Assessing the performance of YOLOv5 algorithm for detecting volunteer cotton plants in corn fields at three different growth stages

Pappu Kumar Yadav [a,*], J. Alex Thomasson [b], Stephen W. Searcy [a], Robert G. Hardin [a], Ulisses Braga-Neto [c], Sorin C. Popescu [d], Daniel E. Martin [e], Roberto Rodriguez [f], Karem Meza [g], Juan Enciso [a], Jorge Solórzano Diaz [h], Tianyi Wang [i]

[a] Department of Biological & Agricultural Engineering, Texas A&M University, College Station, TX, United States of America
[b] Department of Agricultural & Biological Engineering, Mississippi State University, Mississippi State, MS, United States of America
[c] Department of Electrical & Computer Engineering, Texas A&M University, College Station, TX, United States of America
[d] Department of Ecology & Conservation Biology, Texas A&M University, College Station, TX, United States of America
[e] Aerial Application Technology Research, U.S.D.A. Agriculture Research Service, College Station, TX, United States of America
[f] U.S.D.A.- APHIS PPQ S&T, Mission Lab, Edinburg, TX, United States of America
[g] Department of Civil & Environmental Engineering, Utah State University, Logan, UT, United States of America
[h] Texas A&M AgriLife Research & Extension Center, Weslaco, TX, United States of America
[i] College of Engineering, China Agricultural University, Beijing, China

ARTICLE INFO

ABSTRACT

The feral or volunteer cotton (VC) plants when reach the pinhead squaring phase (5–6 leaf stage) can act as hosts for the boll weevil (*Anthonomus grandis* L.) pests. The Texas Boll Weevil Eradication Program (TBWEP) employs people to locate and eliminate VC plants growing by the side of roads or fields with rotation crops but the ones growing in the middle of fields remain undetected. In this paper, we demonstrate the application of computer vision (CV) algorithm based on You Only Look Once version 5 (YOLOv5) for detecting VC plants growing in the middle of corn fields at three different growth stages (V3, V6 and VT) using unmanned aircraft systems (UAS) remote sensing imagery. All the four variants of YOLOv5 (s, m, l, and x) were used and their performances were compared based on classification accuracy, mean average precision (mAP) and F1-score. It was found that YOLOv5s could detect VC plants with maximum classification accuracy of 98% and mAP of 96.3% at V6 stage of corn while YOLOv5s and YOLOv5m resulted in the lowest classification accuracy of 85% and YOLOv5m and YOLOv5l had the least mAP of 86.5% at VT stage on images of size 416 × 416 pixels. The developed CV algorithm has the potential to effectively detect and locate VC plants growing in the middle of corn fields as well as expedite the management aspects of TBWEP.

1. Introduction

The boll weevil (*Anthonomus grandis* L.) is a serious pest that primarily feeds on cotton plants and has cost the U.S. cotton industry more than 23 billion USD in economic losses since it first entered to the U.S. from Mexico in the 1890s (Harden, 2018). The National Boll Weevil Eradication Program (NBWEP) has successfully eradicated the boll weevils from major parts of the U.S.; however southern parts of Texas (the Lower Rio Grande Valley) remain prone to re-infestation each year due to its sub-tropical climatic conditions and proximity to the Mexico border (Roming et al., 2021). The sub-tropical climatic conditions allow cotton plants to grow year-round and therefore the left-over

cotton seeds during the harvest from previous season can grow either at the edges of fields or in the middle of rotation crops like corn (*Zea mays* L.) and sorghum (*Sorghum bicolor* L.) (Wang et al., 2022; Yadav et al., 2019; Yadav et al., 2022a, 2022b). These feral or volunteer cotton (VC) plants when reach the pinhead squaring phase (5–6 leaves) can serve as hosts for the boll weevil pests (Yadav et al., 2022c).

As per the management practices of The Texas Boll Weevil Eradication Foundation (TBWEF), edges of rotation crop fields are inspected for the presence of VC plants and are eliminated if found, to avoid pest re-infestation in future. However, VC plants growing in the middle of fields remain undetected as it is practically not possible to detect and locate them in thousands of acres of fields. Therefore, if any boll weevil is found to be trapped in the pheromone traps of such fields, the whole field is sprayed with chemicals like Malathion ULV (FYFANON® ULV AG) to kill the pests (National Cotton Council of America, 2012). Usually

* Corresponding author.
  E-mail address: pappuyadav@tamu.edu (P.K. Yadav).

spray capable unmanned aircraft systems (UAS) are used for this purpose and the typical spray rate ranges from 0.56 to 1.12 kg/ha (FMC Corporation, 2001). Malathion is classified as toxicity class III pesticide and therefore can be dangerous to humans when not applied wisely (United States Environmental Protection Agency, 2016). Apart from this, excessive spraying of the chemical may kill beneficial insects and pests in the fields of rotation crops. Therefore, it has become a necessity to detect and locate VC plants growing in the middle of rotation crops. On doing this, a spot-spray capable UAS can precisely spray herbicides (before pinhead squaring phase) and pesticides (after pinhead squaring phase) only at the locations of VC plants. Hence, detecting the VC plants is crucial for the management aspects of TBWEF because it can speed up the management practices as well as help minimize the chemical costs associated with herbicides and pesticides.

Detecting VC plants growing in the middle of corn or sorghum fields is a two-step process i.e., classification followed by localization which necessarily means object detection task (Howard et al., 2019; Mustafa et al., 2020; Silwal et al., 2021; Zhao et al., 2019). In our past study, pixel-wise classification method was used with classical machine learning algorithms to detect the regions of VC plants in a corn field (Yadav et al., 2019). This approach resulted in a classification accuracy of around 70% which was far below our expectations and therefore couldn't be used for practical applications. In another approach by Westbrook et al. (2016), traditional image processing method-linear spectral unmixing was used to detect individual cotton plants at an early growth stage with fairly good accuracy. In recent past, many deep learning-based algorithms have been developed and used successfully for object detection tasks. Many of those algorithms use convolution neural networks (CNNs), some of which are YOLOv3 (Redmon and Farhadi, 2018), YOLOv5 (Jocher, 2020), Mask R-CNN (He et al., 2017), etc. The ability of CNNs to extract complex features automatically from images have made them quite popular in classification and detection tasks. This is why they have also been widely used in agricultural applications for varieties of detection tasks like in the case of weed detection in turfgrass (Yu et al., 2019), weed detection in vegetable crops (Jin et al., 2021), plant seedling detection (Jiang et al., 2019) and bunch detection in white grapes (Sozzi et al., 2022).

In this paper, we have shown the application of YOLOv5 for detecting VC plants growing in the middle of corn fields by using remote sensing multispectral imagery collected by unmanned aircraft systems (UAS).YOLOv5 is a one stage object detection algorithm that was originally released in four different variants (YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x) based on the network depth and number of parameters (Jocher, 2020; Jocher et al., 2021). The letters *s, m, l,* and *x* represent small, medium, large, and extra-large respectively depending upon the network depth and parameter size used. The specific objective of this paper is to do a comparative analysis of the performances of all the four variants of YOLOv5 for detecting VC plants in corn fields at three different growth stages (V3, V6 and VT) of corn plants. This study is an extension of our previous studies in which both YOLOv3 and YOLOv5 were used successfully in detecting VC plants at a single growth stage of corn plants (Yadav et al., 2022a, 2022b).

2. Materials and methods

2.1. Experiment sites

The experiment sites were located at two different corn fields: one was based in Hidalgo County near Weslaco, Texas (97°56′21.538″W, 26°9′49.049″N) while the other was based in Burleson County near College Station, Texas (96°25′45.9″W, 30°32′07.4″N) (Fig. 1). Two types of soil (Harlingen clay and Raymondville clay loam) are primarily present at the experiment plot based in Weslaco, TX while Weswood silty clay loam, Yahola fine sandy loam and Belk clay are present at the experiment site of Burleson county (USDA-Natural Resources Conservation Service, 2020). To mimic the presence of VC plants, 105 cotton seeds

of variety Phytogen 350 W3FE (CORTEVA agriscience, Wilmington, Delaware) were planted in the corn field of Weslaco. Some of these were planted in line with corn plants while the rest were planted in the furrow middles. Similarly, a total of 180 cotton seeds-90 of the variety Phytogen 340 W3FE (CORTEVA agriscience, Wilmington, Delaware) and another 90 of the variety Deltapine 1646 B2XF (Bayer AG, Leverkusen, North Rhine-Westphalia, Germany) were planted at the test field in Burleson County.

2.2. Image data acquisition

At the experiment site located in Weslaco, Texas, a three band (RGB: red, green, and blue) FC6310 camera (Shenzhen DJI Sciences and Technologies Ltd., Shenzhen, Guangdong, China) integrated on DJI Phantom 4 Pro quadcopter (Shenzhen DJI Sciences and Technologies Ltd., Shenzhen, Guangdong, China) was used to collect aerial imagery from an altitude of 18.3 m (60 ft) above ground level (AGL). The images acquired by the FC6310 camera had a resolution of 5472 × 3648 pixels and a spatial resolution of 0.5 cm/pixel (0.20 in./pixel). Data were collected on April 7, 2020, between 10 a.m. and 2 p.m. Central Standard Time (CST) at 80% sidelap and 75% overlap. The corn plants were at V3 stage when the data were acquired.

