A CNN-BASED APPROACH TO DETECT COVER DAMAGE OF ROUND COTTON MODULES

M. Z. Iqbal
R. G. Hardin IV
Department of Biological & Agricultural Engineering, Texas A&M University
College Station, TX
J. K. Ward
Department of Biological & Agricultural Engineering, North Carolina State University
Raleigh, NC
J. D. Wanjura
USDA-ARS Cotton Production and Processing Research Unit
Lubbock, TX

Abstract

Round cotton modules covered with engineered plastic film are increasingly popular because they largely automate the cotton harvest. Cover damage may occur during handling of the round cotton modules, resulting in plastic contamination in the processed cotton fiber. Plastic contamination costs millions of dollars every year to the US cotton industry. To uphold its reputation for providing some of the cleanest cotton to the international market, the US cotton industry is prioritizing the removal of plastic contamination from cotton. To produce contamination-free fiber, it is important to determine detailed information about plastic cover damage of cotton modules. Farmers and ginners need to identify damaged covers to be able to repair or handle carefully; consequently, an automatic cover damage identification system would be useful. Therefore, the objective of this research was to use different convolutional neural network (CNN) models to classify the cotton modules with damaged covers during the handling process and identify the best-performing one. A single-board computer-based system was developed and installed on a loader to collect images of the cotton modules during the handling process. This system collected images of the cotton modules from different directions during the handling process and stored them with a unique RFID number. The YOLOv5 CNN model was most accurate in detecting cover damage, with training, validation, and testing accuracy of 75.6%, 65.9%, and 78%, respectively. This model was able to provide information about the status of each module from the images collected by the system with 73% accuracy. The model developed in this study might be a useful tool to reduce plastic contamination in processed cotton.

Introduction

Round cotton modules covered with engineered plastic are formed with a similar mechanism to a round hay baler. These modules were first produced using a John Deere 7760 self-propelled cotton picker, released in 2008 to the farmers (Van Der Sluijs et al., 2015). The modules are ejected from the picker after applying plastic wrap with RFID tags, where the information of that particular module is stored with the tag number. This technology largely automates the cotton harvesting process by eliminating extra labor, and the plastic cover protects the seed cotton and provides compressive force to maintain the module density. These features have led to the growers' rapid uptake of this technology. With the increasing use of this technology, plastic contamination is also increasing. Over 84% of extraneous matter calls for plastic contamination in the 2019 ginning season were associated with round module covers (Wanjura et al., 2020). Most likely, the damaged plastic covers are responsible for much of this contamination. Plastic contamination in processed cotton became a concern to the US cotton industry because it accounts for an estimated annual discount of more than $750 million (Pelletier et al., 2020a) due to a diminished reputation of US cotton in the international market.

The ginning industries have already adopted computer vision techniques with low-cost color cameras to detect the plastic contaminations in the ginned fiber, where traditional machine learning methods (i.e., SVM, KNN) were used. While the system generally worked well, certain colors of plastics are challenging to identify from a cotton background. To improve the accuracy of this system, Pelletier et al. (2020b) developed a dataset could be used to improve the existing system by utilizing the advantages of convolutional neural networks (CNNs). Pelletier et al. (2020a) also developed a plastic contamination inspection system where low-cost IP cameras were used to inform the gin personnel about the plastic contamination status of the specific module feeder. If modules with damaged covers can be identified before entering the module feeder and handled more carefully, then there will likely be less plastic that has to be detected for removal by the plastic inspection system, thus reducing contamination (Iqbal et al., 2021).

To detect the cover damage status of modules during the handling process, a monitoring system is needed for handling equipment. After harvest, modules are staged in the field and often loaded on a trailer. The modules are unloaded from the trailer and again staged in the gin yard. Finally, a gin yard truck moves them to the module feeder for ginning. Alternatively, a module truck may be used to transport modules from the staging location in the field to the gin yard. Although a tractor with a 3-point hitch attachment is sometimes used for staging and loading in the field, loaders and module trucks are typically used in these operations. An effective monitoring system for loaders and module trucks would allow automated detection of cover damage before it reaches the module feeder.

