



# Cucurbitaceous family flower inferencing using deep transfer learning approaches: CuCuFlower UAV imagery data

S. Mithra<sup>1</sup> · T. Y. J. Nagamalleswari<sup>1</sup>

Accepted: 1 March 2023 / Published online: 25 April 2023

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

## Abstract

Cucumbers, pumpkins, and muskmelons use flowering as one of the most important pollination processes. Natural pollinators such as bees and butterflies are declining, drastically reducing harvests. The aim of this research is to use UAV images to build deep transfer learning to detect floral patterns in autonomous pollination of cucurbit plants. The images taken by the drone equipped with an OpenMV camera from farm in various situation are collected and stored in the CuCuFlower database for flower identification. For self-pollination, pollen area and other identifiable characteristics for classification of flower gender and family are observed. In particular, the fast detection and high accuracy of small, densely packed objects of a large farm in real time is still an open challenge. In this work, we address this challenge by finding and identifying small objects in a bloom of a complete image using the method of transfer learning based on YOLOv structure. However, the cheek component of the flower was relatively small compared to the total relative environmental image. The results revealed that the proposed model performs best in detecting the pollinating flower and the AP50 Momentum Factor with 0.95 is 91% and 92%, respectively, which are better than other detection methods.

**Keywords** Accuracy · Aerial application · Agricultural robot · Data collection · Transfer learning · Pollination · Productivity

## 1 Introduction

Flowering is one of the most significant processes in pollination for agricultural crops such as squash, cu-cumbers, pumpkins, eggplant, okra, watermelons, and muskmelons, as well as other vegetables Amasino et al. (2017). These types of plants are pollinated by insects that carry pollen from one plant to another. Pollination is required for the development of fruit in vegetable crops that generate fruit Halder et al. (2019). The pollination of some vegetable plants is accomplished through the production of separate male and female flowers that are pollinated by insects or the wind. The decline in the number of natural

pollinators has a negative impact on the productivity of vegetables Althaus et al. (2021). Cucurbits or gourds are members of the cucurbitaceous family, which includes most pollinating vegetables. The cucurbitaceous family is one of the oldest vegetable families that are cultivated worldwide in tropical and subtropical parts of the world. The cucurbitaceous family, which includes most pollinating vegetables, is often known as the cucurbits or gourd family since it includes most pollination vegetables Hosain et al. (2018). UAVs are widely utilized in a variety of industries, including traffic monitoring, surveillance, inspection, and surveys. Because of their increased movement and flexibility, they have mostly supplanted choppers in recent years. Unmanned aircraft have proven to be one of the technologies that can be used to research air quality, soil component physical properties, and crop growth in a rapid and nondestructive manner Mithra and Malleswari (2021). The goal of this research is to lay a theoretical and experimental framework for the creation of an autonomous aerial vehicle that can supply farmers with crucial information about the plant development medium. The

✉ S. Mithra  
mithrasamuel@gmail.com

✉ T. Y. J. Nagamalleswari  
nagamalt@srmist.edu.in

<sup>1</sup> Department of Networking and Communications, SRM Institute of Science and Technology, Kattankulathur, Chennai, Tamilnadu 603203, India

development of real-time deep learning algorithms with UAVs is significantly faster and more accurate in transforming modern life Ipaté et al. (2015). UAVs have recently come to dominate aerial sensing research, thanks to the deployment of deep neural networks in urban, environmental, and agricultural environments. Convolutional neural networks are a form of machine learning network commonly used to handle image visualization problems such as classification. Image categorization with convolutional neural networks is prevalent, although it has drawbacks. Although a human can recognize the contents of photographs considerably faster than a machine, CNNs have demonstrated a 97.6% success rate in facial recognition (Nayeem 2018). Deep convolutional neural networks (CNNs) are cutting-edge object detectors, yet they are slow and resource intensive. This research compares a lightweight CNN to two deeper CNNs on an item recognition and image classification challenge using a dataset Gogul and Kumar (2017). A classification technique based on deep learning and depth information was developed for rating floral quality based on bloom maturation phase. The classification results from four convolutional neural network models, VGG16, ResNet18, MobileNetV2, and InceptionV3, were adjusted for a four-dimensional (4D) RGBD input and compared with and without depth information to categorize flowers Sun et al. (2021). Detecting items in agricultural contexts is a critical feature for agricultural task automation. Algorithmic tweaks and hyperparameter adjustments are explained to complete the task. The Faster R-CNN algorithm is a mix of a deep convolutional neural network and a region proposal network Oppenheim et al. (2020).

Object detection is a two-part computer vision problem: object localization and object classification. It is used in a lot of other things, such image captioning, object tracking, instance segmentation, and scene understanding, thus it is one of the most important topics Xiong et al. (2021). The detection accuracy of the YOLOv4 model is comparable to the current state of the art, but it is substantially faster. Cotton and other flowering plants, which indicate the change from vegetative to reproductive growth, are vital for crop productivity and adaptation. The purpose of this research was to create a deep learning-based method for Sangaiah et al. (2023) identifying flowering patterns in cotton plants Jiang et al. (2020). Litchi is typically collected by clamping and cutting the little branches, which are readily injured by the picking robot. To semantically segment the litchi branches, a fully convolutional neural network-based semantic segmentation approach is suggested. The DeepLabV3+ model with the Xception 65 feature extraction network produced the greatest results, according to the findings. The concatenated two YOLOv4 model showed the best performance in recognizing floral

features, according to the findings Bochkovskiy (2020). Agriculture is also influenced by a variety of elements such as pests, disease, fertilizers, and other factors that can be managed by properly treating crops. As a result, the primary goal is to create an agriculture drone for spraying pesticides and to discuss various architectures based on unmanned aerial vehicles (UAVs) Kurkute (2020).

Small object recognition is a challenging task in computer vision problem with applications including autonomous driving, Zhang et al. (2022) UAV-based photography, and surveillance Tong et al. (2020). Despite the fact that it is a critical tool for a variety of computer vision jobs in a wide range of industries, current techniques do not perform as well as they could. Despite this, most object identification algorithms have difficulty detecting small objects due to variables such as insufficient information acquired from the area covered by small objects on the image, a high likelihood of location for small objects, and being optimized for medium and large objects Wei (2020). This work used the YOLOv4 model's primary and modified learning methods, namely one-stage and two-stage or transfer learning methodologies, to find and classify a small proportion of a single flower in a whole image recorded from a UAV. The proposed approach for determining the genus and gender of cucurbitaceous flowers is believed to be a practical and successful method. For the development of a successful autonomous pollination strategy, automatic genus and gender classification is required Newstrom-Lloyd et al. (2010). Identification of diverse genders and sex ratios is a critical goal in increasing vegetable yield. Although a human may quickly classify flower gender, in the case of large-scale farming, automated equipment can assist people from the local community in performing detailed assessments. Enabling farmers to monitor the production of cucurbitaceous plants/vegetables in their farming region using this framework. Even though an individual may readily classify the flower gender, in the event of large-scale farming, automatic equipment can assist members of the local community in doing exact evaluations. Farmers can use this framework to track the production of cucurbitaceous plants and vegetables in their area.

