







Using the New YoLo Models in Detecting Small-Sized Objects in the Case of Rice Grains on Branche

Khang Nguyen Quoc , Anh Nguyen Quynh , Hoang Tran Ngoc ,
and Luyl-Da Quach^(✉) 

FPT University, Can Tho City, Vietnam

{KhangNQCE160710, AnhNQCE160515}@fpt.edu.vn,
hoang2531992@gmail.com, luyldaquach@gmail.com

Abstract. Identifying a small-sized object is of interest to many studies, especially the rice grain on the branch. Due to its significance to the evaluation of the grain quality. In identifying rice seeds, there are also some difficulties in separating components such as branches and spikes. The study uses a dataset of images of rice branches with differences in shape, state, and size. After the pre-processing steps, the obtained data has images of relatively small rice grains. The research applies to object recognition advances such as new YOLO models (including YOLOv5, YOLOv6, and YOLOv7). On a dataset of 150 images and nearly 6000 instances, and the results are evaluated on many different epochs, the results show that the highest accuracy belongs to YOLOv7, which is 89.93% (Precision), 87.96% (Recall), and 91.33% (mAP). The study also opens up further studies in detecting diseases on rice, such as grain blight, cotton neck blast, etc.

Keywords: Yolo models · small-sized · count objects · count rice · rice on the branch

1 Introduction

Seed quality plays an essential role in assessing the quality of growth and development of seeds, which is a decisive factor for yield, nutritional composition, taste, and edible quality [1]. However, the evaluation of polymorphic seed quality is still carried out by manual method, which is costly in terms of labor as the assessment is based on chemical analysis in the extraction process. This method was performed through solid phase microextraction, spectroscopic techniques, infrared (IR) spectroscopy, and nuclear magnetic resonance (NMR) besides e-tongue and e-nose for the detection of flavor [2]. Realizing this disadvantage, several studies have been carried out to support the automation process in agriculture.

First, much research has been done in detecting, counting, and locating fruit trees in the garden. Research [3] uses the advancement of deep learning models applied to the data set with 1000 images and 41000 labeled instances of apple. The results obtained the highest accuracy of 92.68% with the CNN algorithm, a semi-supervised clustering

method based on Gaussian Mixture Models, with relatively satisfactory results. With similar work, research [4] uses fewer datasets with algorithms GMM and ResNet50 with the highest accuracy of 95.1%. Research [5, 6] builds MangoNet and MangoYolo networks to detect and count mangoes with the highest accuracy of 98.3%. Research [7] used Mask-RCNN to detect grapes with a camera mounted on an automatic grape harvester with green berries by imaging with an accuracy of 94% in vertical bud positioning and 85.6% in cutting minimal pruning on a dataset with 5700 patches from 38 images. Research [8] used InceptionV2, MobileNet, and Single Shot Multibox Detector (SDD) to detect and count avocados, apples, and lemons with the highest accuracy of 93%. Research [9] detects and counts pear fruit by YOLOv4 models with 98% accuracy. Another fruit of interest is the tomato, with much research focused on supporting the automation of detection, counting, and harvesting. Specifically, the research [10] used the MaskRCNN model on the dataset with different heights with the highest accuracy of 88%. Besides, research [11, 12] performed tomato flower counting by Faster R-CNN thresholding in computer vision with the highest accuracy of 96.02%. Through research, the problem in implementation is dominated by environmental factors such as trees covered by foliage, camera height, light, Etc., which have affected fruit detection and counting.

Similar to fruit detection, much research has focused on seed quality to determine product value. Research [13] counted plots as they affected the quality of tomato cultivars using LocAnalyzer. Research [14] created a TasselNetv2 model that counts wheat spikes based on context-augmented local regression networks with an accuracy of 91.01%. Research [15] used SeedGerm correlation and manual method to count the number of germinated grains of barley, cabbage, corn, pepper, and tomato with minimal error. Research [16] to detect and count the number of sorghum heads using CNN with the final efficiency R2 between humans and machines of 0.88. Research [17] determined the length, width, and height of panicle evaluated clustering with a high accuracy of 89.3% with a 10.7% omission and 14.3% commission rate. Research [18] counted rice and wheat grains using an Android device with an error of less than 2%.