For monitoring system development, cameras have become a popular option in various applications in agriculture because of their low cost, reliability, and measurement speed (Pelletier et al., 2020; Condotta et al., 2020). Using a set of cameras to monitor the module during the handling process would be effective and economical. Developing models using neural networks creates new possibilities in agriculture for data-intensive research and precision analysis. Identifying diseases, fertilizer requirements, and fruit for harvesting has been done using machine vision and Convolutional Neural Networks (CNN's) (Agarwal et al., 2020, Indira et al., 2021). This technology should also be able to detect plastic wrap damage on round cotton modules from the captured images during the handling process. The combination of RFID and machine vision technology might be an effective solution for plastic cover damage detection of the individual modules during the handling process. Over the last decades, RFID has become one of the most promising technological innovations in manufacturing and industry (Urso et al., 2020). This technology also provides significant advantages in smart agriculture by reducing the possibility of human error in data storage with comparatively low cost and labor (Ruiz-Garcia and Lunadei, 2011).

Therefore the goal of this research was to develop a system to automatically identify the cover damage of the round cotton modules. The specific objectives were:

- Improve the module image collection system (Wang et al., 2022) attachable to a loader or module truck
- Compare damage identification accuracy of different CNN models to select the best performing one

Materials and Methods

Image Collection System Development

To achieve the first objective, a portable system was developed using three basic units, which were the upgraded central processing unit (CPU) (Raspberry Pi 4B, Raspberry Pi Ltd., Cambridge, UK), RFID reader (ThingMagic Micro-LTE, Jadak Technologies Inc., NY, USA), and a set of better USB cameras (UC50MPD, Spinel USA LLC, CA, USA) with 120 DB wide dynamic range (Figure 1a). The cameras and RFID reader were connected with the CPU through serial communication, and images from the cameras were stored in the memory of the CPU. To ensure the power supply and serial communication of the cameras, a USB extender (Shenzhen Sgeyr Electronics CO., Ltd., Guangdong, China) was used which was able to transfer the images to the CPU with high speed through a Cat6 cable.

The RFID reader and the cameras were connected to the CPU using the USB serial port. The electronic components were mounted in a metal enclosure. A display was attached to the enclosure, used to indicate system status and provide commands to the system. A Python program was created to regulate the cameras and RFID module during the operation. The cameras turn on and take pictures if the RFID module finds at least one RFID tag. If there is no RFID tag present, the cameras remain off. The system turns on automatically and starts collecting data with the engine key switch of the loader or the truck and saves them with the unique RFID of the round module. The whole system was powered using the batteries of the loader or truck.

Two separate systems were attached to a loader and a module truck at a commercial gin. A Caterpiller 930M loader with a spear-type holder for the round cotton modules was equipped with the system (Figure 1b). Two cameras were placed on the left and right sides, and one camera was attached under the chassis of the loader to capture the images of the modules from three different angles. Two RFID antennas were also attached in the left and right mast of the loader. The masts were built using square steel tubing and bolted to the loader. A Kenworth T480 yard truck was also equipped with the system (Figure 1c). Here two cameras were attached in the top, and two cameras were attached in the bottom of the truck to capture the images of the module from four different angles. The RFID antennas were placed near the top cameras to identify the RFID tags of the module accurately. Using these systems, images of the round cotton modules were collected from the cotton gin in the 2021 ginning season.