At the second site located in Burleson County, RedEdge-MX (AgEagle Aerial Systems Inc., Wichita, Kansas) camera was mounted on a custom UAS (Fig. 2) for aerial data collection. The first set of data were collected on May 5, 2021, between 11:00 a.m. and 2:00 p.m. central daylight-saving time (CDT) at an altitude of nearly 4.6 m (15 ft) above ground level (AGL) when the corn plants were at V6 growth stage. The second set of data were collected on May 14, 2022, between 11:00 a.m. and 2:00 p.m. CDT from an altitude of 4.6-m (15 ft) AGL when the corn plants had reached the VT stage. The acquired images on both the days had approximate spatial resolution of 0.34 cm/pixel.

2.3. YOLOv5

YOLOv5 is the fifth version of YOLO series of object detection algorithm which was released in June of 2020 (Jocher, 2020). Just like its predecessors, it is a single stage detector network which makes it faster compared to other object detection algorithms (Yan et al., 2021; Zhou et al., 2021). The simple architecture of YOLOv5 can be seen in Fig. 3 which was generated by a neural network visualization software tool Netron version 4.8.1 (Lutz, 2017). The overall architecture comprises of 25 nodes which are named from *model 0* to *model 24*. Nodes 0 to 9 (i.e., *model/0* to *model/9*) form the backbone network while nodes 10 to 23 (i.e., *model/10* to *model/23*) represent the neck network and the last node i.e., *model/24* forms the detection network. The last node comprises of three layers to make detections at three different scales with bounding boxes and confidence scores around the detected objects.

The backbone network is comprised of focus, convolution, bottleneckCSP (Cross Stage Partial) and spatial pyramid pooling (SPP) modules. The focus module accepts input images of shape 3 × 640 × 640, where 3 represents the three channels (normally Red, Green, and Blue) while 640 × 640 represent image width and height in pixels. The focus module can process images of sizes other than 640 × 640 pixels and can be customized for different numbers of channels as well. The main objective of the focus module is to enhance the training speed as it makes use of the "hard-swish" activation function. This activation function is a modified version of the swish activation function, replacing the sigmoid function with a piecewise-linear "hard" equivalent (eqs. 1 and 2) (Howard et al., 2019).

$$swish(x) = x\,\sigma(x) = \frac{x}{1+e^{-x}} \qquad (1)$$

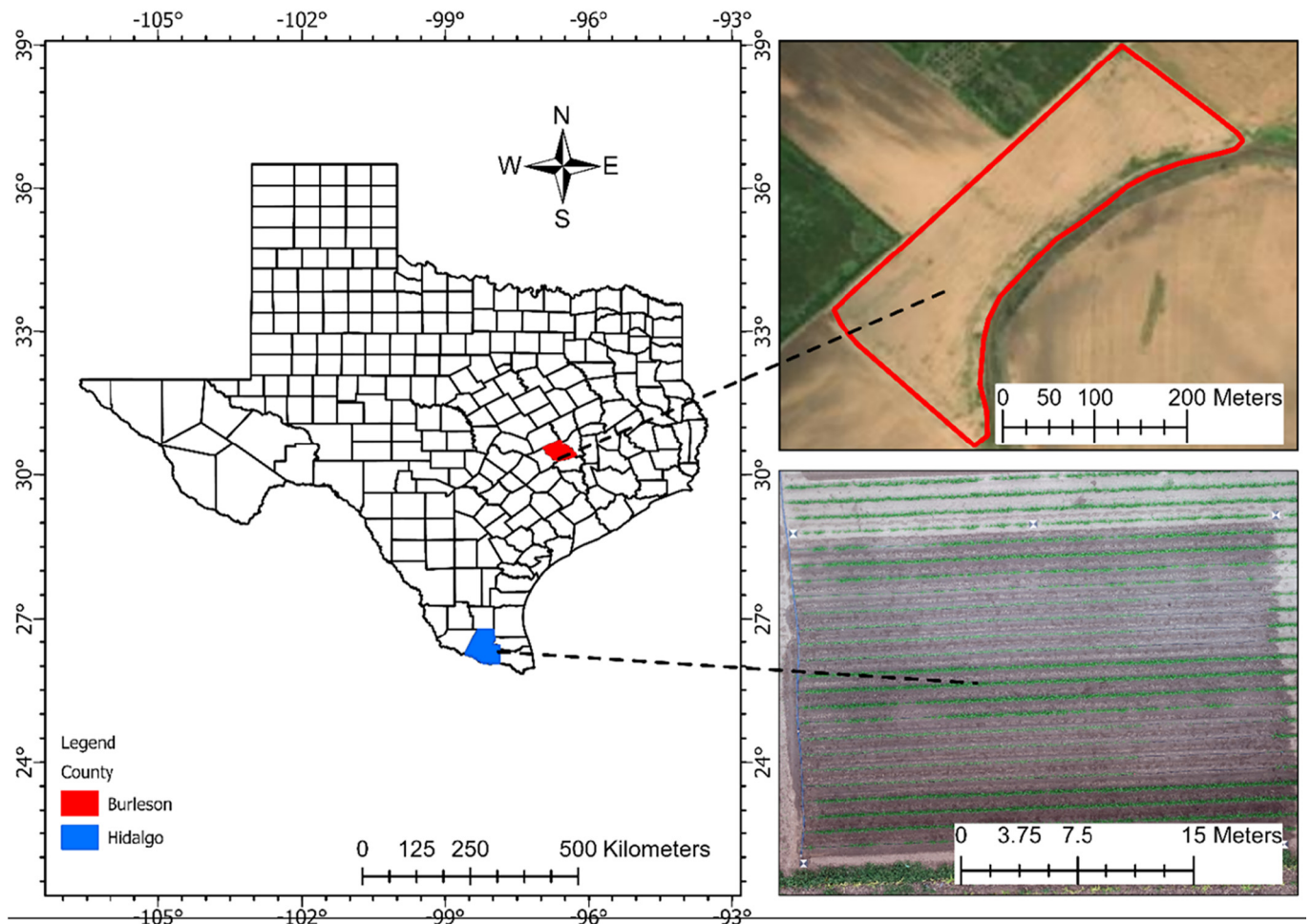$$h-swish(x) = x\,\frac{ReLU6(x+3)}{6} \qquad (2)$$

Fig. 1. Experiment site of two corn fields located in Burleson and Hidalgo counties.

where σ is the sigmoid function and ReLU6 is a modified version of rectified linear unit (ReLU) activation function in which the maximum value is limited to 6. The bottleneckCSP module enhances the feature extraction process by making use of a residual network, i.e., concatenating deeper features with shallow features. The SPP module is used to ensure that a fixed-size feature vector is generated from any different size feature maps by using three parallel maxpooling layers.

The neck network is used to ensure that detection accuracy of objects is not compromised due to different scales and sizes. It makes use of a path aggregation network (PANet) to enhance feature propagation irrespective of scale (Liu et al., 2018). Modules 17, 20 and 23, which are part of the PANet and belong to the Neck network (also called P3, P4 and P5), output three feature maps belonging to objects at scales of small, medium, and large, respectively. P3 outputs a feature map of 80×80 pixels, while P4 and P5 output feature maps of 40×40 and 20×20, respectively.

The detect network is comprised of three detect layers to make detections at three scales corresponding to the features output from P3, P4 and P5. This network applies three anchor boxes at each scale on the three feature maps to output a vector containing information
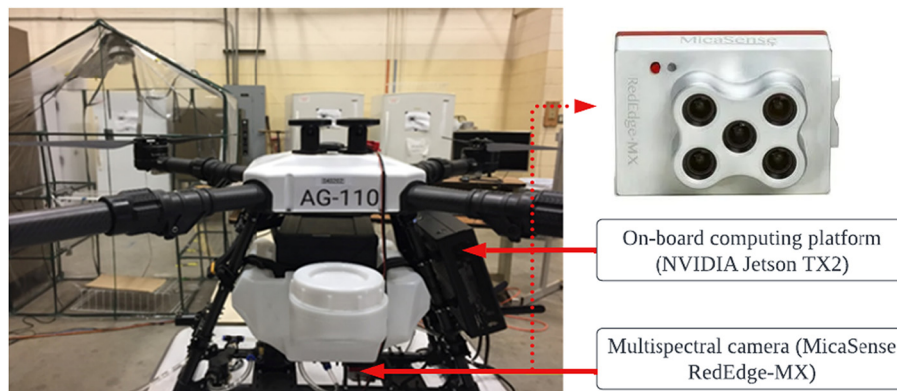