The research shows that the implementation of seed counting has yet to be engaged in much research, but this is a decisive factor for the quality and yield of seed-harvesting crops. Besides, the research using the advancements of the YOLO model has been of great interest to object counting studies. Therefore, the study takes the first step in detecting rice seeds on rice panicles to serve as a basis for further studies. The research is divided into sections. Section 2 explores related studies in rice seed detection with its advancements and challenges to be made. Section 3 provides knowledge related to YOLO models; Section 4 proposes methods and assessments for data made through the actual collection of the research. Section 5 the results obtained from implementation. Section 6 discusses the results of the research, Sect. 7 conclusions related to the study, and suggestions for further research directions.

2 Related Works

Recent technological advances play an essential role in seed detection. Research [19] shows potential in using near-infrared spectroscopy, hyperspectral and multispectral imaging, Raman spectroscopy, infrared thermography, and soft X-ray imaging methods.

Research [20] classified 14 rice varieties of *Oryza sativa* with 3500 samples with nearly 50000 seeds. The machine learning methods used in this research include VGG16, VGG19, Xception, InceptionV3, InceptionResNetV2, LR, LDA, k-NN, and SVM. The highest accuracy is 95.15%. In addition, the research [21] used advances in technology using hyperspectral imaging and chemometrics on individually isolated particles with an overall accuracy of 88%. Similarly, the research [22] using hyperspectral imaging with deep learning algorithms gave an accuracy of 99.19% on the oat seeds dataset, and research [23] used transfer learning with the highest accuracy of 97.2% on soybean seed varieties. Another approach in particle detection is built on optical micrographs using image processing with relatively satisfactory results. In addition, the advancement of deep learning and transfer learning models has shown substantial development in agriculture, such as diagnosing diseases in chickens [24, 25], shrimp [26], foliar diseases [27, 28], fish [29], palm trees diseases [30].

The research in this section demonstrates a remarkable development in technological advances in the formulation and development of techniques for particle detection. However, the remaining problem of the studies is that they are being carried out with photos in the laboratory, using high equipment, so the cost problem is still limited. In addition, the tiny size of the seeds on the cotton when separated is analyzed in Section IV. Therefore, this study performing rice seed detection in paddy ears will create a challenge in formulating and developing follow-up studies that help support farmers through mobile devices (Fig. 1).



Fig. 1. The image illustrates the sample of the research.

3 YOLO Models

3.1 YOLO Network and Algorithms Brief

The YOLO algorithm divides the trained image into $S \times S$ grids of cells, where each cell has different detection tasks. This network structure is created from 24 convolutional layers and connected layers described (Fig. 2). In which the convolutional layers will extract the features of the image, and the full-connected layers will predict the probability and the coordinates of the object. After the fully connected layer, the tensor of $S \times S \times (B \times 5 + C)$ is output, where B represents the number of predicted targets in each mesh and C represents the number of object types. The final result is obtained by regressing the box

object’s position and evaluating the tensor data’s type probability. Therefore, the YOLO algorithm can detect targets quickly but cannot detect small targets, or its detection efficiency is not good [31]. So many later versions of this algorithm since version 5 have tried to overcome this by changing the structure and adding image processing methods. In this study, we will evaluate the latest version of the YOLO models, including versions 5, 6, and 7.

