

Systematic literature review: application of deep learning processing technique for fig fruit detection and counting

Ahmad Shukri Firdhaus Kamaruzaman¹, Adi Izhar Che Ani¹, Mohammad Afiq Hamdani Mohammad Farid¹, Siti Juliana Abu Bakar¹, Mohd Ikmal Fitri Maruzuki¹, Samsul Setumin¹, Mokh. Sholihul Hadi²

¹Centre for Electrical Engineering Studies, Universiti Teknologi MARA, Cawangan Pulau Pinang, Permatang Pauh Campus, Malaysia

²Department of Electrical Engineering, Faculty of Engineering, State University of Malang Indonesia

Article Info

Article history:

Received Jul 22, 2022

Revised Aug 28, 2022

Accepted Oct 27, 2022

Keywords:

Deep learning

Detection and counting

Model performance

Systematic literature review

ROSES

ABSTRACT

Deep learning has shown much promise in target identification in recent years, and it's becoming more popular in agriculture, where fig fruit detection and counting have become important. In this study, a systematic literature review (SLR) is utilised to evaluate a deep learning algorithm for detecting and counting fig fruits. The SLR is based on the widely used 'Reporting Standards for Systematic Evidence Synthetics' (ROSES) review process. The study starts by formulating the research questions, and the proposed SLR approach is critically discussed until the data abstraction and analysis process is completed. Following that, 33 relevant research involving the agriculture sector, fruit, were selected from many studies. IEEE, Scopus, and Web of Sciences are three databases to investigate. Due to the lack of fig fruit research, fruit and vegetable studies have been included because they use similar methods and processes. The SLR found that various deep learning algorithms can count fig fruit in the field. Furthermore, as most approaches obtained acceptable results, deep learning's performance is acceptable in F1-score and average precision (AP), higher than 80%. Moreover, improvements can be produced by enhancing the existing deep learning model with the personal dataset.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Adi Izhar Che Ani

Centre for Electrical Engineering Studies, Universiti Teknologi MARA Penang Branch

Permatang Pauh Campus

13500 Permatang Pauh, Malaysia

Email: adiizhar@uitm.edu.my

1. INTRODUCTION

In the field [1] of detecting things in images, computer vision and pattern recognition are tools in this growing field. Object detection methods have numerous applications, including the detection of vegetables and fruits [2], [3]. The recent explosion in the capability of both AI algorithms and image sensor technology has led to the rise of the automated fruit detection system [4], [5]. Fruit detection in orchards has traditionally relied largely on manual visual inspection, which is both time-consuming and labor-intensive [6]–[10]. Human perception is used in conventional orchards to record agricultural data, however there can be large discrepancies between the two sets of data because farmers have widely diverse levels of competence [11]. Hence, for developing autonomous harvesting, targeted medicine applications, and many other intelligent farm machinery technologies, an effective automatic detection approach for the agriculture sector is essential [12]–[14]. Deep learning approaches based on machine vision can extract hidden patterns from agricultural datasets to construct and build a prediction framework that can help agriculturists diagnose [15]. The fruits' size, colour, and form are used to train neural networks [16].

In machine learning and deep learning, computer vision algorithms have enhanced the efficiency of image recognition and detection tasks [17]. The findings [18] indicate that machine learning can enhance the performance of a picking robot's detection technique. Deep learning's progress has turned computer vision into an agricultural vision for target detection and image semantic segmentation, giving the best results [19]. Because of its ability to handle large amounts of data, deep learning has proven to be a very effective tool [20]. The interest in using hidden layers has exceeded traditional techniques in terms of popularity, particularly in object recognition, classification, and detection [21]. One of the most popular types of deep neural networks is known as a convolutional neural network (CNN) [22]–[24].

Due to the fast growth of CNN in recent years, the detection accuracy and speed of CNNs are frequently superior to traditional object detection methods [25]. CNNs in their many versions have been put to use in the detection of a wide variety of fruits [26]–[28]. CNN is built on a convolution-shaped pyramid and pooling layers, which minimise picture width and height while enhancing depth measurement [29]. As a result of this, the classifier is positioned atop a pyramid, which serves the purpose of connecting the many nodes that make up the neural network.

This paper critically examines the detection and counting of fig fruits using deep learning by conducting a systematic literature review. Literature indicates that deep learning outperforms conventional fruit identification, recognition, and counting techniques. Aside from that, deep learning in agriculture is both time and cost-efficient. In this research, we implemented the SLR based on the standard review methodology called "Reporting Standards for Systematic Evidence Syntheses" (ROSES). The SLR method used in this study is looked at from the start of the research question to the end of gathering and analyzing data.

2. METHOD

2.1. The review protocol-ROSES

The current study's SLR follows the ROSES review process [30]–[33] as a reference. ROSES review protocols, as can be seen illustrated in Figure 1, are specifically created for SLR. Haddaway *et al.* [34] stated that ROSES also includes a comprehensive set of reporting requirements for the conservation and environmental management research synthesis community. Formulation of the topic of interest, systematic searching techniques, quality evaluation, and data abstraction and analysis are the four main steps in conducting SLR according to ROSES. Three sub-processes are required in systematic searching strategies: identification, screening, and eligibility of obtained articles. Only high-quality articles related to the main research question are selected and reviewed throughout the SLR procedure.

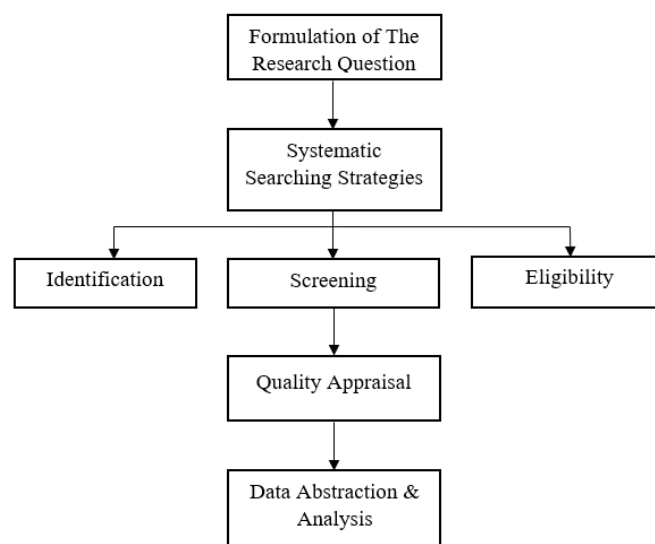


Figure 1. ROSES review protocol modified

2.2. Formulation of research question

The preparation [35] of the research topic or question is based on the population, intervention, and Context (PICO) tool, defined based on three main elements: population, interest, and context [36]. This tool

was adopted from Pollock and Berge [37] to evolve research questions suitable for the review. Based on this tool, the population of this research was deep learning, the interest is detection and counting, and the context will be fig fruit in the wild. The three main aspects of this review, based on this technique, were then utilized as a guideline in generating the primary findings, which are:

RQ1: Which deep learning model approaches were used to detect and count fruit in the wild?

RQ2: What is the dataset preparation process used?

RQ3: What is the performance of each deep learning model overall?