The objective of the proposed work is to locate and detect a flower in a complete image using one-stage and concatenated YOLOv4 with transfer learning approach. The images used in the proposed model are acquired using unmanned aerial vehicle (UAV) in the poly houses and farms. The proposed method included (1) developing a platform that used a deep learning strategy combined with transfer learning to recognize flowers in the farm and detect annotated images at three different resolutions, and (2) evaluating the platform's performance in cucurbitaceous

flower tests. Metrics are used to assess the performance, which is then compared to standard approaches.

## 2 Generating CuCuFlower dataset

Datasets of cucurbitaceous family plant images labeled with pertinent odometer and surrounding data by UAV are not yet generally available because to the high cost of obtaining and annotating this type of data. Furthermore, any supervised machine learning method, including deep learning, necessitates that the data used to train the model be representative of the data used during testing Mohanty et al. (2016). Artificial pollination applications are likely to have problems with incomplete training data due to the challenges of generating a data collection. First, the raw field data of cucurbitaceous family flower images are acquired from the agricultural farms and poly houses using drone. Then, the original flower images are expanded by data augmentation operations and further refined by expert annotation. Finally, the dataset is split into three sections: the training set is used to train the YOLOv4 one- and two-stage models, the validation set is used to tune the parameters and evaluate the model, and the testing set is used to verify the model's generalization Barth et al. (2018).

### 2.1 UAV imagery data collection

UAV technology has grown in popularity and is now widely employed in both military and civilian applications. UAV technologies can reach every nook and Ramachandran et al. (2021) cranny of the earth in smart farming López et al. (2021). The potential for using UAVs in intelligent agriculture is huge. UAVs may be able to obtain more reliable and faster NDVI data than satellite photography for more immediate review Stankovic et al. (2022). A DJI Mavic Mini captured the photographs of the cucurbitaceous farm, and the drone was around 100 cm above the flower canopy, thus the flower canopy is mostly in the center of the photograph. The images were collected from different agricultural farms and poly houses near Ambathur, Chennai. The Mavic Mini weighs less than 0.55lbs/250 g, making it almost as light as a smartphone. The Mavic Mini can shoot 12MP aerial photographs and record 2.7 K High-Definition video Wang et al. (2022). The Mavic Mini's weight allows it to fly for longer than similar consumer Fly Cams; it can fly for up to 30 min on a fully charged battery. Drone-captured raw data must go through the *data handling* procedure, which includes enhancing, annotating, and finally handing over to flower recognition Sowmya and Radha (2021).

A total of 3042 original images were obtained using drone camera as shown in Fig. 1, which includes nine

classes of cucurbitaceous family and 971 images of non-flower plant with grown vegetables. Generally, the raw data consist of odometry data for drone and surrounding data of flowers for flower recognition in artificial pollination.

### 2.2 CuCuFlower UAV imagery data augmentation

Flower recognition is one of the most important processes in artificial pollination. It recognizes flowers from photographs with the location of the flower in the image using computer vision data from a drone. Image augmentation is the process of manipulating photographs to create different versions of the same content to provide the model with more training examples. Randomly changing the orientation, brightness, or scale of an input image, for example, need a model that considers how an image topic appears in a number of conditions Wan and Goudos (2019). Figure 2 shows Mosaic data augmentation, DropBlock regularization, and CIoU loss using YOLOv4, cross-stage partial connections (CSP), cross mini-batch normalization (CmBN), self-adversarial-training (SAT), weighted-residual connections (WRC) and Mish activation. They are called as universal features since they should work independently of the computer vision tasks, datasets, or models. In short, YOLOv4 results in a more advanced object detection network architecture as well as new data augmentation techniques.

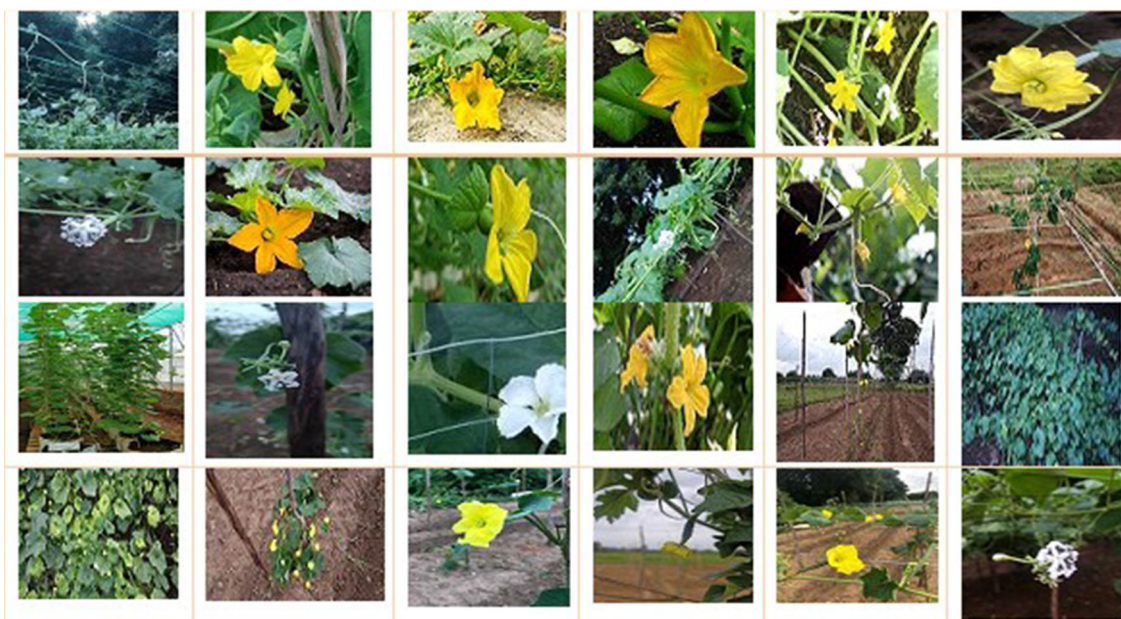
The dataset has been tenfold enlarged from raw data of total 3042 photographs to generate the custom dataset for this application as indicated in Table 1.

