2021 6th International Conference for Convergence in Technology (I2CT)
Pune, India. Apr 02-04, 2021

Real-Time Weed Detection using Machine Learning and Stereo-Vision

Siddhesh Badhan, Kimaya Desai, Manish Dsilva, Reena Sonkusare, Sneha Weakey
Department of Electronics and Telecommunication, Sardar Patel Institute of Technology
Mumbai, India
*siddhesh.badhan@spit.ac.in, †kimaya.desai@spit.ac.in, ‡manish.dsilva@spit.ac.in, §reena_kumbhare@spit.ac.in,
¶sneha_15weakey@spit.ac.in

2021 6th International Conference for Convergence in Technology (I2CT) | 978-1-7281-8876-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/I2CT51068.2021.9417989

***Abstract*—Weeds are a problem as they compete with desirable crops and use up water, nutrients, and space. Some weeds also get entangled in machinery and prevent efficient harvesting. Hence weed removal systems are a necessity. Development of a successful weed removal system involves correct identification of the unwanted vegetation. The paper proposes a Real-time weed detection system that uses machine learning to identify weeds in crops and stereo-vision for 3D crop reconstruction. Structure from motion technique is utilized on a video of a farm to generate a 3D point cloud. The machine learning model is trained on two manually created datasets of cucumber and Onion crop. Convolutional Neural Networks (CNN) and ResNet-50 algorithms are used to train the machine learning models. It is seen that the ResNet-50 model outperforms the Convolution Neural Networks model. ResNet-50 model gives an overall accuracy of 84.6% for the cucumber dataset while it gives an accuracy of 90% for the Onion crop dataset.**

***Index Terms*—Weed detection, agriculture, Machine Learning, Stereo-Vision**

I. Introduction

Weeds cause a loss of $11 billion in India for 10 major crops. As per ICAR "Unlike the visible impact of diseases and insects, the impact of weeds goes unnoticed". Manual weeding holds a major contribution of weed control in India. The increasing scarcity of labour and ever-rising costs result in farmers adopting new options which help them reduce the labour and cost. With the continuous rise in demand for food grains and the land for cultivation of crops decreasing; such a situation deserves the emergence.

Existing systems include image processing based approaches.One of the weed detection methods uses morphological features while identifying weeds in lawns [1]. The assumption considered here is that the grass in the lawn contains a group of edges while weeds do not which may result in some bias in prediction. False weed detection rates are higher in this method. Also, it requires higher processing time considering the morphological computations involved.

Image processing is used to identify weeds in a ragi plantation [2]. It involves the use of a threshold-based approach to creating a region of interest to spray the herbicide. The difference in leaf size of ragi crops and weeds is the deciding factor. One of the major shortcomings in this approach is the dependency on lighting conditions for correct results. One of the existing methodologies involves the use of machine learning algorithms namely Support Vector Machine (SVM), Convolutional Neural Network (CNN) and Artificial Neural Network (ANN). The algorithms are used for training the model for detecting weeds in four different crops containing weeds of two types [3]. Soil masking is performed to find the region of interest. Features like shapes are considered while differentiating between weeds and crops. With SVM and ANN, there is the false identification of crops as weeds in comparison to CNN. Factors like background, lighting conditions and similarity in shape feature weeds and crops.

While existing frameworks for 3D reconstruction involve use of multiview stereo techniques like Clustering Views for Multi-View Stereo (CMVS) for automatic leaf phenotyping [4]. This method is seen to have errors while calculating length, width and leaf inclination angles. With larger leaf sizes like maize plants, the proposed methodology is not reliable enough.It can be seen that techniques based on the identification of weeds and crops by constructing a 3D-world map also exist [5]. It involves the use of motion technique along with changing the zoom to recover the 3D local map and fusing it into the world-map. The system has height as the major differentiating factor between weed and crop. System drawbacks may include errors in determining weeds when crops are of the same height as weeds. Existing systems use Delaunay triangulation algorithm in order to reconstruct the 2D frames. The area, length, width and inclination angle for each leaf is calculated and later compared with the actual values [4]. Infrared-laser triangulation sensor along with high-resolution smart camera is used for recreating the 3D scene of the environment. The only drawback is a higher camera cost [6].Using 3D Robots is also existing where a system is used for analyzing the growth of plant in indoor conditions. A Gantry robot system which consists of a 3D Laser scanner is used which helps in capturing the point cloud data of the plant from multiple views [7]. Reconstruction of a dense 3D point cloud using a single image is proposed. Extraction of shape information from an input image is done followed by embedding the shape information into point cloud [8].

