Yield Estimation of High-Density Cotton Fields Using Low-Altitude UAV Imaging and Deep Learning

Fei Li
Shihezi University
Jingya Bai
Shihezi University
Mengyun Zhang
Shihezi University
Ruoyu Zhang (zryzju@gmail.com)
Shihezi University https://orcid.org/0000-0003-0370-3249

Research Article

Keywords: yield estimation, unmanned aerial vehicle, CD-SegNet, densely planted cotton

Posted Date: January 11th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1217302/v1

Yield Estimation of High-Density Cotton Fields Using Low-altitude UAV Imaging and Deep Learning

Fei Li [1,2†], Jingya Bai [1,2†] , Mengyun Zhang [1,2] , Ruoyu Zhang [1,2*]

*Correspondence: zrjzju@gmail.com

†Fei Li and Jingya Bai contributed equally.

[1] College of Mechanical and Electrical Engineering, Shihezi University, Shihezi, 832000, China;

[2] Key Laboratory of Northwest Agricultural Equipment, Ministry of Agriculture, P.R. China, Shihezi, 832000

Abstract

Background: Different from other parts of the world, China has its own cotton planting pattern. Cotton are densely planted in wide-narrow rows to increase yield in Xinjiang, China, causing the difficulty in the accurate evaluation of cotton yields using remote sensing in such field with branches occluded and overlapped.

Results: In this study, low-altitude unmanned aerial vehicle (UAV) imaging and deep convolutional neural networks (DCNN) were used to estimate the yields of densely planted cotton. Images of cotton field were acquired by an UAV at the height of 5 m. Cotton bolls were manually harvested and weighted afterwards. Then, a modified DCNN model was developed by applying encoder-decoder recombination and dilated convolution for pixel-wise cotton boll segmentation termed CD-SegNet. Linear regression analysis was employed to build up the relationship between cotton boll pixels ratio and cotton yield. Yield estimations of four cotton fields were verified after

machine harvest and weighting. The results showed that CD-SegNet outperformed the other tested models including SegNet, support vector machine (SVM), and random forest (RF). The average error of the estimated yield of the cotton fields was 6.2%.

Conclusions: Overall, the yield estimation of densely planted cotton based on low-altitude UAV imaging is feasible. This study provides a methodological reference for cotton yield estimation in China.

Keywords: yield estimation; unmanned aerial vehicle; CD-SegNet; densely planted cotton

1 Background

Xinjiang is China's most important cotton planting base and occupies a pivotal position in the world's cotton industry. In 2020, the total cotton output in China was 5.91 million tons, of which 87.3% (5.16 million tons) was produced in Xinjiang. Fast and reliable estimation of cotton yields prior to harvest is essential for crop management, cotton trade, and policy making. At present, Xinjiang has widely adopted the dense planting pattern of “short-dense-early”. This pattern employs alternate wide and narrow rows (66 and 10 cm), and the number of plants per hectare is between 240,000 and 270,000. Although this pattern has obvious advantages for withstanding natural disasters and increasing yields [1], the plant density in this pattern is relatively high. Moreover, the narrow rows of cotton plants are staggered and severely occluded, which poses certain difficulties for imaging-based yield estimation.

Currently, traditional cotton yield estimation methods are labour-intensive and inefficient. Thus, they cannot meet the needs of the rapidly developing cotton industry

[2]. Within a cotton field, there may be spatial differences in yields, which may produce large errors in estimation. As space technology has continued to develop, crop yield estimation methods based on satellite remote sensing technology have been widely used [3–5]. Cotton yields can be accurately predicted by using yield estimation models constructed with remote sensing data as well as vegetation index[6]. However, satellite remote sensing images are limited by the effects of temporal and spatial resolution along with cloud cover. Moreover, they are usually insufficient for accurate estimation of field-scale crop yields. In contrast, unmanned aerial vehicles (UAVs), due to their flexibility and low-altitude flight capability, have quickly become an ideal tool for the precision monitoring of crops [7,8]. UAV-based low-altitude remote sensing platforms can obtain images with very high temporal and spatial resolution that are free from atmospheric interference [9-11]. Akash et al. [12] developed a machine learning framework for estimating cotton yields using multitemporal remote sensing data collected with unmanned aerial systems (UASs), and obtained more reliable crop yield estimates. Stroppiana et al. [13] used low-altitude UAV remote sensing to predict the yields of wheat and soybeany. However, this type of research is mainly based on the satellite remote sensing.

Vegetation indices are suitable for estimating cotton yields only during the mid-growth stage. Their estimation capability is limited during the mature stage. Due to the boll opening and background objects, such as cotton bolls, branches, and leaves during the later growth stage of cotton, obvious differences are always found in visual characteristics such as colour and morphology. Therefore, cotton yields can be directly

estimated by the remote sensing images of a defoliated cotton field and background segmentation. Huang et al. [14] used UAV images to estimate the yield of cotton based on cotton boll coverage and plant height. Feng et al. [15,16] comprehensively evaluated the performance of different growth stages and different image characteristics when estimating cotton yields and determined that the plant height and cotton fibre index are features for estimating cotton yields before harvest. Xu et al. constructed a cotton yield estimation model based on UAV remote sensing data [17]. However, in the abovementioned yield estimations, the density of cotton plants was lower than that in Xinjiang, China. Furthermore, the problem of interlacing between cotton plants was relatively minor. In addition, the abovementioned research obtained orthomosaic images of the entire cotton field. Generating these types of images is complicated and time-consuming. Many scholars have tried to develop various ground-based sensing systems. For example, images obtained by a digital camera installed on a robot platform have been used to estimate the number of cotton bolls based on the images collected by either a 3D sensor system, the estimated value of cotton bolls, or the lint obtained from the point cloud [18]. With the calculation of the number of cotton bolls in the field, accurate cotton yield predictions can be provided [19]. However, in accordance with the actual conditions of fields with high-density cotton, the movement of the ground sensing platform and image acquisition are difficult, and the estimation accuracy may be affected. Using UAVs at lower flying altitudes to acquire images as ground-based sensing platforms can solve the problem of restricted movement.