### 2.3 CuCuFlower UAV imagery data annotation

The classifying and labeling of data for AI applications is known as data annotation. To improve AI implementations, training data must be correctly categorized and annotated for a given use case. YOLO label, open labeling, YOLO mark, BBox-Label-Tool, and other open-source GUI applications can let you simply generate label files from images Ushasukhanya and Karthikeyan (2021). LabelImg, a python-based open-source application that stores annotations as XML files and organizes them into PASCAL VOC format, was used for image annotation of target classes in the custom dataset as shown in Fig. 3. During annotation for images in the training dataset, each XML comprises information about the target class and the related bounding coordinate Tuda and Luna-Maldonado (2020). It is a free open-source image annotator that can be used to make YOLOv4 annotations. In order to train with YOLOv4, our custom dataset must adhere to the YOLO format. Each image in the collection would be connected to

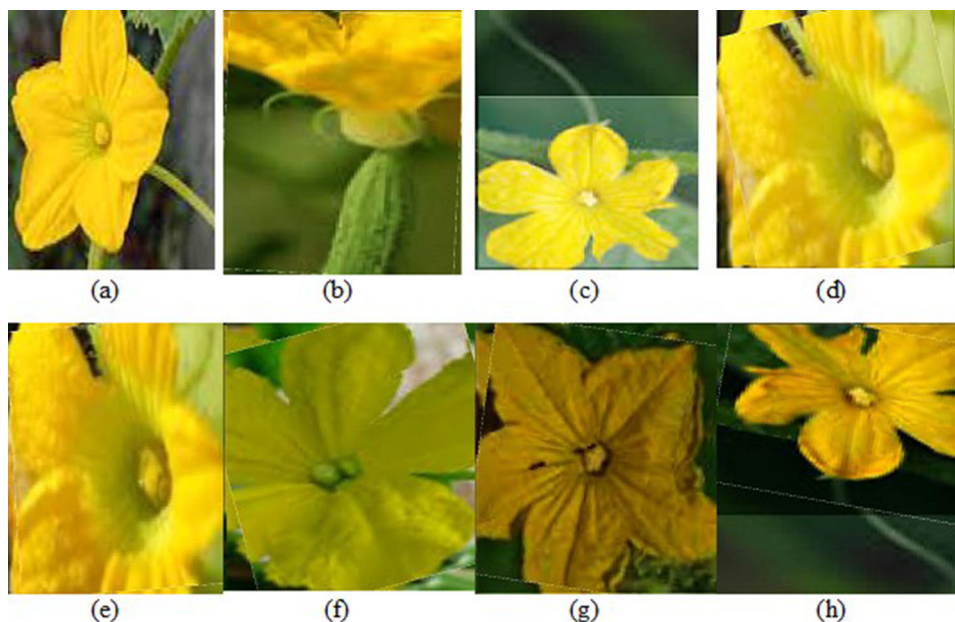


< object - class > < x center > < y center > < width > < height >



**Fig. 1** Examples of cucurbitaceous flower images acquired through DJI Mavic Mini in farms and poly houses can be used for flower classification studies

**Fig. 2** Image data enhancement methods include: (a) original image, (b) crop, (c) blur processing, (d) 90-degree ACW rotation, (e) horizontal mirror projection, and (f) color processing. Transformation of (g–h) brightness



a.txt file with the same name, which would have the object classes and coordinates in the format below:

The customized *CuCuFlower dataset* has been generated for artificial pollination. From the customized *CuCu-Flower dataset*, a total of 4042 images were randomly selected for constructing the network, respectively.

**2.4 CuCuFlower dataset classification**

The purpose of this research is to develop a real-time, high-performance flower detection model based on an improved version of the cutting-edge YOLOv4 algorithm. The cucurbitaceous family was chosen since it is one of the most productive plant families in terms of the quantity and percentage of species used for human consumption Li et al.

**Table 1** CuCuFlower database categorized into various characteristics

S. no	Family name	Genus	Color	Gender	Pollinating system	Flower ratio	#images
1	Cucumber (c.sativus)	Cucumis	Bright yellow	Thin stalk, budding fruit	Self and cross pollination	10:1	442
2	Muskmelon (c.melo)	Cucumis	Yellow	3 Stamens, 3 stigma	Self and cross pollination	12:1	480
3	Squash (c.maxima)	Cucurbitae	Yellowish orange	Long stalks in middle of stigma	Self or cross pollination	10:1	435
4	Pumpkin (c.moschata)	Cucurbitae	Creamy white	Single stamen, budding fruit	Self or cross pollination	10:1	420
5	Sponge Gourd (c.Loofah)	Luffa	Yellow	Thin stalk, budding fruit	Self or cross pollination	20:1	435
6	Calabash (c.siceraria)	Lagenaria	Creamy white	Thin stalk, budding fruit	Self or cross pollination	15:1	483
7	Watermelon (c.Lanatus)	Citrullus	White	Thin stalk, inferior ovaries	Self or cross pollination	13:1	433
8	Snake Gourd (c.cucumerina)	Trichosanthe	White, fringes	Thin, stalk, budding fruit	Self or cross pollination	10:1	428
9	Bitter Gourd (c.charantia)	Momordica	Yellow	Thin, stalk, budding fruit	Self or cross pollination	25:1	486

**Fig. 3** sample image annotations: (a) labeled as cucurbit female flower, (b) labeled as cucurbit male flower



(2020). It consists of 965 species spread throughout 95 genera. Initially, a total of 3042 original images have been collected from farms and poly houses using DJI Mavic Mini drone in different scenarios in contrast to the data from current vision datasets. Utilizing different augmentation procedures, we present the custom *CuCuFlower dataset* which consists of a total 4042 annotated images of nine major cucurbitaceous family classes. Table 1 represents the customized *CuCuFlower dataset*, which consist

of 9 number of cucurbitaceous family plant and it is classified into their genus, gender, color, pollinating system, flowering patterns, and ratio.

From the *CuCuFlower dataset*, a total of 4042 annotated photographs were utilized in this study, with 3433 (80%) serving as training integrated validations and 609 (20%) serving as testing image sets. The LabelImg tool was used to split nine cucurbitaceous classes into distinct bounding boxes and label them. The bounding boxes, that were

generated out of images of flower features in the smallest possible area, constituted a potential region of interest (ROI). The labels were utilized to train the model and relate to the genus and gender of each bloom. 4042 augmented pictures were used to train the model, each with a three-step rotation and brightness/darkness increments.