Considering the drawbacks in the existing methodologies, the proposed approach uses machine learning and deep learning based models for detecting weeds in two different field environments with crops namely cucumber and onion with varying weed types. Also parameters like system accuracy, computation time are discussed when the frames extracted from the video data are modelled over Convolutional Neural Networks and RESNET-50. To recreate the environment, 3D reconstruction is done using Structure from Motion technique over video data.

II. Methodology

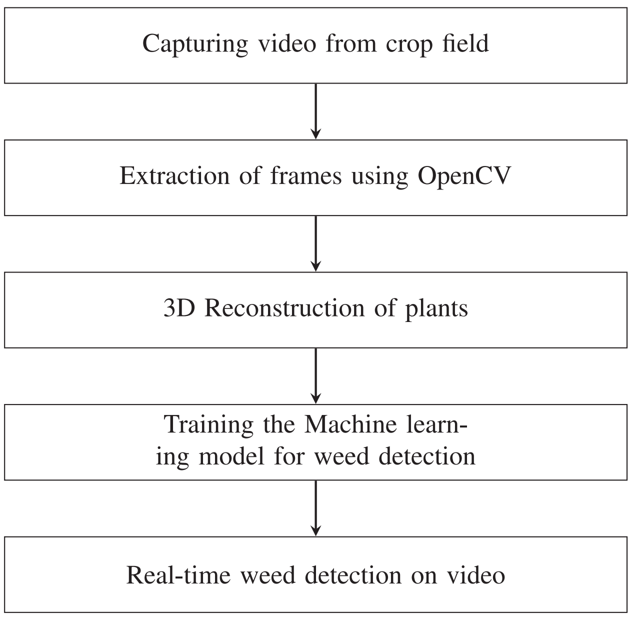


Fig. 1. Flow process of the system

The first step of the proposed system is reconstruction of the 3D model of crops from the video of the crops and the second step is the detection and classification of weeds and crops. The 3D reconstruction and the weed detection are performed on the actual video from the crop field containing the crops and weeds. The working of the system can be seen in Figure 1. The 3D reconstruction consists of capturing video followed by frame extraction, feature matching, sparse and dense reconstruction. After 3D reconstruction is performed, the next part is detecting the weeds in the crops. The weed detection part mainly consists of Data acquisition, Pre-processing, Image segmentation, Training the model, Testing the model and real-time weed detection in crops.

III. 3D Reconstruction of Plants

Flow of 3D reconstruction of plants is shown in Figure 2. A video of plant is taken using a mobile camera with 360 degrees view of the plant. The video is further processed using OpenCV for extraction of a number of frames from the video. Frames are extracted on an average of 10 frames per second. VisualSFM software [9], [10] is used for reconstruction of 3D model of the plants. Feature matching is the first step performed in the VisualSFM software for reconstruction of 3D model of the plants. Scale-Invariant Feature Transform (SIFT) algorithm [11] is used for feature extraction. Detection and description of features is done using the SIFT algorithm. These detected and extracted frames are further used for feature matching where the same features/ points are matched in two different consecutive images.