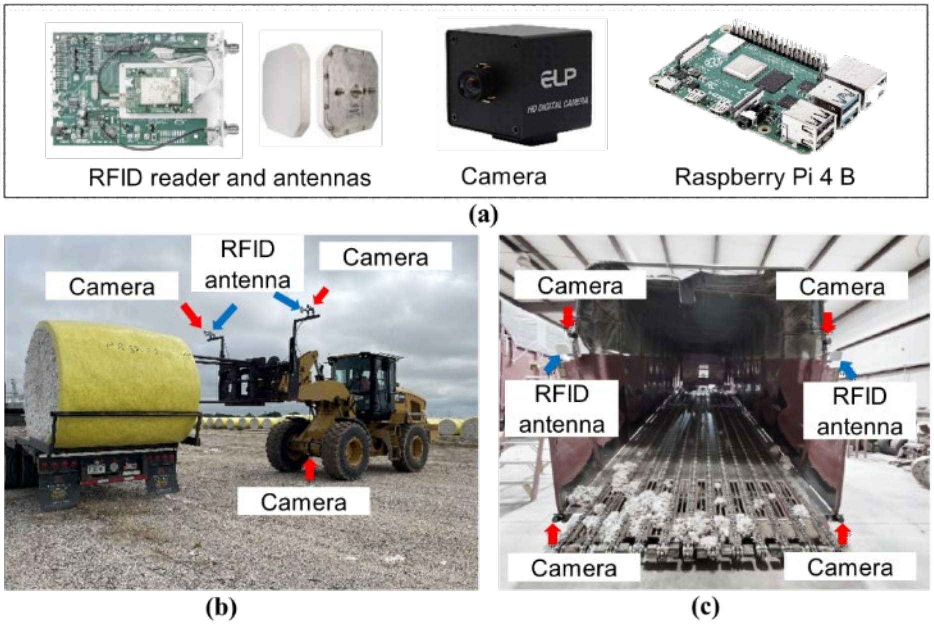
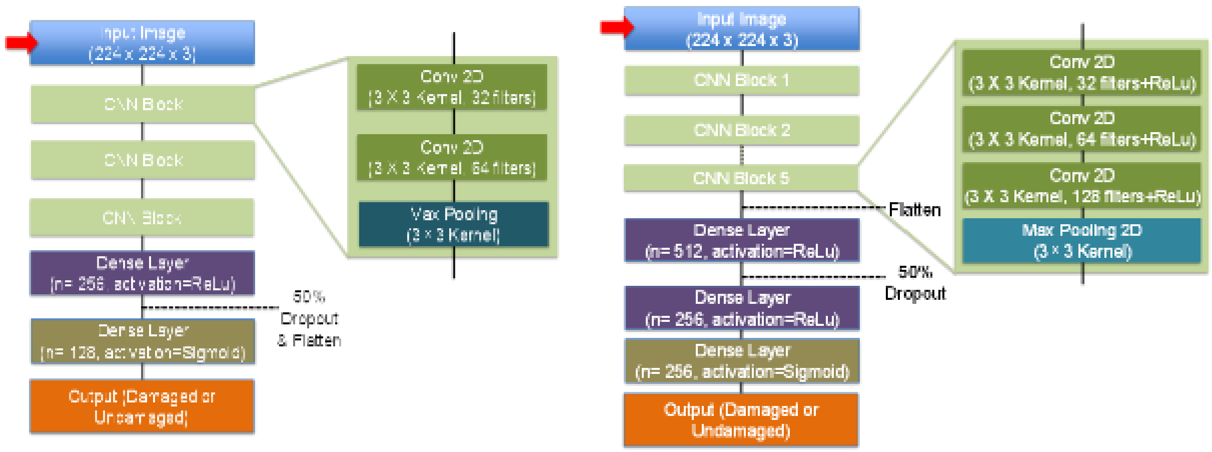


Figure 1. Components used in the development of the image collection (a) and system attached to a Caterpillar 930M loader (b) and a Kenworth T480 yard truck (c).