In addition, significant advances in data collection and computing in recent years have promoted the rapid development of deep learning (DL). As a powerful feature learning algorithm, DL is superior to traditional feature extraction methods in many fields. Li et al. [20] used a full convolutional network (FCN) and interference region removal module to segment the remote sensing data of cotton in the field. Ma et al. [21] proposed the EarSegNet semantic segmentation method, which can achieve accurate segmentation of wheat ears from canopy images acquired during the flowering period. However, the images used in previous studies were all taken from a fixed platform, and UAV images were not included.

In this study, a method based on DL and UAV low-altitude imaging was proposed to estimate the yield of density planted cotton after defoliation. Low-altitude UAV imaging was used for image collection, and semantic segmentation was then directly applied to pixel-level classification of the original UAV images to acquire the images of cotton bolls. Finally, a yield estimation model was constructed based on the pixels ratio of opening cotton bolls. The specific objectives were as follows: 1) to directly perform efficient and accurate segmentation of cotton boll images collected using UAVs during the defoliation period; 2) to construct and verify a yield estimation model based on a single image feature; and 3) to evaluate the effect of multiple cotton fields on yield estimation.

2 Results and Discussion

2.1 Performance evaluation

In this study, four segmentation models were recombined using the designed compilation and decoding blocks, namely, Model1, Model2, Model3, and Model4. The

trained models were used to test the concentration of 800 images in the test sets and we present the results of the experimental evaluation of the selected segmentation approaches in terms of mIoU, Recall, Precision and F1-Score. Based on the results of the testing dataset(Table 1), Model1 has the lowest accuracy, while its recall was the best. Model4 had the highest accuracy among the tested models and it reached on average mIoU=77.13%, recall=84.71%, precision=90.82%, F1-Score=87.93%. Under the same number of convolution blocks, the models performed better after dilated convolution was added. The results for the training and testing of other algorithms on the same image data are shown in Table 2. The test results show that the modified model superior to the original SegNet model and traditional machine learning algorithms (support vector machine (SVM) and random forest (RF)) in cotton field image segmentation. This is probably due to that the deep learning model has a series of convolution structure to extract additional features without needing to design it manually.

Fig. 1 shows the segmentation results of the abovementioned network models on cotton field images with complex backgrounds. Model 1 has the worst segmentation effect, especially for ground images under illumination, for which a large area of incorrect segmentation is exhibited. Model 2 performs better than Model1, However, there are still some errors in the segmentation at the edges of the image. Model 3 and Model4 have better segmentation effects; in particular, the Model 4 exhibits a better segmentation effect on reflective ground, and within the cotton bolls, it can classify cotton leaves as bolls, thereby demonstrating better segmentation logic.

Table 1 Comparison of results of different encoder and decoder methods

CD-SegNet	Encoder-block +Decoder-block	mIoU (%)	Recall (%)	Precision (%)	F1-score (%)
Model 1	a+a	74.63	88.36	81.35	84.45
Model 2	b+b	74.85	88.77	83.84	85.29
Model 3	c+c	73.48	84.52	86.61	85.35
Model 4	d+d	77.13	84.71	90.82	87.93

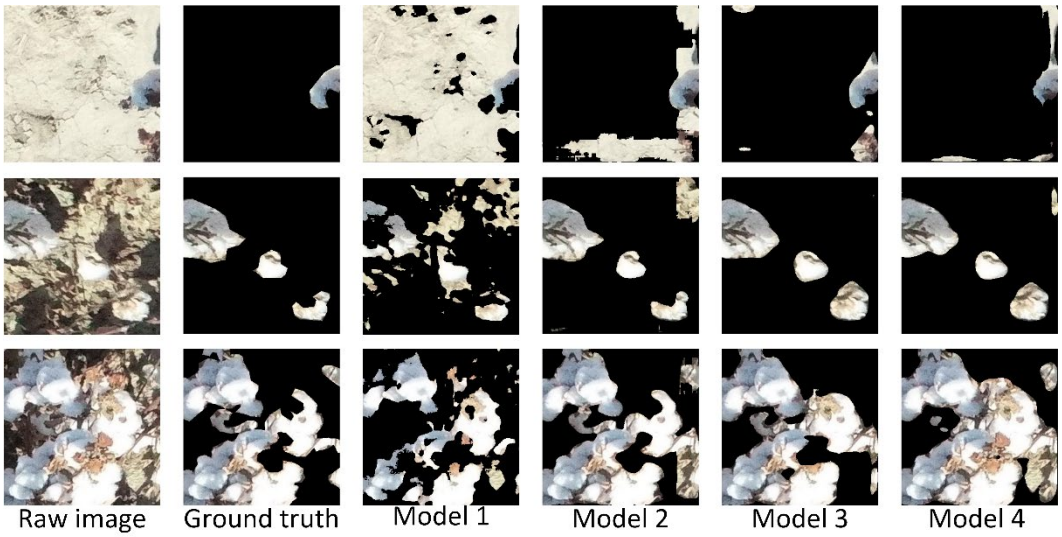

Fig.1 Display of the segmentation of complex background images with different models

