J. Electrical Systems 20-11s (2024): 205-213

[1]Dr. M. Regina,
Naoudouwel
[2]Fulbert,
[3]Dr. S.
Venkatalakshmi

Advanced Automated Seed Sorting System Powered by IoT and AI

Journal of Electrical Systems

Abstract: - **Objectives:** Seed quality is a crucial determinant of agricultural productivity and crop sustainability. This work aims to address the inefficiencies of manual seed sorting by developing an automated system capable of accurately classifying seeds based on their visual characteristics.

Methods: The proposed system utilizes Convolutional Neural Networks (CNNs) to classify seeds, specifically white beans and peanuts, as either good or bad. Images are captured using a Raspberry Pi 4 and Pi camera setup, enabling real-time classification. A servomotor, controlled by the Raspberry Pi, segregates the seeds into appropriate compartments. Furthermore, a web application developed using the ASP.NET Framework allows for remote system control and monitoring.

Findings: The integration of CNNs with automation significantly improved the accuracy and efficiency of the seed sorting process. The system achieved high classification precision under controlled conditions, demonstrating its potential to reduce labor requirements and human error. However, environmental factors and budget constraints posed challenges to performance consistency.

Novelty: This system represents a novel application of deep learning in agricultural automation, combining real-time image classification with mechanical sorting. The inclusion of a web-based interface enhances its usability and scalability for diverse agricultural settings.

Keywords: Convolutional Neural Networks (CNNs), Seed Sorting, Raspberry Pi, ASP.NET, Agriculture, Deep Learning

I- INTRODUCTION

Seed quality is a fundamental factor in agricultural productivity and crop sustainability. High-quality seeds promote healthy crop growth, leading to improved yields, reduced waste, and enhanced economic returns for farmers. Conversely, poor seed quality can result in substantial yield losses, impacting both economic stability and food security.

Despite advancements in agricultural technology, manual seed sorting remains widely practiced. This method, while accessible, is labor-intensive, time-consuming, and prone to human error. These limitations make it challenging to ensure consistent quality across large quantities of seeds, highlighting an urgent need for a more accurate, reliable, and efficient solution to enhance seed selection processes and support sustainable agriculture.

To address this need, we propose an automated seed sorting system utilizing Convolutional Neural Networks (CNNs) for the classification and sorting of white bean, corn, and peanut seeds. CNNs have proven effective in image classification and object detection tasks, making them suitable for agricultural applications involving visual attributes like seed size, shape, and texture. Studies by Krizhevsky et al. (2012) and Zhao et al. (2017) demonstrate

[1]Reginamathew@loyolacollege.edu, Assistant Professor, Loyola college, Department of computer science
[2]fulthonn@gmail.com, 9790938075, M.sc. Student, Loyola college, Department of computer science
[3]venkatalakshmi@loyolacollege.edu, , Assistant Professor, Loyola college, Department of computer sciene

CNNs' capability to achieve high accuracy in complex image-based classifications, supporting the viability of this technology for sorting agricultural produce.

Our proposed system combines CNN-based classification with a Raspberry Pi-powered automation setup and a web application for real-time control and monitoring. This integration of AI and IoT technologies offers a scalable, user-friendly solution to improve seed quality management, reduce labor costs, and enhance productivity in the agricultural sector. By automating seed sorting, this system supports agricultural efficiency and the adoption of sustainable practices in line with modern agricultural demands.