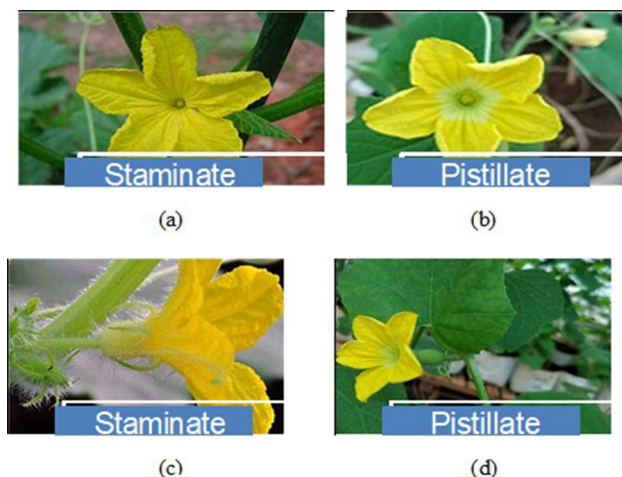
In Fig. 4, it describes the features of the cucumber flower where *cucurbit sativa* is the family name belongs to the cucumis genus. The Botanical Term *staminate* mentions the male flower and *pistillate* describes the female flower of *cucurbit sativa* plant Pandey and Choudhary (2020). In addition, 971 photographs of non-flowering plants, including grown cucumber and squash in the plant, were utilized to train the deep learning approach.

### 3 The proposed flower recognition model

One of the most essential skills for artificial pollination is flower recognition. The recognition of flowers is done from raw photographs, with the first step being to locate the flower in the image, followed by identifying and extracting the bloom's distinctive properties. We employed many versions and setups of the YOLO algorithm to complete the object recognition procedure in this investigation Du et al. (2020). The flower recognition using YOLOv4 model is summarized into flower detection, flower segmentation, feature extraction, and flower classification with genus and gender.

#### 3.1 Generating one- and two-stage YOLOv4 learning method

The recorded images were forwarded to a specialized deep learning server for labeling and model training. By



**Fig. 4** Cucurbit flower gender features are classified as: (a) staminate/stigma, (b) pistillate/anther, (c) staminate stalk, (d) pistillate stalk

manually labeling augmented flower images with their associated species, genus, and genders, the *CuCuflower* dataset was developed. As illustrated in Figure 5, one-stage and concatenated transfer learning approaches were investigated for training the model. The prepared *CuCuflower* dataset is fed into the YOLOv4 model to complete the learning process in the one-stage learning technique. In the following process, the learnt parameters and weights were used in the run time testing. The output of the one-stage YOLOv4 learning algorithm categorizes floral species and gender with a higher ratio of attributes Padmanabula et al. (2020).

The *CuCuflower* dataset was processed in two stages using the two-stage YOLOv4 learning approach. The main task of proposed model is to fast detect and high accuracy of small objects in image and densely packed images in farm. The small object in this research specifies the stigma part of the flower Wu et al. (2020). To identify the flower genus, the flower color and petals are important. The existing approaches do not work as well as they should in determining the gender of a flower. Uses two-stage learning techniques to increase the performance of YOLOv4, while using a resolution on remote sensing photographs to tackle the inadequacy of information problem for small item detection. In the first stage of training, the *CuCuflower* dataset is used as input data to implement the learning process using the parameters and standard weights. One of the model's outputs was information on the bounding box of a floral item Chen and Li (2019). This bounding box was used to determine the position and size of the flower photographs for cropping. The model is trained, is then generates the first level of detection and classifies the flower genus. Subsequently, the output of first-stage learning is processed as input in second-stage training method with the same learned parameter and weights. This model is then trained to detect the stigma part the flower and classifies the gender of the image.

Figure 5 depicts the proposed model, which comprises the YOLOv4 algorithm's one- and two-stage learning approaches. The one-stage learning YOLOv4 model was used to see if the ground-truth input contained every single bloom. The photograph was labeled with a rectangle bounding box that indicated the relative genus and gender above it if each bloom was included. A flowerless image of plants, on the other hand, was one that was not identified Roy (2022). During the two-stage learning procedure, two streams of training were utilized. The YOLOv4 network's first training step was utilized to see if the entire image included any flowers. The flower's clipped image was used in the subsequent processing. The independent YOLOv4 network was used in the second training stage to identify a related genus and gender within the cropped image derived



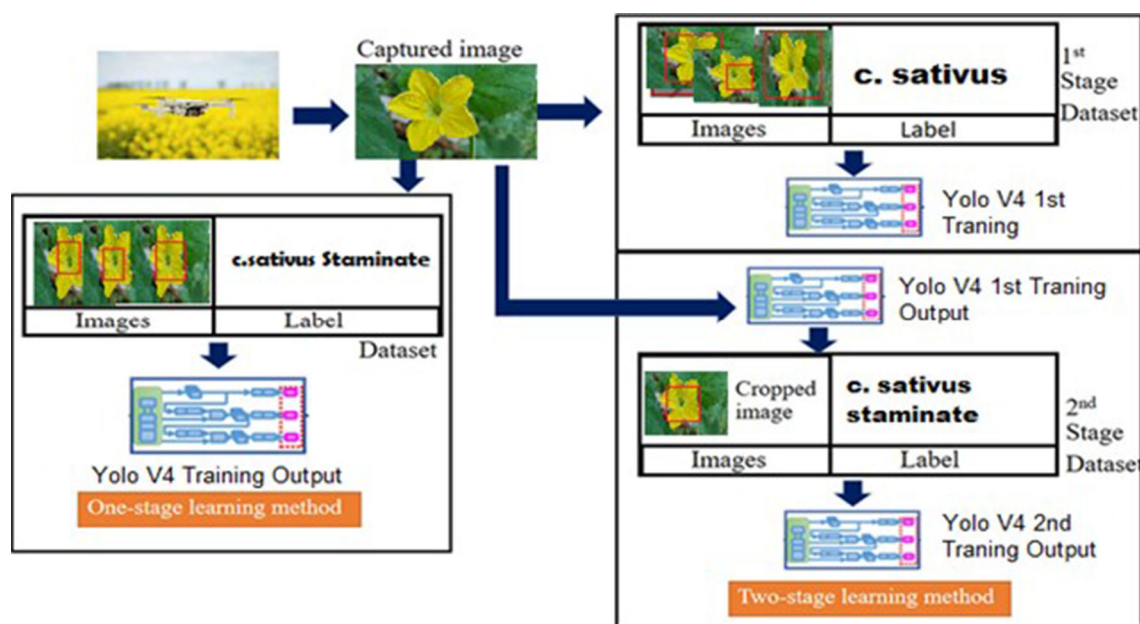


Fig. 5 One- and two-stage flower recognition method

from the first step's image including the flower Sivakumar and Nagamalleswari (2022).