Automatic Cover Damage Detection

Three different CNN models were trained to automatically detect wrap damage in collected images of modules, and their detection accuracy performance was compared. AlexNet was used in this study because it already proved an excellent performance (over 90% detection accuracy) in different aspects of agriculture (i.e., weed detection (Beeharry and Bassoo, 2020), pesticide residue detection in fruits (Jiang et al., 2019), disease detection in a plant leaf (Arya and Singh, 2019)). This model contained three CNN blocks with two convolutional layers, a max-pooling layer, and two dense layers (Figure 2a). A rectified linear unit (ReLU) activation function was used in the first dense layer, and a sigmoid activation function was used in the final dense layer for a binary output (damaged or undamaged). After the first dense layer, 50% dropout was used to flatten the model. Module images were collected using the smartphone camera from different gin yards to train the model. The dataset was created with 2,178 images where 50% were in the damaged class, and the rest of the 50% were in the undamaged class. From this data set, 1,694 random images were used to train the model, and the remaining 484 were used to validate it. Before entering the images into the model, they were resized in 224×224×3 format.

The VGG16 CNN architecture is also popular because of its high accuracy in image classification. It was successfully implemented for detecting rice and wheat leaf diseases (Jiang et al., 2021), weeds in corn and soybean (Ahmad et al., 2021), apples for robotic harvesting (Fu et al., 2020), pest damage in corn (Ishengoma et al., 2021), and other image detection problems in agriculture. VGG16 is an improvement over the AlexNet with five CNN blocks and three dense layers. The first two CNN blocks contain two convolutional layers with a max-pooling layer in each, and the other three CNN blocks have three convolutional layers with a max-pooling layer in each (Figure 2b). Like AlexNet, 50% dropout was used after the first dense layer. This model used ReLU and sigmoid activation functions in the first and final dense layers, respectively. For training and validating this model, the same dataset was used as with AlexNet.



(a) **(b)**

Figure 2. The architecture of the (a) AlexNet and (b) VGG16 models used to classify the module images with or without cover damage.

YOLO (You-Only-Look-Once) is a one-stage object detection approach that uses CNN's as the backbone. YOLO was proposed to increase the detection speed over the other popular CNN-based object detection methods (i.e., R-CNN, Fast R-CNN, Faster R-CNN). YOLO quickly became popular because it provides good accuracy with higher detection speed and can be applied in real-time object detection applications. Currently, the YOLO series (Redmon et al., 2016; Redmon and Farhadi, 2017; Redmon and Farhadi, 2018; Bochkovskiy et al., 2020; Joseph Nelson, 2020) has evolved from YOLOv1 to YOLOv5 by increasing the efficiency. All versions of YOLO were successfully implemented in different studies related to agriculture with acceptable accuracy (Wang et al., 2022; Wu et al., 2020; Tian et al., 2019). The cover damage detection network structure used in this study is illustrated in Figure 3, where CSPDarknet (Cross Stage Partial Network), PANet (Path-Aggregation Network), and YOLOv5 were used as the backbone, neck, and head of the model, respectively. The backbone network works to extract the high-level features of the images, the neck network helps to better fuse the extracted features of the image, and the head detects the objects from the image. Spatial Pyramid Pooling (SPP) was used in this model to fuse the multiscale features of the images. From the module image dataset collected with smartphone cameras, 970 images were used for training and 240 images were used to validate the model. All of these images were annotated with bounding boxes using a labelling tool and assigned a label as "damaged" for each box.