Fig. 2. A customized unmanned aircraft systems (UAS) with on-board computing platform and MicaSense RedEdge-MX multispectral camera.
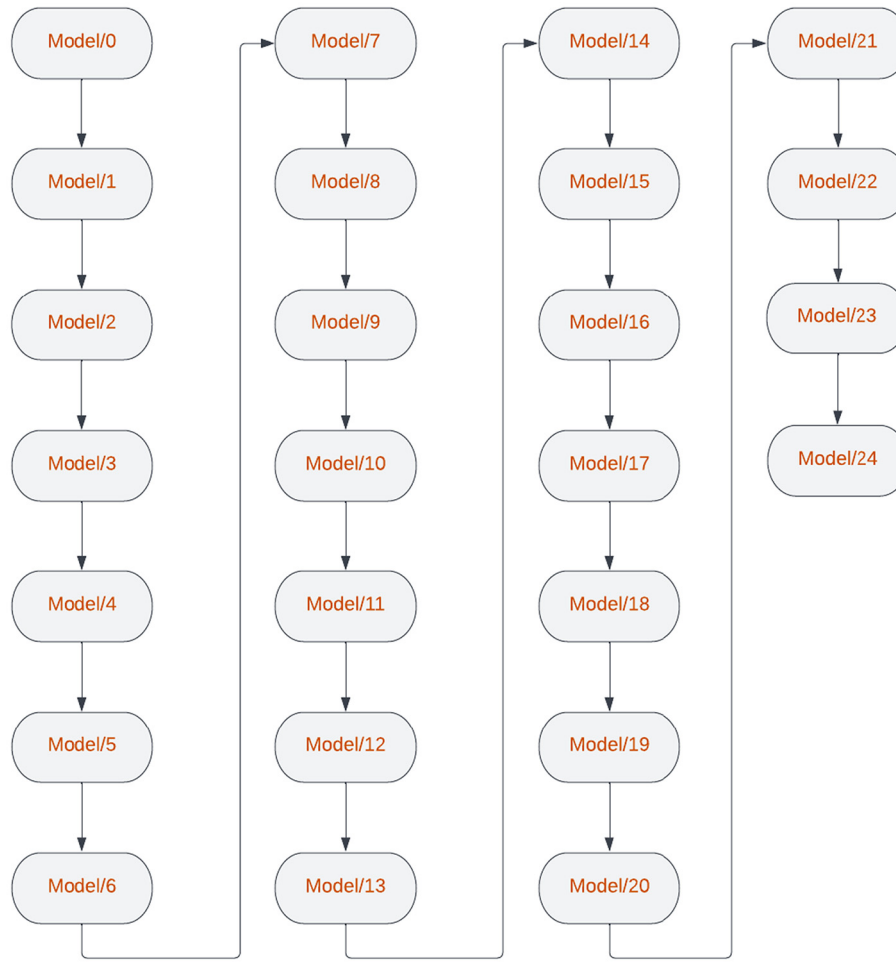
```
Model/0          Model/7          Model/14         Model/21
   ↓                ↓                ↓                ↓
Model/1          Model/8          Model/15         Model/22
   ↓                ↓                ↓                ↓
Model/2          Model/9          Model/16         Model/23
   ↓                ↓                ↓                ↓
Model/3          Model/10         Model/17         Model/24
   ↓                ↓                ↓
Model/4          Model/11         Model/18
   ↓                ↓                ↓
Model/5          Model/12         Model/19
   ↓                ↓                ↓
Model/6          Model/13         Model/20
```

Fig. 3. Simplified version of YOLOv5 architecture generated by Netron visualization software tool.

about the class probability, objectness score and predicted bounding box (BB) coordinates.

The original architecture of YOLOv5 was customized to detect a single class of VC plants and accept input images of shape $3 \times 416 \times 416$ for this study. YOLOv5 was originally released in four different variants: YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x based on network depth and number of parameters used. Here, s, m, l, and x represent small, medium, large, and extra-large respectively. All these four variants were used in this study.

2.4. Performance metrics of YOLOv5

The performances of each trained YOLOv5 model were assessed by using the metrices accuracy, precision (P), recall (R), mean average precision (mAP) and F1-score as calculated in Eqs. 3, 4, 5, 6 and 7.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$P = \frac{TP}{TP + FP} \tag{4}$$

$$R = \frac{TP}{TP + FN} \tag{5}$$

$$F1-\text{Score} = 2 \times \frac{P \times R}{P + R} \tag{6}$$

$$\text{mAP} = \frac{1}{n} \sum_{k=1}^{k=N} AP_k \tag{7}$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, FN is the number of false negatives, AP_k is the average precision (AP) of class k (in our case $N = 1$ i.e., VC class), and n is the number of thresholds ($n = 1$ in our case i.e., 0.50 or 50%). This essentially means, the mAP and AP are same for a single class case like in this study. The mAP values were calculated at 50% threshold value for intersection over union (IoU) meaning, all the predicted bounding boxes that resulted in ratios of overlapping areas to the union areas with ground truth bounding greater than 50% were considered and the remaining were discarded (Gan et al., 2021; Padilla et al., 2021; Sharma, 2020; Wu et al., 2021). mAP is a metric based on the area under precision-recall curve (PRC) that is preprocessed to eliminate zig-zag behavior (Padilla et al., 2021). Therefore, in case of class imbalance where there are more instances of one class than the other, mAP is a more reliable and powerful metric to analyze performance of a classifier (Padilla et al., 2021; Saito and Rehmsmeier, 2015).

2.5. Dataset preparation for training YOLOv5

The RGB images collected at experiment plot located in Weslaco, TX, were split into 416×416-pixel after which augmentation techniques were applied to images that contained at least a VC plant using

Augmentor Python library (Bloice, 2017) that generated a total of 409 images. To accomplish this, *rotate*, *flip_left_right*, *zoom_random* and *flip_top_bottom* operations were used with probability values of 0.80, 0.40, 0.60 and 0.80 respectively as explained by Bloice et al. (2017). Roboflow, a web-based platform was used to annotate ground truth bounding boxes for VC plants (Roboflow, 2022) by first dividing the images in the ratio 16:3:1 for training, validation and test dataset and then exporting them in YOLOv5 PyTorch format. A total of more than 1750 instances for VC class were obtained in the training, validation, and test datasets (Fig. 4A). It was also found that majority of VC plants were of 4 to 6 cm in height and 5 cm in width (Fig. 4A).

The five-band multispectral images collected at experiment site located in Burleson County, Texas were radiometrically corrected and then some preprocessing methods were applied using the source code from GitHub (GitHub, Inc., San Francisco, CA, U.S.A.) repository of MicaSense (MicaSense Incorporated, 2022). This generated RGB images of size 1207 × 923 pixels. Among the generated RGB images, only the ones which had VC plants were chosen and then split into size 416 × 416 pixels. This again resulted in many images without VC plants which were again discarded. The resulting images with at least one VC plants were then augmented to generate a total of 387 images at V6 stage and 280 at VT stage. Roboflow was then used to annotate the ground truth
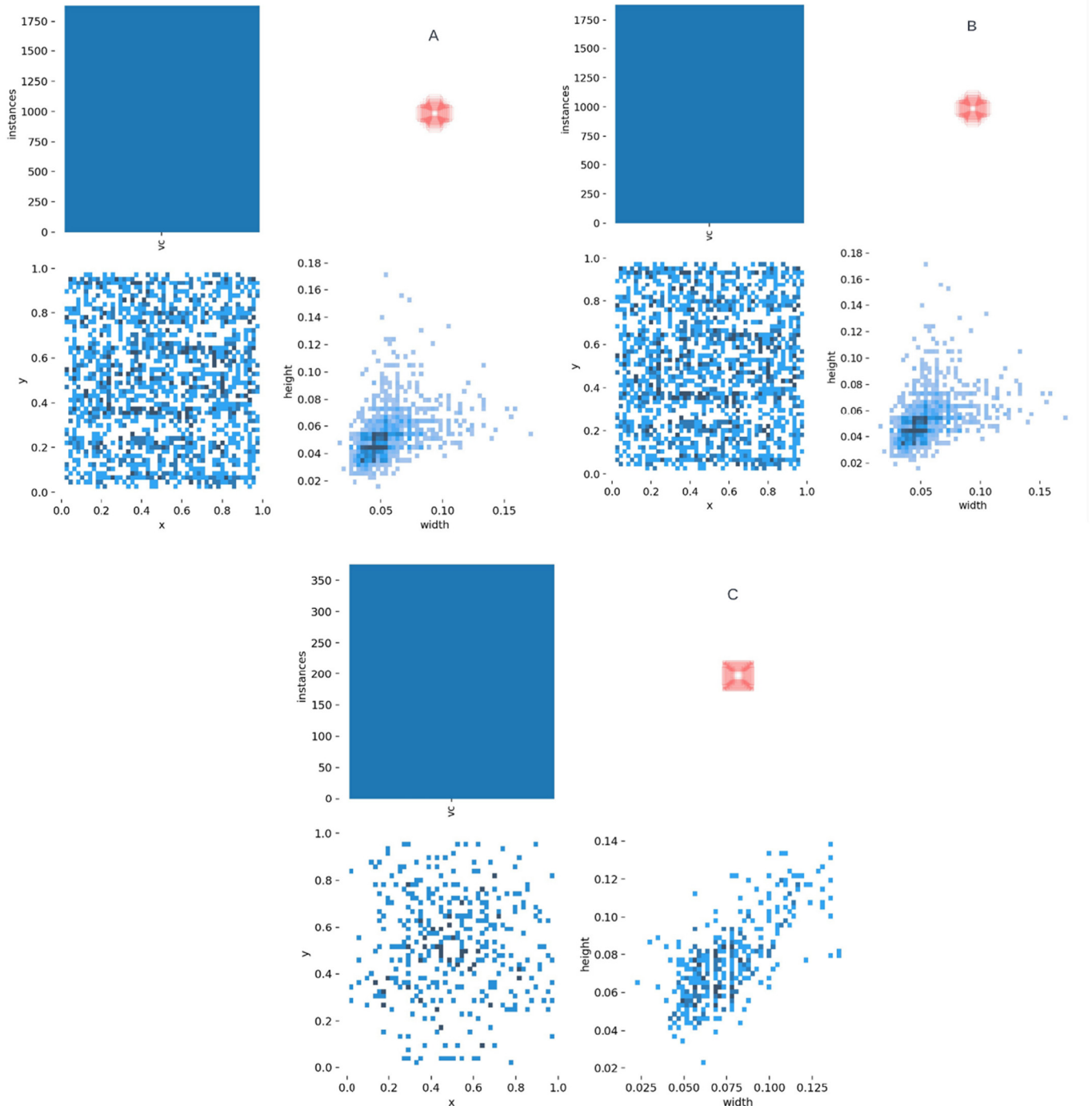