2.3. Systematic searching strategies

The next section will discuss the three sub-processes that make up the second SLR approach. These sub-processes are the identification, screening, and eligibility procedures that are used to locate relevant and related articles for this review. Through this section, the number of articles can be limited from a large number to a limited number, from which only the best and most appropriate papers will be used and focused.

2.3.1. Identification

The method for identification involves looking for words that have the same meaning as the phrases, are related to them, and serve as the primary keywords for the investigation. The objective is to expand the search capabilities of the chosen database so that related articles may be found more easily for the review. The keywords are selected in accordance with the study's objective and problems [38], and they are retrieved with the help of three main indexed databases: IEEE, Scopus, and Web of Science. The keywords used to discover related articles were derived from the study question. Furthermore, to avoid a low number of papers being retrieved, the study widened the scope by including fruits and vegetables to collect more papers. Table 1 displays the exhaustive search string built by the researchers using the Boolean operators "AND" and "OR", phrase searching, compression, and wild cards in both databases. This approach successfully retrieved 1032 IEEE articles, 417 Scopus articles, and 206 Web of Science (WOS) articles.

Table 1. Advanced search string

Database	Advance search string
IEEE	("detect* AND count*") AND ("Fig*" OR "Fruit*" OR "Vegetable*") AND ("Artificial Intelligence" OR "Deep Learning" OR "Neural Network" OR "Machine Learning")
Scopus	ABS ((detect* AND count*) AND (("Fig") OR ("Fruit") OR ("Vegetable"))) AND (("Artificial Intelligence") OR ("Deep Learning") OR ("Neural Network") OR ("Machine Learning"))
Web of Science	ALL=(ALL ((detect* AND count*) AND (("Fig") OR ("Fruit") OR ("Vegetable"))) AND (("Artificial Intelligence") OR ("Deep Learning") OR ("Neural Network") OR ("Machine Learning")))

2.3.2. Screening

The 1655 papers in this study were filtered using the criteria for article selection. The method was carried out automatically using the database's sorting function. The research question produced in the preceding procedure served as the basis for the selection standard [39], [40]. Then, the publications were collected between 2017 and 2021, and only English-language published articles were chosen from IEEE, Scopus, and WOS. Furthermore, this study only reviews published journal and conference papers to ensure that acceptable scientific papers relating to our study are included. Other kinds of papers, like books and article reviews, did not make the cut. During this process, about 1249 papers were not included because they did not meet the criteria. Table 2 shows the inclusion and exclusion articles based on timeline, language and type of source.

Table 2. The criteria for inclusion and exclusion

Criteria	Inclusion	Exclusion
Timeline	2017-2021	Before 2017
Language	English	Non-English
Type of source	Journal and Conference	Other Than Journal and Conference

2.3.3. Eligibility

The third step is the eligibility process. In this study, all of the articles that were found were checked manually to make sure that after the screening process, all of the remaining articles met the criteria [41]. The procedure was completed by reading the titles and abstracts of the selected papers. During this procedure, 373 articles were eliminated since their principal goal was unclear and they did not have a major impact on

climate change. Besides, for papers that have duplicates in different resources, one of the papers will be removed to ensure that the study does not refer to the same paper. The remaining 33 papers were chosen for the quality assessment step of the procedure. Based on Figure 2, the flow diagram shows that the research method had narrowed down the focus of this study to only 33 papers to be reviewed from 1655 papers identified by using a systematic literature review.

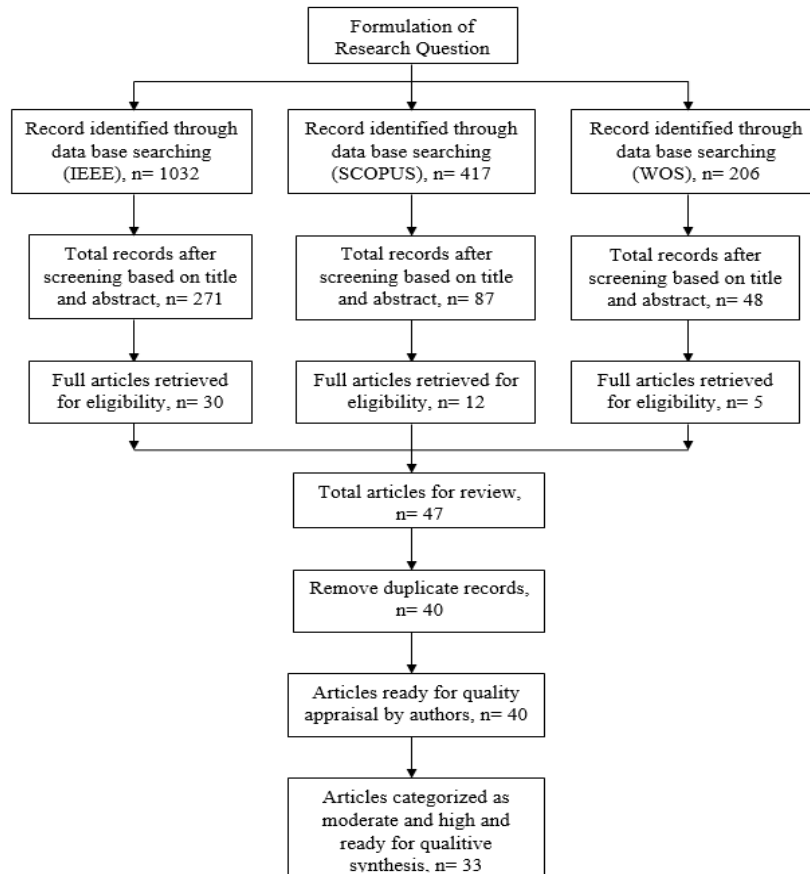


Figure 2. The flow diagram of research methodology

2.4. Quality appraisal

The articles were displayed utilizing the quality assessment to choose the high-quality content [42]. The papers were categorized into three quality levels during this process: high, moderate, and low. To determine the rating of quality, high and medium articles were examined based on the methodology and results of the articles. The research was evaluated using the following quality criteria:

QA1: Is the study related to the research objectives?

QA2: Is the deep learning models technique mentioned in the study?

QA3: Is there a description of the data preparation method?

QA4: Is the research methodology stated in detail?

QA5: Has the effectiveness of the proposed methodology been analysed?

2.5. Data abstraction and analysis

In-depth, the researcher screened the abstract, methodology [42], results, and discussion parts for all 33 papers. The research questions were used to keep track of the data abstractions. This meant that any data from the studies that could help answer the research questions was abstracted and put into a table.

3. RESULTS AND DISCUSSION

A total of 33 papers were reviewed based on the research method. Several aspects were developed based on the systematic review, including publication year, deep learning model technique, dataset

preparation, and deep learning model performance evaluation. The results presented were based on a review of previously published research on these topics.

3.1. Selected primary studies

Utilizing IEEE, Scopus, and WOS, 33 papers were selected as the primary study for review based on the research method. All the selected papers discussed the deep learning model approach. The selected primary study's identification (I.D.), publication titles, authors, and publication year are presented in Table 3.

