Cotton Detection Using YOLOv5

Rahul Mapari
Department of E&TC, Pimpri Chinchwad College of Engineering and Research, Ravet,
Pune, India
rahul.mapari@pccoer.in

Anshu Varghese
Department of E&TC, Pimpri Chinchwad College of Engineering and Research, Ravet,
Pune, India
anshu.varghese_etc2020@pccoer.in

Aditya Deshmukh
Department of E&TC, Pimpri Chinchwad College of Engineering and Research, Ravet,
Pune, India
aditya.deshmukh1_etc2020@pccoer.in

Abhimanyu Kanase
Department of E&TC, Pimpri Chinchwad College of Engineering and Research, Ravet,
Pune, India
abhimanyu.kanase_etc2020@pccoer.in

Atharva Junonikar
Department of E&TC, Pimpri Chinchwad College of Engineering and Research, Ravet,
Pune, India
atharva.junonikar_etc2020@pccoer.in

Ayushi Singh
Department of E&TC, Pimpri Chinchwad College of Engineering and Research, Ravet,
Pune, India
ayushi.singh_etc2020@pccoer.in

Prajit Nair
Department of E&TC, Pimpri Chinchwad College of Engineering and Research, Ravet,
Pune, India
prajit.nair_etc2020@pccoer.in

Abstract— Cotton Harvesting has always been a labor-intensive and time-consuming task which involved significant challenges. Manual harvesting resulted in inconsistencies in yield and quality. In traditional systems the cotton blooms could not be detected accurately due to obstruction due to leaves, or detecting the sky instead of the cotton based on features like color, etc. To address this issue Cotton Harvesting Rover implements a robotic system which can detect cotton blooms accurately using computer vision technology. The system autonomously identifies the cotton blooms amongst the fields based on particular features and then picks them using a robotic arm. This study reviews the state-of-the-art deep learning methods for object detection. This review paper will provide a comprehensive study of various deep learning methods which can be useful for automatic detection and various image processing methodologies. Additionally, it offers performance indicators for every technique used, including F-1 score, accuracy, and precision. The climatic condition for cotton productivity in various regions has been discussed. This paper gives a review about the alternative to traditional methods which needed prior knowledge, processing of large data, and reducing difficulties caused due to computer hardware.

Keywords—Cotton Picking, Computer vision, Deep learning, Object detection, YOLOv5.

I. INTRODUCTION

Traditionally cotton picking was done by using methodologies such as hand picking, chemical defoliation, mechanical stripping, manual picking machines, etc. The idea of the project is to develop an advanced cotton-picking robot that incorporates computer vision to enhance and upgrade the efficiency and accuracy of cotton harvesting.

Computer vision is a field that allows computers to extract information from images and other visual inputs. Based on this information it takes necessary actions. It trains the machines to perform according to the functionality that is indicated to be done. Computer vision can be used in various applications such as the detection of faces, video capturing, and tracking moving objects. Using image processing and computer vision the cotton bolls can be detected based on their features and the robotic arm aims to pick them and store them in a container. A deep learning object detection model called YOLOv5 (You Only Look Once version 5) is employed for cotton blossom detection in real-time. The process for successful cotton boll detection includes data collection, labeling the data, selecting the version of the model, training the model by providing the annotated dataset, fine-tuning, and then evaluating the results based on test datasets.

In India and other emerging nations, farmers have an efficient method for harvesting cotton. Despite recent substantial developments, there are still challenges with its adoption in India. In affluent nations, machinery handles all aspects of cotton picking. The Indian agriculture sector has become more mechanized due to rising labor costs and a shortage of labor [1]. According to Haolu Li, remote sensing imageries were used for accurate identification of cotton crop. The deep-learning model used was DenseNet, which is a multidimensional densely connected convolutional network. Remote sensing techniques benefit viewpoints like growth monitoring, disease identification, area estimation and multiple parameters. SVM can do agricultural classification in which the support vectors are limited and the training dataset can be reduced without altering the accuracy of classification. Remote sensing images are obtained from satellites, unmanned air vehicles and unmanned ground vehicles [2]. Xu et al detail the procedure for gathering aerial photos of cotton fields, analyzing the photos to determine possible bloom locations, and employing a convolutional neural network to categorize the possible blooms as either non-blooming objects or cotton flowers [3].