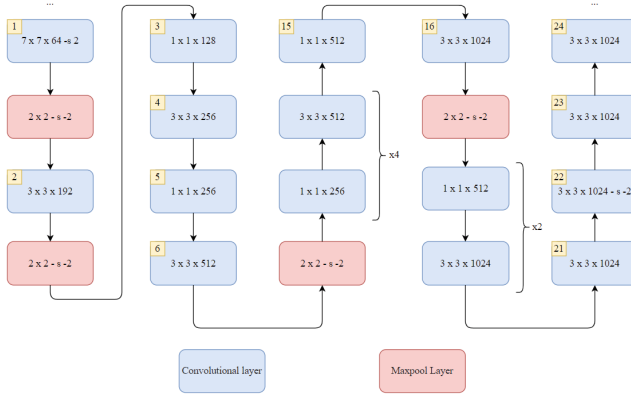


Fig. 2. Illustrating the Convolution layer network structure of the YOLOv1 models.

3.2 Overview of YOLO Models

YOLOv5 is an upgraded architecture from its predecessors to improve recognition efficiency. It has significantly improved the processing time of deeper networks [32]. This will become important with the project into larger datasets containing small objects and real-time detection. With the structure being changed with some properties:

- Backbone: from CSPResidualBlock (version 4) switch to using module C3.
- Neck: SPP + PAN --> SPPF + PAN. Adopting a module similar to SPP, but twice as fast and calling it SPP - Fast (SPPF).
- Data Augmentation: Mosaic Augmentation, Copy-paste Augmentation, MixUp Augmentation.
- Anchor Box: Using the technique of applying genetic algorithm (GA) to the Anchor Box after the k-means step, the Anchor Box works better with the user’s custom datasets but no longer works well with each Common Objects in Context (COCO).

YOLOv6 is a single-stage object detection model based on YOLO architecture [33]. This version is researched and developed into open source by the author Meituan. YOLOv6 achieves a more robust performance than Yolov5 when benchmarking against the MS Coco dataset. This model has evolved the backbone and neck layer with a new structure called EfficientRep Backbone and a Rep-PAN Neck. This change makes the number of parameters of v6 larger than v5, even two times. However, the training time

between the two models is mainly about the same with the same amount of data, and v6 gives a higher probability of object type demonstrated on the COCO dataset.

YOLO-v7 is the latest model of the YOLO family published and studied by WongKin Yiu [34]. The author confirms that this model is superior to the object development models. The change and many new structures applied make Yolov7 achieve high efficiency and faster time. Some new techniques to improve the efficiency of the above model such as the Label Assignment technique, Model Scaling, Backbone use Efficient Layer Aggregation Network (ELAN), Neck with SPPCSPC architecture developed from YOLOv4 model, Re-parameterization in YOLOv5 model.

4 Methodology

The research evaluated the three latest models of the YOLO family on a dataset containing small objects in the Object Detection domain, for YOLO models, SOTA one-stage object detection models, is popular set of object detection models used for real-time object detection and classification in computer vision. The proposed research method uses YOLO models to separate each rice grain on the rice smudging data set to evaluate the effectiveness of each model. The actual object separation results of each model compare the productivity of versions with each other.

4.1 Data Collection and Preprocessing

The dataset used in the rice grain analysis was provided by the Mekong Delta Rice Institute and captured through a Samsung Galaxy Note 10 phone. The rice image dataset used through selection includes 150 images. Image data will be processed through steps including data labeling and resizing (640×640 px) when included in the YOLO model. Using the above-sized image to make the objects reach a small size on each image. Nearly 6,000 objects were labeled using LabelImg and RoboFlow processing and labeling tools. The processing is illustrated below (Fig. 3).

The data is labeled in the YOLO format using a self-designed program that can provide a bounding box and x, y, height, and width label coordinates. On the other hand, for efficient training, the images are labeled with the open-source LabelImg tool and using the YOLO format. Afterward, the processed data will be pushed to RoboFlow to export two complete datasets with a train and valid ratio of 80:20 suitable for 3 YOLO models. When training and testing, the dataset will be exported by RoboFlow for each YOLO version.

Based on the results of the analysis, the research visualized the distribution of object box occurrences and sizes in the data set (Fig. 4). In Fig. 4a represents the position of the centroids of the object boxes in the image after scaling to (0,1) the image, and it can be observed that the objects are primarily concentrated in the center of the image; Fig. 4b represents the ratio of the size of the box object to the size of the image, where it can be observed that not only there are many objects of the same size but also some objects of different sizes.