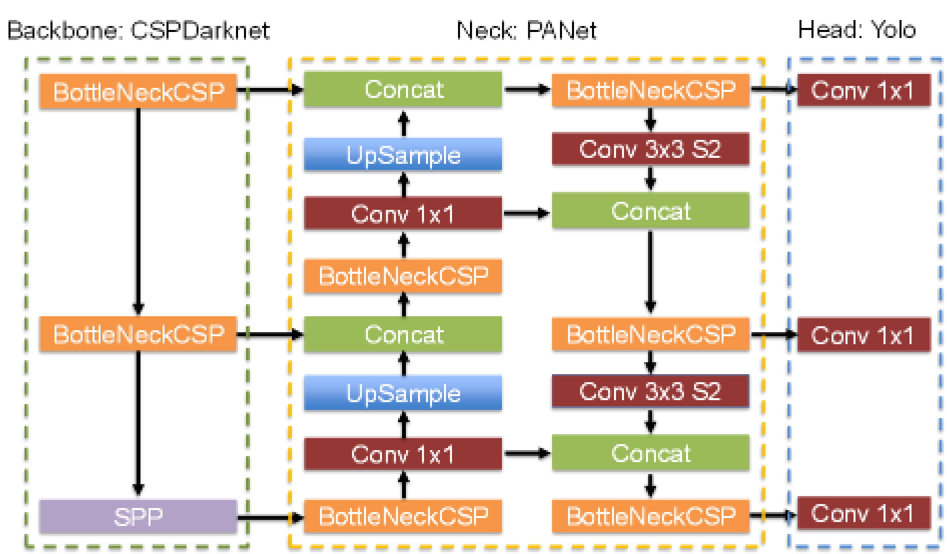


Figure 3. The architecture of the YOLOv5 model for detecting the cover damage in the picture of round cotton modules.

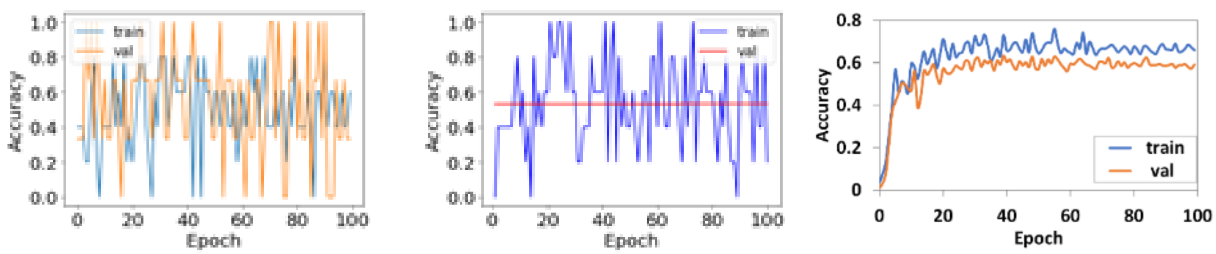
All CNN models were trained using the same set of hyper-parameters: image size = 224×224, batch size = 6, strides = 2×2, learning rate = 0.001, epochs = 100. Two different image sets for testing were created from additional images captured by smartphone (135 images) and the image collection system (46 images), containing both damaged and undamaged covers of the round cotton modules. The performance of all CNN models was evaluated using this same testing dataset. All CNN models were constructed and tested in Google Colaboratory, where a Tesla T4 GPU was used (15 GB GPU memory).

Results and Discussion

Comparison of Models

The training and validation accuracy of the AlexNet was not steady. Both for training and testing, the accuracy randomly varied from 0 to 100% (Figure 4a). The model's average training and validation accuracy was 54% and 51%, respectively. This result means the model did not learn about the nature of the images. In the VGG16 model, the training accuracy gave random output with increasing epochs like the AlexNet model, but the validation accuracy for this model remained the same with an increasing number of epochs. This model's average training and validation accuracy was 56% and 50%, respectively (Figure 4b).

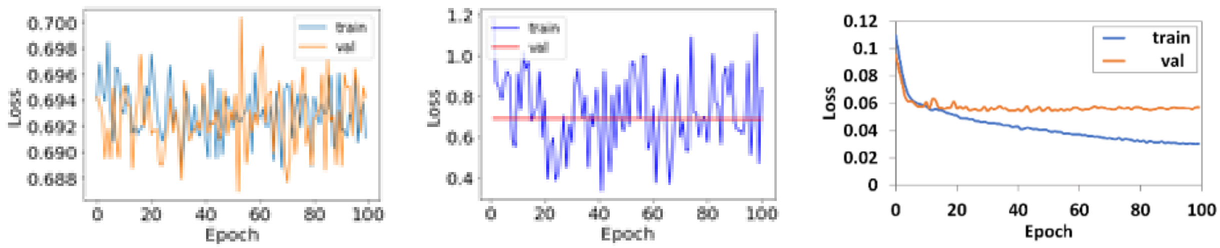
On the other hand, the training and validation accuracy graphs in the YOLOv5 model indicated that the model learned image features with an increasing number of epochs. The training and validation accuracy was maximum at the 47th epoch. After 64 epochs, accuracy decreased to a certain level, and after 65, the accuracy became constant with the number of epochs (Figure 4c). The maximum training and validation accuracy was 75.6% and 65.9%, respectively, after 56 epochs. This training time was considered optimum for this model, and this model was used for testing.