[6] Has described how to adjust picture color pixel values and create upper/lower boundaries for white cotton bolls utilizing a fuzzy reasoning-based approach in conjunction with RGB and HSV color channels for cotton detection. Amanda Isaac et al. conclude that by combining the bitwise-masking and PCA methods on diced images, they were successful in reducing both the size of the file and its dimensions. This method outperformed the normal JPEG compression and generated a notable proportion of smaller files while maintaining high-quality reconstructions [7]. Adalberto I. S. Oliveira developed a robot whose purpose is to monitor the cotton crop and move between rows of crops. The control system of the robot consists of an image-based visual servoing method and a fuzzy logic-based controller [8]. According to Zhang, Yan and Yang etc an Improved version of YOLOv5 was presented which integrated DenseNet, attention mechanism, and bi-frequency frequency prediction to correctly and economically identify closed cotton bolls in the field. According to the experiment results, the suggested approach outperforms the original YOLOv5 model as well as alternative approaches like YOLOv3, SSD, and When weighing the simultaneous factors of precision of detection, computational price, model size, and speed, FasterRCNN [10].

Yong Wang and the research team present a novel approach to cotton recognition based on information from color subtraction of different cotton components. Dynamic Freeman chain coding is operated to reduce noise and improve recognition accuracy in order to raise the accuracy rate of cotton detection [11]. Nimkar Amey Sanjay concludes that to create sustainable and successful agriculture, the goal of the project is to identify, harvest, and store cotton using a cotton harvester that makes use of strong image processing, robotics, and digital analytics. The device would use several robotic arms constructed in the harvest zone to pick numerous crops. This machine is ideal for women because it is remote-controlled and electrically operated, eliminating the need for a driver [12]. This work by W. M. Porter explores and validates the optical detection of cotton rows. The row depth is detected using a stereo camera, and the top and bottom portions of cotton row canopy are then calculated using a pixel-based method that assumes a distribution for the high and low pixels. Pixel-based sliding window algorithms and perspective transform are used to detect left and right rows. For the cotton-picking robot to navigate smoothly, the center of the rows is calculated along with the Bayesian score of the detection [13]. The review also covers the many kinds of AI algorithm and sensors that are employed in these systems. The creation of an enhanced YOLOv5 for the identification of damage based solely on appearance in cotton seeds is the main objective of this work. [16]. The creation of a computer vision algorithm that uses YOLOv5m to identify volunteer cotton plants in cornfields is covered in the document [17]. This paper used Point annotation and large-scale fusion-based cotton boll localization approach [19]. An overview of sensors and systems in navigation systems, recognition and detection algorithm classification, and crop row techniques are given in this work [24].

II. FINDINGS FROM LITERATURE SURVEY

The following holes that needed to be focused on were discovered from our extensive deep learning assessment of various cotton-picking approaches.

From the literature survey following gaps are identified:

1. Obstruction caused due to leaves, branches or other materials can slow the picking time.
2. Implementation cost due to hardware, software and maintenance.
3. Autonomous systems must have less maintenance time.
4. Only identifies specific cotton blooming patterns and does not specify other use cases.
5. Could not identify cotton bolls which were hidden behind some branches leaving behind some bolls.

III. GAP IDENTIFICATION

According to the existing literature and the state ofknowledge in object detection, there are some limitations thatmay arise in the methodology, observed findings and some questions which may be unanswered. The below table contains all the gaps/limitations which were identified duringthe analysis of the papers. This paper describes the existing methodology, the design and formulation, methods for data collection and their drawbacks. Based on this analysis it is possible to address issues that had been faced in the previousresearch papers. Around 26 papers were studied, out of which25 gaps were identified as mentioned in below:

The cost and availability of labor are two major issues that Indian cotton farmers must deal with [1]. The use of supervised classification, which necessitates a large quantity of training data is a challenge [2]. The lower surface resolution at higher flight altitudes may prevent the pipeline from precisely detecting cotton flowers [3]. Machine learning techniques do not use automatic feature extraction and need careful tuning of hyperparameters [4]. In real field obstruction caused by branches and other leaves can result in longer picking time. Implementation cost due to hardware, software, and maintenance is high [5]. Detects only white cotton bolls therefore may not be directly applicable to other types of crops or monitoring scenarios. Requires specialized equipment [6]. Specifies the outlines the application case for identifying flowering patterns in cotton blooms, and does not specify other use cases. Does not specify the computational requirement and time complexity [7]. This paper specifies only crop monitoring and using a mobile robot whose accuracy is 8/10 use cases [8]. Potential challenges faced can be field obstacles, theft, cost, etc. Due to the system being autonomous, they must require minimal management time

[9]. Techniques like YOLOv3, SSD, and Faster-RCNN showed less performance in terms of detection, computational costing, model size, and speed [10]. Requires specific lighting conditions, does not work well with certain types of cotton [11]. This paper focuses on the development of a cotton harvester using ML and IP techniques and not on the practical use [12]. New automated machines will have to match the reliability, self-sufficiency, and economic competitiveness of the existing harvesters used in the U.S [13]. The system was struggling to get a boll if it is located behind the stem or branch. Some-times it causes the system to leave behind some bolls [14]. Need for specialized sensors and algorithm calibration, difficulty in comparing AI algorithm performances among papers using different data sets [15]. It is deficient in thorough assessment, comparative analysis, and specific details on a few key areas, including the computational efficiency and the algorithm's performance with coated cotton seeds [16]. ACO algorithm used for generating the optimal path and spot-spray is stochastic in nature and may not provide a globally optimal solution [17]. It's possible that not all of the computational resources needed for MCBLNet training and implementation particularly in real-time applications have been taken into consideration [19]. Testing equipment for robotic arms are costly [20]. Challenges involved in complexity in robotic arm programming [21]. Possible drawbacks could be limited to only color and shape recognition, environmental conditions could impact on the reliability of computer vision [22]. The challenges of differentiating between crop species, the requirement for extreme precision and accuracy in data collection, and the possibility of interference from meteorological elements like wind and rain [23]. Accurate direction finding and self-governing agricultural equipment are challenged by the unstructured and complicated character of the agricultural environment [24].

IV. Image Dataset

The image dataset was produced in June and July of 2023 utilizing images taken from local farmers' farms in Marathwada and information gathered from the internet. The internet-based dataset was gathered from two distinct websites: figshare and Roboflow. The Roboflow dataset contained more than 1421 photographs, of which we used all the data. In a similar vein, a dataset with almost 6000 photos was removed from the figshare site since the model's training accuracy was below 0.60. The preprocessing section includes details on the five classes that is, the five categories into which the Roboflow dataset is divided that make up the dataset. The preparation section contains thorough information on augmentation. The pre-labeled and annotated online dataset we collected aided in the model creation process. The model was tested exclusively using the second dataset, which was gathered from nearby farms. The images were taken from various perspectives, postures, and backgrounds, ensuring that the model's testing would be appropriate for a variety of environments.

Figure 1. displays the real photos from our dataset, which consists of 614 total images. After the augmentation process which is covered in the preprocessing section the number of images increased to 1421. The image resolution for the dataset before preprocessing is 640x640 pixels. The data set is divided in to 85% training,10% validation and 5% testing[27]. Photographs of individual cotton bolls on the plant are included in the dataset to aid the model in identifying cotton bolls up close, while photographs of several cotton bolls on a single plant are included to aid in the detection of cotton bolls at a distance.

Fig. 1. Image Dataset

V. Data Preprocessing

As part of the data analysis process, data preprocessing transforms unprocessed data into a format that can be understood and interpreted by computers and machine learning. To get raw data for our study, we annotated the photos by placing an anchor box around the check that has to be detected. In this instance, there are various stages of the cotton boll's growth, including split, mature, early, and flower. This is divided into five categories by first defining all classes for every growth stage and then, during the annotating process, choosing a certain class each time a bounding box is made. Once every image has been annotated, the next step is data augmentation. This process involves increasing the size of the data by doing things like rotating the images 90 degrees clockwise and counterclockwise, flipping them vertically and horizontally, and adjusting the exposure of the photos between -12% and +12%.