Fig. 4. VC class instances in training, validation and test datasets used for the YOLOv5 model at V3 (A), V6 (B) and VT (C) growth stages of corn plants.

bounding boxes for VC plants. Of these, training, validation, and test images were divided in the ratio 16:3:1. A total of more than 800 instances (Fig. 4B) of VC class were obtained in the datasets for the V6 stage and 400 for the VT stage (Fig. 4C). Most of the VC plants at V6 stage of corn were of 6 to 8 cm in height and width while at the VT stage, they were mostly of height 8 cm and width 8 to 10 cm (Fig. 4B and C). The VC plants at V3 and VT stage of corn were almost equally distributed at all the regions of the images while they were mostly located at the four corners of the images at V6 growth stage of corn (Fig. 4A, B and C).

2.6. YOLOv5 training

The source code for YOLOv5 was obtained from the GitHub repository of Ultralytics Inc. (Jocher, 2020). All the four variants of YOLOv5 were trained on Tesla P100-PCIE-16 GB (NVIDIA, Santa Clara, CA, U.S.A.) GPU using the Google Colaboratory (Google LLC, Melno Park, CA, U.S.A.) AI platform. Each of them was trained with initial learning rate of 0.01, final learning rate of 0.1, momentum of 0.937, weight decay of 0.0005 for a total of 200 in Fig. 5.

3. Results

3.1. VC detection in a field with corn at V3 growth stage

The results for different performance metrics that were obtained during the training process using the training and validation datasets are shown in Fig. 6.

At V3 stage dataset, the convergence was achieved within the 200 training iterations by using transfer learning approach in the training process i.e., by starting training from the pretrained weights of YOLOv5 (COCO dataset weights). It was found that the maximum values of precision, recall, mAP, and F1-score reached 87%, 91%, 89% and 83% respectively. Similarly, 94% of the VC plants were correctly classified while 6% were mistaken for background class i.e., either corn plants or weeds. However, none of the instances from the background class were mistaken for VC plants (Fig. 7-left). Almost every image had more instances of background class (i.e., corn and weeds) than the VC class, which means class instances were imbalanced. From past studies, it has been found that classifier's performance can be biased towards the majority class, therefore it has been recommended to analyze performance based on the PRC as seen in Fig. 6 (Saito and Rehmsmeier, 2015). Some of the detection results of VC plants are shown in Fig. 8 where the detected VC plants are enclosed within the red bounding boxes with corresponding confidence scores.

3.2. VC detection in a field with corn at V6 growth stage

In Fig. 9, the graphs for different performance metrics can be seen that were obtained after training YOLOv5s by using datasets belonging to corn plants at V6 growth stage.

Due to lack of improvement in performance for 100 iterations, there was early stopping of the training around the 130th iteration. This means that convergence was achieved well before the 130th iteration. The maximum values of precision, recall, mAP, and F1-score were found to be 97%, 97%, 96% and 93% respectively. In terms of classification accuracy, 98% of the VC plants were correctly classified while only 2% of them were misclassified as corn or weeds and none of the background class (i.e., corn and weeds) were misclassified as VC plants (Fig. 10-left).
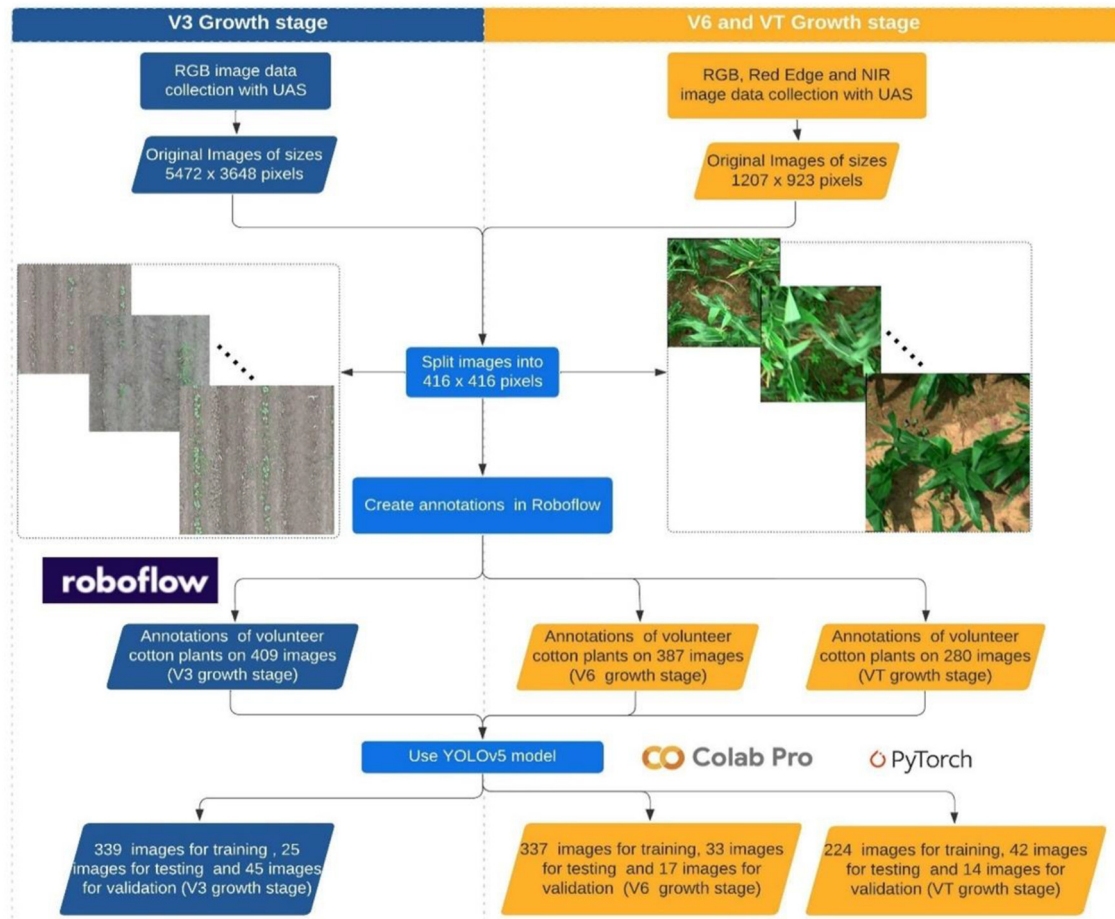