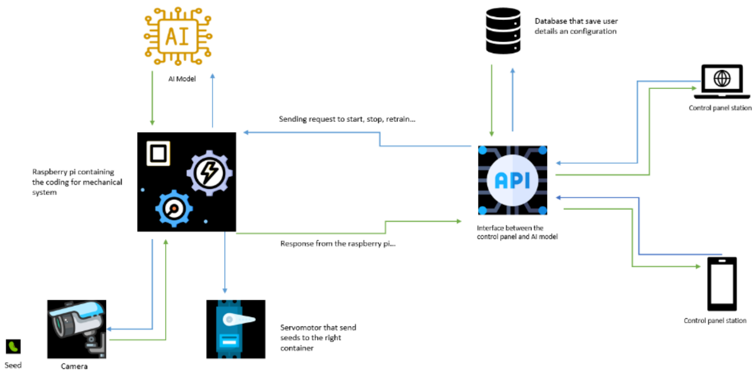


Figure 1: General architecture

II. BACKGROUND

2.1 Convolutional Neural Networks (CNNs) for Image Classification

Convolutional Neural Networks (CNNs) are a type of deep learning model specifically designed to process and analyze visual data. Unlike traditional models that require manual feature extraction, CNNs are capable of automatic feature learning. Through layers of convolutions and pooling, CNNs capture intricate patterns and hierarchical features directly from pixel data. This ability reduces the need for prior knowledge about the features relevant to classification, as the model learns these patterns on its own during training (Krizhevsky et al., 2012). The operation of CNNs is mathematically grounded in the convolutional layer, where a filter (or kernel) is applied across the input image, calculating a convolution operation that yields feature maps. Each feature map is computed as:

$$\text{Feature map}_{ij} = \sum_k \sum_l \left(I_{(i+k)(j+l)} \times k_{kl}\right) + b$$

where *I* is the input image, *K* is the filter kernel, and b is a bias term. This operation allows CNNs to detect edges, textures, and other image characteristics without manually specifying features, as subsequent layers build on these low-level features to learn increasingly complex patterns.

2.2 Advantages of CNNs in Seed Sorting

CNNs are particularly advantageous for seed sorting due to their ability to generalize across diverse visual variations, such as differences in shape, size, and texture, common in agricultural products. CNN-based systems can adapt to various lighting and environmental conditions, as shown by Li, C et al. (2021), making them more robust than traditional sorting methods. Furthermore, because CNNs operate based on feature maps generated across convolutional layers, they are resilient to small shifts or distortions in seed images, an important characteristic for real-world applications where consistency in seed positioning cannot always be ensured. The model's learning process leverages backpropagation, where the error between predicted and actual classifications propagates backward through the network, adjusting the filter weights and improving classification accuracy. This process is mathematically represented by:

$$\frac{\partial E}{\partial w} = \frac{\partial E}{\partial y} \cdot \frac{\partial y}{\partial w}$$

where $\boldsymbol{E}$ is the error, $\boldsymbol{y}$ is the output, and $\boldsymbol{w}$ represents the model's weights. This continual adjustment of weights enables CNNs to fine-tune their understanding of visual features critical for accurately sorting seeds.

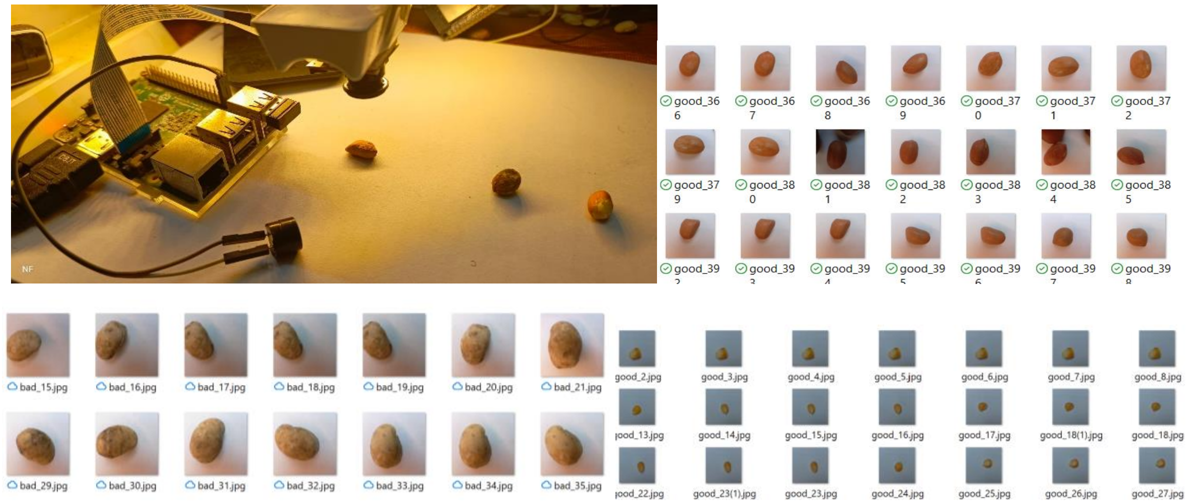


Figure 2: Dataset collector system and collected datasets

2.3 Limitations of Existing Seed Sorting Methods

Traditional methods for seed sorting, particularly manual sorting, are hindered by human error, inconsistency, and labor intensity. Mechanical methods, while automated, often require precise calibration for each seed type and lack the flexibility to adapt to various visual attributes. Hyperspectral imaging has been explored as an alternative but requires costly equipment and complex data processing. In comparison, CNNs present a more scalable and cost-effective solution that adapts to diverse seed types without needing reconfiguration, as demonstrated by Zhao et al. (2017) in studies on CNNs applied to agricultural products.