Table 3. Summary of selected primary study

ID	Title	Author	Year
S1	Deep learning implementation using convolutional neural network in mangosteen surface defect detection	Azizah <i>et al.</i> [10]	2017
S2	Deep fruit detection in orchards	Bargoti and Underwood [43]	2017
S3	Real-Time Vegetables Recognition System based on Deep Learning Network for Agricultural Robots	Zheng <i>et al.</i> [3]	2018
S4	Faster R-CNN Implementation Method for Multi-Fruit Detection Using Tensorflow Platform	Basri <i>et al.</i> [9]	2018
S5	Apple recognition based on Convolutional Neural Network Framework	Liang <i>et al.</i> [19]	2018
S6	A Detection Method for Tomato Fruit Common Physiological Diseases Based on YOLOv2	Zhao and Qu [12]	2019
S7	Tomato Fruit Image Dataset for Deep Transfer Learning-based Defect Detection	Luna <i>et al.</i> [7]	2019
S8	Pitaya detection in orchards using the MobileNet-YOLO model	Li <i>et al.</i> [8]	2019
S9	Cucumber Fruits Detection in Greenhouses Based on Instance Segmentation	Liu <i>et al.</i> [26]	2019
S10	Automatic Detection of Single Ripe Tomato on Plant Combining Faster R-CNN and Intuitionistic Fuzzy Set	Hu <i>et al.</i> [23]	2019
S11	A Computer Vision System for Guava Disease Detection and Recommend Curative Solution Using Deep Learning Approach	Haque <i>et al.</i> [15]	2019
S12	Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of 'MangoYOLO'	Koirala <i>et al.</i> [44]	2019
S13	Intra Class Vegetable Recognition System using Deep Learning	Duth and Jayasimha [2]	2020
S14	Fast and Accurate Detection of Banana Fruits in Complex Background Orchards	Fu <i>et al.</i> [4]	2020
S15	A fruit detection algorithm based on R-FCN in natural scene	Jian <i>et al.</i> [13]	2020
S16	Visual Perception and Modeling for Autonomous Apple Harvesting	Kang <i>et al.</i> [14]	2020
S17	A Deep Neural Network based disease detection scheme for Citrus fruits	Kukreja and Dhiman [27]	2020
S18	Fruit Detection in the Wild: The Impact of Varying Conditions and Cultivar	Halstead <i>et al.</i> [11]	2020
S19	Deep Learning for Assessing Unhealthy Lettuce Hydroponic Using Convolutional Neural Network based on Faster R-CNN with Inception V2	Pratama <i>et al.</i> [45]	2020
S20	IHDS: Intelligent Harvesting Decision System for Date Fruit Based on Maturity Stage Using Deep Learning and Computer Vision	Faisal <i>et al.</i> [18]	2020
S21	Automated Sorter and Grading of Tomatoes using Image Analysis and Deep Learning Techniques	Bautista <i>et al.</i> [28]	2020
S22	Deep learning image segmentation and extraction of blueberry fruit traits associated with harvestability and yield	Ni <i>et al.</i> [46]	2020
S23	Tomato Fruit Detection and Counting in Greenhouses Using Deep Learning	Afonso <i>et al.</i> [21]	2020
S24	Deep learning models compression for agricultural plants	Fountsop <i>et al.</i> [29]	2020
S25	A Novel Greenhouse-Based System for the Detection and Plumpness Assessment of Strawberry Using an Improved Deep Learning Technique	Zhou <i>et al.</i> [47]	2020
S26	Grape detection, segmentation, and tracking using deep neural networks and three-dimensional association	Santos <i>et al.</i> [48]	2020
S27	Intact Detection of Highly Occluded Immature Tomatoes on Plants Using Deep Learning Techniques	Mu <i>et al.</i> [49]	2020
S28	Fig Fruit Recognition Method Based on YOLO v4 Deep Learning	Yijing <i>et al.</i> [50]	2021
S29	Implementation of Deep Learning Methods to Identify Rotten Fruits	Chakraborty <i>et al.</i> [17]	2021
S30	Deep Learning for improving the storage process: Accurate and automatic segmentation of spoiled areas on apples	Stasenko <i>et al.</i> [24]	2021
S31	Easy domain adaptation method for filling the species gap in deep learning-based fruit detection	Zhang <i>et al.</i> [5]	2021
S32	In-field automatic detection of grape bunches under a totally uncontrolled environment	Ghiani <i>et al.</i> [51]	2021
S33	Strawberry Yield Prediction Based on a Deep Neural Network Using High-Resolution Aerial Orthoimages	Cheng <i>et al.</i> [52]	2021

3.2. Publication years

This study's publication year has been set between 2017 and 2021, covering the study within the last 5 years. Since deep learning is one of the recent studies, especially in the agriculture sector, it is best to review or study the recent studies. Figure 3 depicts the number of studies conducted in agriculture over the last five years, from 2017 to 2021, including the use of deep learning for object detection. Most of the studies

were published in 2020, a total of 16. In 2017, only 2 studies were found to do deep learning in the agricultural sector, which increased to 3 studies in 2018.

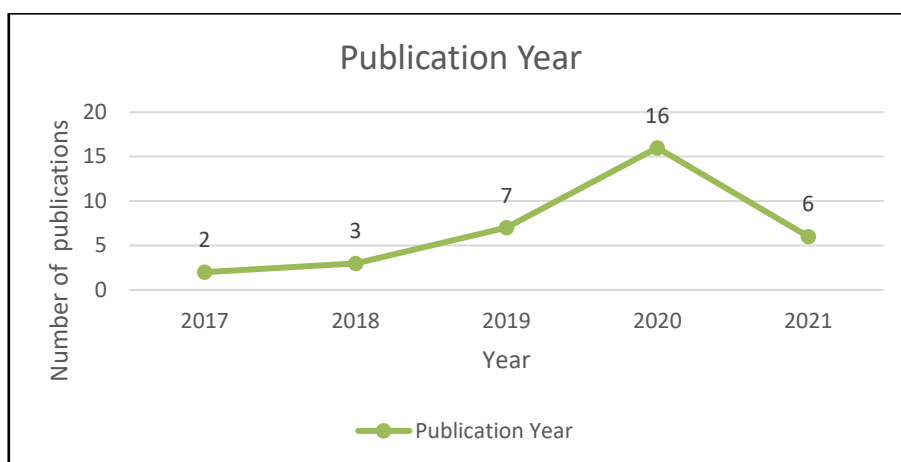


Figure 3. Number of selected studies in the past five years

In addition, there has been a rise in the number of publications published thus far this year. During 2019, seven studies were released. When compared to 2019 and 2020, there was a considerable drop in 2021. This decrease could be attributed to the effects of the Covid-19 pandemic, which resulted in fewer conferences due to sanitary regulations. However, since this study was conducted in 2021, it is normal that only 6 studies be published in 2021 since they might not be readily available for analysis. It shows that the study of deep learning in agriculture is getting more and more popular every year in the world we live in now.

3.3. Quality appraisal result

Based on the Q.A. question explained in section 2.4, 33 studies have been chosen and analyzed. To measure the quality of the articles, the researchers modified the scoring technique used by Alsolai *et al.* [53]. The following was the quality evaluation score system: i) One-point equals Yes, ii) 0.5 points = Partially, and iii) 0 equals No. The scoring points separated the articles into three categories: i) zero to two points were regarded weak, ii) two to five points were considered moderate, and iii) three to five points were rated strong. Table 4 shows the summaries of the Q.A. analysis results. 21 studies have been rated as strong, with five having a full score that meets all the criteria in the question. Next, 12 studies were rated as moderate, with most scoring three points, and only one study scored two-point-five.

