near-infrared spectroscopy has been widely studied in the qualitative analysis of textile fibers, traceability, impurity content of raw cotton, qualitative analysis of blended fabrics, etc. The accuracy of using near-infrared spectroscopy to estimate micronaire value exceeds 97% [41], which is consistent with the results of the correlation analysis of this study. All spectral indexes that have significant influence on the estimation results are calculated according to the NIR (Figure 10). However, there are complex relationships between cotton fiber quality and spectra, structural parameters and texture information. Traditional linear regression models may be difficult to express in terms of their mapping relationships. Compared with the traditional least squares linear regression (Figure 13), machine learning can achieve higher precision regression through training parameters (Figure 15). In this study, we used deep learning to segment the cotton boll to get BOP and remove the soil pixels to obtain a more accurate spectral index. The results show that a more accurate spectral index, using a neural network algorithm and time series data, can improve the estimation accuracy of cotton fiber quality parameters.

4.2. Limitations and Prospects

Because this study did not carry out large-scale, cross-regional sampling, regional differences were not considered. Therefore, the next step of this study is to collect more time series data from different regions and constantly adjust and optimize the model. In addition, the hyperspectral full-frame imaging system can obtain canopy spectral reflection data with a wider band range and higher spectral resolution, which may help to improve the accuracy of fiber quality monitoring. The research results have important practical significance for improving the efficiency of cotton breeding research, especially in breeding experiments involving thousands of samples.
5. Conclusions

In this study, time series RGB and multispectral UAV images were used to obtain relevant cotton canopy attributes and establish a machine learning model for monitoring cotton fiber quality parameters. The ability to combine high-resolution RGB images and multispectral information to predict cotton fiber quality is very powerful. The model can predict well the upper-half mean length, uniformity index, and micronaire value with an $R^2$ value of 0.8250, 0.8014, and 0.7722, respectively. Redundant input variables are eliminated through sensitivity analysis to obtain the optimal subset of input variables, which reduces the cost of collecting multispectral data. In addition, when using the method of removing soil pixels to calculate the spectral index for prediction, $R^2$ is 4.01% higher than the method of using the average value to calculate the spectral index. In this research, a large- and small-scale cotton fiber quality parameter prediction model is established, which provides a valuable tool for cotton breeding research. The results of this study enable UAVs to replace most of the manual inspection work, greatly improve the efficiency of cotton breeding research, and accelerate the breeding of high-quality cotton varieties.

Author Contributions: Conceptualization, data curation, formal analysis, methodology, supervision, investigation, writing—original draft, review and editing, W.X.; data curation, formal analysis, methodology, software, validation, visualization, writing—original draft, review and editing, W.Y.; conceptualization, methodology, supervision, L.Z.; methodology and investigation, P.C.; methodology and investigation, Y.Z.; methodology and investigation, P.C.; conceptualization, funding acquisition, project administration, software, supervision, writing—review and editing, Y.L.; equal contribution to this work, W.X. and W.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Laboratory of Lingnan Modern Agriculture Project (NT2021009), the Leading Talents of Guangdong Province Program (2016J06G689), 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).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study are available from the corresponding author upon reasonable request.

Acknowledgments: We sincerely thank Cui Lihua from Shandong Lvfeng Agriculture Group Co. Ltd. for providing practical field assistance.

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

References

38. Xie, C.Q.; Yang, C. A review on plant high-throughput phenotyping traits using UAV-based sensors. Comput. Electron. Agric. 2020, 178, 105731. [CrossRef]


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