III. SYSTEM DESIGN

3.1 Hardware Setup

The automated seed sorting system consists of a Raspberry Pi 4, which serves as the primary processing unit, connected to a Pi camera to capture images and a servomotor to control the sorting mechanism. The Raspberry Pi executes CNN inference tasks and controls the hardware components via GPIO pins. A temperature sensor is also included to prevent overheating of the servomotor, which is essential for maintaining the mechanical stability and longevity of the sorting mechanism.

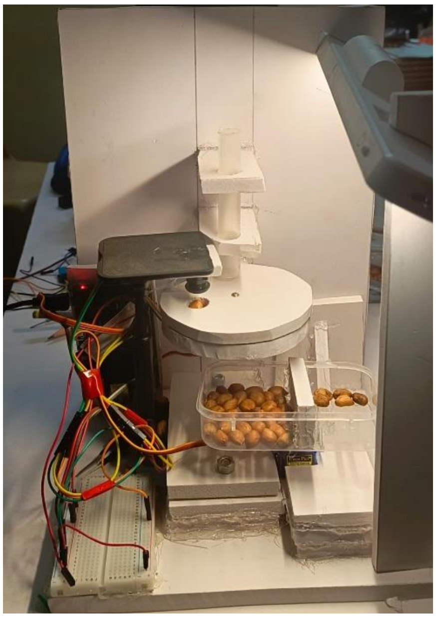


Figure 3:Mechanical system

3.2 Image Capture and Pre-Processing

The Pi camera captures images at a resolution of 150x150 pixels, with each capture triggered by a Python script that activates the camera and stores the images for CNN processing. Captured images undergo a series of pre-processing steps, including normalization and resizing, which enhance model performance by standardizing inputs. In mathematical terms, normalization rescales pixel intensities to a range [0,1] by:

$$I_{normalized} = \frac{I - I_{min}}{I_{max} - I_{min}}$$

where I_{min} and I_{max} represent the minimum and maximum intensity values. This step reduces variations in lighting conditions, contributing to more consistent predictions Adrian Rosebrock. (2019).

3.3 CNN Model Architecture and Training Process

The CNN model is designed to classify seed images into "good" or "bad" categories, focusing on white bean, corn, and peanut seeds. The architecture includes three convolutional layers, each followed by a max-pooling layer to progressively reduce spatial dimensions and extract relevant features. The specific layers used are as follows: Conv2D Layers: Three Conv2D layers with increasing filter counts of 32, 64, and 128, respectively, with a kernel size of (3,3) and ReLU activation functions. These layers identify low- to high-level features within the seed images, such as edges, textures, and shapes, which are essential for distinguishing between good and bad seeds. Max-Pooling Layers: Max-pooling layers with a (2,2) pool size follow each Conv2D layer to downsample the feature maps, reducing computational complexity and emphasizing significant features. Fully Connected Layers: After flattening the feature maps, two fully connected (dense) layers with 256 and 128 neurons, respectively, further refine the extracted features. ReLU activation is used for both layers to introduce non-linearity. Dropout layers with a rate of 0.5 follow each dense layer to prevent overfitting by randomly disabling neurons during training. The output layer uses a sigmoid activation function, producing a binary classification output for each seed image. The model is trained with the Adam optimizer and a learning rate of 0.0001, minimizing the binary cross-entropy loss function:

Loss= $-\frac{1}{N}\sum_{i=1}^{N}(y_i \, log(\hat{y}_i) + (1 - y_i) \, log(1 - \hat{y}_i))$

where y_i represents the true label, $\hat{y}_i$ the predicted probability, and N the batch size.