3.4. Deep learning approach

There are two main types of deep learning detection techniques used for image detection systems. The first type is the detection method based on region generation, also known as two-stage target detection, in which an algorithm first generates a series of candidate frames, and then classifies the targets in those frames. The region convolution neural network (RCNN), the mask RCNN, the fast RCNN, and the faster RCNN are all good examples of this type of network. Though effective, these techniques are too time-consuming and cumbersome to be used in real-time detection settings [50].

The second type is regression-based methods, which simultaneously perform target localization and target category prediction (hence the name "one-stage target detection"). Among the many types of networks available, single shot detection (SSD) and the you only look once (YOLO) series stand out as particularly effective applications. Santos *et al.* [48] mentioned that this family of methods has a quick recognition speed and can meet real-time requirements.

Table 5 summarises the results of a meta-analysis that found 10 studies had used a single-stage object detection approach. One study employed YOLOv2, two employed YOLOv3, two employed YOLO v4, and one employed SSD. Additionally, three studies have modified the architecture or network of the YOLO to create their own YOLO deep learning model and improved the performance. The studies from S8 and S31 have modified the YOLOv3 model and improved the performance, while the study from S12 has created its own YOLO model, named MANGO-YOLO, based on the architecture of the YOLO network.

Table 4. Methodological quality assessment of studies

Study ID	Q1	Q2	Q3	Q4	Q5	Score	Rating
S1	0.5	1.0	0.5	1.0	1.0	4.0	Strong
S2	1.0	1.0	1.0	1.0	1.0	5.0	Strong
S3	0.0	0.5	0.5	0.5	1.0	2.5	Moderate
S4	1.0	1.0	0.5	0.5	1.0	4.0	Strong
S5	0.5	1.0	1.0	1.0	1.0	4.5	Strong
S6	0.0	1.0	0.5	1.0	1.0	3.5	Strong
S7	0.5	1.0	0.5	1.0	0.5	3.5	Strong
S8	0.5	1.0	1.0	1.0	0.5	4.0	Strong
S9	0.5	0.5	0.5	1.0	0.5	3.0	Moderate
S10	0.5	1.0	1.0	1.0	0.0	3.5	Strong
S11	0.0	0.5	1.0	0.5	1.0	3.5	Strong
S12	1.0	1.0	1.0	1.0	1.0	5.0	Strong
S13	0.0	0.5	1.0	0.5	1.0	3.0	Moderate
S14	1.0	1.0	1.0	0.5	1.0	4.5	Strong
S15	0.5	1.0	1.0	1.0	0.5	4.0	Strong
S16	0.0	0.5	0.5	1.0	1.0	3.0	Moderate
S17	0.5	0.5	1.0	0.0	1.0	3.0	Moderate
S18	1.0	0.5	1.0	1.0	0.5	4.0	Strong
S19	0.5	1.0	0.5	1.0	1.0	4.0	Strong
S20	0.0	0.5	0.5	1.0	1.0	3.0	Moderate
S21	0.0	0.5	1.0	1.0	1.0	3.5	Strong
S22	0.0	1.0	0.5	0.5	1.0	3.0	Moderate
S23	1.0	1.0	1.0	1.0	1.0	5.0	Strong
S24	0.5	1.0	0.5	0.5	0.5	3.0	Moderate
S25	1.0	1.0	1.0	0.0	1.0	5.0	Strong
S26	0.0	0.5	0.5	1.0	1.0	3.0	Moderate
S27	1.0	0.5	0.5	1.0	0.5	3.5	Strong
S28	1.0	1.0	1.0	1.0	1.0	5.0	Strong
S29	0.5	1.0	1.0	0.0	1.0	4.5	Strong
S30	0.0	1.0	1.0	0.5	0.5	3.0	Moderate
S31	0.0	0.5	1	1	0.5	3.0	Moderate
S32	1.0	1.0	0.5	1.0	0.5	4.0	Strong
S33	0.0	0.5	0.5	1	1	3.0	Moderate

Also, 23 research have utilised two-stage target detection; 12 of them have employed mask RCNN, 9 have used Faster RCNN, and 3 have employed the modified deep learning model based on the current Faster RCNN model in terms of the backbone. One study used a modified mask RCNN. Table 5 clearly shows that the majority of the studies used two-stage target detection, particularly mask RCNN, as deep learning models in their studies.

Table 5. Deep learning model approach

Object detection method	Deep learning model	Study ID
One stage target detection	YOLOv2	S6
	YOLOv3	S3, S16
	YOLOv4	S14, S28
	Modified YOLO	S8, S12, S31
	SSD	S5
Two stage target detection	Mask RCNN	S1, S11, S7, S13, S15, S17, S20, S22, S23, S24, S30, S32
	Faster RCNN	S2, S4, S5, S10, S18, S19, S21, S29, S33
	Modified faster RCNN	S25, S26, S27
	Modified mask RCNN	S9

Figures 4 and 5 display the architecture or network of one-stage and two-stage detection, respectively. The two-stage detector can be split by a region of interest (RoI) pooling layer [52]. The region proposal network (RPN) is the initial stage that predicts possible bounding boxes [54]. For the subsequent classification and bounding box regression tasks, features are pooled from each candidate box using the RoI pooling technique, which is the focus of the second phase. However, a one-stage detector makes bounding box predictions in a step, without the need for region predictions. It uses a grid box and anchors to limit the shape of the item and pinpoint where it is in the picture [44].

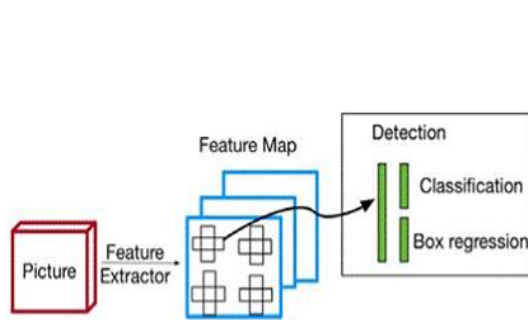


Figure 4. One stage target detection

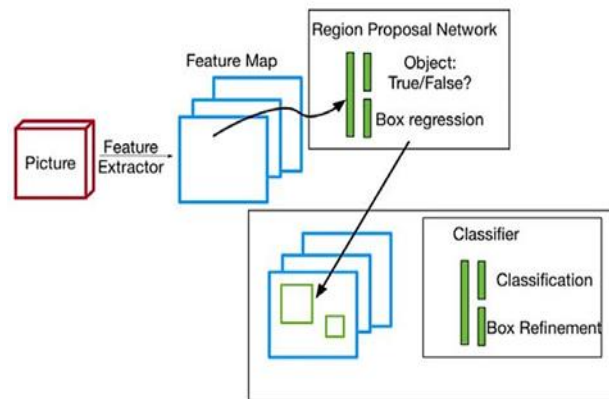


Figure 5. Two stages target detection

3.5. Dataset preparation

Preparing a dataset is one of the must-do steps or processes in object detection, recognition and counting. There were several processes of dataset preparation, such as the collection of datasets, resizing, auto-orientation, annotation, data augmentation and data splitting. Annotation was a critical step in object detection. Every study on object detection, recognition, and counting must go through an annotation process [46]. Table 6 shows all the studies from S1-S33 were doing annotation processes for the data preparation. Annotation is done with the help of professional human annotators using specified labels. In simple terms, image annotation entails adding metadata to a dataset that allows machines to recognize certain items in the image. Thus, the dataset that has been annotated must be verified by the expert.