Figure 2. shows the platform where the data annotation and augmentation was done and the dataset was exported in the v5 PyTorch compactible format [27].

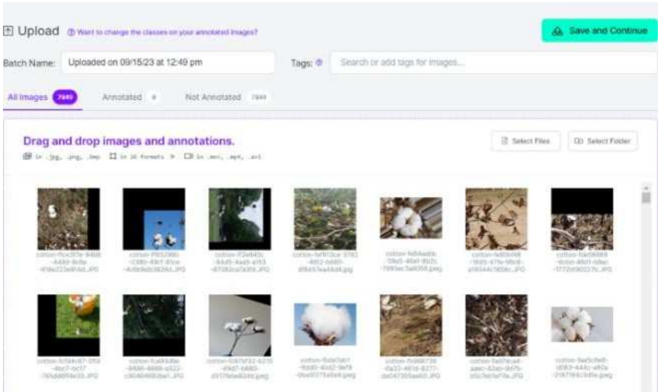

Fig. 2. Annotated Dataset

VI. GENERALIZED METHODOLOGY

Explains in detail how YOLOv5, a pre-trained model based on deep learning, can be specifically trained for a dataset of cotton and, once deployed, be able to identify cotton in any environment. The selection of YOLOv5 or any other ML model is contingent upon the particular tasks at hand and the availability of data for the model's training. YOLO models are employed for high-performance object detection. An image is divided into grid systems by YOLO, and every grid knows what's inside of it.

Several YOLO versions are contrasted in Table 1. Among them, v5 is the most suitable for our use in cotton detection since it is open source, lighter than v7, and more accurate than v4. As can be seen in the table below, version 4 and 7 solely use the COCO dataset, while version 5 employ a variety of D5 datasets for data training. The 's' model is used in the research for cotton detection.

TABLE I. COMPARISON OF YOLO VERSIONS

Parameter	YOLOv4	YOLOv5	YOLOv7
Speed	Varies based on the model selected.	Has good balance	Fastest
Accuracy	varies according to the selected mode	Offers good balance	Highest
Open-source	No	Yes	No
Training data	COCO dataset	Diverse D5 dataset	COCO dataset
Parameters and Computation	Higher generally	Low than v4	Lowest

The explanation of the YOLO cotton detecting model is provided by Fig. 3 :

Cotton Image: Gathering data is the first step in training a model. We have gathered this data from a variety of sources, such as cotton cultivated in a lab or a greenhouse, etc. The training dataset comes from Roboflow.

Image Annotation: An image is annotated using machine learning and artificial intelligence algorithms. Frequently, image annotation is done by human annotators using animage annotation tool to tag relevant information or label images, for as by giving appropriate classes to various items in an image. As stated in the section on image datasets, we developed two classes for our model. Pre-trained models are frequently available for fundamental tasks like classification and segmentation, and with the aid of Transfer Learning and little data, these models may be customized to particular use cases. The data preprocessing section goes into detail about the platform called Roboflow, where the annotation was completed.

Annotated Dataset: Following the platform's annotation process, the data is given to the Colab for model training, with the received data divided into three sections: test, validation, and train.

Image augmentation and preprocessing: The dataset is first enlarged using the augmentation procedure, which is described in detail in the data pretreatment section, prior to exporting it to Colab.

YOLOv5 training: We begin by exporting our previously processed information from the Roboflow platform in v5 Pytorch compatible format, which is necessary to train our v5 model on our custom dataset. Upon completion, a repository code is generated. After using the code to import the dataset into Colab, we set the parameters to begin training. First, we set the image size, to the same as the export size. Next, we set the epochs for this model. We experimented with epochs ranging from 50 to 200 to train our model, however, 150 epoch and batch size of 32 produced the best results. Using this input parameter, we were able to get an accuracy of 0.83 [16] [17].