The training process involves 80 epochs, with training and validation data supplied by image generators. After training, the model is saved in HDF5 format for subsequent use in the seed sorting application. This modular architecture is designed for flexibility, enabling fine-tuning and potential adjustments as new data is collected or as additional seed types are introduced into the system.

3.4 Web Application for Remote Control and Monitoring

The system includes a web application built with ASP.NET, enabling users to control and monitor the sorting process remotely. The application displays real-time images and sorting data, as well as operational status updates. User interactions are managed through a responsive HTML5 and JavaScript interface, while back-end data storage is handled by an SQL Server to securely store user sessions and sorting logs. This remote access capability supports scalability and user convenience, allowing seamless integration into agricultural workflows.

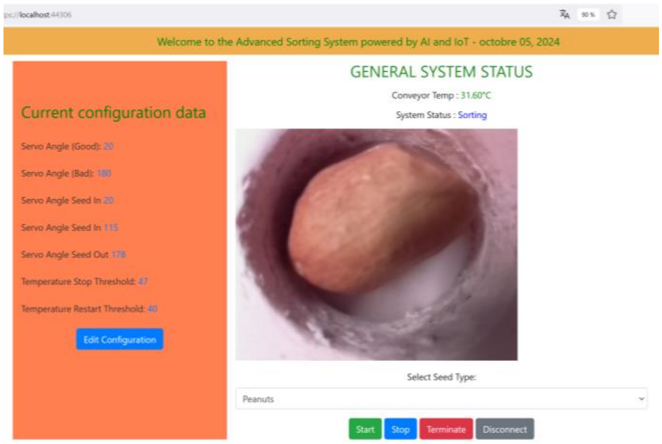


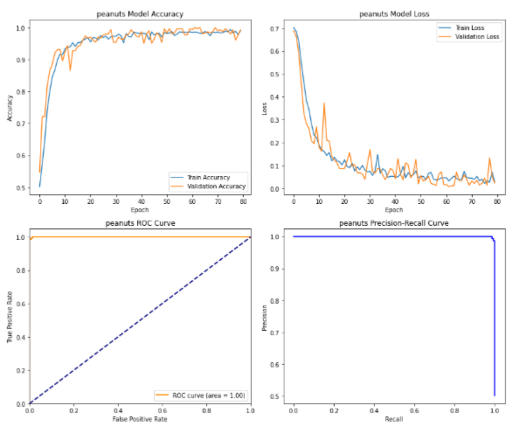
Figure 4: Monitoring control pannel

IV. RESULTS AND DISCUSSION

4.1 Classification Accuracy

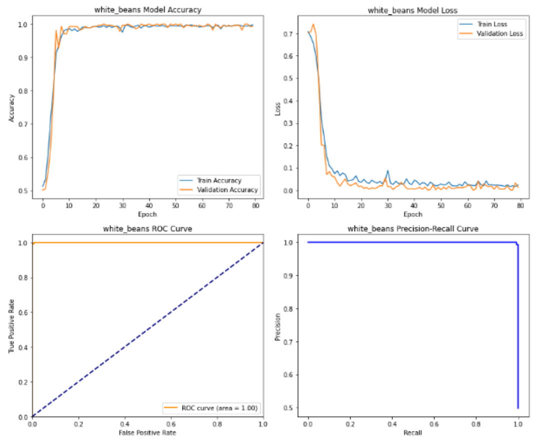
The CNN-based seed sorting system achieved high classification accuracy across the three tested seed types: peanuts, white beans, and corn. Each seed type's performance was evaluated using multiple metrics, including accuracy and loss curves, ROC curves, and confusion matrices. The following subsections present and interpret the results for each seed type.