(a) (b) (c)

Figure 4. Training and validation accuracy of the CNN models used in this study: (a) AlexNet, (b) VGG16, and (c) YOLOv5.

Training and validation loss of the AlexNet model was gradually reduced for the first few epochs, and after that was unstable (Figure 5a). This model's minimum training and validation loss were 0.68 and 0.65, respectively. Although this is a reasonable minimum model loss, the model is unsuitable due to its instability. For the VGG16 model, the training loss was random and did not decrease with additional training epochs, while the validation loss remained the same (Figure 5b). These results indicated that the model could not predict the image's specific features. The training and validation loss for the YOLOv5 model decreased with additional training. The training loss reduced with an exponential pattern and The validation loss also initially decreased exponentially, but after 24 epochs was constant (Figure 5c). This result means the training dataset may not have sufficient information to learn some features of damaged module cover images. Some types of cover damage in the validation set may not be represented in the training set. Also, the model might be overfit from learning unrelated features of the training data set that are coincidentally present in some of the damaged module images. There is a significant gap between the training and validation loss graph because the training images contain features slightly different than the validation images.



(a) **(b)** **(c)**

Figure 5. Training and validation loss of the CNN models used in this study: (a) AlexNet, (b) VGG16, and (c) YOLOv5.

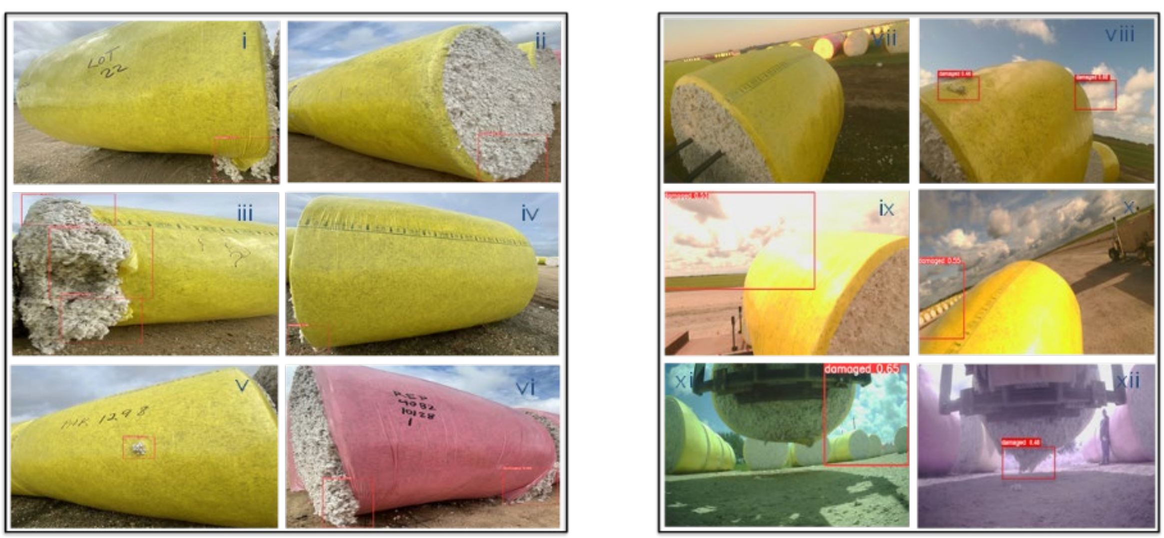
All models were tested with the same dataset containing 135 images of damaged and undamaged cotton modules. The testing accuracy for the AlexNet, VGG167, and YOLOv5 was 37%, 42%, and 78%, respectively. Table 1 summarizes the quantitative results of all three models. Based on the training, validation, and testing accuracy, YOLOv5 seems more promising to detect cover damage of round cotton modules.