Table 6. Dataset preparation analysis

Dataset	Dataset preparation	Study ID
Collected manually	Resize and annotation	S3, S6, S8, S10, S12, S14, S19, S21, S23, S27, S28, S30, S33
	Resize, annotation and data augmentation	S5, S9, S20, S22, S25, S26, S32
Public dataset	Resize and annotation	S1, S2, S4, S7, S13, S15, S16, S18, S27
	Resize, annotation and data augmentation	S11, S17, S24, S29, S31

Next, resizing was also an important process in object detection besides doing an annotation. Every deep learning algorithm has its standard size of input or image to be extracted [7]. Moreover, Liu *et al.* [26] investigated that resizing the image to a smaller size can reduce the training time for the model to recognize or learn the image. However, if the images were resized too small, such as 76 x 76 pixels or 144 x 144 pixels, the input image would not be sharp enough for the network to recognize and learn the input image. This sampling method has two key drawbacks [55].

First, fine-grained visual characteristics essential for detecting small, abstract objects like balls may be lost due to possible image subsampling. The elongation of items in the image caused by resizing to a squared format results in a change in the characteristic shape of the fruit in particular and contributes to additional distortion of the fruit's aesthetic qualities. The common sizes of the input image for object detection, recognition, and counting were 416 x 416 pixels and 512 x 512 pixels [8], [10]. If the image size were larger than this standard size, it would cause the image's resolution to be high, thus increasing the training time for the model to learn and recognize the input image [27]. The higher the image's resolution, such as 1080 x 1080 pixels, the sharper the image will be for the model to recognize even the small objects in the image, causing the longer training time. Hence, following the standard size of 416 x 416 pixels and 512 x 512 pixels that other researchers have recommended is the best option.

Based on Table 6, the dataset used in object detection, recognition, and counting studies can be divided into two types. The first is to use public datasets, which have been published by other researchers and have given permission to the other researchers to use their datasets. There were 14 studies based on tables that used public datasets. Typically, public datasets may be found in the MS COCO dataset, which contains about 330 thousand photos and 80 object categories [13], [56], ImageNet is a database of over a million photos and one thousand different kinds of objects [43] and the Kaggle dataset [57], which has been collected manually by the researcher and published on it. Besides, Roboflow was also one of the public datasets that new researchers or other users could access. Roboflow provides users access to public datasets and the ability

to submit their own custom data. Roboflow supports several different annotation and export formats [58]. All the data in these public datasets have been resized and readily used by the user.

The second uses the personal dataset, which is manually collected at the orchard. From our survey, 19 studies collected their datasets manually. There were several reasons why the researchers used their dataset, such as limited resources for the dataset they needed to use for their study. Furthermore, employing a personal dataset might provide a selection of photos that match the researcher's requirements. The public dataset did not have all the images that could satisfy the researcher's requirements. For example, the researcher needs data for fruit under different weather, location, and background [11].

There were 12 studies doing data augmentation, including 7 studies on personal datasets and 5 on public datasets. Data augmentation is a technique for solving the problem of overfitting in the training stage of CNNs [59]. Overfitting occurs when random noise or errors are presented instead of the underlying relationship. The researchers used several data augmentation techniques, including brightness adjustment, blur, cropping, rotation, flip, zoom, and noise disturbances [27]. Bargoti and Underwood [43] have investigated that after adding more photos to the data, the model can learn as many irrelevant patterns as possible during the training phase, which helps it avoid overfitting and makes it better at its job. Last but not least, Xu *et al.* [60] mentioned that splitting the dataset correctly into training, testing, and validation was also an important task for better performance. The whole database was separated into two datasets, the training set and the testing set, with the images being selected at random. When it comes to neural network applications, the most common training/testing dataset splitting ratio is 80/20 [61], although other similar splitting ratios, such as 70/30, should not significantly affect the performance of the resulting models [50].

3.6. Deep learning model performance

The effectiveness of the deep learning model will be discussed in this section. Researchers in the 33 studies used 10 different deep learning models. The effectiveness of the deep learning model must be assessed. It is because, by evaluating the performance, the deficiency of the model to detect the object can be improved to achieve the study objective. It is proven that the studies were successful in achieving their objectives. The highest the performance of each model is scored for every metric, the better the deep learning model.

It is essential to evaluate the efficacy of a deep learning model by looking at its accuracy, precision, recall, F1-score, and average precision (AP). Table 7 illustrates the data taken from it about how accurate, precise, recallable, or sensitive it is. There are 4 main ideas that are utilized to evaluate performance metrics: true positive (TP), true negative (TN), false positive (FP), and false negative (FN) [14]. It is true when the prediction is right and the predicted label matches the ground-truth label, and it is false when the predicted label does not match the ground-truth label [62]. The label that was predicted was either negative or positive.

Overall, if the prediction is wrong, the first word will be false. If not, then it is definitely correct. True positive and negative rates should be maximised, while false positive and false negative rates should be kept to a minimum. And for detection purposes, the prediction needs to account for where in the image an instance of the class is [55]. The amount of overlap between the discovered bounding box and the ground truth object was applied to evaluate the accuracy of the detection [55]. For a detection to be considered accurate, there should be more than a 50% overlap in between predicted bounding box (Bp) and the ground truth bounding box (Bgt) [63].

Table 7. Deep learning performance algorithm

Metrics of performance	Description	Algorithm
Accuracy	The percentage of the actual result that corresponds to the prediction made with high accuracy.	$A = \frac{TP + TN}{TP + TN + FP + FN}$
Precision	The percentage of accurately predicted positive results as a percentage of the total number of positive results predicted.	$P = \frac{TP}{TP + FP}$
Recall	The percentage of positive outcomes that were accurately predicted compared to the total outcome.	$P = \frac{TP}{TP + FN}$
F1-score	The weighted average of precision and recall if the class is not evenly distributed.	$F1 - Score = \frac{2(R \times P)}{R + P}$
Average precision	The sum of the precisions at each threshold, weighted by the increase in recall.	$AP = \int_0^1 P(R) dR$

Based on Table 8, 17 studies focused on evaluating their deep learning performance on F1-score, 12 evaluated their model performance through AP, and 9 evaluated the model's accuracy. In contrast, the study from paper (S10) did not detail the performance of their deep learning models. The researcher evaluated the performance based on the mean relative error (MRE). Next, most of the researchers' deep

learning model performance approaches achieve more than 80% accuracy, precision, recall, F1-score and AP. However, one paper (S19) has accuracy and recall below 80%, but it's still acceptable since their F1-score was 80%. It is because Bresilla *et al.* [64] stated that precision and recall are so closely linked that we can only utilize the F1-score, which considers both precision and recall when calculating the score and how well the forecast matches the ground truth. For a comprehensive evaluation of a model's efficacy, researchers utilise the F1-score, which gives equal importance to accuracy and recall [26]. A F1-score is calculated by averaging the classification model's recall and accuracy.