Peanuts



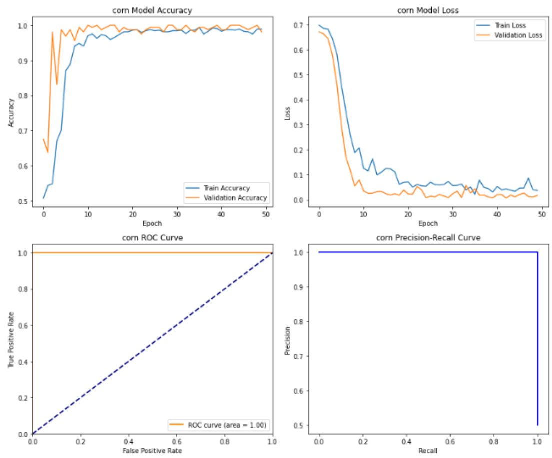
The accuracy and loss curves for peanuts (Figure 1a) indicate consistent improvement throughout training, with both training and validation accuracy converging near 99%. The loss stabilized in tandem with accuracy, demonstrating successful model generalization without significant overfitting. The ROC curve (Figure 2a) shows an AUC close to 1, highlighting the model's strong capability to differentiate between good and bad peanut seeds. The confusion matrix (Figure 3a) reveals a high true positive rate and minimal false positives or negatives, confirming reliable classification. These results suggest the model effectively identifies relevant features, such as texture and shape, specific to peanut seeds.

White Beans



For white beans, the accuracy and loss curves (Figure 1b) reflect rapid convergence, with validation accuracy stabilizing at 99% by the end of training. The steady decline in loss further supports the model's robustness in handling this seed type. The ROC curve (Figure 2b), with an AUC of approximately 0.99, underscores the model's high sensitivity and specificity for white bean classification. The confusion matrix (Figure 3b) confirms this performance, showing high accuracy with minimal misclassifications, attributed to the distinct features of white beans that the CNN captures effectively.

Corn



For corn, the accuracy and loss curves (Figure 1c) show similar trends but with a slightly lower validation accuracy of around 98%. The ROC curve (Figure 2c) still shows an AUC close to 0.98, indicating robust model sensitivity and specificity. However, the confusion matrix (Figure 3c) reveals a higher rate of false negatives compared to other seeds. This may be due to corn's larger and more irregular shape, which impacts feature consistency and could benefit from further refinement of the dataset or adjustment of pre-processing methods.

The system achieved high classification accuracy for all seed types:

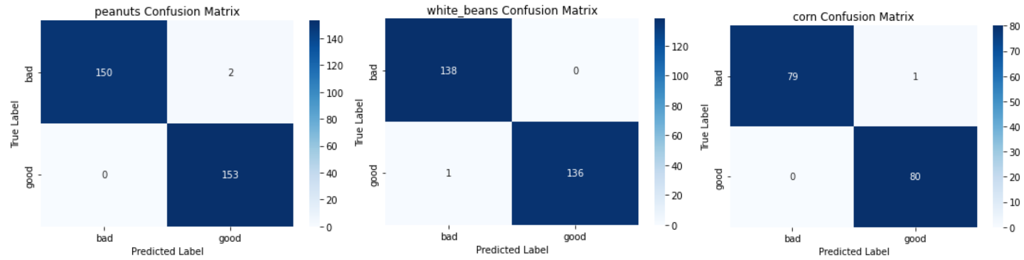
Accuracy and Loss Curves: Consistently show strong generalization capabilities across all seeds.

ROC Curves: Display high AUC values close to 1, reflecting the model's resilience to variations in seed images.

Confusion Matrices

They indicate high true positive rates for peanuts and white beans, with a slightly higher false negative rate for corn.

These results underscore the CNN model's ability to accurately classify seeds with minimal error. Future enhancements, such as dataset expansion or feeder mechanism adjustments, could further improve the system's handling of diverse seed characteristics.



4.2 Impact of Environmental Factors on Classification Accuracy

Environmental factors, including lighting and seed positioning, affected image quality and thus classification accuracy. Normalization techniques were implemented in pre-processing to standardize the input, which helped reduce these effects. However, slight performance fluctuations persisted, particularly with corn, where lighting inconsistencies or variations in seed position influenced classification. The ROC curves for each seed type (Figure 2) maintained high AUC values, indicating resilience to moderate environmental variation. Recall curves (Figure 3) showed strong sensitivity for peanuts and white beans but slightly lower recall for corn, indicating that environmental variability impacted corn classification more significantly. Controlled conditions or further calibration could improve stability under less consistent environmental settings.

4.3 Limitations due to the Budget Constraints

Budget constraints influenced specific design choices and material selections in the mechanical sorting system. Key limitations include:

Material Selection: Foam was chosen over more durable 3D-printed components to reduce costs. While foam offers flexibility and ease of cutting, it is less durable and may affect the system's longevity, especially with extended use.