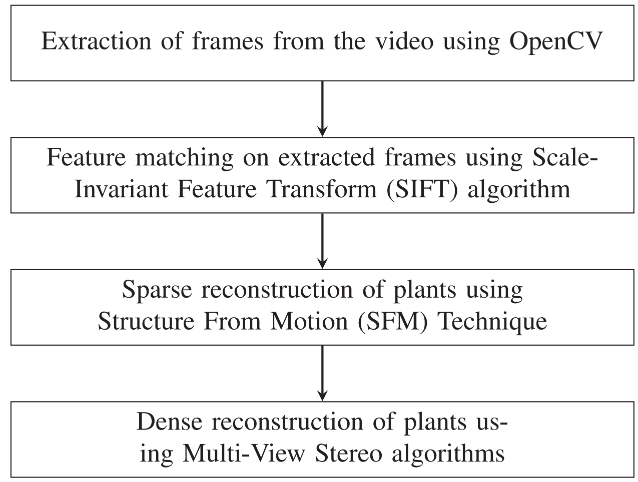


Fig. 2. Flow chart of 3D reconstruction of plants

The second step performed using the VisualSFM software is the generation of 3D Point Cloud. This is done using the Structure from Motion (SFM) technique which helps in getting 3D Point Cloud using the 2 dimensional images from a single camera and the Red, Green, Blue (RGB) and the Depth values of the image. This helps in getting a single 3D image from a series of 2D images with all the characteristics of each point in the image. Sparse reconstruction is achieved from this step. The output of Feature Matching is used for creation of a Stereo Dataset useful for training a Machine Learning model to efficiently classify between weeds and crops. The third step performed in the VisualSFM software is making the reconstruction denser. This is done using the Multi-View Stereo algorithms namely the Clustering views for Multi-View Stereo (CMVS) algorithm, Patch-based Multi-View Stereo (PMVS) algorithm [12]. CMVS decomposes the input images into clusters of images of magnable size. PMVS takes images and its parameters as input and reconstructs a 3D structure of only the rigid structure visible which makes the reconstruction denser.

IV. Weed Detection in Crops

The most important task of the weed detection system is the recognition of weeds in the crops which is a difficult task considering the intra-class variations in plants [13]. The classification of the video into crops and weeds is done in two parts in training the model and testing the trained model. The accuracy of the trained model is directly dependent on the dataset used to train the model. Creating a dataset which replicates the crops in a crop field is the most important thing for any machine learning model to achieve better results in performing real-time weed detection. The trained model is finally deployed for detecting weeds in crops in real-time.

1) Data Acquisition: Acquiring a relevant dataset is the most important task in machine learning. Sometimes this task becomes much tedious and time-consuming depending on the domain of the application. Low-resolution mobile cameras can also be used for capturing the videos of the crop fields containing the crops and the weeds. Different lighting conditions and different angles are used for increasing the accuracy of the classifier. For this study, several videos were taken from fields containing different types of crops and weeds. These videos served as a base for the generation of the training and the testing dataset of the model. Cucumber and Onion crops field videos were taken for creation of the dataset. A readily available dataset of 17500 images containing crops and weeds was also used to check the accuracy of the model [14].

2) Pre-processing: The videos are used for extraction of a number of frames using OpenCV. The desired number of frames containing the crops and weeds are obtained from the video by varying the number of frames extracted per second from the video. Manual tagging of weeds and crops is done on the extracted frames obtained from the video for creating a training dataset for the classifier. Manual tagging helps in segregation of crops and weeds, which is required for creating the training dataset. All the tagged frames are processed for making the frames of the same size and resolution. This is done using OpenCV. Cucumber and Onion crop videos were considered for creating the dataset. Once the pre-processing of the dataset is done, it is ready to be used for training the machine learning model.

3) Training the model: After labelling the dataset, the machine-learning and deep-learning-based classifier model is trained. The model trained needs to be robust and efficient in order to correctly classify between weeds and crops in the presence of noise and different climatic conditions. Since CNN gives better performance than other classifiers such as SVM and ANN, CNN and ResNet-50 are the two deep neural networks that are used to train the model [3].

The model is trained first on the readily available 17500 images as well as the data collected from the crop fields by us. The results of both the models are compared in the results section. CNN takes in the input image, assigns weights and biases to it which helps in differentiating the input image and the other images. Since the deep CNN architectures give better accuracy as feature extractors, ResNet-50 is used [15]. ResNet-50 is a convolutional neural network that has 177 layers [16]. Training the model is eased in ResNet-50 as the partial data of the input does not pass through the neural network and directly goes to the output [17]. The model is trained later on the manually created cucumber and onion dataset. The cucumber dataset has 177 unique images whereas the onion dataset has 160 unique images.

4) Testing the model: For finding out how efficient the model is, testing needs to be done. For testing the model, a different testing dataset is created. The testing dataset is created from the frames extracted from the video which are not used for training the model. The testing dataset comprises of crops and weeds. The extracted frames undergo thresholding followed by object identification which is done using OpenCV. Clustered contours were created in the frames using PlantCV [18]. Cropping of frames and resizing of images with contours is performed as the final step of creation of the testing dataset. The testing dataset is taken as input and is given to the trained model for classification of weeds and crops. This helps in finding out the accuracy of the model.