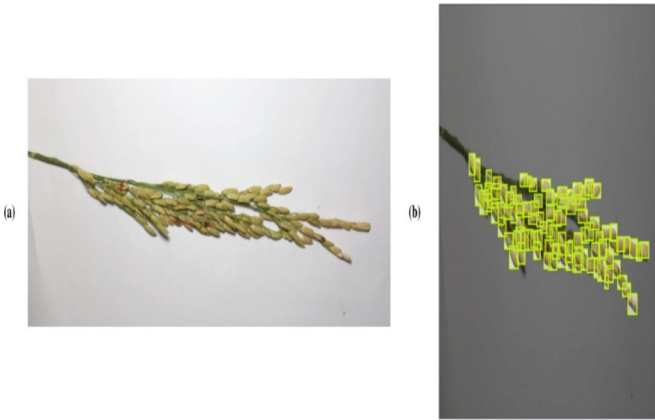


Fig. 3. Illustrate the rice image in the dataset: (a) Original image (b) Processed image.

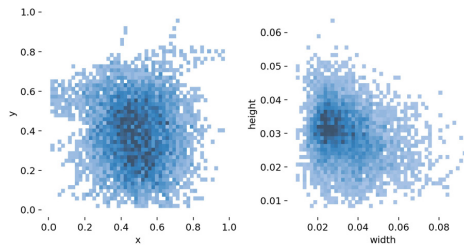


Fig. 4. Characteristics of the dataset used: (a) Location of objects in the box, (b) Size of the box.

4.2 Experimental Environment

The software and hardware parameters used for the model training in this document are shown under Table 1.

Table 1. Configuration parameters

Device	Configuration
System	Windows 10 Home Single Language
CPU	Intel(R) Core (TM) i7-1065G7
GPU	16 GB GPU Tesla T4
GPU accelerator	CUDA 11.2, Cudnn 11.0
Frames	Pytorch
Compilers	Google. LLC. Collab and Anaconda
Python version	3.8

4.3 Proposed Method

After the data collection and processing are completed, they will be fed into YOLO models for object detection training. YOLO models are initialized on the Google Colab Cloud platform through open source code developed on GitHub with some fixed parameters such as Table 2.

Table 2. Fixed training parameters

Hyperparameter	Value
Batch size	4
Image size	640 (px)
Epochs	200
Epochs evaluation interval	1
Learning rate	0.01

The version with little complexity will be used in which model parameters YOLOv5, YOLOv6, and YOLOv7. At the same time, these models also fit the dimensions of the trained images. The provided models use transfer learning techniques from pre-trained models on the COCO dataset of 80 classes to achieve results with small epochs, with the mean average in validation dataset by IoU (threshold equals 50%) and number of params are detailed in Table 3.

Table 3. Model properties

Model	Params (M)	mAP ^{val} ₅₀ (%)
Yolov5s	7.2	56.8
Yolov6s	17.2	60.4
Yolov7	36.9	69.7

The process of our method is illustrated simply through the steps that include data processing, model training, evaluation of the model's effectiveness on the test set, and comparison of each model's results based on metrics (Fig. 5).

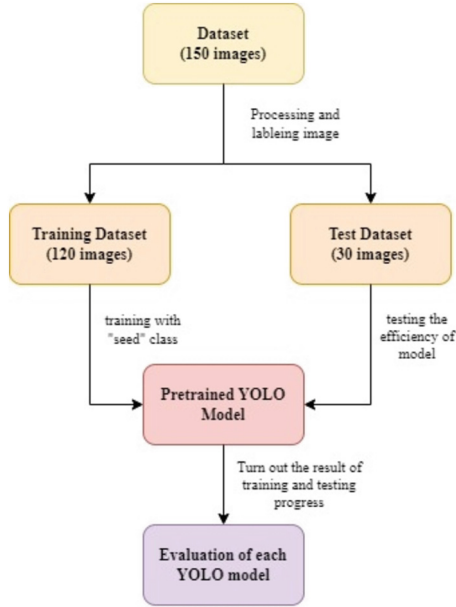


Fig. 5. Illustrate progress in research.