Train Model Evaluation: Following the completion of testing on unseen data, we apply the trained model to the task of object detection in test set images. Following object detection, the performance is assessed, and measures such as mAP, precision, and the accuracy and efficacy of the model in object detection are analyzed [10].

Detection: Following cotton detection, a box surrounding the cotton appears, seemingly indicating the level of confidence. This level simply indicates how much confidence that the object being recognized is cotton or not (as an example, if the level is 0.8, it indicates an 80% confidence on detected cotton).

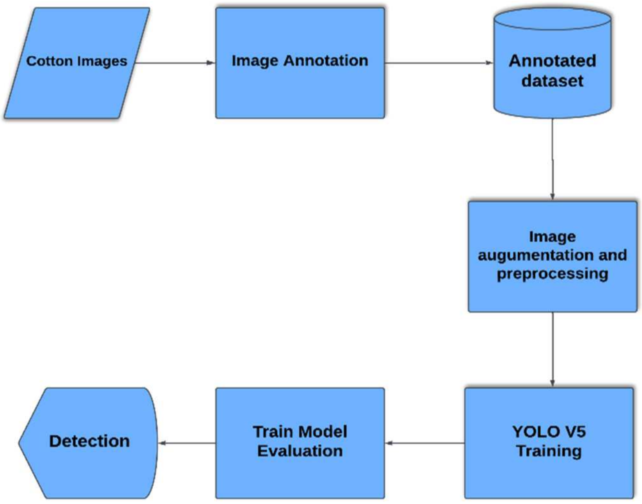

Fig. 3. Generalized flow of YOLOv5 cotton detection model.

VII. PERFORMANCE METRICS

A confusion matrix, also called an error matrix, is a specific table arrangement used in the field of machine learning, specifically with regard to statistical classification problems. It makes the performance of an algorithm easier to visualize. Typically, this algorithm is supervised learning-related; in

unsupervised learning, it is commonly referred to as a matching matrix.

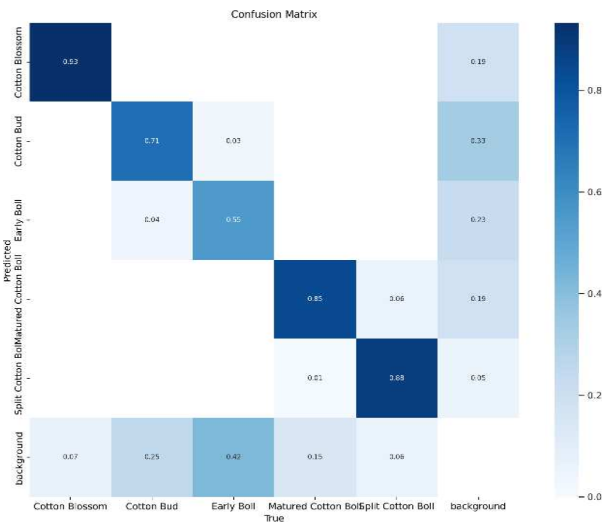

Fig. 4. Confusion Matrix

From Figure 4. we can see confusion matrix that true value for the cotton blossom matches the projected values by 0.93 times, the true value for the bud matches by 0.71 times, and the true value for the early boll matches by 0.55 times. Both the split cotton forecast and the matured boll prediction match the true value by 0.88 times and 0.85, respectively. The background is clearly influencing the values of the cotton buds and the early boll.