### 3.2 YOLOv4 network structure

In the computer onboard UAV, it is important to include the fastest object detection algorithm with the least amount of processing load. The YOLO algorithm is used to detect the object as quickly as possible while also locating it. Applying a single neural network over the entire image and treating identified item localization as a regression problem is what fastest detection entails. This neural network's architecture is known as Dark net: a form of CNN. It consists of 24 convolutional layers that operate as feature extractors and two dense layers that do predictions. It is a type of computer technology that searches for instances of semantic items belonging to a specific class (such as individuals, buildings, cars, plants, and animals) in digital photographs and videos. Object detection is divided into two parts: the first is object localization, and the second is object classification. The YOLOv4 model is a cutting-edge real-time object detection model Liu et al. (2021).

The next two aspects are the backbone and neck that are in responsible of feature extraction and aggregation. The detection neck and detection head are the most important parts of the object detector. Finally, the head oversees detection and prediction Padilla, Netto, and Silva (2020). The Head is primarily in charge of detection (both localization and classification). Dense detection refers to the fact that YOLO is a one-stage detector that does both at the same time. Sparse detection is a two-stage detector that

executes them separately and aggregates the results as shown in Fig. 6.

### 3.3 Initial parameters of the proposed model

A total of 2720, 661, and 661 samples were chosen at random from the custom dataset for the training, validation, and test sets, respectively. The inputted dataset images of size 512X512 were considered to improve the accuracy of the recommended detection model for different flower development phases. The basic configuration settings for the YOLOv4 model (for instance, starting learning rate, momentum value, number of channels, decay regularization, and so on) were retained the same Rehman et al. (2021). The most important starting configuration parameters for the upgraded YOLOv4 model are supplied as shown in Table 2.

### 3.4 Proposed Model's Performance Metrics

Intersection over union (IoU), F1 score, average precision (AP), precision (P), recall (R), and mean average precision (mAP) are some of the most prominent statistical measures of matrices used to evaluate the performance of deep learning-based object recognition models Rezatofighi et al. (2019). YOLOv4 now includes a scale-invariant evaluation metric. The accuracy of target object detection is determined by the IoU. The overlap area ratio between the bounding box prediction from the model and the genuine bounding area of the item is used by IoU to calculate the

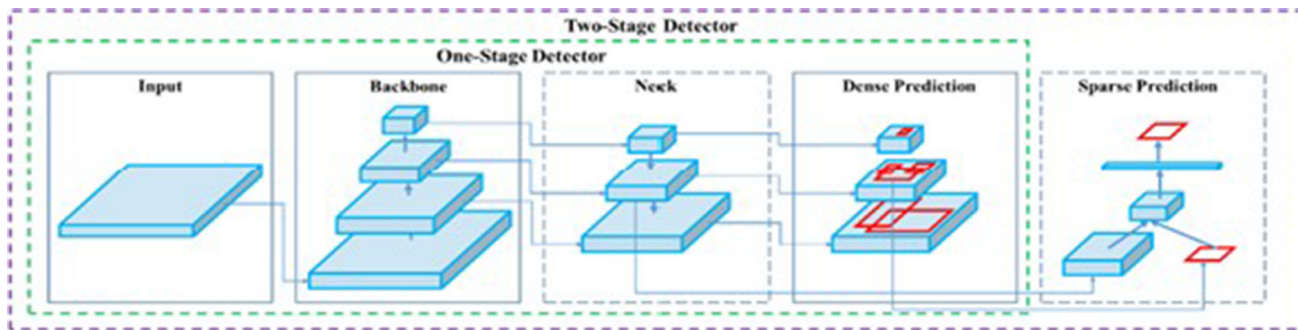


Fig. 6 YOLOv4 network architecture

Table 2 Initial configuration parameters of YOLOv4 model

Input size	Batch	Subdivision	Channels	Momentum
512 × 512	16	4	3	0.95
Learning rate	Decay	Classes	Filters	Training steps
0.001	0.005	9	27	1000

efficiency and performance of the supplied model, which can be represented as Equation 1,

$$\text{Intersection\_over\_Union} = \frac{A_{\text{overlap}}}{A_{\text{union}}}, \tag{1}$$

$A_{\text{overlap}}$  is the intersection region between the model’s bounding box prediction and the true bounding box of the object.  $A_{\text{union}}$ , on the other hand, is the intersection of the two bounding boxes indicated earlier. If IoU is greater than 0.5 in binary classification, the classified object class can be defined as true positive (TP). If the IoU is less than 0.5, the corresponding class might be classified as false positive (FP). Using the definitions of TP, FP, and FN, the performance parameters P and R can be expressed as equation 2:

$$\text{Precision} = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Positive})} \tag{2}$$

$$\text{Recall} = \frac{\text{True Positive}}{(\text{True Positive} + \text{False Negative})} \tag{3}$$

From the above Eq. 3, models with a higher P score are better at differentiating negative datasets, while models with a higher R score are better at detecting positive datasets. F1 score can be calculated using Eq. 4 to determine the degree of test accuracy precision:

$$\text{F 1 Score} = \frac{2 \text{ Precision Recall}}{(\text{Precision} + \text{Recall})} \tag{4}$$

The F1 score is a derived that reconciles the model’s precision value and memory by combining the mean of accuracy and recall. The model is more trustworthy if the

F1 score is higher. In most cases, the area under a PR curve equals AP, which can be written as equation 5,

$$\text{Average Precision} = \int_0^1 \text{Precision}(\text{Recall}) d \text{Recall} \tag{5}$$

A broader area under the precision–recall curve correlates with a higher average precision, signifying better object class prediction accuracy, whereas mean average precision is the average of all average precision values and it can be stated as equation 6,

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^N \text{Average Precision}_i \tag{6}$$

Bounding box regression is a prominent method for predicting localization boxes on input images in dense object recognition models Lin et al. (2017). The suggested model includes complete IoU to improve the accuracy and speed of convergency for the target bounding box prediction process (CIoU). The CIoU loss has been defined using the aspect ratio consistency parameter  $\gamma$  and the positive trade-off value  $\alpha$ , which can be stated as the following equations 7, 8, 9:

$$L_{\text{CIoU}} = 1 - \text{IoU} + \frac{\rho(b, b_{gt})}{c^2} \alpha \gamma \tag{7}$$

$$\gamma = \frac{4}{\pi} \tan^{-1} \frac{w_{gt}}{h_{gy}} - \tan^{-1} \frac{w^2}{h} \tag{8}$$

$$\alpha = \frac{\gamma}{(1 - \text{IoU}) + \gamma} \tag{9}$$

The heights and widths of the ground-truth and prediction bounding boxes are represented as  $h_{gt}$ ,  $h$  and  $w_{gt}$ ,  $w$ .