Table 2 Segmentation results comparing CD-SegNet with SegNet, SVM, and RF

Model Name	mIoU (%)	Recall (%)	Precision (%)	F1-score (%)
CD-SegNet	77.13	84.71	90.82	87.93
SegNet	74.52	81.36	89.71	85.47
SVM	64.27	78.28	73.42	75.58

RF	58.63	66.84	78.51	72.16

2.2 Sampled Image Segmentation

The segmentation network composed of CD-SegNet was used to segment the 20 original images to calculate the pixels ratio of the cotton bolls. These results were then compared with the actual results. Fig. 2a shows the correlation between the actual results and the segmentation results of the CD-SegNet method. The absolute coefficient R^2 was 0.97. Fig. 2b shows that the relative error of the boll pixels ratio obtained by using this segmentation network in each sample image is 0.27%-14.35%, and the overall relative error is 4.77%. Fig. 2b also shows that when the boll pixels ratio is less than 25%, the relative error is larger, and the actual value is less than the segmentation value; when the boll pixels ratio is greater than 25%, the relative error decreases, and the actual value was greater than the segmentation value. By reviewing the segmented image, we found that this phenomenon was caused by the misalignment of the exposed ground and the boundary of cotton bolls. Therefore, the results show that the CD-SegNet method can accurately segment the cotton boll pixels and calculate its area ratio. However, in some cases, its performance may be limited by the light and background conditions of the image.

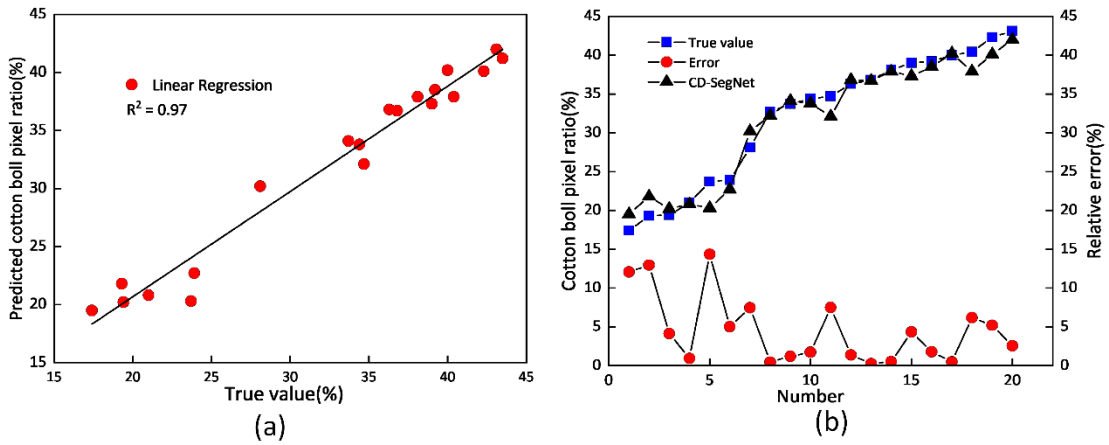

Fig.2 Comparison of the CD-SegNet network segmentation results with the actual area ratio of cotton bolls in the sampled image: (a) correlation between the actual data and the results of the CD-SegNet method; (b) relative error analysis

2.3 Yield estimation of Cotton Field

Using the cotton boll pixels ratio in the cotton field images calculated from the CD-SegNet segmentation, the yield of each cotton field is estimated, and the estimated yield value is compared with the harvested yield of the cotton picker. See Table 3. The relative error of the yield estimate is 0.67%- 10.5%. The relative error of the estimated yield increases with the increase of the actual yield. The UAV images obtained in this study are vertical orthographic images. In the vertical view, the lower layer of cotton bolls may be obscured by the upper layer of cotton bolls along with other sticks and leaves. In the same area, a higher yield means that more cotton bolls are obscured.

Table 3 Use of SegNet segmentation to estimate the area ratio of cotton bolls in each cotton field

Cotton field number	Actual yield (kg ha^{-1})	Estimated yield (kg ha^{-1})	Difference between the actual and estimated yields (kg ha^{-1})	Relative error of yield estimation (%)
1	5090	5124	34	0.67
2	6480	7158	678	10.5
3	5350	5116	234	4.4
4	5843	6391	548	9.4
		Average error		6.2

3 Conclusion

In this work, we proposed and evaluated a cotton field yield estimation model that used DL image processing technology to segment cotton field images acquired by low-altitude UAV platform and then uses the calculation of the segmented cotton boll pixels ratio as an input variable for cotton field yield estimation. The established model can segment the cotton boll pixels with a relative error of 0.27%-14.35% and an R-square of 0.97 with 20 cotton field images of 38 m^2 and can estimate the yield of 4 cotton fields

of 38 hectares with an average error of 6.2%. The results of this study verified the feasibility of estimating cotton field yield using low-altitude UAV imaging. The proposed method will help realize the estimation of cotton yield on the scale of plots and urban areas while improving the efficiency of cotton yield statistics in Xinjiang. This will provide agricultural scientists, agricultural government departments, and crop managers with more accurate crop information, enabling them to make scientific decisions. In the future, we will try to apply an approach of layered yield to reduce the yield estimation error of the cotton field with high density.