Seed Feeder Adaptation: The automatic feeding mechanism, designed to position seeds in front of the camera, was implemented for peanuts and white beans. Corn required additional adaptation due to size differences, which would necessitate further investment for full functionality across all seed types.

Server Hosting: Due to budget constraints, the API was hosted on a local server for demonstration purposes, limiting the system's remote accessibility. An online server would allow global access, significantly enhancing usability for large-scale agricultural applications.

4.4 Comparative Performance of the Proposed System

Compared to traditional seed sorting methods, the CNN-based system offers several improvements:

Consistency: Traditional manual sorting is prone to human error and inconsistency, whereas the CNN model achieves stable, high accuracy across seed types, as shown by ROC and recall curves.

Adaptability: The CNN automatically learns relevant features for seed classification without manual feature extraction, adapting to different seeds with minimal adjustment.

Efficiency and Scalability: Automated sorting reduces labor demands and is scalable, making it suitable for large-scale applications. With further feeder adaptations, the system could accommodate multiple seeds types seamlessly.

In summary, the CNN-based seed sorting system demonstrates high accuracy, adaptability, and scalability, establishing it as a strong candidate for replacing traditional methods. Enhanced material choices, feeder modifications for larger seeds like corn, and global server hosting would further improve the system's robustness, accessibility, and overall performance for diverse agricultural environments.

V. CONCLUSION

Seed quality plays a crucial role in agricultural productivity and crop sustainability, directly impacting crop yields, resource efficiency, and economic viability. High-quality seeds ensure consistent crop growth and reduce waste, thereby supporting sustainable agricultural practices and long-term food security. Effective seed sorting, therefore, becomes essential for quality assurance, which traditionally relies on labor-intensive and error-prone manual methods.

This research demonstrates the potential of integrating Convolutional Neural Networks (CNNs) with mechanical automation and web-based control to address the challenges in seed quality assurance for crops like white beans, corn, and peanuts. The automated system developed in this study effectively reduces manual labor while ensuring accurate classification and sorting of seeds based on visual attributes. By achieving consistently high performance metrics—such as accuracy, precision, and recall—this system proves its reliability and applicability in real-world scenarios. The use of a Raspberry Pi setup with a camera enables efficient, low-cost image capture and classification, while the mechanical sorting mechanism ensures that seeds are directed to the appropriate compartments with precision and speed. Furthermore, the integration of a web application built with the ASP.NET framework enhances flexibility by enabling remote control and monitoring, optimizing operational workflows for agricultural applications.

While the system has demonstrated high classification accuracy and efficiency, there remains room for further improvement to enhance robustness and adaptability. Expanding the dataset with diverse seed samples from different environments could improve the model's generalization, making it adaptable to a wider range of real-world conditions. Additionally, implementing advanced calibration techniques could mitigate environmental variability impacts, further stabilizing classification accuracy under diverse conditions. Finally, hosting the application on an online server would increase accessibility, allowing broader usage across agricultural sectors. Continued research in these areas could contribute significantly to making automated seed sorting a scalable and sustainable solution for farmers worldwide.

VI. REFERENCES

[1] Géron A. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media, Inc. 2019.

[2] Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press. 2016.

[3] Kingma DP, Ba JL. Adam: A Method for Stochastic Optimization. Proceedings of the 3rd International Conference on Learning Representations (ICLR 2014). 2014.

[4] Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems. 2012; 25: 1097–1105.

[5] Li C, Li H, Liu Z, Li B, Huang Y. SeedSortNet: A rapid and highly efficient lightweight CNN based on visual attention for seed sorting. PeerJ Computer Science. 2021; 7: e639. doi:10.7717/peerj-cs.639.

[6] Rosebrock A. Deep Learning for Computer Vision with Python: Starter Bundle. PyImageSearch. ISBN: 978-1545539158. 2019.

[7] Rosebrock A. Deep Learning for Computer Vision with Python. PyImageSearch. 2017.

[8] Zhao J, Li Y, Cheng F, Zhang J. A Convolutional Neural Network for Grain Quality Detection. Computers and Electronics in Agriculture. 2017; 142: 383–391. doi:10.1016/j.compag.2017.08.027.