5 Result

5.1 Evaluation Metrics

To evaluate and compare the results of YOLO models, we use some standard metrics developers use in Deep Learning, such as Precision used to measure the quality of true prediction (TP) is performed by (1), Recall is the ratio of the number of true positives among those that are False Negative (TP + FN) (2). These are two measurement methods commonly used to assess understanding between models. The Precision Average (3) and Mean Precision Average (4) by N number of classes are the indicators used to evaluate the trained parameters of the models of the YOLO family.

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$R = \frac{TP}{TP + FN} \quad (2)$$

$$AP = \frac{1}{11} \sum_{R_i} PR_i \quad (3)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (4)$$

In particular, the research needs to determine whether the object prediction is true or false based on an intersection over Union (IoU) concept. IoU is the ratio between the intersection and union of the predicted bounding and the ground truth; the formula and illustration of IoU are shown in Fig. 6. The research uses the usual threshold to make predictions right or wrong is 50%.

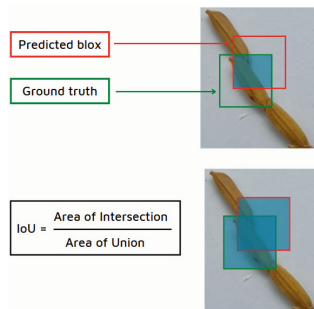


Fig. 6. First Section Illustrated IoU with rice samples on the dataset. Area of Intersection: the intersection of the predicted bounding box and ground truth. Area of Union: the combination of both the area of prediction bound box and ground truth.

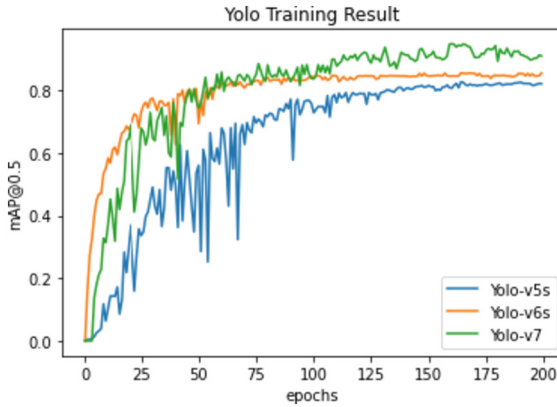
5.2 Training Result

Through the steps from data processing, model structure, and configuration to training. The statistical research found that the results of the models based on the valid data set achieved the highest Mean Average Precision ($\text{mAP}^{0.5}$), with YOLOv7 reaching 94.22% at the 188th epoch. The results of the training process between 3 models are evident in Table 4 and Fig. 7. In which the highest result is YOLOv7 with evaluation parameters such as Precision (93.2%), Recall (84.27%), and F1 Score (88.51%) compiled in Table 4.

On the other hand, the study found that the YOLOv7 model significantly improved compared to the previous models (YOLOv5 and YOLOv6). As for the results obtained by YOLOv5 and YOLOv6 models during small object recognition training, YOLOv7 only needed half the time to reach the same result, at epoch 60 reached 85.13% (compared to 82.68% in YOLOv5) and 62 reached 85.81% (compared to 85.51% in YOLOv6). However, the results also show that with the number of epochs less than 100, YOLOv6 gives stable results over time. However, this model still needs to drift to get high results and only gradually increase the effect more slowly. Between the YOLOv5 and YOLOv6 models not too much difference in results when training in the last epochs and almost saturating at nearly 85% mAP. In the research, YOLO models still make a difference and develop significantly between YOLO models.

Table 4. Results of the yolo models.