The precision of a prediction is measured by how closely it matches the actual outcome [22]. Accuracy, however, is of use only when the dataset or sample is well-balanced. Accuracy [60] is defined as the ratio of true positive samples to all anticipated positive samples. Recall [25] measures how well a model is able to predict real positive samples relative to the total number of actual positive samples. In general, recall decreases as precision improves [49]. Precision and recall levels can be shown by plotting a P-R graph [23]. In order to weigh the relative importance of accuracy and recall in a model's overall evaluation. Yijing *et al.* [50] proposed utilising AP as a complete evaluation indicator. Better model performance is indicated by a bigger value of AP, which is the area under the P-R curve [52].

Table 8. Result based on performance metrics

Study ID	Model	Result				
		Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Average precision (%)
S1	Mask RCNN	97.5				
S2	Faster RCNN				90	
S3	Improved YOLOv3					87.89
S4	Improved faster RCNN	99				
S5	SSD-500				89	
S6	Improved YOLOv2			97		97.24
S7	Mask RCNN	95.75				
S8	Modified YOLO	97		90	81.2	
S9	Improved mask RCNN		90.68	88.29	89.47	
S10	Faster RCNN					
S11	Mask RCNN	95.61	97.98	97	97.49	
S12	Modified YOLO				96.8	98.3
S13	Mask RCNN	95.5				
S14	YOLOv4					99.29
S15	Mask RCNN	94.23				
S16	YOLOv3		87	85.8	86.4	
S17	Mask RCNN	89.1				
S18	Faster RCNN				90	
S19	Faster RCNN	70	97	68	80	
S20	Mask RCNN				94.8	
S21	Faster RCNN					88
S22	Mask RCNN		91	89	90	
S23	Mask RCNN				90	
S24	Mask RCNN				86.97	
S25	Improved faster RCNN				86	
S26	Faster RCNN				91	
S27	Modified faster RCNN				83.67	87.83
S28	YOLOv4					93
S29	Faster RCNN					89
S30	Mask RCNN		88.9	92	90.4	90
S31	Improved YOLOv3					87.5
S32	Mask RCNN					91
S33	Faster RCNN					83

4. CONCLUSION

An SLR was performed on the development of AI in agriculture recently, with special focus on the use of a deep learning algorithm for detecting and counting fig fruits. However, due to the scarcity of articles on fig fruits, the review expanded the scope of the study to include fruits and vegetables. This systematic review study is meant to add to what is known by giving an overview of the available deep learning models used to identify, recognise, and count fruits, as well as the procedure for preparing datasets and evaluating deep learning model performance metrics. The best deep learning models for detecting and counting fruit in the wild will be identified, along with their advantages and disadvantages, the goals of the various dataset preparation processes, the most effective performance metric for evaluating the model, and the research gaps that need to be explored. It would be simpler to identify all relevant alternatives of the relevant search terms

to cover practically all associated information, and to gain a comprehensive enough reading to develop a grasp of the problem, from review articles rather than books.

In conclusion, after examining 33 publications published between 2017 and 2021, most studies used the faster RCNN and mask RCNN two-stage object detection methods. However, while comparing the performance of each deep learning model, it was shown that one-stage object detection, YOLO, outperformed two-stage object detection. In conclusion, the majority of publications advocate for data augmentation and resizing to combat overfitting and increase performance and decrease training time.

ACKNOWLEDGEMENT

The authors appreciate the chance to undertake this research, which involves technical and financial support, from the Centre for Electrical Engineering Studies Universiti Teknologi MARA (UiTM), Cawangan Pulau Pinang, Malaysia.

REFERENCES




- [1] C. Wiesener, S. Ruppiner, and T. Schauer, "Robust Discrimination of Flexion and Extension Phases for Mobile Functional Electrical Stimulation (FES) Induced Cycling in Paraplegics," *IFAC-PapersOnLine*, vol. 49, no. 32, pp. 210–215, 2016, doi: 10.1016/j.ifacol.2016.12.216.
- [2] P. S. Duth and K. Jayasimha, "Intra Class Vegetable Recognition System using Deep Learning," in *Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2020, pp. 602–606, doi: 978-1-7281-4876-2/20.
- [3] Y.-Y. Zheng, J.-L. Kong, X.-B. Jin, T.-L. Su, M.-J. Nie, and Y.-T. Bai, "Real-Time Vegetables Recognition System based on Deep Learning Network for Agricultural Robots," in *2018 Chinese Automation Congress (CAC)*, 2018, pp. 2223–2228, doi: 10.1109/CAC.2018.8623610.
- [4] L. Fu, J. Duan, X. Zou, J. Lin, L. Zhao, J. Li, and Z. Yang, "Fast and Accurate Detection of Banana Fruits in Complex Background Orchards," *IEEE Access*, vol. 8, pp. 196835–196846, 2020, doi: 10.1109/ACCESS.2020.3029215.
- [5] W. Zhang, K. Chen, J. Wang, Y. Shi, and W. Guo, "Easy domain adaptation method for filling the species gap in deep learning-based fruit detection," *Hortic. Res.*, vol. 8, no. 1, p. 119, 2021, doi: 10.1038/s41438-021-00553-8.
- [6] Venkatesh, N. Y., S. U. Hegde, and S. S., "Fine-tuned MobileNet Classifier for Classification of Strawberry and Cherry Fruit Types," in *2021 International Conference on Computer Communication and Informatics (ICCCI)*, 2021, pp. 1–8, doi: 10.1109/ICCCI50826.2021.9402444.
- [7] R. G. de Luna, E. P. Dadios, A. A. Bandala, and R. R. P. Vicerra, "Tomato Fruit Image Dataset for Deep Transfer Learning-based Defect Detection," in *2019 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM)*, Nov. 2019, pp. 356–361, doi: 10.1109/CIS-RAM47153.2019.9095778.
- [8] X. Li, Y. Qin, F. Wang, F. Guo, and J. T. W. Yeow, "Pitaya detection in orchards using the MobileNet-YOLO model," in *2020 39th Chinese Control Conference (CCC)*, 2020, pp. 6274–6278, doi: 10.23919/CCC50068.2020.9189186.
- [9] H. Basri, I. Syarif, and S. Sukaridhoto, "Faster R-CNN Implementation Method for Multi-Fruit Detection Using Tensorflow Platform," in *2018 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC)*, 2018, pp. 337–340, doi: 10.1109/KCIC.2018.8628566.
- [10] L. M. Azizah, S. F. Umayah, S. Riyadi, C. Damarjati, and N. A. Utama, "Deep learning implementation using convolutional neural network in mangosteen surface defect detection," in *2017 7th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, 2017, pp. 242–246, doi: 10.1109/ICCSCE.2017.8284412.
- [11] M. Halstead, S. Denman, C. Fookes, and C. McCool, "Fruit Detection in the Wild: The Impact of Varying Conditions and Cultivar," in *2020 Digital Image Computing: Techniques and Applications (DICTA)*, 2020, pp. 1–8, doi: 10.1109/DICTA51227.2020.9363407.
- [12] J. Zhao and J. Qu, "A Detection Method for Tomato Fruit Common Physiological Diseases Based on YOLOv2," in *2019 10th International Conference on Information Technology in Medicine and Education (ITME)*, 2019, pp. 559–563, doi: 10.1109/ITME.2019.00132.
- [13] L. Jian, Z. Mingrui, and G. Xifeng, "A fruit detection algorithm based on R-FCN in natural scene," in *2020 Chinese Control And Decision Conference (CCDC)*, 2020, pp. 487–492, doi: 10.1109/CCDC49329.2020.9163826.
- [14] H. Kang, H. Zhou, and C. Chen, "Visual Perception and Modeling for Autonomous Apple Harvesting," *IEEE Access*, vol. 8, pp. 62151–62163, 2020, doi: 10.1109/ACCESS.2020.2984556.
- [15] A. S. M. F. Al Haque, R. Hafiz, M. A. Hakim, and G. M. Rasiqul Islam, "A Computer Vision System for Guava Disease Detection and Recommend Curative Solution Using Deep Learning Approach," in *2019 22nd International Conference on Computer and Information Technology (ICIT)*, 2019, pp. 1–6, doi: 10.1109/ICIT48885.2019.9038598.
- [16] R. Yamparala, R. Challa, V. Kantharao, and P. S. R. Krishna, "Computerized Classification of Fruits using Convolution Neural Network," in *2020 7th International Conference on Smart Structures and Systems (ICSSS)*, 2020, pp. 1–4, doi: 10.1109/ICSSS49621.2020.9202305.
- [17] S. Chakraborty, F. M. J. M. Shamrat, M. M. Billah, M. Al Jubair, M. Alauddin, and R. Ranjan, "Implementation of Deep Learning Methods to Identify Rotten Fruits," in *2021 5th International Conference on Trends in Electronics and Informatics (ICOEI)*, 2021, pp. 1207–1212, doi: 10.1109/ICOEI51242.2021.9453004.
- [18] M. Faisal, M. Alsulaiman, M. Arafah, and M. A. Mekhtiche, "IHDS: Intelligent Harvesting Decision System for Date Fruit Based on Maturity Stage Using Deep Learning and Computer Vision," *IEEE Access*, vol. 8, pp. 167985–167997, 2020, doi: 10.1109/ACCESS.2020.3023894.
- [19] Q. Liang, J. Long, W. Zhu, Y. Wang, and W. Sun, "Apple recognition based on Convolutional Neural Network Framework," in *2018 13th World Congress on Intelligent Control and Automation (WCICA)*, 2018, pp. 1751–1756, doi: 10.1109/WCICA.2018.8630705.
- [20] H. Younis, M. Z. Khan, M. U. G. Khan, and H. Mukhtar, "Robust Optimization of MobileNet for Plant Disease Classification with Fine Tuned Parameters," in *2021 International Conference on Artificial Intelligence (ICAI)*, 2021, pp. 146–151, doi:

- 10.1109/ICAIS2203.2021.9445261.
- [21] M. Afonso, H. Fonteijn, F. S. Fiorentin, D. Lensink, M. Mooij, N. Faber, G. Polder, and R. Wehrens, "Tomato Fruit Detection and Counting in Greenhouses Using Deep Learning," *Front. Plant Sci.*, vol. 11, 2020, doi: 10.3389/fpls.2020.571299.
 - [22] J. F. V. Oraño, E. A. Maravillas, and C. J. G. Aliac, "Jackfruit Fruit Damage Classification using Convolutional Neural Network," in *2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, 2019, pp. 1–6, doi: 10.1109/HNICEM48295.2019.9073341.
 - [23] C. Hu, X. Liu, Z. Pan, and P. Li, "Automatic Detection of Single Ripe Tomato on Plant Combining Faster R-CNN and Intuitionistic Fuzzy Set," *IEEE Access*, vol. 7, pp. 154683–154696, 2019, doi: 10.1109/ACCESS.2019.2949343.
 - [24] N. Stasenkov, E. Chernova, D. Shadrin, G. Ovchinnikov, I. Krivolapov, and M. Pukalchik, "Deep Learning for improving the storage process: Accurate and automatic segmentation of spoiled areas on apples," in *2021 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, 2021, pp. 1–6, doi: 10.1109/I2MTC50364.2021.9460071.
 - [25] K. Alibabaci, P. D. Gaspar, T. M. Lima, R. M. Campos, I. Girão, J. Monteiro, and C. M. Lopes, "A Review of the Challenges of Using Deep Learning Algorithms to Support Decision-Making in Agricultural Activities," *Remote Sens.*, vol. 14, no. 3, 2022, doi: 10.3390/rs14030638.
 - [26] X. Liu, D. Zhao, W. Jia, W. Ji, C. Ruan, and Y. Sun, "Cucumber Fruits Detection in Greenhouses Based on Instance Segmentation," *IEEE Access*, vol. 7, pp. 139635–139642, 2019, doi: 10.1109/ACCESS.2019.2942144.
 - [27] V. Kukreja and P. Dhiman, "A Deep Neural Network based disease detection scheme for Citrus fruits," in *2020 International Conference on Smart Electronics and Communication (ICOSEC)*, 2020, pp. 97–101, doi: 10.1109/ICOSEC49089.2020.9215359.
 - [28] J. F. Bautista, C. D. Oceña, M. J. Cabrerós, and S. P. L. Alagao, "Automated Sorter and Grading of Tomatoes using Image Analysis and Deep Learning Techniques," in *2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, 2020, pp. 1–6, doi: 10.1109/HNICEM51456.2020.9400055.
 - [29] A. N. Fountsop, J. L. E. K. Fendji, and M. Atemkeng, "Deep Learning Models Compression for Agricultural Plants," *Appl. Sci.*, vol. 10, no. 19, 2020, doi: 10.3390/app10196866.
 - [30] N. A. K. Rosili, N. H. Zakaria, R. Hassan, S. Kasim, F. Z. C. Rose, and T. Sutikno, "A systematic literature review of machine learning methods in predicting court decisions," *IAES Int. J. Artif. Intell.*, vol. 10, p. 1091, 2021, doi: 10.11591/ijai.v10.i4.
 - [31] H. A. M. Shaffril, N. Ahmad, S. F. Samsuddin, A. A. Samah, and M. E. Hamdan, "Systematic literature review on adaptation towards climate change impacts among indigenous people in the Asia Pacific regions," *J. Clean. Prod.*, vol. 258, p. 120595, 2020, doi: <https://doi.org/10.1016/j.jclepro.2020.120595>.
 - [32] D. A. Cook, R. Brydges, B. Zendejas, S. J. Hamstra, and R. Hatala, "Mastery Learning for Health Professionals Using Technology-Enhanced Simulation: A Systematic Review and Meta-Analysis," *Acad. Med.*, vol. 88, 2013, doi: 10.1097/ACM.0b013e31829a365d.
 - [33] C. L. Edridge, A. J. Dunkley, D. H. Bodicoat, T. C. Rose, L. J. Gray, M. J. Davies, and K. Khunti, "Prevalence and Incidence of Hypoglycaemia in 532,542 People with Type 2 Diabetes on Oral Therapies and Insulin: A Systematic Review and Meta-Analysis of Population Based Studies," *PLoS One*, vol. 10, 2015, doi: 10.1371/journal.pone.0126427.
 - [34] N. R. Haddaway, B. Macura, P. Whaley, and A. S. Pullin, "ROSES RepOrting standards for Systematic Evidence Syntheses: pro forma, flow-diagram and descriptive summary of the plan and conduct of environmental systematic reviews and systematic maps," *Environ. Evid.*, vol. 7, no. 1, pp. 4–11, 2018, doi: 10.1186/s13750-018-0121-7.
 - [35] P. Wang, K. H. Low, and A. H. McGregor, "A subject-based motion generation model with adjustable walking pattern for a gait robotic trainer: NaTure-gaits," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2011, pp. 1743–1748, doi: 10.1109/IROS.2011.6095049.
 - [36] Z. Munn, C. Stern, E. Aromataris, C. Lockwood, and Z. Jordan, "What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences," *BMC Med. Res. Methodol.*, vol. 18, no. 1, pp. 1–9, 2018, doi: 10.1186/s12874-017-0468-4.
 - [37] A. Pollock and E. Berge, "How to do a systematic review," *Int. J. Stroke*, vol. 13, no. 2, pp. 138–156, 2018, doi: 10.1177/1747493017743796.
 - [38] E. O. Igbekele, A. A. Adebisi, F. A. Ibikunle, M. O. Adebisi, and O. O. Oludayo, "Research trends on CAPTCHA: A systematic literature," *Int. J. Electr. Comput. Eng.*, vol. 11, no. 5, pp. 4300–4312, 2021, doi: 10.11591/ijece.v11i5.
 - [39] E. Vilvert, L. Stridh, B. Andersson, A. Olson, L. Aldén, and A. Berlin, "Evidence based disease control methods in potato production: a systematic map protocol," *Environ. Evid.*, vol. 11, no. 1, pp. 1–8, 2022, doi: 10.1186/s13750-022-00259-x.
 - [40] I. Odun-Ayo, T.-A. Williams, M. Odusami, and J. Yahaya, "A systematic mapping study of performance analysis and modelling of cloud systems and applications," *Int. J. Electr. Comput. Eng.*, vol. 11, pp. 1839–1848, May 2021, doi: 10.11591/ijece.v11i2.pp1839-1848.
 - [41] E. Raheem, "Insight on Face Liveness Detection: A Systematic Literature Review," *Int. J. Electr. Comput. Eng.*, vol. 9, May 2019.
 - [42] S. K. C. Osmia, A. Ngadeninb, N. M. Nora, H. Husena, M. A. Yahyaa, and N. M. Dauda, "Bonding Strength of Bamboo Reinforcement in Concrete – A Systematic Literature," *J. Kejuruter. SI*, vol. 4, no. 1, pp. 1–11, 2021, doi: 10.17576/jkukm.
 - [43] S. Bargoti and J. Underwood, "Deep fruit detection in orchards," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, May 2017, pp. 3626–3633, doi: 10.1109/ICRA.2017.7989417.
 - [44] A. Koirala, K. B. Walsh, Z. Wang, and C. McCarthy, "Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of 'MangoYOLO,'" *Precis. Agric.*, vol. 20, no. 6, pp. 1107–1135, 2019, doi: 10.1007/s11119-019-09642-0.
 - [45] I. Y. Pratama, A. Wahab, and M. Alaydrus, "Deep Learning for Assessing Unhealthy Lettuce Hydroponic Using Convolutional Neural Network based on Faster R-CNN with Inception V2," in *2020 Fifth International Conference on Informatics and Computing (ICIC)*, Nov. 2020, pp. 1–6, doi: 10.1109/ICIC50835.2020.9288554.
 - [46] X. Ni, C. Li, H. Jiang, and F. Takeda, "Deep learning image segmentation and extraction of blueberry fruit traits associated with harvestability and yield," *Hortic. Res.*, vol. 7, no. 1, p. 110, 2020, doi: 10.1038/s41438-020-0323-3.
 - [47] C. Zhou, J. Hu, Z. Xu, Y. Jibo, H. Ye, and Y. Guijun, "A Novel Greenhouse-Based System for the Detection and Plumpness Assessment of Strawberry Using an Improved Deep Learning Technique," *Front. Plant Sci.*, vol. 11, 2020, doi: 10.3389/fpls.2020.00559.
 - [48] T. Santos, L. de Souza, A. dos Santos, and S. Avila, "Grape detection, segmentation and tracking using deep neural networks and three-dimensional association," *Comput. Electron. Agric.*, vol. 170, p. 105247, 2020, doi: 10.1016/j.compag.2020.105247.
 - [49] Y. Mu, T.-S. Chen, S. Ninomiya, and W. Guo, "Intact Detection of Highly Occluded Immature Tomatoes on Plants Using Deep Learning Techniques," *Sensors (Basel)*, vol. 20, no. 10, p. 2984, 2020, doi: 10.3390/s20102984.
 - [50] W. Yijing, Y. Yi, W. Xue-fen, C. Jian, and L. Xinyun, "Fig Fruit Recognition Method Based on YOLO v4 Deep Learning," in *2021 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, 2021, pp. 303–306, doi: 10.1109/ECTI-CON51831.2021.9454904.