4 Methods

4.1 Data Acquisition and Experimental Platform

The experimental site including four fields were located at Tuanjie Farm (44°13′09.3″N, 88°16′27.3″E) in Fukang City, Xinjiang Uygur Autonomous Region, China (Fig. 3a). The densely planting pattern (alternate wide-narrow rows (66 cm and 10 cm, respectively)), shown in Fig. 4, is widely used in Xinjiang, combined with plastic film and drip irrigation. The planting density was approximately 263,000 plants/ha. Images were collected by an industry-grade quadcopter (MATRICE200 V1, DJI Inc., Shenzhen, China) equipped with a cloud platform ZENMUSE X4S and a FC6510 camera. The camera has a fixed focal length of 8.8 mm, F/208-11 focal ratio, and field of view (FOV) of 84°. The resolution of image was 5472×3078 pixels (JPG format). Data collection was from October 11 to 18, 2020, after cotton defoliation.

In natural conditions, to maximize the proximity to cotton while avoiding disturbance from the UAV rotor airflow on the cotton plants, the flying height was set to 5 m. The image resolution at this point was 0.15 cm/pixel. There were 20 sampling

areas set up in cotton field No. 1. Each measuring point was 23,000 cm^2 (230 cm × 100 cm). Coloured flags were inserted at the four corners of the sampling area to label the boundary. To accurately connect the sampling area images and the output data, waypoint photos were taken for the sampling area images (Fig. 5a) based on the designed flight route shown in Fig. 3b, and the cotton was manually picked and measured with a gram-level precision balance. The cotton in the entire field was harvested using a cotton harvester (John Deere CP690, USA) and transported to a cotton factory for weighing.

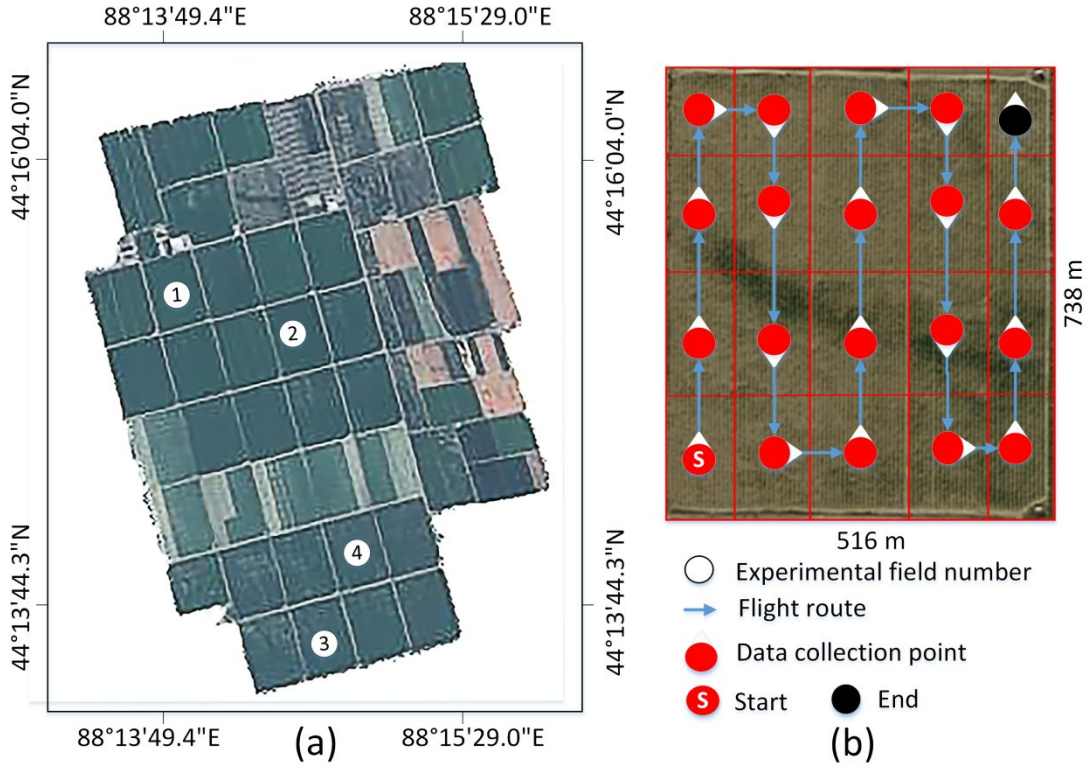

Fig.3 Experimental cotton field location:(a) Study area; (b) image collection design