Table 1. Comparison of different CNN models used in this study

Model	Training Accuracy	Validation Accuracy	Testing Accuracy
AlexNet	0.54 (1694 images)	0.51 (484 images)	0.37 (135 images)
VGG16	0.56 (1694 images)	0.50 (484 images)	0.42 (135 images)
YOLOv5	0.756 (970 images)	0.659 (240 images)	0.78 (135 images)

Performance Evaluation of the YOLOv5 Model

The trained YOLOv5 model was tested using both images captured with a smartphone and the developed image collection system. The round module cover damage detection accuracy was 78% (tested with 135 images) for the images collected by smartphone and 73% (tested with 46 images) for the images collected by the developed system.



(a) **(b)**

Figure 6. Example of cotton module images after testing with the YOLOv5 CNN model: (a) images captured by smartphone and (b) images collected by the developed system installed on a loader.

Figure 6a shows that the model detected small sections of damaged wrap (i, v) as well as covers with multiple damaged locations (iii, vi) from the images of round cotton modules. One cause of false positive results was a rough seed cotton surface on the face of the round module or loose cotton on the outside of the module. During the performance evaluation of the model using the round cotton modules picture captured by the system (Figure 6b), it was noticed that some false positive results occcurred when clouds in the background looked like loose cotton (viii, x, xi). Over-exposed images (ix) often had false positives because they resembled a damaged module.

Summary

An automated image collecting system for the round cotton modules was developed, which was attached to module handling equipment. The system was able to capture images from different angles and save them with the module's unique RFID number. To identify the cover damage of the modules automatically, three CNN models were evaluated to identify the best performer. YOLOv5 performed best, detecting cover damage of the round cotton modules with an accuracy of 78% for the smartphone-collected images and 73% for the images captured by the monitoring system. However, the detection accuracy of the model was not as high as desired. More images will be collected to better train the model. The images collected with the monitoring system have decreased detection accuracy due to the complex background. Pre-processing images to remove the background should improve damage detection accuracy.

Acknowledgments

This work was supported by Cotton Incorporated under agreement No. 18-496. The authors are also thankful to the cooperating producers and ginners.

References

Agarwal, M., Gupta, S. K., Biswas, K. K. (2020). Development of Efficient CNN model for Tomato crop disease identification. Sustainable Computing: Informatics and Systems, 28: 100407.

Ahmad, A., Saraswat, D., Aggarwal, V., Etienne, A., Hancock, B. (2021). Performance of deep learning models for classifying and detecting common weeds in corn and soybean production systems. Computers and Electronics in Agriculture, 184: 106081.

Arya, S., Singh, R. (2019). A Comparative Study of CNN and AlexNet for Detection of Disease in Potato and Mango leaf. In IEEE International Conference on Issues and Challenges in Intelligent Computing Techniques, ICICT 2019: 1–6.

Beeharry, Y., Bassoo, V. (2020). Performance of ANN and AlexNet for weed detection using UAV-based images. In 2020 3rd International Conference on Emerging Trends in Electrical, Electronic and Communications Engineering, ELECOM 2020: 163–167.

Bochkovskiy, A., Wang, C.-Y., Liao, H.-Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. ArXiv Preprint, arXiv: 2004.10934.

Condotta, I. C. F. S., Brown-Brandl, T. M., Pitla, S. K., Stinn, J. P., Silva-Miranda, K. O. (2020). Evaluation of low-cost depth cameras for agricultural applications. Computers and Electronics in Agriculture, 173: 105394.