5) Real-time weed detection in crops: The trained model is made to classify weeds from crops in real time on the input video of the crop fields by creating the bounding boxes around the weeds detected in the input crops field video. This helps in real-time classification of weeds from crops in the video itself, thus mitigating the extra efforts of storing the detected weeds and crops separately on cloud or any storage device.

V. Results & Discussion

To create the training dataset, frames are extracted from the crop video. Manual cropping of weeds and crops is done on these frames. Table 1 given below shows manually cropped crops and weeds for cucumber dataset and onion dataset.

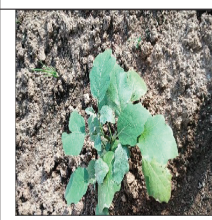
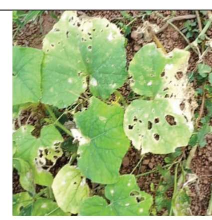
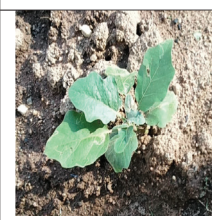
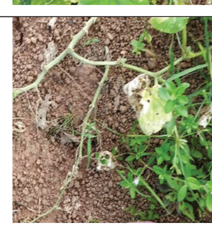
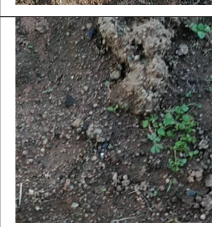
‘	**Cucumber**	**Onion**
Frame		
Crop		
Weed		

Table 1: Generated dataset for classifying weeds in Cucumber and Onion

To create the testing dataset, the extracted frames undergo thresholding. This is followed by detection of clustered contours of plants using PlantCV library. By drawing and cropping the bounding boxes around these clustered contours, the plants(both weed and crops) are detected in each frame.

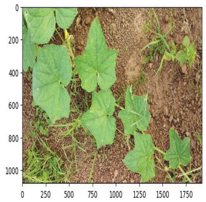
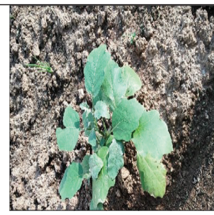
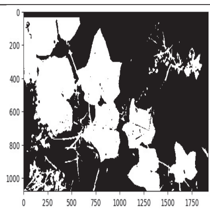
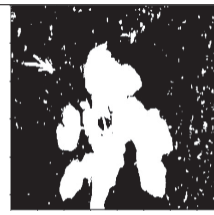
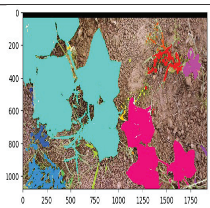
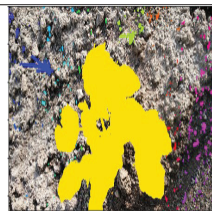
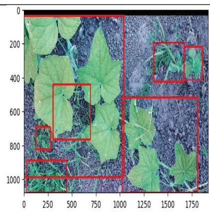
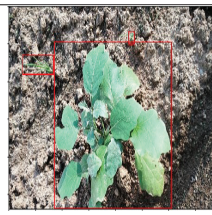
	Cucumber	Onion
Frame		
Threshold		
Contours		
Detected		

Table 2: Contours and Segmentation of Cucumber and Onion

Table 2 consists of Cucumber and Onion frames. It shows the different phases involved while creating the testing dataset.

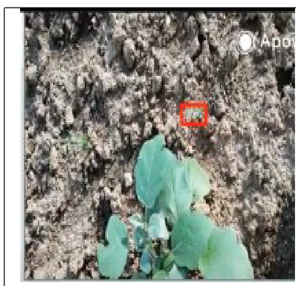
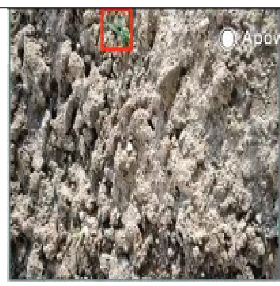
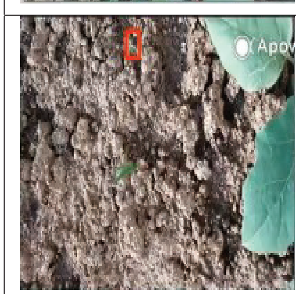
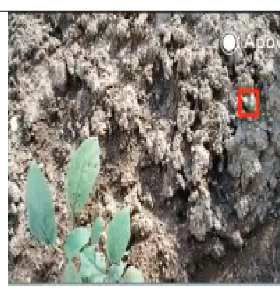
	