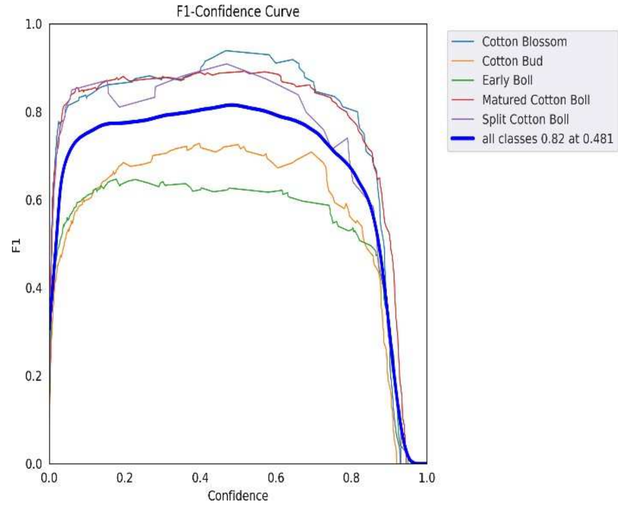

Fig. 5. F1 confidence curve

In Figure. 5. the F1 curve as a reference for all classes, the confidence value of 0.481 maximizes both precision and recall. A greater confidence value is preferred in many situations. This model's confidence value must be 0.8, which isn't far from its highest confidence of 0.82.

$$F1 = 2\left[\frac{precision.recall}{precisio+recall}\right] \quad (1)$$

Where,

$$precision = \frac{TP}{TP+FP} \quad (2)$$

$$recall = \frac{TP}{TP+FN} \quad (3)$$

$$accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (4)$$

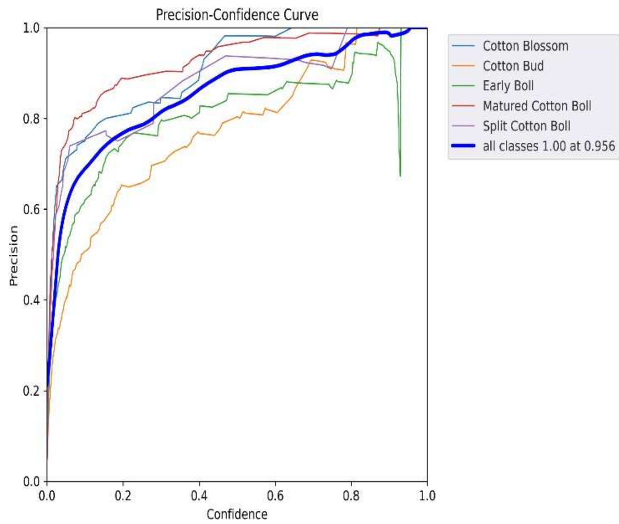

Fig. 6. Precision confidence curve

Figure 6 illustrates the relationship between the model's confidence and precision.

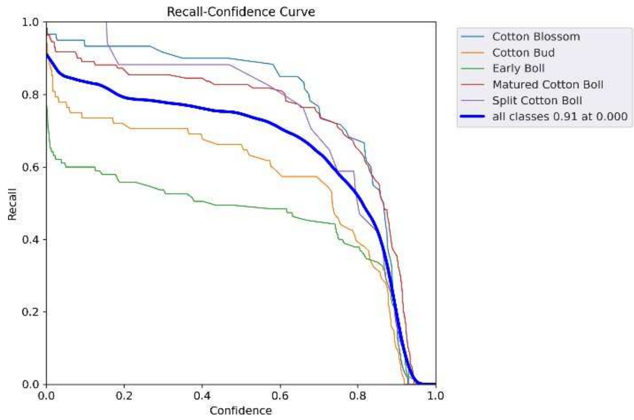

Fig. 7. Precision recall confidence curve

Figure 7 shows the relationship between the trained model's recall and confidence.

VIII. Conclusion

Using the PyTorch compactible dataset and YOLOv5, we were able to successfully build a cotton detection model using the Google Colab platform. Figures 8 and 9 illustrate the model's detection.