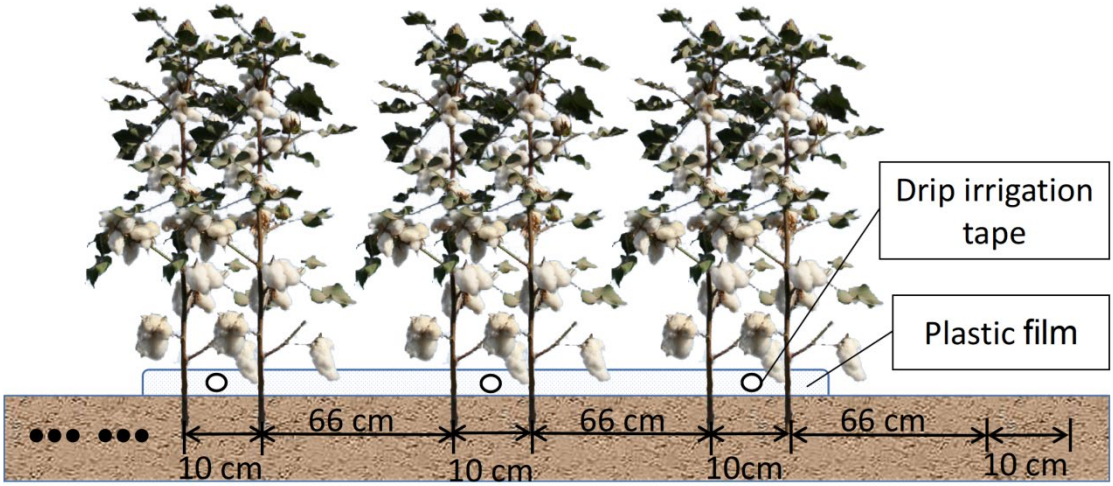

Fig.4 The densely planting pattern of cotton with wide-narrow row in Xinjiang, China

To properly train the developed DL model, a large number of cotton field images were obtained (initial dataset). If the training data are insufficient and non-representative, the model could not learn the potential discriminative features of cotton, resulting in poor generalization performance. Therefore, after image collection (Fig. 5a), flight shooting outside the sampling area was carried out, and a total of approximately 400 high-resolution RGB images were obtained (Fig. 5b), which were used to train the cotton images of the model. Limited by computer capabilities, the original images were too large to directly train the DL model. Based on factors such as different backgrounds, lighting, and growth conditions, 20 images were selected from the acquired images, and an area of 300×300 pixels was extracted (Fig. 6). The images of 4000 cotton plants were obtained in total. All image data were labelled using Python's labelme application. Each pixel of image was labelled as cotton or background. The labelled images were binary images.

Fig.5 Partial original images of cotton fields taken by a UAV over (a) the sampling area; (b) other areas

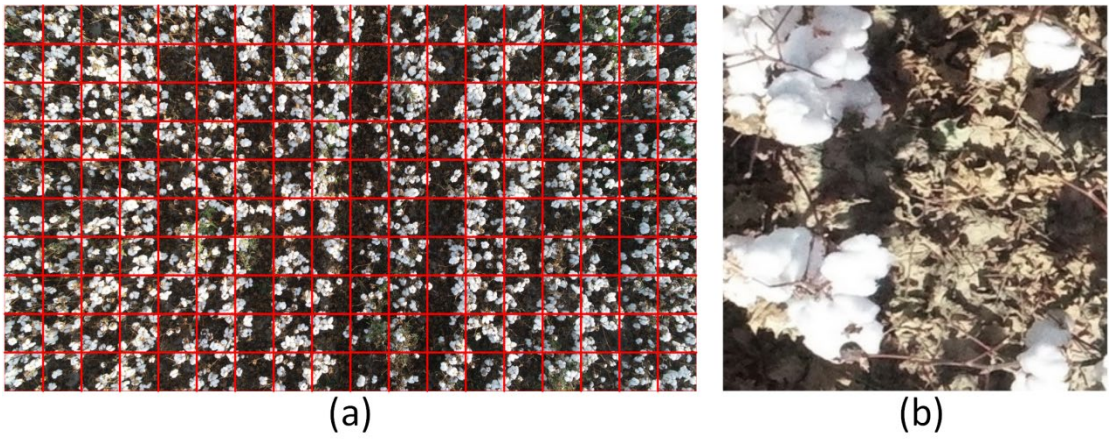

Fig.6 Image cropping (a) Cropping guides on the original image; (b) cropped image (300×300 pixels)

4.2 Image Feature Analysis

Li et al.[20] believed that image segmentation in natural environments was a specific computer vision task that could not be simply attributed to the factors such as shape, texture, colour, and pattern recognition. The background of a cotton field image was changeable and complex mainly due to three issues. First, the sunlight in Xinjiang is strong during the day, and cotton field images can be saturated (Fig. 7a and b), which make the cotton bolls appear very similar to the background, leading to the difficulty in distinguishing them by a single feature (colour, shape, and texture). Moreover,

backgrounds such as cotton leaves (Fig. 7c), cotton hull (Fig. 7d), cotton stem (Fig. 7e), and weeds (Fig. 7f) partially occlude the cotton area, and the occluded area becomes part of the background. Finally, orthographic imagery was employed for the UAV images in this study. Therefore, the lower bolls (Fig. 7g) and the ground (Fig. 7h) are blocked by the upper layer of the cotton plant, resulting in uneven illumination.

According to the above analysis, the cotton feature extraction method needs to meet the following requirements:

(1) Shallow feature information and high-level semantic information can be extracted simultaneously;

(2) Multiscale local information is included;

(3) The extracted features are insensitive to changes in light intensity.

Usually, the manual extraction of brightness, edges, texture, colour, and other shallow visual features from images cannot meet these requirements very well. Therefore, this study used semantic segmentation to resolve this issue.