Model	Precision(%)	Recall(%)	F ₁ Score	mAP(%)	Best Epoch
Yolov5s	81.53	80.89	81.2	82.68	165
Yolov6s	84.17	83.65	83.9	85.51	165
Yolov7	93.2	84.27	88.51	94.22	188

**Fig. 7.** The graph shows the mAP ratio of 3 YOLO models for each epoch.

6 Discussion

In object recognition, more specifically the YOLO family, the Yolov7 model is their authors claim a breakthrough with a performance and comprehension increase of more than 120% compared to the previous version. Moreover, the research also thought this model would achieve the best results compared to the generally tested models. The results are similar to the hypothesis that YOLOv7 achieved outstanding results compared to two models, YOLOv5s and YOLOv6s, respectively. This proves that YOLOv7 is a relatively new model and should be exploited, especially for small, hard-to-extract objects characterized by previous structures. However, this model still needs to perform better on the test dataset, as we can see the mAP object recognition is up to 94%. However, the correctness and accuracy depend on metrics such as Recall or Precision, and F1 Score is at most 90%. This can be explained by the fact that the size of our data set needs to be larger, specifically only 150 images, and only using the same structure with the image size of 640×640 px. In addition, because the main purpose of the study is to compare the models, all the basic parameters mentioned above will be used, so there will be some limitations on the accuracy of the model. However, because the model is trained on a dataset that is close to reality, applying this model will also bring a system to help farmers identify rice seeds and then separate them from each other diseased and disease-free, serving the larger systems to bring up the disease status of rice quickly. For us, the research on comparing models and optimization algorithms still need to

be explored. However, this topic contains potential positive solutions in this computer vision area.

7 Conclusion

In object detection, research shows that YOLO7v is the best of the last three versions of the YOLO family. The YOLOv7 model can identify small entities, even with less than 200 epochs. However, these models still cannot achieve too high accuracy because the data set needs to be larger, and the primary purpose is not. of the study is to compare the models with each other. Therefore, in the following studies, we may increase the dataset, apply many methods of image processing, evaluate other structures in the field of Object Detection, and propose a structure or method. The new algorithm for detecting small-sized objects. At the same time, drones will be applied to apply the Yolo model to identify objects on a large scale. The research also shows that the detection of diseases on rice, such as rice smudging and rice blast will be developed in the future.

References

1. Feng, L., Zhu, S., Liu, F., He, Y., Bao, Y., Zhang, C.: Hyperspectral imaging for seed quality and safety inspection: a review. *Plant Methods* **15**(1), 91 (2019). <https://doi.org/10.1186/s13007-019-0476-y>
2. Farghal, H.H., Mansour, S.T., Khattab, S., Zhao, C., Farag, M.A.: A comprehensive insight on modern green analyses for quality control determination and processing monitoring in coffee and cocoa seeds. *Food Chem.* **394**, 133529 (2022). <https://doi.org/10.1016/j.foodchem.2022.133529>
3. Hani, N., Roy, P., Isler, V.: MinneApple: a benchmark dataset for apple detection and segmentation. *IEEE Robot. Autom. Lett.* **5**(2), 852–858 (2020). <https://doi.org/10.1109/LRA.2020.2965061>
4. Häni, N., Roy, P., Isler, V.: A comparative study of fruit detection and counting methods for yield mapping in apple orchards. *J. Field Robotics* **37**(2), 263–282 (2020). <https://doi.org/10.1002/rob.21902>
5. Kestur, R., Meduri, A., Narasipura, O.: MangoNet: a deep semantic segmentation architecture for a method to detect and count mangoes in an open orchard. *Eng. Appl. Artif. Intell.* **77**, 59–69 (2019). <https://doi.org/10.1016/j.engappai.2018.09.011>
6. Koirala, A., Walsh, K.B., Wang, Z., McCarthy, C.: Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of ‘MangoYOLO.’ *Precision Agric.* **20**(6), 1107–1135 (2019). <https://doi.org/10.1007/s11119-019-09642-0>
7. Zabawa, L., Kicherer, A., Klingbeil, L., Töpfer, R., Kuhlmann, H., Roscher, R.: Counting of grapevine berries in images via semantic segmentation using convolutional neural networks. *ISPRS J. Photogramm. Remote. Sens.* **164**, 73–83 (2020). <https://doi.org/10.1016/j.isprsjsrs.2020.04.002>
8. Vasconez, J.P., Delpiano, J., Vougioukas, S., Auat Cheein, F.: Comparison of convolutional neural networks in fruit detection and counting: a comprehensive evaluation. *Comput. Electron. Agric.* **173**, 105348 (2020). <https://doi.org/10.1016/j.compag.2020.105348>
9. Parico, A.I.B., Ahamed, T.: Real time pear fruit detection and counting using YOLOv4 models and deep SORT. *Sensors* **21**(14), 4803 (2021). <https://doi.org/10.3390/s21144803>
10. Afonso, M., et al.: Tomato fruit detection and counting in greenhouses using deep learning. *Front. Plant Sci.* **11**, 571299 (2020). <https://doi.org/10.3389/fpls.2020.571299>