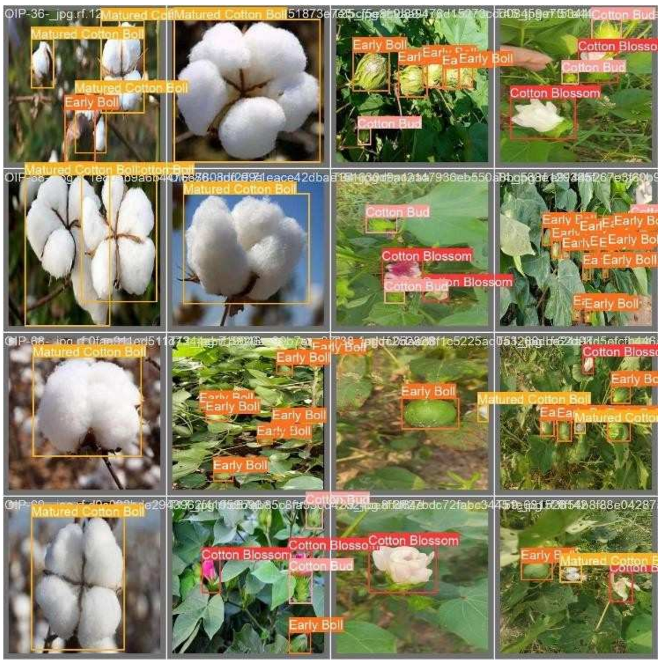

Fig. 8. Labeled image

These annotated photos were uploaded to Google Colab to train the model, as seen in Figure 8. As stated in the pre-processing section, the photos were categorized into five categories. With the help of the bounding boxes, the model was able to identify the characteristics in the fed data.

The model was able to predict the existence of the cotton boll after completing its training. When fig. 8 was entered into the model for testing, the expected output (fig.9.) resembled as the same image which was used for training.

Fig. 9. Predicted image

IX. Future Scope

Integration of a robotic arm with real-time detection: Based on real-time detection from the video feed, the robotic arm can use sensors and actuators to pluck the cotton boll.

Combining Various Sensor Systems: Use a range of sensors, including LiDAR and infrared sensors, to provide the robot with a thorough grasp of the surroundings in the cotton field.

Personalized Path Planning and Self-governing Navigation: To improve the robot's agility in the field, create complex algorithms for autonomous navigation and dynamic path planning.

Technological Developments in Image Processing: Investigate and apply state-of-the-art image processing techniques to improve cotton bloom detection accuracy and speed.

X. References

1. A.S. Deshmukh and A. Mohanty, “Cotton mechanisation in india and across globe: a review,” International Journal of Advance Research in Engineering, Science & Technology, vol. 3, no. 1, pp. 2393–9877, 2016.

2. Li, Haolu, et al. “Identifying cotton fields from remote sensing images using multiple deep learning networks.”Agronomy 11.1 (2021): 174.

3. Xu, Rui, et al. “Aerial images and convolutional neural network for cotton bloom detection.” Journal Of Display Technology, Vol. 8, No. 1,January 2012

4. Saleem, Muhammad Hammad, Johan Potgieter, and Khalid Mahmood Arif. ”Automation in agriculture by machine and deep learning techniques: A review of recent developments.”Precision Agriculture 22 (2021): 2053-2091.

5. K. Fue, W. Porter, E. Barnes, C. Li, and G. Rains,“Center-articulated hydrostatic cotton harvesting rover using visual-

servoing control and a finite state machine," Electronics, vol. 9, no. 8, p. 1226, 2020.

6. Z. Xu, M. A. Latif, S. S. Madni, A. Rafiq, I. Alam, and M. A. Habib,"Detecting white cotton bolls using high-resolution aerial imagery acquired through unmanned aerial system," IEEE Access, vol. 9, pp. 169068–169081, 2021.

7. A. Issac, H. Yadav, G. Rains, and J. M. Velni, "Dimensionality reduction of high-throughput phenotyping data in cotton fields," IFAC-PapersOnLine, vol. 55, no. 32,pp. 153–158, 2022.

8. A. I. Oliveira, T. M. Carvalho, F. F. Martins, A. C. Leite, K. T. Figueiredo, M. M.Vellasco, and W. Caarls, "On the intelligent control design of an agricultural mobile robot for cotton crop monitoring," in 2019 12th International Conference on Developments in eSystems Engineering (DeSE), pp. 563–568, IEEE, 2019.

9. E. Barnes, G. Morgan, K. Hake, J. Devine, R. Kurtz, G. Ibendahl, A. Sharda,G. Rains, J. Snider, J. M. Maja, et al., "Opportunities for robotic systems and automation in cotton production,"AgriEngineering, vol. 3,no. 2, pp. 339–362, 2021.