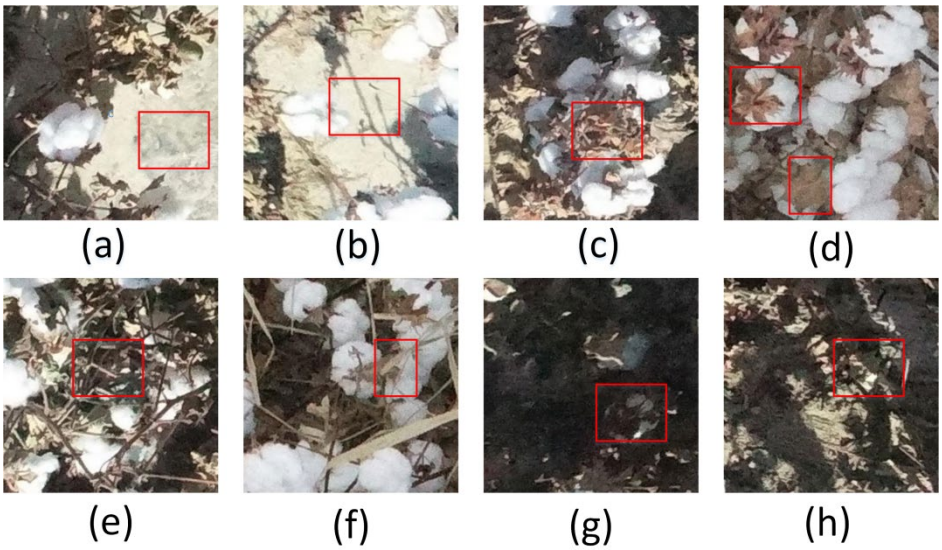

Fig.7 Some typical backgrounds in cotton segmentation: (a) weeds; (b) cotton shells; (c) cotton sticks; (d) uneven light; (e) shadows; (f) film; (g) cotton leaves; (h) hardened ground

4.3 SegNet Network Architecture

SegNet is based on pixel-level semantic segmentation in a convolutional neural network (CNN), consisting of an encoder and a decoder to form a symmetric network [22]. The encoder comprises 5 coding blocks including a convolutional block and a pooling layer. The convolutional block is composed of a convolutional layer, a batch normalization (BN) layer, and a rectified linear unit (ReLU) layer. Each encoder layer corresponds to a decoder layer. The decoder upsamples the feature map. The upsampling part has more feature channels. The network is used to transfer the context feature information to the higher resolution layer, and ultimately, the feature map size is consistent with the original input image size. The softmax layer is the layer that normalizes the input vector to the probability distribution. Fig. 8 is the diagram of SegNet.

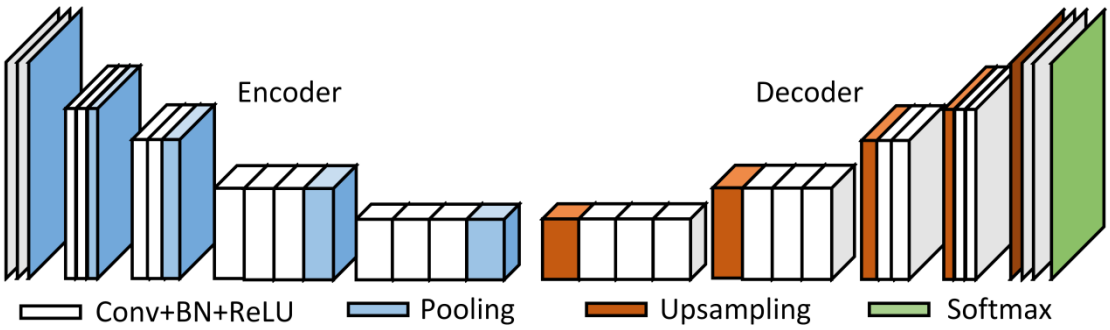

Fig.8 SegNet network

4.4 CD-SegNet Network

The cotton boll dilated convolution SegNet (CD-SegNet) network redesigns and combines the coding and decoding blocks based on the SegNet framework. The number

of convolutional blocks in the encoding block and the decoding block is reduced, and dilated convolution is adopted. The reduction in the number of convolutional blocks can effectively reduce the parameters and improve segmentation efficiency. However, the corresponding receptive field is reduced. Based on a traditional CNN, the dilated convolutional network uses additional magnified two-dimensional filters to increase the size of the receptive field without increasing the calculation parameters. The dilation rate is an important parameter in dilated convolution and represents the degree of expansion. The model in this paper uniformly uses a 3×3 convolution block to replace the 7×7 and 5×5 convolution kernel blocks, which require more memory. As shown in Fig. 9(a), when r=1, the receptive field of the input image corresponding to the feature map is 3×3 without dilation. When (b) r=2, the receptive field increases to 5×5, and (c) has the same receptive field size, However, the number of parameters is reduced by half. For the problem of image segmentation, since it is necessary to predict the pixels, the feature map must be upsampled to obtain a feature map with the same size as that of the original image. This process inevitably results in the loss of some information. As shown in Fig. 10, we also use the addition of dilated convolution to reduce information loss.