11. Rahim, U.F., Mineno, H.: Tomato flower detection and counting in greenhouses using faster region-based convolutional neural network. *J. Image Graph.* **8**(4), 107–113 (2020). <https://doi.org/10.18178/joig.8.4.107-113>
12. Afonso, M., Mencarelli, A., Polder, G., Wehrens, R., Lensink, D., Faber, N.: Detection of tomato flowers from greenhouse images using colorspace transformations. In: Oliveira, P.M., Novais, P., Reis, L.P. (eds.) *Progress in Artificial Intelligence: 19th EPIA Conference on Artificial Intelligence, EPIA 2019, Vila Real, Portugal, September 3–6, 2019, Proceedings, Part I*, pp. 146–155. Springer International Publishing, Cham (2019). https://doi.org/10.1007/978-3-030-30241-2_13
13. Spetale, F.E., Murillo, J., Vazquez, D.V., Cacchiarelli, P., Rodríguez, G.R., Tapia, E.: LocAnalyzer: a computer vision method to count locules in tomato fruits. *Comput. Electron. Agric.* **173**, 105382 (2020). <https://doi.org/10.1016/j.compag.2020.105382>
14. Xiong, H., Cao, Z., Lu, H., Madec, S., Liu, L., Shen, C.: TasselNetv2: in-field counting of wheat spikes with context-augmented local regression networks. *Plant Methods* **15**(1), 150 (2019). <https://doi.org/10.1186/s13007-019-0537-2>
15. Colmer, J., et al.: SeedGerm: a cost-effective phenotyping platform for automated seed imaging and machine-learning based phenotypic analysis of crop seed germination. *New Phytol.* **228**(2), 778–793 (2020). <https://doi.org/10.1111/nph.16736>
16. Ghosal, S., et al.: A weakly supervised deep learning framework for sorghum head detection and counting. *Plant Phen.* **2019**, 2019/1525874 (2019). <https://doi.org/10.34133/2019/1525874>
17. Malambo, L., Popescu, S.C., Horne, D.W., Pugh, N.A., Rooney, W.L.: Automated detection and measurement of individual sorghum panicles using density-based clustering of terrestrial lidar data. *ISPRS J. Photogramm. Remote. Sens.* **149**, 1–13 (2019). <https://doi.org/10.1016/j.isprsjprs.2018.12.015>
18. Liu, T., et al.: Rice and wheat grain counting method and software development based on Android system. *Comput. Electron. Agric.* **141**, 302–309 (2017). <https://doi.org/10.1016/j.compag.2017.08.011>
19. Xia, Y., Xu, Y., Li, J., Zhang, C., Fan, S.: Recent advances in emerging techniques for non-destructive detection of seed viability: a review. *Artific. Intell. Agric.* **1**, 35–47 (2019). <https://doi.org/10.1016/j.aiia.2019.05.001>
20. Kiratiratanapruk, K., et al.: Development of paddy rice seed classification process using machine learning techniques for automatic grading machine. *J. Sens.* **2020**, 1–14 (2020). <https://doi.org/10.1155/2020/7041310>
21. Bao, Y., Mi, C., Wu, N., Liu, F., He, Y.: Rapid classification of wheat grain varieties using hyperspectral imaging and chemometrics. *Appl. Sci.* **9**(19), 4119 (2019). <https://doi.org/10.3390/app9194119>
22. Wu, N., et al.: Variety identification of oat seeds using hyperspectral imaging: investigating the representation ability of deep convolutional neural network. *RSC Adv.* **9**(22), 12635–12644 (2019). <https://doi.org/10.1039/C8RA10335F>
23. Zhu, S., et al.: A rapid and highly efficient method for the identification of soybean seed varieties: hyperspectral images combined with transfer learning. *Molecules* **25**(1), 152 (2019). <https://doi.org/10.3390/molecules25010152>
24. Quach, L.-D., Pham-Quoc, N., Tran, D.C., Hassan, M.F.: Identification of chicken diseases using VGGNet and ResNet models. In: Vo, N.-S., Hoang, V.-P. (eds.) *Industrial Networks and Intelligent Systems: 6th EAI International Conference, INISCOM 2020, Hanoi, Vietnam, August 27–28, 2020, Proceedings*, pp. 259–269. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-63083-6_20
25. Quach, L.-D., et al.: An AI-based chicken disease management system. In: Rosdiazli Ibrahim, K., Porkumaran, R.K., Nor, N.M., Prabakar, S. (eds.) *International Conference on Artificial Intelligence for Smart Community: AISC 2020, 17–18 December, Universiti Teknologi*