10. Zhang, Yan, et al. "An improved YOLO network for unopened cotton boll detection in the field." Journal of Intelligent & Fuzzy Systems 42.3 (2022): 2193-2206.

11. Y. Wang, X. Zhu, and C. Ji, "Machine vision based cotton recognition for cotton harvesting robot," in Computer And Computing Technologies In Agriculture, Volume II: First IFIP TC 12 International Conference on Computer and Computing Technologies in Agriculture (CCTA 2007), Wuyishan, China, August 18-20, 2007 1,pp. 1421–1425, Springer, 2008.

12. N. A. Sanjay, N. Venkatramani, V. Harinee, and V. Dinesh, "Cotton harvester through the application of machine learning and image processing techniques", Materials Today: Proceedings, vol. 47, pp. 2200–2205, 2021.

13. K. Fue, T. Georgia, W. Porter, E. Barnes, and G. Rains, "Visual row detection using pixel-based algorithm and stereo camera for cotton-picking robot," 2019 ASABE Annual International Meeting, Boston, Massachusetts, 2019.

14. K. G. Fue, E. M. Barnes, W. M. Porter, and G. C. Rains, "Visual control of cotton-picking rover and manipulator using a ros-independent finite state machine," in 2019 ASABE Annual International Meeting, p. 1, American Society of Agricultural and Biological Engineers, 2019.

15. S. Fountas, I. Malounas, L. Athanasakos, I. Avgoustakis, and B. Espejo-Garcia, "Ai-assisted vision for agricultural robots," AgriEngineering, vol. 4, no. 3, pp. 674–694,2022.

16. Liu, Z., Wang, L., Liu, Z., Wang, X., Hu, C., & Xing, J. (2023). Detection of Cotton Seed Damage Based on Improved YOLOv5. Processes, 11(9), 2682.

17. Yadav, Pappu Kumar, et al. "Computer Vision for Volunteer Cotton Detection in a Corn Field with UAS Remote Sensing Imagery and Spot Spray Applications." arXiv preprint arXiv:2207.07334 (2022).

18. Bhavan, S. Abhinand. "YOLOv5 Crop Detection Deep Learning Model using Artificial Intelligence (AI) and Edge Computing." 2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS). IEEE, 2022.

19. Sun, Ming, et al. "Cotton boll localization method based on point annotation and multi-scale fusion." Frontiers in Plant Science 13 (2022): 960592.

20. Shah, S.K., Mishra, R., Prasanna, D. and Mohapatro, G., 2019. Fabrication and deviation analysis of a 5-axis robotic arm using image processing technique. Materials Today: Proceedings, 18, pp.4622-4629.

21. Reddy, R. and Nagaraja, S.R., 2014, December. Integration of robotic arm with vision system. In 2014 IEEE International Conference on Computational Intelligence and Computing Research (pp. 1-5). IEEE.

22. Rai, Neerparaj, Bijay Rai, and Pooja Rai. "Computer vision approach for controlling educational robotic arm based on object properties." In 2014 2nd International Conference on Emerging Technology Trends in Electronics, Communication and Networking, pp. 1-9. IEEE, 2014.

23. Micheletto, Matias J., Carlos I. Chesñevar, and Rodrigo Santos. "Methods and Applications of 3D Ground Crop Analysis Using LiDAR Technology: A Survey." Sensors 23.16 (2023): 7212.

24. Shi, Jiayou, et al. "Row Detection BASED Navigation and Guidance for Agricultural Robots and Autonomous Vehicles in Row-Crop Fields: Methods and Applications." Agronomy 13.7 (2023): 1780.

25. Gharakhani, Hussein, J. Alex Thomasson, and Yuzhen Lu. "Integration and preliminary evaluation of a robotic cotton harvester prototype." Computers and Electronics in Agriculture 211 (2023): 107943.

26. Adke, Shrinidhi, et al. "Supervised and weakly supervised deep learning for segmentation and counting of cotton bolls using proximal imagery." Sensors 22.10 (2022): 3688.

27. Cotton Boll https://universe.roboflow.com/proj1-gyxkz/cotton_data/dataset/2/images, 31 Jan 2024.