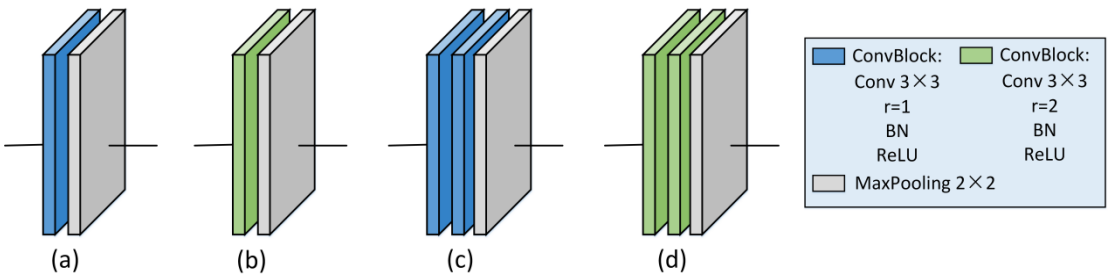

Fig.9 Coding block architecture

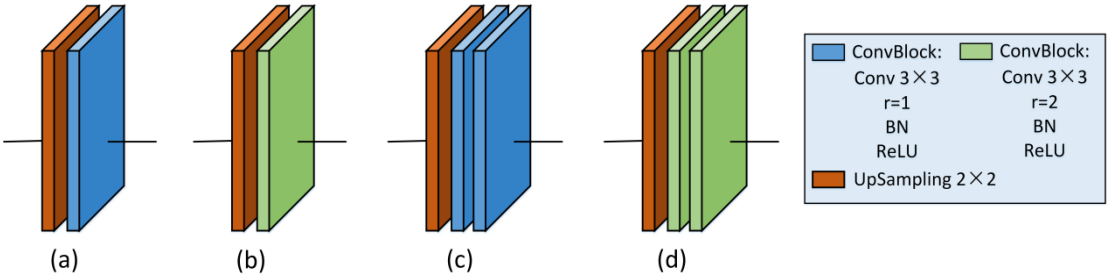

Fig.10 Decoding block architecture

4.5 Model Training

This experiment was conducted using the Windows 10 desktop operating system on an Intel(R) Gold6126 CPU processor, with a default frequency of 2.60 GHZ and memory of 64 GB. The graphics card used was an NVIDIA GeForce RTXTM2060 (with 6G video memory), and the Python version was 3.6, compiled on Jupyter in Anaconda. The DL framework used Pytorch, and a combination of Cuda10.0 and cudnn7.4.1.5 was used for GPU acceleration to improve the model training speed. The model gradient descent adopted the adaptive momentum stochastic optimization method (Adam). The learning rate was 0.001, and the beta first-order and second-order attenuation coefficients were set to 0.9 and 0.98. The training-related parameters are shown in Table 4.

Table 4 Relevant parameters of the deep learning segmentation model training

GSD	Image size	Epoch	Learning rate	Batch size	Sample number of training set	Sample number of validation set
0.15	300×300	50	0.001	64	3200	800

4.6 Evaluation Metrics

In this study, the pixel accuracy (PA), recall, mean intersection over union (MIOU), and F1 score (F1-score) were used to evaluate the segmentation accuracy. All the evaluation indicators used were calculated from the parameters in the confusion matrix

(Table 5). In the model accuracy evaluation, the confusion matrix was mainly used to compare the predicted values and the actual values and was calculated by comparing the position of each actual pixel with the position of the predicted pixel.

PA refers to the percentage of correctly classified pixels in the total pixels (Equation 2). Recall is the proportion of all actual values where the predicted value is the actual value (Equation 4). The intersection ratio (intersection over union, IOU) is a standard metric used to evaluate the accuracy of semantic segmentation (Equation 1). MIOU refers to the average of all categories of IOU (Equation 5). F1 -score was the harmonic mean of the precision and recall and was used in statistics and as indicator to measure the accuracy of a binary classification (Equation 3).

$$IOU = \frac{TP}{TP + FP + FN} \tag{1}$$

$$CPA = \frac{TP}{TP + FP} \tag{2}$$

$$F1\ Score = \frac{2CPA * Recall}{CPA + Recall} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$MIOU = \frac{1}{k+1}\sum_{i=0}^{k}\frac{TP}{TP + FP + FN} \tag{5}$$

Where TP and TN stand for the number of pixels correctly classified for the cotton and non-cotton classes, and FP and FN stand for the number of misclassified pixels. Coefficient of determination (R^2) were used to quantify the accuracy of model segmentation.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(t_i - c_i)^2}{\sum_{i=1}^{n}(t_i - \overline{t_i})^2} \tag{6}$$

Where t_i and c_i are the number of pixels of cotton bolls segmented by the model and the actual number of pixels of cotton bolls in the image, respectively, and $\overline{t_i}$ is the

341 average value of the actual number of pixels of cotton bolls in the image.

342 Table 5 Confusion matrix

Confusion matrix		True value	
		Positive	Negative
Predictive value	Positive	True positive (TP)	False positive (FP)
	Negative	False negative (FN)	True negative (TN)

343

344 **4.7 Yield estimation analysis**

345 Regression analysis is a statistical analysis method for determining the
346 interdependent quantitative relationship between two or more variables. Linear
347 regression is one of the most widely used regression analysis methods, and it is also the
348 preferred regression analysis method. Linear regression analysis model with SciPy
349 calculation library was used to analyse the relationship between the boll pixels ratio in
350 the sample's pixels and the actual yield of the cotton field. A total of 20 samples were
351 used in this study, 15 of which were used for linear regression model construction, while
352 the remaining 5 were used for model testing. Once the regression model was obtained,
353 the output of the test samples was calculated through the regression model, and the
354 relative error that exists between the actual yield was calculated. The Equation 7 was
355 used for calculating the pixels ratio of cotton bolls.

356 $$\text{Cotton boll pixels ratio}=\frac{\text{Number of cotton boll pixels in the image}}{\text{Total number of pixels in the image}} \quad (7)$$

357 ## Declarations

358 ### Acknowledgements

359 The authors gratefully acknowledge the financial support provided by the National
360 Natural Science Foundation of China (31760343). The authors also would like to thank
361 Hao Li, Zhiqiang Zhai and Yangyang Li for their assistance in the experiment.