Fu, L., Majeed, Y., Zhang, X., Karkee, M., Zhang, Q. (2020). Faster R–CNN–based apple detection in dense-foliage fruiting-wall trees using RGB and depth features for robotic harvesting. Biosystems Engineering, 197: 245–256.

Indira, D. N. V. S. L. S., Goddu, J., Indraja, B., Challa, V. M. L., Manasa, B. (2021). A review on fruit recognition and feature evaluation using CNN. Materials Today: Proceedings.

Iqbal, Z., Hardin, R. G., Wang, T., Ward, J. K., Wanjura, J. D. (2021). Round modules: Handling logistics and cover damage, 2nd year. In 2021 Beltwide Cotton Conferences: 623–629.

Ishengoma, F. S., Rai, I. A., Said, R. N. (2021). Identification of maize leaves infected by fall armyworms using UAV-based imagery and convolutional neural networks. Computers and Electronics in Agriculture, 184: 106124.

Jiang, B., He, J., Yang, S., Fu, H., Li, T., Song, H., He, D. (2019). Fusion of machine vision technology and AlexNet-CNNs deep learning network for the detection of postharvest apple pesticide residues. Artificial Intelligence in Agriculture, 1: 1–8.

Jiang, Z., Dong, Z., Jiang, W., Yang, Y. (2021). Recognition of rice leaf diseases and wheat leaf diseases based on multi-task deep transfer learning. Computers and Electronics in Agriculture, 186: 106184.

Joseph Nelson, J. S. (2020). YOLOv5 is Here: State-of-the-Art Object Detection at 140 FPS. Retrieved January 21, 2022, from https://blog.roboflow.com/yolov5-is-here/

Pelletier, M. G., Holt, G. A., Wanjura, J. D. (2020a). A Cotton Module Feeder Plastic Contamination Inspection System. AgriEngineering, 2(2): 280–293.

Pelletier, M. G., Holt, G. A., Wanjura, J. D. (2020b). A Plastic Contamination Image Dataset for Deep Learning Model Development and Training. AgriEngineering, 2(2): 317–321.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016: 779–788).

Redmon, J., Farhadi, A. (2017). YOLO9000: Better, faster, stronger. In Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017: 6517–6525.

Redmon, J., Farhadi, A. (2018). YOLOv3: An Incremental Improvement. ArXiv Preprint, arXiv:1804.02767.

Ruiz-Garcia, L., Lunadei, L. (2011). The role of RFID in agriculture: Applications, limitations and challenges. Computers and Electronics in Agriculture, 79(1): 42–50.

Tian, Y., Yang, G., Wang, Z., Wang, H., Li, E., Liang, Z. (2019). Apple detection during different growth stages in orchards using the improved YOLO-V3 model. Computers and Electronics in Agriculture, 157: 417–426.

Urso, O., Chiacchio, F., Compagno, L., D'Urso, D. (2020). An RFID application for the process mapping automation. In Procedia Manufacturing, 42: 8–15.

Van Der Sluijs, M. hj, Long, R. L., Bange, M. P. (2015). Comparing cotton fiber quality from conventional and round module harvesting methods. Textile Research Journal. SAGE PublicationsSage UK: London, England.

Wang, T., Hardin IV, R. G., Ward, J. K., Wanjura, J. D., Barnes, E. M. (2022). A smart cotton module tracking and monitoring system for handling logistics and cover damage. Computers and Electronics in Agriculture, 193, 106620.

Wang, Z., Jin, L., Wang, S., Xu, H. (2022). Apple stem/calyx real-time recognition using YOLO-v5 algorithm for fruit automatic loading system. Postharvest Biology and Technology, 185: 111808.

Wanjura, J., Pelletier, M., Ward, J., Hardin, B., Barnes, E. (2020). Prevention of Plastic Contamination When Handling Cotton Modules. Cotton Incorporated, 1–10.

Wu, D., Lv, S., Jiang, M., Song, H. (2020). Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments. Computers and Electronics in Agriculture, 178:105742.