### 4 Experimental results and analysis

The goal of this study is to create a deep learning-based method for recognizing plant blossoming patterns. This paper adopts the CuCuFlower genus classification and detection custom dataset, which is structured into nine categories and has nine classes and 4042 annotated



examples. In contrast to the data from current vision datasets, the photographs were obtained with unmanned aerial vehicle (UAV) cameras in poly homes and farms, and they were captured in a wide variety of circumstances. This study used one- and two-stage learning algorithms from the YOLOv4 network models as a detection algorithm to identify pollination flowers.

#### 4.1 Overall training results of one-stage and two-stage learning method

To compare the overall performances of the proposed detection model, the values of IoU, F1-score, mAP, final validation loss, and average detection time were compared with one-stage YOLOv4 and two-stage YOLOv4 learning. In the one-stage methodology, the computer was trained using a dataset to predict the flower genus. The two-stage methodology, on the other hand, employed the same dataset for model training in two rounds to identify the relative genus and identify the flower of interest. The model is trained on a GeForce GTX 1080ti GPU in a CUDA 10.0 experimental environment using the Pytorch framework. The training parameters are as follows: The initial learning rate is 0.001, the input image pixel size is 512\*512, the batch size is 16, and the input image pixel size is 512\*512. The weight decays to 0.005, the epoch is 1000, and the momentum factor is 0.95. Precision, recall, and mAP following model training are depicted in the graph. When comparing IoU values, the proposed model was found to have the highest IoU value of 0.922, which is 6.1 percent higher than the original YOLOv4 model (see Figure 7). Therefore, the suggested detection model has a higher accuracy in detecting bounding boxes than the other two models. The model improved YOLOv4 in terms of detection efficiency and accuracy, improving by 7.6% and 7.3%, respectively, with an F1 score of 0.959 and a mAP of 0.912.

A 50 percent and 95 percent confidence range is used to evaluate the YOLO networks' model-wise object

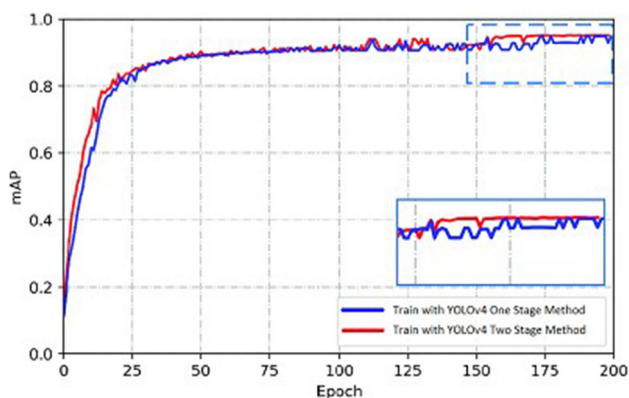


Fig. 7 Intersection over Union (IoU) overlapping rate

identification performance measures. As mentioned bold in Table 3, it is the maximum value of the proposed statistical parameters by model Ushasukhanya et al. (2022). This type of target is regarded identified when the intersection over union (IOU) overlapping rate is greater than 50% and the true-value frame area is greater than 50%; otherwise, it is undiscovered.

Figure 7 depicts the experimental outcomes of accuracy. As demonstrated in Table 3, both models have an accuracy of 0.839 and 0.912, respectively. Differentiate and classify floral interference, overlapping flower photographs, and fuzzy flower images. The proposed learning model has a high mean average precision of 0.912, according to the ROC and AUC evaluations.

Figure 8 shows a comparison of the precision–recall (PR) curves for these three models. Only when area under the PR curve for a particular recall is the highest among all three models, it can be assumed that the recommended model's accuracy value for that recall is higher. In terms of detection accuracy, the proposed model surpasses the one-stage YOLOv4 learning model and the two-stage YOLOv4 learning model.

The validation loss curves of two models are compared in Figure 9. The loss in YOLOv4 began to decrease significantly after around 500 training steps in the initial phase, but the loss in the proposed model began to decrease after around 1000 training steps, suggesting that the recommended model has stronger convergence characteristics than YOLOv4. Loss saturated with a final loss value of 1.65 after around 1000 training steps, whereas the final loss values for the one-stage YOLOv4 and two-stage YOLOv4 learning models were 4.31 and 1.65, respectively, as shown in Table 3. The recommended model has a faster convergence rate and enhanced convergence properties than the original YOLOv4 model, demonstrating superior performance and detection accuracy.

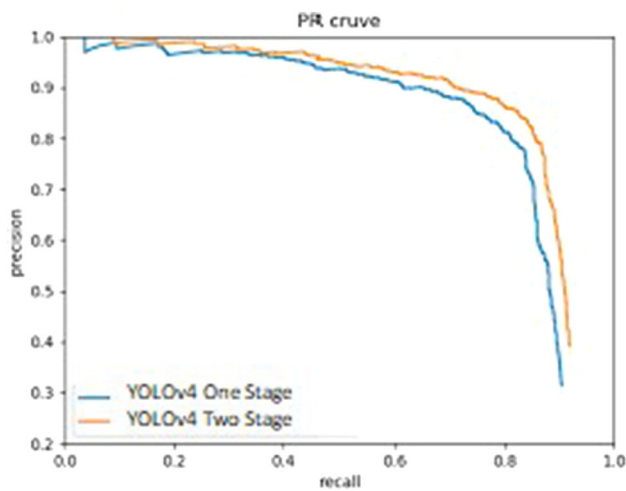
#### 4.2 Detection results for CuCuflower dataset

Figure 10 illustrates the detection results of both one-stage and two-stage learning in the cucurbitaceous plant, with gender and genus belonging to each of the nine different classes obtained and compared to YOLOv4 models. Detection data for each cucurbitaceous flower were received, including the detected flower genus and gender. As shown in Table 3, the detection results for each identified flower are described and compared between the one-stage learning model and the two-stage learning model. The suggested model's bounding box prediction is more accurate than one-stage learning and two-stage learning YOLOv4 detection models for all classes, according to the detection results.