	

Table 3: Real-Time Weed Detection on Onion dataset

Table 3 shows the real-time weed detection performed on the Onion dataset. It comprises of frames with weeds marked with red bounding boxes captured in real time.

VI. Comparative Study

Three datasets (Cucumber, Onion and the External dataset) are trained on the CNN and Resnet-50 models. The model behaviour is determined on parameters like training time and accuracy.

Dataset	Parameter	CNN	Resnet-50
Research Paper (17.5K img) [14]	Accuracy	88%	88.8%
External Dataset (our model)	Accuracy	75%	86%
	Training Time	32min	6hrs
Cucumber Dataset (60-70 images)	Accuracy	67.8%	84.6%
	Training Time	1min	3min
Onion Dataset (60-70 images)	Accuracy	73.1%	90%
	Training Time	1min	3min

Table 4: Evaluation of our Model

As per Table 4 the results obtained in training on a set of 17500 images is 75% for Convolution Neural Network(CNN) Model and 86% for Resnet-50 Model. The dataset considered is purely of images. The parameters considered are R,G,B of these images. However, when we tried the same considering the Depth parameter as well with our own captured data of cucumbers, we have achieved an accuracy of 67.8% for CNN and 84.6% for Resnet-50. We also tried to validate the model for Onion Dataset as well which was also captured by us for which the accuracy is 73.1% for CNN and 90% for Reset-50.

	Detected Crops	Detected Weeds
Actual Crops	40	0
Actual Weeds	21	116

Table 5: Cucumber Confusion Matrix

As per Table 5 we can see that out of 40 actual crops, all have been correctly classified as weeds. And out of 137 weeds, 116 have been correctly classified as weeds and 21 have been wrongly classified as crops.

	Detected Crops	Detected Weeds
Actual Crops	100	16
Actual Weeds	0	44

Table 6: Onion Confusion Matrix

As per Table 6 we can see that out of 44 actual weeds, all have been correctly classified as weeds. And out of 116 crops, 100 have been correctly classified as crops and 16 have been wrongly classified as weeds.

VII. Conclusion

This paper delineates our extensive work on Real-time weed detection in plants. The existing weed detection mechanisms use common image processing techniques and hence, do not give efficient results. The system designed performs 3D Reconstruction of the plants and Real-time weed detection in crops. Reconstruction of a 3D environment is done using Feature matching and Structure from Motion techniques. Machine learning models like CNN (Convolutional Neural Network) and ResNet-50 are trained and tested on 3 different datasets. The datasets for training and testing the models are manually created using the videos taken from the crop fields. The accuracy observed with ResNet-50 model is found to be 90%. Thus the overall process of detecting weeds is improved and made real-time.

As India has a vast agricultural sector and various types of crops are grown, the system can be widely extended by including as many crops in the system as possible and making it suitable for all the types of crops. The system can be deployed on the cloud for faster processing and real-time weed detection and thus making it easier to deploy in a practical on-field scenario. The system can find practical applications in the existing weed detection and removal robots and systems.

References

[1] U. Watchareeruetai, Y. Takeuchi, T. Matsumoto, H. Kudo and N. Ohnishi, "Computer Vision Based Methods for Detecting Weeds in Lawns," 2006 IEEE Conference on Cybernetics and Intelligent Systems, Bangkok, 2006, pp. 1-6, doi: 10.1109/ICCIS.2006.252275.

[2] R. Aravind, M. Daman and B. S. Kariyappa, "Design and development of automatic weed detection and smart herbicide sprayer robot," 2015 IEEE Recent Advances in Intelligent Computational Systems (RAICS), Trivandrum, 2015, pp. 257-261, doi: 10.1109/RAICS.2015.7488424.