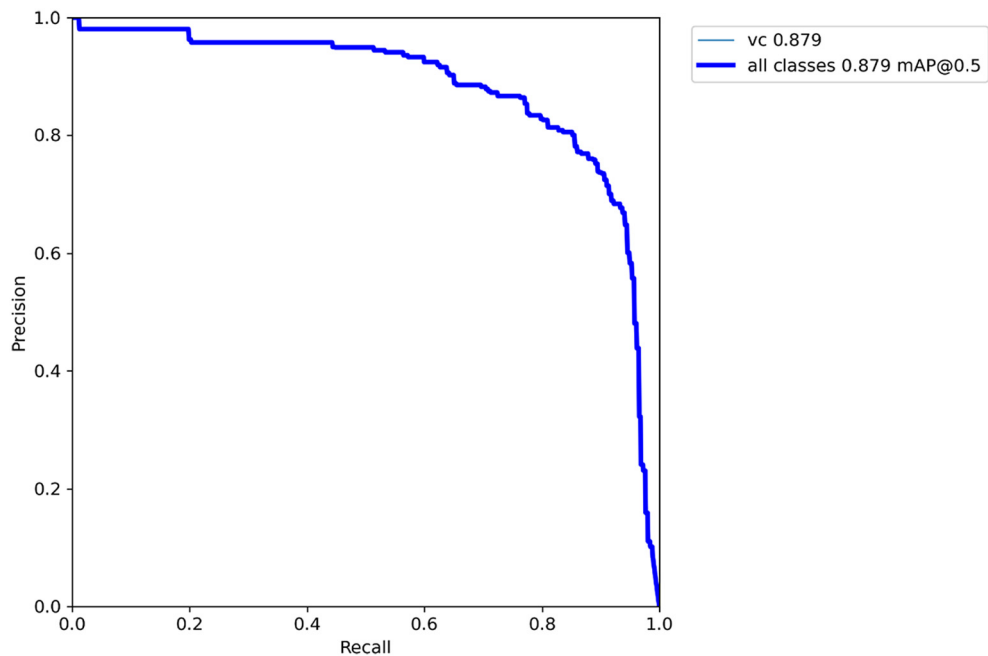**Fig. 5.** Work flowchart showing entire processes from data acquisition to YOLOv5 training.

Fig. 6. Example graphs showing performance metrics along with PR-curve (PRC) obtained after training YOLOv5s using the dataset belonging to the V3 growth stage of corn plants.
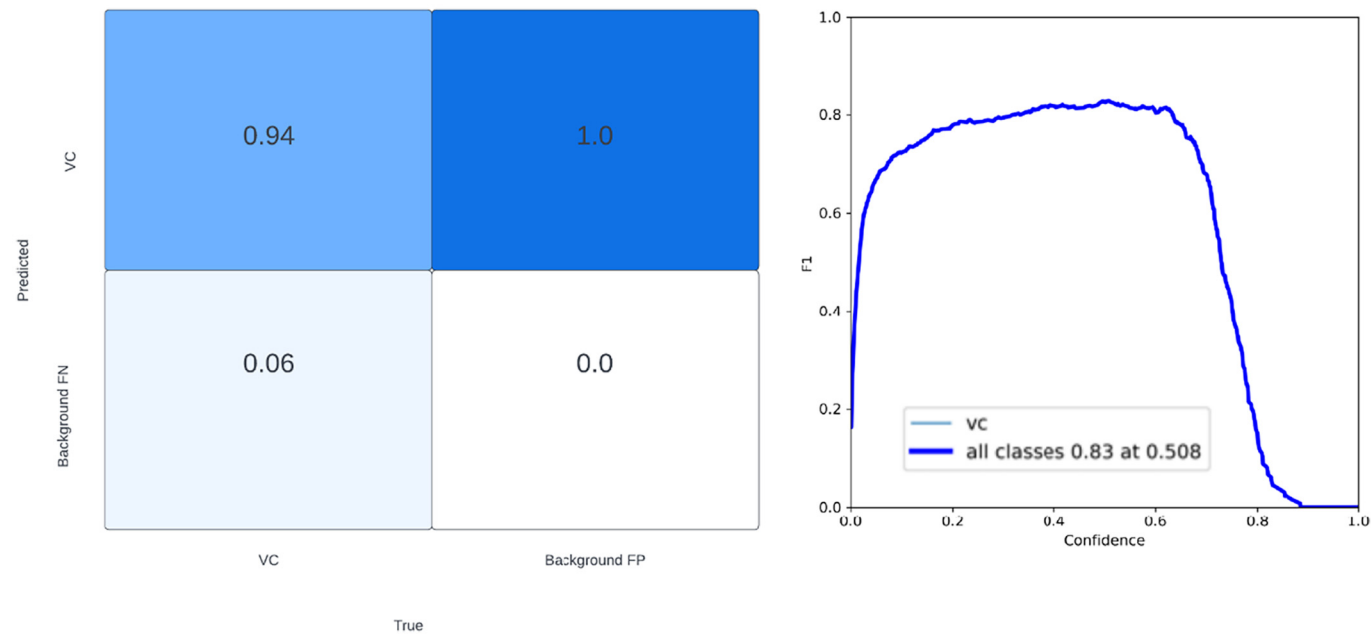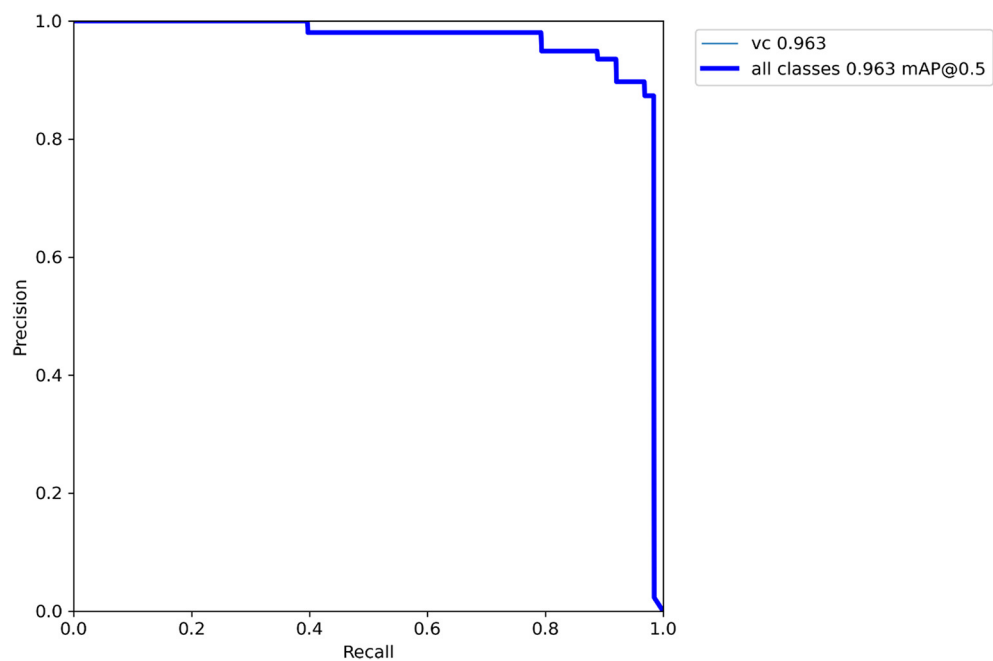

Fig. 7. Classification results shown in confusion matrix (left) and F1-score calculated over a range of confidence scores (right) of VC plants after YOLOv5s was trained on dataset belonging to the V3 growth stage of corn plants.

Fig. 8. YOLOv5s detection results of VC plants in a field of corn at V3 growth stage.

Fig. 9. Example graphs showing performance metrics along with PR-curve (PRC) obtained after training YOLOv5s using the dataset belonging to the V6 growth stage of corn plants.
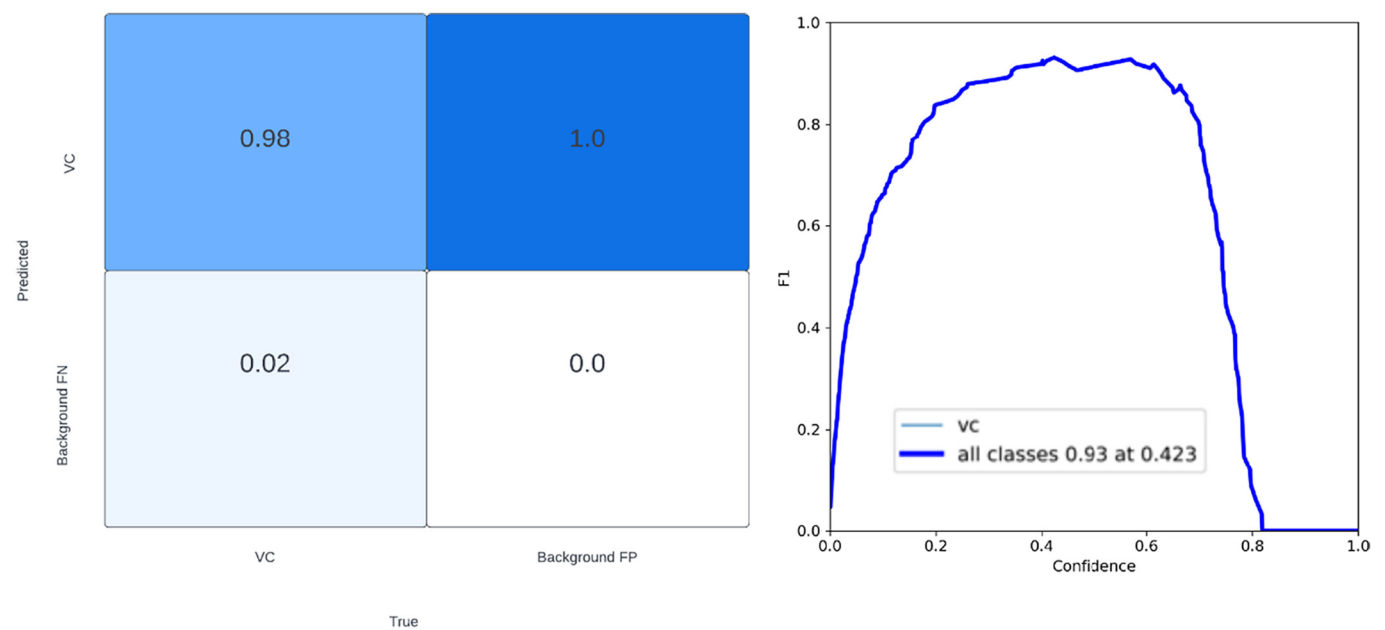