Author's contributions

Fei Li and Jingya Bai managed UAV flight for aerial imagery, analysed the data and wrote this paper under supervision of Ruoyu Zhang and Mengyun Zhang.All authors reviewed and revised the manuscript.

Author's details

[1] College of Mechanical and Electrical Engineering, Shihezi University, Shihezi, 832000, China; [2] Key Laboratory of Northwest Agricultural Equipment, Ministry of Agriculture, P.R. China, Shihezi, 832000.

Funding

This project has received funding from The National Natural Science Foundation of China project 31760343. International Science and Technology Cooperation Promotion Program of Shihezi University project GJHZ201705.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

[1] Deng YH, Ning S. The stasus quo of the development of machine-picked cotton in Xinjiang and the solutions and prospects for several issues[J]. Cotton Science, 2020, 42(05):26-29.

[2] Cao WB, Liu JD; Zhao LB, et al. Research on optimal time selection for cotton yield estimation by remote sensing in northern Xinjiang[J]. China Cotton, 2007(03): 10-11.

[3] Meng L, Zhang X-L, Liu H, et al. Estimation of Cotton Yield Using the Reconstructed Time-Series Vegetation Index of Landsat Data[J]. Canadian Journal of Remote Sensing, 2017, 43(3): 244-255.

[4] He L, Mostovoy G. Cotton Yield Estimate Using Sentinel-2 Data and an Ecosystem Model over the Southern US[J]. Remote Sensing, 2019, 11(17).

[5] Dalezios N R, Domenikiotis C, Loukas A, et al. Cotton yield estimation based on NOAA/AVHRR produced NDVI[J]. Physics and Chemistry of the Earth Part B-Hydrology Oceans and Atmosphere, 2001, 26(3): 247-251.

[6] Meng L, Liu H, Zhang X, et al. Assessment of the effectiveness of spatiotemporal fusion of multi-source satellite images for cotton yield estimation[J]. Computers and Electronics in Agriculture, 2019, 162: 44-52.

[7] Duan T, Chapman S C, Guo Y, et al. Dynamic monitoring of NDVI in wheat agronomy and breeding trials using an unmanned aerial vehicle[J]. Field Crops Research, 2017, 210: 71-80.

[8] Jin X, Liu S, Baret F, et al. Estimates of plant density of wheat crops at emergence from very low altitude UAV imagery[J]. Remote Sensing of Environment, 2017, 198: 105-114.

[9] Cuaran J, Leon J. Crop Monitoring using Unmanned Aerial Vehicles: A Review[J]. Agricultural Reviews, 2021, 42(2): 121-132.

[10] Turner D, Lucieer A, Malenovsky Z, et al. Spatial Co-Registration of Ultra-High Resolution Visible, Multispectral and Thermal Images Acquired with a Micro-UAV over Antarctic Moss Beds[J]. Remote Sensing, 2014, 6(5): 4003-4024.

[11] Zhang C, Kovacs J M. The application of small unmanned aerial systems for precision agriculture: a review[J]. Precision Agriculture, 2012, 13(6): 693-712.

[12] Ashapure A, Jung J, Chang A, et al. Developing a machine learning based cotton yield estimation framework using multi-temporal UAS data[J]. Isprs Journal of Photogrammetry and Remote Sensing, 2020, 169: 180-194.

[13] Stroppiana D, Migliazzi M, Chiarabini V, et al. RICE YIELD ESTIMATION USING MULTISPECTRAL DATA FROM UAV: A PRELIMINARY EXPERIMENT IN NORTHERN ITALY[C]. IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2015: 4664-4667.

[14] Huang Y, Brand H J, Sui R, et al. COTTON YIELD ESTIMATION USING VERY HIGH-RESOLUTION DIGITAL IMAGES ACQUIRED WITH A LOW-COST SMALL UNMANNED AERIAL VEHICLE[J]. Transactions of the ASABE, 2016, 59(6): 1563-1574.

[15] Feng A, Sudduth K, Vories E, et al. Cotton yield estimation based on plant height from UAV-based imagery data[C]. 2018 ASABE Annual International Meeting, 2018:

1.
[16] Feng A, Zhou J, Vories E D, et al. Yield estimation in cotton using UAV-based multi-sensor imagery[J]. Biosystems Engineering, 2020, 193: 101-114.
[17] Xu W, Chen P, Zhan Y, et al. Cotton yield estimation model based on machine learning using time series UAV remote sensing data[J]. International Journal of Applied Earth Observation and Geoinformation, 2021, 104.
[18] Dube N, Bryant B, Sari‐Sarraf H, et al. Cotton boll distribution and yield estimation using three‐dimensional point cloud data[J]. Agronomy Journal, 2020, 112(6): 4976-4989.
[19] Sun S, Li C, Paterson A H, et al. Image processing algorithms for infield single cotton boll counting and yield prediction[J]. Computers and Electronics in Agriculture, 2019, 166.
[20] Li Y, Cao Z, Xiao Y, et al. DeepCotton: in-field cotton segmentation using deep fully convolutional network[J]. Journal of Electronic Imaging, 2017, 26(5).
[21] Ma J, Li Y, Liu H, et al. Improving segmentation accuracy for ears of winter wheat at flowering stage by semantic segmentation[J]. Computers and Electronics in Agriculture, 2020, 176: 105662.
[22] Badrinarayanan V, Kendall A, Cipolla R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation[J]. Ieee Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.