[3] S. T., S. T., S. G. G.S., S. S. and R. Kumaraswamy, "Performance Comparison of Weed Detection Algorithms," 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2019, pp. 0843-0847, doi: 10.1109/ICCSP.2019.8698094.

[4] D. Li, G. Shi, W. Kong, S. Wang and Y. Chen, "A Leaf Segmentation and Phenotypic Feature Extraction Framework for Multiview Stereo Plant Point Clouds," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 2321-2336, 2020, doi: 10.1109/JSTARS.2020.2989918.

[5] A. J. Sachez and J. A. Marchant, "Fusing 3D information for crop/weeds classification," Proceedings 15th International Conference on Pattern Recognition. ICPR-2000, Barcelona, Spain, 2000, pp. 295-298 vol.4, doi: 10.1109/ICPR.2000.902917.

[6] D. Šeatović, H. Kutterer and T. Anken, "Automatic weed detection and treatment in grasslands," Proceedings ELMAR-2010, Zadar, 2010, pp. 65-68.

[7] A. Chaudhury et al., "Machine Vision System for 3D Plant Phenotyping," in IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 16, no. 6, pp. 2009-2022, 1 Nov.-Dec. 2019, doi: 10.1109/TCBB.2018.2824814.

[8] S. Choi, A. Nguyen, J. Kim, S. Ahn and S. Lee, "Point Cloud Deformation for Single Image 3d Reconstruction," 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 2019, pp. 2379-2383, doi: 10.1109/ICIP.2019.8803350.

[9] C. Wu, "Towards Linear-Time Incremental Structure from Motion," 2013 International Conference on 3D Vision - 3DV 2013, Seattle, WA, 2013, pp. 127-134, doi: 10.1109/3DV.2013.25.

[10] C. Wu, S. Agarwal, B. Curless and S. M. Seitz, "Multicore bundle adjustment," CVPR 2011, Providence, RI, 2011, pp. 3057-3064, doi: 10.1109/CVPR.2011.5995552.

[11] Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision 60, 91–110 (2004).

[12] Y. Furukawa and J. Ponce, "Accurate, Dense, and Robust Multiview Stereopsis," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 8, pp. 1362-1376, Aug. 2010, doi: 10.1109/TPAMI.2009.161.

[13] M. Alam, M. S. Alam, M. Roman, M. Tufail, M. U. Khan and M. T. Khan, "Real-Time Machine-Learning Based Crop/Weed Detection and Classification for Variable-Rate Spraying in Precision Agriculture," 2020 7th International Conference on Electrical and Electronics Engineering (ICEEE), Antalya, Turkey, 2020, pp. 273-280, doi: 10.1109/ICEEE49618.2020.9102505.

[14] A. Olsen, D. A. Konovalov, B. Philippa, P. Ridd, J. C. Wood, J. Johns, W. Banks, B. Girgenti, O. Kenny, J. Whinney, B. Calvert, M. Rahimi Azghadi, and R. D. White, "DeepWeeds: A Multiclass Weed Species Image Dataset for Deep Learning," Scientific Reports, vol. 9, no. 2058, 2 2019. [Online]

[15] Liu, B., Bruch, R. Weed Detection for Selective Spraying: a Review. Curr Robot Rep 1, 19–26 (2020). https://doi.org/10.1007/s43154-020-00001-w

[16] M. Abdulsalam and N. Aouf, "Deep Weed Detector/Classifier Network for Precision Agriculture," 2020 28th Mediterranean Conference on Control and Automation (MED), Saint-Raphaël, France, 2020, pp. 1087-1092, doi: 10.1109/MED48518.2020.9183325.

[17] Y. Zheng, J. Kong, X. Jin, X. Wang, T. Su, and M. Zuo "CropDeep: The Crop Vision Dataset for Deep-Learning-Based Classification and Detection in Precision Agriculture" sensors(Basel), vol. 19, no. 5, March 2019.

[18] Gehan MA*, Fahlgren N*, Abbasi A, Berry JC, Callen ST, Chavez L, Doust AN, Feldman MJ, Gilbert KB, Hodge JG, Hoyer JS, Lin A, Liu S, Lizárraga C, Lorence A, Miller M, Platon E, Tessman M, Sax T. 2017. PlantCV v2: Image analysis software for high-throughput plant phenotyping.