Fig. 10. Classification results shown in confusion matrix (left) and F1-score calculated over a range of confidence scores (right) of VC plants after YOLOv5s was trained on dataset belonging to the V6 growth stage of corn plants.
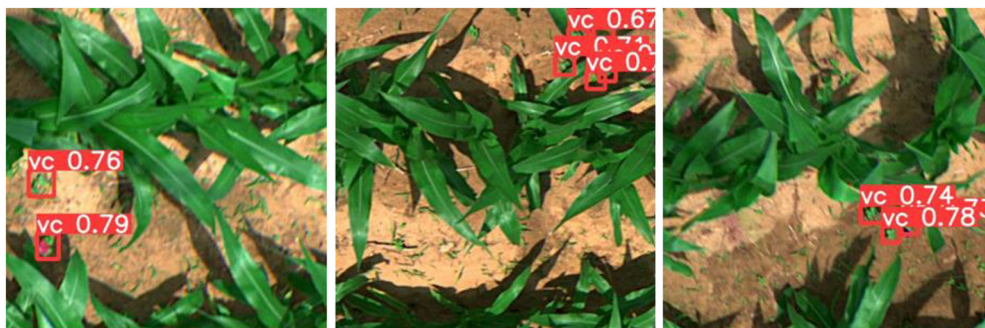

Fig. 11. YOLOv5s detection results of VC plants in a field of corn at V6 growth stage.
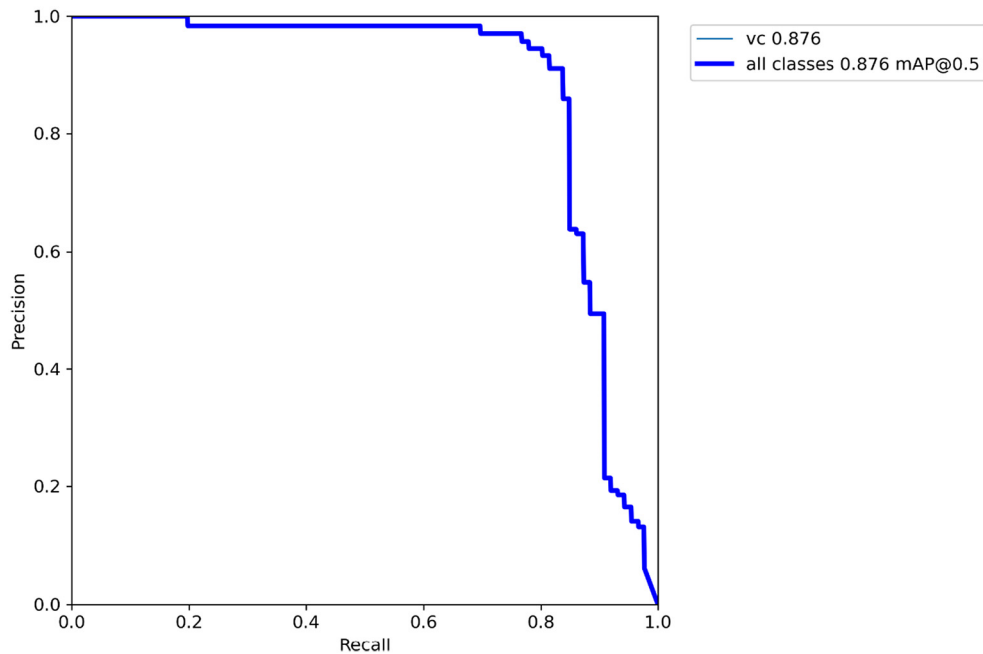
Fig. 12. Example graphs showing performance metrics along with PR-curve (PRC) obtained after training YOLOv5s using the dataset belonging to the VT growth stage of corn plants.

Some of the detection results of VC plants are shown in Fig. 11 where the detected VC plants are enclosed within the red bounding boxes with their corresponding confidence scores.

3.3. VC detection in a field with corn at VT growth stage

Fig. 12 shows the graphs for different performance metrics that were obtained after training YOLOv5s by using datasets belonging to corn plants at VT growth stage.

As in the V6 case, there was early stopping in the training process around 165th iteration for the VT case too as no improvement was observed in the last 100 iterations. This implies that convergence was achieved well before the 165th iteration. The maximum values of precision, recall, mAP, and F1-score were found to be 99%, 86%, 89% and 87% respectively. In terms of classification accuracy, 85% of the VC plants were correctly classified while 15% of them were misclassified as corn or weeds and none of the background class (i.e., corn and weeds) were misclassified as VC plants (Fig. 13-left).

Some of the detection results of VC plants are shown in Fig. 14 where the detected VC plants are enclosed within the red bounding boxes with their corresponding confidence scores.
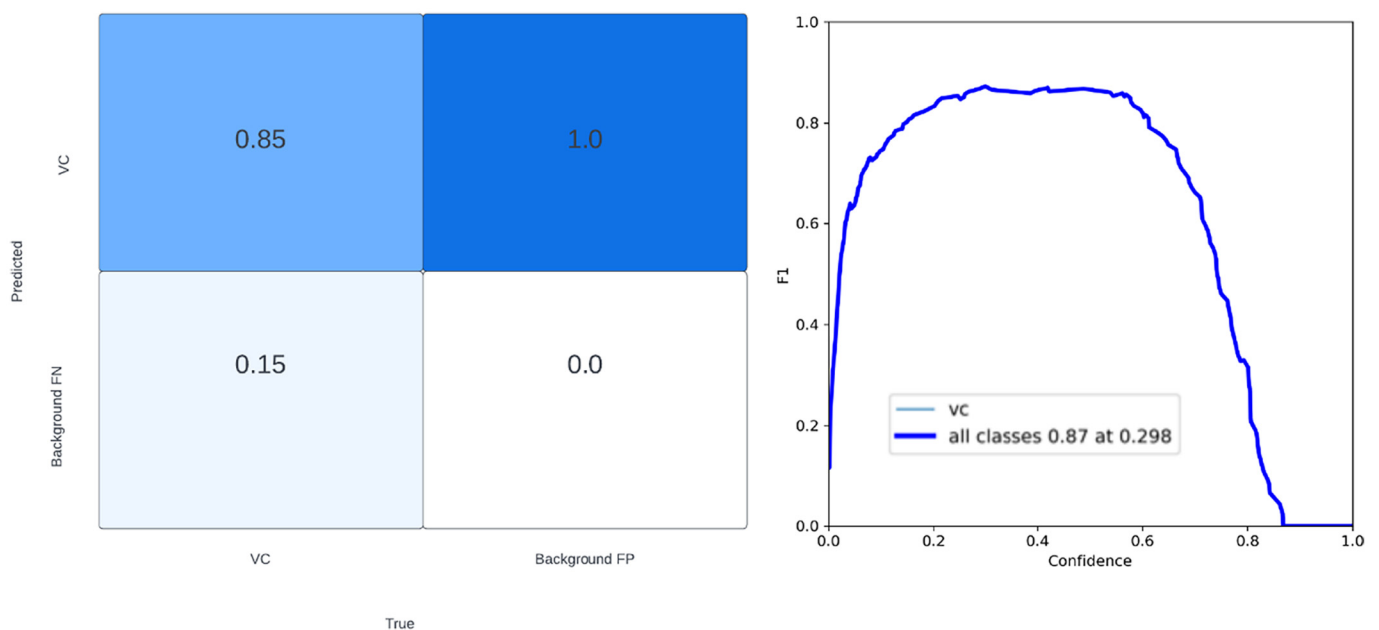

Fig. 13. Classification results shown in confusion matrix (left) and F1-score calculated over a range of confidence scores (right) of VC plants after YOLOv5s was trained on dataset belonging to the VT growth stage of corn plants.
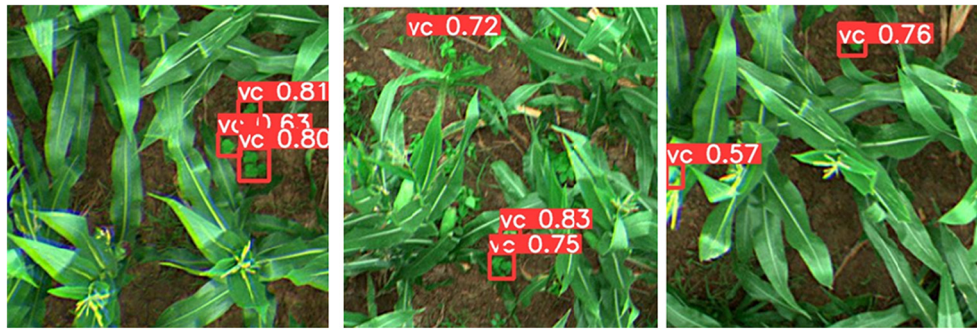
Fig. 14. YOLOv5s detection results of VC plants in a field of corn at VT growth stage.