**Table 3** IoU, F1 score, mAP, validation loss, detection speed (FPS), and detection time (ms) are compared

Learning method	Mean Average precision (mAP)	Intersection over union (IoU)	F1 score	Validation loss	Detection speed (FPS)	Detection Time (ms)
One-stage YOLOv4 method	0.839	0.861	0.883	4.31	65.22	15.301
Two-stage YOLOv4 method	<b>0.912</b>	<b>0.922</b>	0.959	1.65	56.89	17.577

Bold signifies the proposed result compared with other results

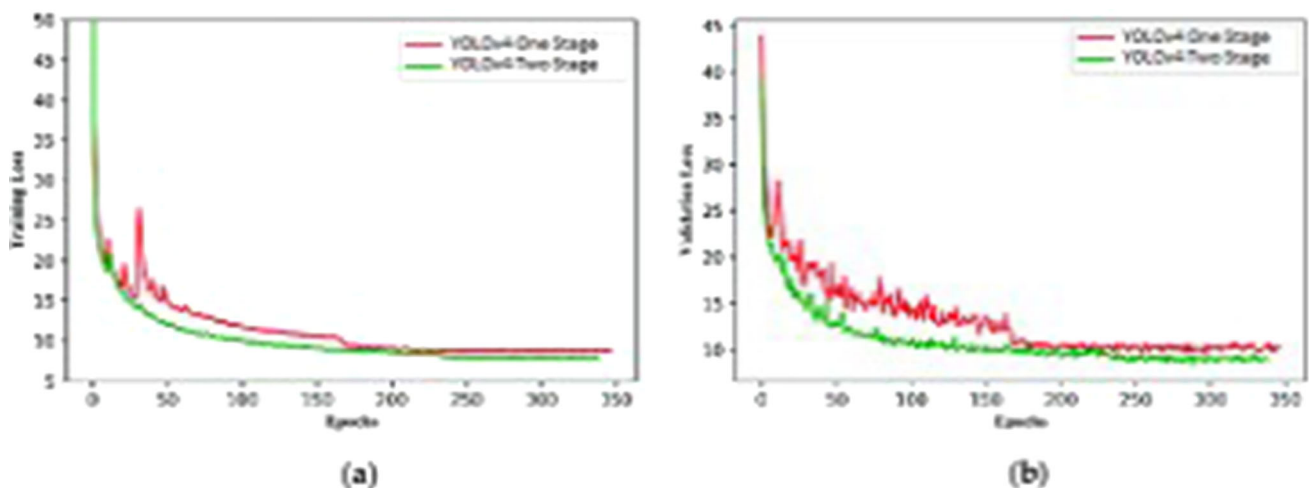


**Fig. 8** Comparison of PR curves between both one-stage and two-stage learning method

## 5 Conclusion

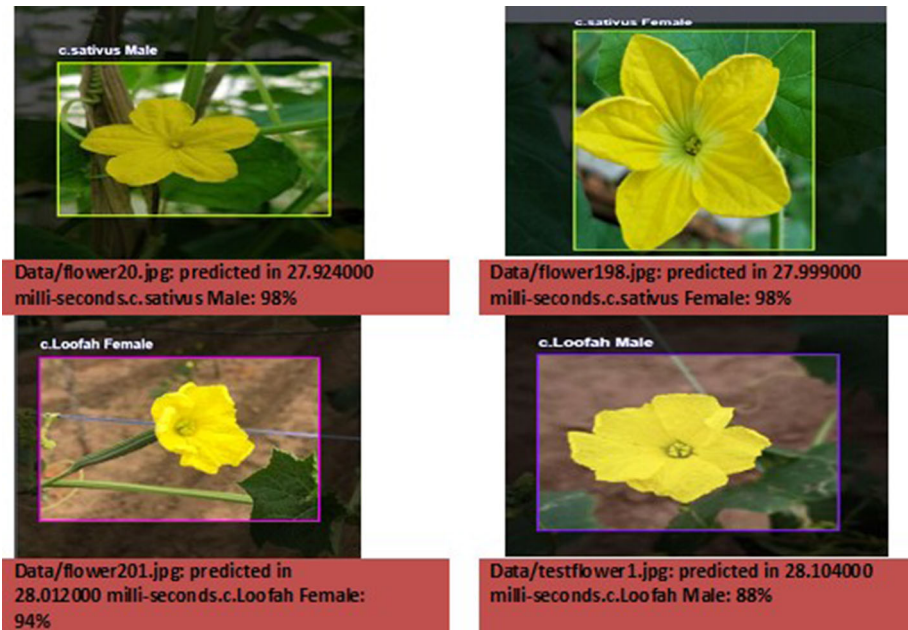
This research suggested a transfer learning object detection approach based on identifying and recognizing microscopic sections of cucurbitaceous pollination flowers. The

flower photographs were taken as a separate dataset called *CuCuflower* dataset in agricultural farms and poly homes using the DJI Mavic Mini drone. To categorize the genus and gender of blooming flowers in cucurbitaceous family plants, the YOLOv4 model uses a one- and two-stage deep learning strategy. The program was trained using the *CuCuflower* dataset to predict the flower species and gender in the one-stage learning approach. The two-stage learning approach (concatenated), on the other hand, is trained in two rounds with the same dataset, with the first step being to train the algorithm to recognize the properties of different cucurbitaceous family flowers as well as non-related flowers. The first step's output is then analyzed to determine the flower's genus and gender. To distinguish the species and gender of cucurbitaceous family plants, the training dataset included 4042 augmented photographs arranged into nine groups. According to the detection results, the two-stage learning model outperforms the one-stage learning model, with a mAP accuracy of 0.912, allowing it to recognize and categorize flower photographs with floral interference, overlapping ones, and fuzzy flower images. The suggested transfer learning method is an effective and fast way for recognizing small items in densely packed situations in real time.



**Fig. 9** Comparison between one-stage and two-stage learning method: (a) Training loss, (b) validation loss

**Fig. 10** Examples show the results of some flower genus and gender detection



**Funding** This study and all authors have received no funding.

**Data availability** The data used in this research study is available upon request due to privacy and confidentiality concerns. Interested parties may contact [mithrasamuel@gmail.com] to request access to the data. We are committed to ensuring the reproducibility and transparency of our research, and we will make every effort to provide access to the data in a timely and responsible manner.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interests.