- Petronas, Malaysia, pp. 707–716. Springer Nature Singapore, Singapore (2022). https://doi.org/10.1007/978-981-16-2183-3_68
26. Duong-Trung, N., Quach, L.-D., Nguyen, C.-N.: Towards classification of shrimp diseases using transferred convolutional neural networks. *Adv. Sci. Technol. Eng. Syst. J.* **5**(4), 724–732 (2020). <https://doi.org/10.25046/aj050486>
 27. Quach, L.-D., Quoc, K.N., Quynh, A.N., Ngoc, H.T.: Evaluation of the efficiency of the optimization algorithms for transfer learning on the rice leaf disease dataset. *Int. J. Adv. Comput. Sci. Appl.* **13**(10) (2022). <https://doi.org/10.14569/IJACSA.2022.0131011>
 28. Taujuddin, N.S.A.M., et al.: Detection of plant disease on leaves using blobs detection and statistical analysis. *IJACSA* **11**(8) (2020). <https://doi.org/10.14569/IJACSA.2020.0110852>
 29. Sikder, J., Sarek, K.I., Das, U.K.: Fish disease detection system: a case study of freshwater fishes of Bangladesh. *Int. J. Adv. Comput. Sci. Appl.* **12**(6) (2021). <https://doi.org/10.14569/IJACSA.2021.01206100>
 30. Abu-zanona, M., Elaiwat, S., Younis, S., Innab, N., Kamruzzaman, M.M.: Classification of palm trees diseases using convolution neural network. *Int. J. Adv. Comput. Sci. Appl.* **13**(6) (2022). <https://doi.org/10.14569/IJACSA.2022.01306111>
 31. Wu, W., et al.: 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 (2021). <https://doi.org/10.1371/journal.pone.0259283>
 32. Jocher, G., et al.: ultralytics/yolov5: v4.0 - nn.SiLU() activations, Weights & Biases logging, PyTorch Hub integration. Zenodo (2021). <https://doi.org/10.5281/ZENODO.4418161>
 33. Li, C., et al.: YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications (2022). <https://doi.org/10.48550/ARXIV.2209.02976>
 34. Wang, C.-Y., Bochkovskiy, A., Liao, H.-Y. M.: YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors (2022). <https://doi.org/10.48550/ARXIV.2207.02696>