Table 1
Comparison of performance metrics of four variants of YOLOv5 in detecting VC plants growing in the middle of fields with corn plants at three different growth phases.

Corn	Classification_Acc				mAP@50				F1-Max			
	YOLOv5				YOLOv5				YOLOv5			
Growth Phase	s	m	l	x	s	m	l	x	s	m	l	x
V3 (early)	94	92	90	93	87.9	86.5	86.5	87.4	83	82	83	83
V6 (mid)	98	95	92	95	96.3	93.5	91.6	92.0	93	92	91	91
VT (tassel)	85	85	91	86	87.6	90.3	89.0	87.6	87	88	85	88

3.4. Comparison of detection results at three different growth stages of corn with all the four variants of Yolov5

Table 1 shows the comparative results of all the four variants of YOLOv5 in detecting VC plants growing in fields with corn plants at three different growth phases i.e., at V3, V6 and VT.

It is evident that VC plants were most accurately classified from the V6 growth stage corn than the V3 and VT growth stages by using all the four variants of YOLOv5. In terms of detection accuracy based on mAP at 50% IoU, similar results were observed (Table 1). Between V3 and VT growth stage, VC plants were more accurately classified as well as detected in the V3 than the VT stage.

The average mAP for detecting VC plants in fields with corn plants at V3, V6 and VT growth stages by using all the four variants of YOLOv5 were found to be 87%, 93% and 89% with standard deviations of 0.69, 2.13 and 1.3 respectively. Similarly, the average classification accuracy of VC plants in fields with corn plants at V3, V6 and VT growth stages were found to be 92%, 95% and 87% with standard deviations of 1.7, 2.4 and 2.9 respectively.

4. Discussion

The images acquired at V3 growth stage of corn plants were captured by high resolution camera but at the same time they were captured from an altitude that is four times more than the images that were captured at V6 and V8 growth stages. This resulted in a lower spatial resolution (0.5 cm/pixel) than the images captured at the other two growth stages (0.34 cm/pixel). At this stage, there were not many weeds in the background of images which is why the average classification accuracy of 92% was promising (Fig. 15A). Even though this was an increment of more than 31% than our previous approach (Yadav et al., 2019), the average classification accuracy at the V6 stage was found to be even more at 95% (an increment of nearly 36%). It was expected that the accuracy might be lower at V6 stage as compared to the V3 stage due to larger canopy size of corn plants and relatively more weeds in the background. However, the higher spatial resolution in the imagery by capturing images at lower altitude (4.6 m /15 ft) resulted in a better classification accuracy (Table 1, Fig. 15A). The other reason for higher accuracy at V6 stage could be due to the fact that the images were preprocessed by using affine transform, unsharp mask and gamma correction (MicaSense Incorporated, 2022). However, these effects were not enough to improve classification accuracy at the VT growth stage (Fig. 15A). We assume this was because at this stage the weeds in the background were prominent and caused in more misclassifications than at the V6 stage.

In this study, at all the three growth stages of corn, the instances of VC plants as compared to the instances of background class (i.e., corn with weeds) was always less which means there was imbalance between the positive (VC plants) and negative (background) classes. In such cases, performance of any classifiers are found to be biased towards the majority class and therefore one cannot
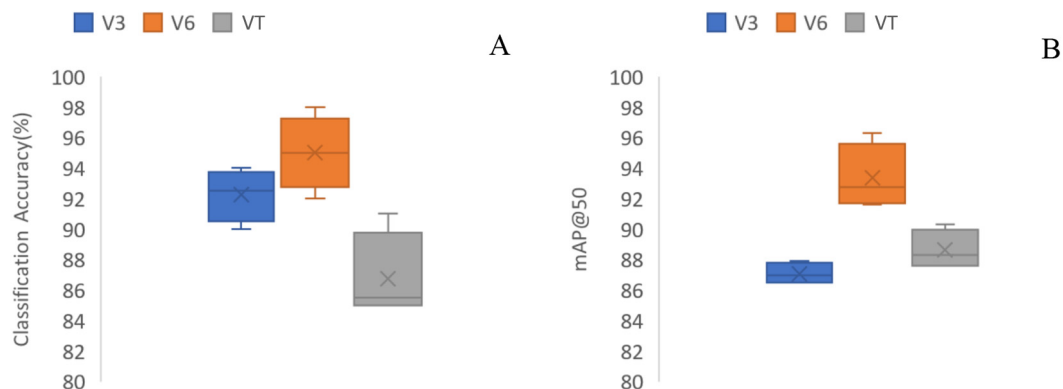

Fig. 15. (A) Classification accuracy distribution with all the four variants of YOLOv5 at all the three growth stages of corn plants. (B) mAP distribution with all the four variants of YOLOv5 at all the three growth stages of corn plants.

accurately make inference about the performance (Saito and Rehmsmeier, 2015; Sofaer et al., 2019). In such cases, area under the PRC and AP/mAP are found to be more effective measure of the classifier performance (Miao and Zhu, 2021). Therefore, to make fair comparison without results being affected by class instance imbalance, mAP values are used from which the highest detection accuracy was found at the V6 growth stage followed by VT and V2 (Fig. 15B). This result was expected because images at V6 and VT stages were of higher spatial resolution as well had undergone preprocessing techniques which were absent in the case of images at V3 stage. The presence of prominence in weeds at VT stage resulted in lower detection accuracy as compared to the V6 stage.

The findings reported in this paper has the potential to be used for near real-time VC detection in corn fields at V6 growth stage (for maximum accuracy) by deploying the trained YOLOv5 model on a spot-spray capable UAS (Yadav et al., 2022c) with a computer vision capability.

5. Conclusions and future recommendations

There is a tradeoff between detection accuracy and the amount of area surveyed because to survey larger area, images need to be captured from higher altitude like in the case of V3 stage which eventually decreases spatial resolution and a compromise in detection accuracy. However, if 87% mAP is not too low then we recommend of surveying at V3 growth stage for detecting VC plants as it results in detecting them in a larger area. We also expect the detection accuracy to improve and perhaps surpass that of V6 and VT stage when the images can be preprocessed using the techniques described in the GitHub repository of MicaSense (MicaSense Incorporated, 2022). This way, the system will be more practically viable and applicable by the personnel involved in the boll weevil eradication program.

CRediT authorship contribution statement

Pappu Kumar Yadav: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **J. Alex Thomasson:** Conceptualization, Data curation, Investigation, Methodology, Formal analysis, Project administration, Resources, Supervision, Writing – review & editing. **Stephen W. Searcy:** Formal analysis, Writing – review & editing. **Robert G. Hardin:** Project administration, Formal analysis, Resources, Supervision, Writing – review & editing. **Ulisses Braga-Neto:** Investigation, Methodology, Formal analysis, Writing – review & editing. **Sorin C. Popescu:** Formal analysis, Writing – review & editing. **Daniel E. Martin:** Writing – review & editing. **Roberto Rodriguez:** Data curation, Formal analysis, Writing – review & editing. **Karem Meza:** Resources, Writing – review & editing. **Juan Enciso:** Resources, Writing – review & editing. **Jorge Solórzano Diaz:** Resources, Writing – review & editing. **Tianyi Wang:** Writing – review & editing.

Declaration of Competing Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

References

Bloice, M.D., 2017. Augmentor: Image Augmentation Library in Python for Machine Learning. https://doi.org/10.5281/ZENODO.1041946.

Bloice, M., Stocker, C., Holzinger, A., 2017. Augmentor: An Image Augmentation Library for Machine Learning. *ArXiv Preprint ArXiv:1708.04680.* https://doi.org/10.21105/joss.00432.

FMC Corporation, 2001. FYFANON ULV AG. In *FYFANON ULV AG.* https://doi.org/10.1016/b978-081551381-0.50007-5.

Gan, T., Zha, Z., Hu, C., Jin, Z., 2021. Detection of Polyps during Colonoscopy Procedure Using YOLOv5 Network.

Harden, G.H., 2018. Texas boll weevil eradication foundation cooperative agreement. https://www.usda.gov/sites/default/files/33099-0001-23.pdf.