**Ethical approval** This article does not contain any studies with human participants performed by any of the authors.

## References

- Althaus S, Berenbaum M, Jordan J et al (2021) No buzz for bees: Media coverage of pollinator decline. *Proc Natl Acad Sci* 118:e2002552117
- Amasino R, Cheung A, Dresselhaus T et al (2017) Focus on flowering and reproduction. *Plant Physiol* 173:1–4
- Barth R, Jsselmuiden JI, Hemming J et al (2018) Data synthesis methods for semantic segmentation in agriculture: a capsicum annum dataset. *Comput Electron Agricult* 144(284):296
- Bochkovskiy A, Wang C-Y, Liao H-Y (2020) Yolov4: Optimal speed and accuracy of object detection
- Chen L, Li Y (2019) Intelligent autonomous pollination for future farming—a micro air vehicle conceptual framework with artificial intelligence and human-in-the-loop. *IEEE Access* 7:119706–119717
- Du L, Zhang R, Wang X (2020) Overview of two-stage object detection algorithms. *J Phys: Conf Ser* 1544:012033
- Gogul I, Kumar S (2017) Flower species recognition system using convolution neural networks and transfer learning. pp 1–6
- Halder S, Khan R, Perween T et al (2019) Role of pollination in fruit crops: a review. *Pharma Innov J* 8:695–702
- Hossain M, Yeasmin F, Rahman M et al (2018) Role of insect visits on cucumber (*cucumis sativus* l.) yield. *J Biodiv Conserv Biores Manag* 4:81–88
- Ipate G, Voicu G, Dinu I (2015) Research on the use of drones in precision agriculture. *Univ Politeh Buch Bull Ser* 77:263–274
- Jiang Y, Li C, Xu R et al (2020) Deep flower: a deep learning-based approach to characterize flowering patterns of cotton plants in the field. *Plant Methods* 16:1–7
- Kurkute S (2018) Drones for smart agriculture: a technical report. *Int J Res Appl Sci Eng Technol* 6:341–346
- Li Y, Wang H, Dang LM et al (2020) A deep learning-based hybrid framework for object detection and recognition in autonomous driving. *IEEE Access* 8:395
- Lin K, Goyal H, Girshick R, Dollar P (2017) Focal loss for dense object detection. In: *IEEE international conference on computer vision (ICCV)* 42
- Liu H, Fan K, Ouyang Q et al (2021) Real-time small drones detection based on pruned yolov4. *Sensors* 21:3374
- Lopez A, Jurado JM, Ogayar CJ et al (2021) A framework for registering uav-based imagery for crop- tracking in precision agriculture. *Int J Appl Earth Observat Geoinform* 9:102274
- Stankovic M, Mirza MM, and Karabiyik U (2021) Uav forensics: Dji mini 2 case study. *Drones*
- Mithra S, Malleswari TYJN (2021) A literature survey of unmanned aerial vehicle usage for civil applications. *J Aerosp Technol Manag* 13:e4021
- Mohanty SP, Hughes DP, Salathe M (2016) Using deep learning for image-based plant disease detection”. *Front Plant Sci* 7:1046
- Nayeem MM (2018) flower identification using machine learning this report conferred to the department of CSE of daffodil international. PhD Thesis
- Newstrom-Lloyd L, Neeman GPG, Jürgens A, Dafni A (2010) A framework for comparing pollinator performance: Effectiveness and efficiency. *Biolog Rev* 85:435
- Oppenheim D, Shani G, Edan Y (2020) Tomato flower detection using deep learning
- Padilla R, Netto S, da Silva E (2020) A survey on performance metrics for object-detection algorithms



- Padmanabula S, Puvvada R, Sistla V, et al. (2020) Object detection using stacked yolov. *Ingénierie Des Systèmes d'Inform* **25**: 691–697
- Pandey S, Choudhary B (2014) *Cucumber*
- Ramachandran A, Sangaiah AK (2021) A review on object detection in unmanned aerial vehicle surveillance. *Int J Cogn Comput Eng* **2**:215–228
- Rehman E, Khan M, Algarni F et al (2021) Computer vision-based wildfire smoke detection using uavs. *Math Probl Eng* **27**:1–9
- Rezatofighi H, Tsoi N, Gwak J, et al (2019) Generalized intersection over union: a metric and a loss for bounding box regression. In: 2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR), pp 658–666
- Roy A, Bose R, Bhaduri J (2022) A fast accurate fine-grain object detection model based on yolov4 deep neural network. *Neural Comput Appl* **34**:1–27
- Sangaiah AK, Rezaei S, Javadpour A, Zhang W (2023) Explainable AI in big data intelligence of community detection for digitalization e-healthcare services. *Appl Soft Comput* **136**:110119
- Sivakumar M, Nagamalleswari T (2022) An analysis of deep learning models for dry land farming applications. *Appl Geomat* **1**:1–7
- Sowmya V, Radha R (2021) Heavy-vehicle detection based on yolov4 featuring data augmentation and transfer-learning techniques. *J Phys: Conf Ser* **1911**:012029
- Sun X, Zhenye L, Zhu T et al (2021) Four-dimension deep learning method for flower quality grading with depth information. *Electronics* **10**:2353
- Tong K, Wu Y, Zhou F (2020) Recent advances in small object detection based on deep learning: a review. *Image Vis Comput* **97**:103910
- Tuda M, Luna-Maldonado A (2020) Image-based insect species and gender classification by trained supervised machine learning algorithms. *Ecolog Inform* **60**:101135
- Ushasukhanya S, Karthikeyan M (2021) Automatic human detection using reinforced faster r-cnn for electricity conservation system. *J Intell Autom Soft Comput* **32**:1
- Ushasukhanya S, Jothilakshmi S, Sridhar SS (2022) Development and optimization of deep convolutional neural network using taguchi method for real-time electricity conservation system. *Int J Inform Technol* **14**(1):4
- Wan S, Goudos S (2019) Faster r-cnn for multi-class fruit detection using a robotic vision system. *Comput Netw* **168**:107036
- Wang L, Zhao Y, Liu S et al (2022) Precision detection of dense plums in orchards using the improved yolov4 model. *Front Plant Sci* **13**:1024
- Wei W (2020) Small object detection based on deep learning. In: 2020 IEEE international conference on power, intelligent computing and systems (ICPICS), pp 938–943
- Wu D, Lv S, Jiang M et al (2020) Using channel pruning-based yolo v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments. *Comput Electron Agric* **178**:105742
- Xiong J, Liu B, Zhong Z et al (2021) Litchi flower and leaf segmentation and recognition based on deep semantic segmentation. *Nongye Jixie Xuebao/Trans Chin Soc Agric Mach* **52**:252–258
- Zhang J, Feng W, Yuan T, Wang J, Sangaiah AK (2022) SCSTCF: spatial-channel selection and temporal regularized correlation filters for visual tracking. *Appl Soft Comput* **118**:108485

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.