He, K., Gkioxari, G., Dollar, P., Girshick, R., 2017. Mask R-CNN. IEEE International Conference on Computer Vision (ICCV). pp. 2961–2969.

Howard, A., Sandler, M., Chen, B., Wang, W., Chen, LC., Tan, M., Chu, G., Vasudevan, V., Zhu, Y., Pang, R., Le, Q., Adam, H., 2019. Searching for mobileNetV3. Proceedings of the IEEE/CVF International Conference on Computer Vision. 1314–1324. https://doi.org/10.1109/ICCV.2019.00140.

Jiang, Y., Li, C., Paterson, A.H., Robertson, J.S., 2019. DeepSeedling: deep convolutional network and Kalman filter for plant seedling detection and counting in the field. Plant Methods 15 (1), 1–19. https://doi.org/10.1186/s13007-019-0528-3.

Jin, X., Che, J., Chen, Y., 2021. Weed identification using deep learning and image processing in vegetable plantation. IEEE Access 9, 10940–10950. https://doi.org/10.1109/ACCESS.2021.3050296.

Jocher, Glenn, 2020. YOLOv5. https://github.com/ultralytics/yolov5.

Jocher, G., Kwon, Y., guigarfr, perry0418, Veitch-Michaelis, J., Ttayu, Suess, D., Baltacı, F., Bianconi, G., IlyaOvodov, Marc, Lee, C., Kendall, D., Falak, Reveriano, F., FuLin, GoogleWiki, Nataprawira, J., ... Shead, T.M., 2021. ultralytics/yolov3: v9.5.0 - YOLOv5 v5.0 release compatibility update for YOLOv3. e96031413 https://doi.org/10.5281/ZENODO.4681234.

Liu, S., Qi, L., Qin, H., Shi, J., Jia, J., 2018. Path aggregation network for instance segmentation. Comp. Soc. Conf. Comp. Vision Pattern Recog. 8759–8768. https://doi.org/10.1109/CVPR.2018.00913.

Lutz, R., 2017. Netron. https://github.com/lutzroeder/netron.

Miao, J., Zhu, W., 2021. Precision–recall curve (PRC) classification trees. Evol. Intel. https://doi.org/10.1007/s12065-021-00565-2.

MicaSense Incorporated, 2022. MicaSense RedEdge and Altum Image Processing Tutorials. https://github.com/micasense/imageprocessing.

Mustafa, T., Dhavale, S., Kuber, M.M., 2020. Performance analysis of inception-v2 and Yolov3-based human activity recognition in videos. SN Comp. Sci. 1 (3), 1–7. https://doi.org/10.1007/s42979-020-00143-w.

National Cotton Council of America, 2012. Protocol for the Eradication of the Boll Weevil in the Lower Rio Grande Valley in Texas and Tamaulipas, Mexico. In *cotton.org.* https://www.cotton.org/tech/pest/bollweevil/upload/ITAC-Protocol-d8-Eng-Dec-2012-LS-ed.pdf.

Padilla, R., Passos, W.L., Dias, T.L.B., Netto, S.L., Da Silva, E.A.B., 2021. A comparative analysis of object detection metrics with a companion open-source toolkit. Electronics (Switzerland) 10 (3), 1–28. https://doi.org/10.3390/electronics10030279.

Redmon, J., Farhadi, A., 2018. YOLOv3:An Incremental Improvement. Computer Vision and Pattern Recognition. https://pjreddie.com/media/files/papers/YOLOv3.pdf.

Roboflow, 2022. Roboflow. https://roboflow.com/.

Roming, R., Leonard, A., Seagraves, A., Miguel, S.S., Jones, E., Ogle, S., 2021. Sunset Staff Reports with Final Results. www.sunset.texas.gov.

Saito, T., Rehmsmeier, M., 2015. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS One 10 (3), 1–21. https://doi.org/10.1371/journal.pone.0118432.

Sharma, Vinay, 2020. *Face Mask Detection using YOLOv5 for COVID-19* [California State University-San Marcos]. https://scholarworks.calstate.edu/downloads/wp988p69r?locale=en.

Silwal, A., Parhar, T., Yandun, F., Kantor, G., 2021. A Robust Illumination-Invariant Camera System for Agricultural Applications. http://arxiv.org/abs/2101.02190.

Sofaer, H.R., Hoeting, J.A., Jarnevich, C.S., 2019. The area under the precision-recall curve as a performance metric for rare binary events. Methods Ecol. Evol. 10 (4), 565–577. https://doi.org/10.1111/2041-210X.13140.

Sozzi, M., Cantalamessa, S., Cogato, A., Kayad, A., Marinello, F., 2022. Automatic bunch detection in white grape varieties using YOLOv3, YOLOv4, and YOLOv5 deep learning algorithms. Agronomy 12 (2), 319. https://doi.org/10.3390/agronomy12020319.

UNITED ST ATES ENYLRONMENT AL PROTECTION AGENCY, 2016. Malathion: Human Health Draft Risk Assessment for Registration Review.

USDA-Natural Resources Conservation Service, 2020. Web Soil Survey. https://websoilsurvey.sc.egov.usda.gov/App/HomePage.htm.

Wang, T., Mei, X., Alex Thomasson, J., Yang, C., Han, X., Yadav, P.K., Shi, Y., 2022. GIS-based volunteer cotton habitat prediction and plant-level detection with UAV remote sensing. Comput. Electron. Agric. 193 (December 2021), 106629. https://doi.org/10.13031/aim.202000219.

Westbrook, J.K., Eyster, R.S., Yang, C., Suh, C.P.C., 2016. Airborne multispectral identification of individual cotton plants using consumer-grade cameras. Remote Sens. Appl. Soc. Environ. 4, 37–43. https://doi.org/10.1016/j.rsase.2016.02.002.

Wu, W., Liu, H., Li, L., Long, Y., Wang, X., Wang, Z., Li, J., Chang, Y., 2021. Application of local fully convolutional neural network combined with YOLO v5 algorithm in small target detection of remote sensing image. PLoS One 16 (10), e0259283. https://doi.org/10.1371/journal.pone.0259283.

Yadav, P., Thomasson, J.A., Enciso, J., Samanta, S., Shrestha, A., 2019. Assessment of different image enhancement and classification techniques in detection of volunteer

cotton using UAV remote sensing. SPIE Defense + Commerc. Sens. https://doi.org/10.1117/12.2518721 May 2019, 20.

Yadav, P.K., Thomasson, J.A., Hardin, R.G., Searcy, S.W., Braga-Neto, U.M., Popescu, S.C., Martin, D.E., Rodriguez, R., Meza, K., Enciso, J., Solorzano, J., Wang, T., 2022a. Volunteer cotton plant detection in corn field with deep learning. Autonomous Air Ground Sens. Syst. Agricult. Optimiz. Phenotyp. VII 1211403 (June), 3. https://doi.org/10.1117/12.2623032.

Yadav, P.K., Thomasson, J.A., Hardin, R., Searcy, S.W., Braga-neto, U., Popescu, S.C., Martin, D.E., Rodriguez, R., Meza, K., Enciso, J., Solorzano, J., Wang, T., 2022b. Detecting Volunteer Cotton Plants in a Corn Field with Deep Learning on UAV Remote-Sensing Imagery. *ArXiv Preprint ArXiv:2207.06673.* , pp. 1–38 https://doi.org/10.48550/arXiv.2207.06673.

Yadav, P.K., Thomasson, J.A., Searcy, S.W., Hardin, R.G., Popescu, S.C., Martin, D.E., Rodriguez, R., Meza, K., Diaz, J.S., Wang, T., 2022c. Computer Vision for Volunteer Cotton Detection in a Corn Field With UAS Remote Sensing Imagery and Spot-Spray

Applications. *ArXiv Preprint ArXiv :2207.07334, Vc.* , pp. 1–39 https://doi.org/10.48550/arXiv.2207.07334.

Yan, B., Fan, P., Lei, X., Liu, Z., Yang, F., 2021. A real-time apple targets detection method for picking robot based on improved YOLOv5. Remote Sens. 13 (9), 1–23. https://doi.org/10.3390/rs13091619.

Yu, J., Sharpe, S.M., Schumann, A.W., Boyd, N.S., 2019. Deep learning for image-based weed detection in turfgrass. Eur. J. Agron. 104 (October 2018), 78–84. https://doi.org/10.1016/j.eja.2019.01.004.

Zhao, Z.Q., Zheng, P., Xu, S.T., Wu, X., 2019. Object detection with deep learning: a review. IEEE Trans. Neural Networks Learn. Syst. 30 (11), 3212–3232. https://doi.org/10.1109/TNNLS.2018.2876865.

Zhou, F., Zhao, H., Nie, Z., 2021. Safety helmet detection based on YOLOv5. Int. Conf. Power Electron. Comp. Appl. ICPECA 2021, 6–11. https://doi.org/10.1109/ICPECA51329.2021.9362711.