- [51] L. Ghiani, A. Sassu, F. Palumbo, L. Mercenaro, and F. Gambella, "In-Field Automatic Detection of Grape Bunches under a Totally Uncontrolled Environment," *Sensors*, vol. 21, no. 11, 2021, doi: 10.3390/s21113908.
- [52] Y. Chen, W. S. Lee, H. Gan, N. Peres, C. Fraisse, Y. Zhang, and Y. He, "Strawberry Yield Prediction Based on a Deep Neural Network Using High-Resolution Aerial Orthoimages," *Remote Sens.*, vol. 11, no. 13, 2019, doi: 10.3390/rs11131584.
- [53] H. Alsolai and M. Roper, "A systematic literature review of machine learning techniques for software maintainability prediction," *Inf. Softw. Technol.*, vol. 119, p. 106214, 2020, doi: <https://doi.org/10.1016/j.infsof.2019.106214>.
- [54] A. Hamidinekoo, G. A. Garzo'n-Martínez, M. Ghahremani, F. M. Corke, R. Zwigelaar, J. H. Doonan, and C. Lu, "DeepPod: a convolutional neural network based quantification of fruit number in Arabidopsis," *Gigascience*, vol. 9, May 2020, doi: 10.1093/gigascience/giaa012.
- [55] A. Hiemann, T. Kautz, T. Zottmann, and M. Hlawitschka, "Enhancement of Speed and Accuracy Trade-Off for Sports Ball Detection in Videos—Finding Fast Moving, Small Objects in Real Time," *Sensors*, vol. 21, no. 9, 2021, doi: 10.3390/s21093214.
- [56] G. Devnani, A. Jaiswal, R. John, R. Chaurasia, and N. Tirpude, "Performance Evaluation of Fine-tuned Faster R-CNN on specific MS COCO Objects," *Int. J. Electr. Comput. Eng.*, vol. 9, p. 2548, May 2019, doi: 10.11591/ijece.v9i4.pp2548-2555.
- [57] A. Abdulmunem, Z. Abutiheen, and H. J. Aleqabie, "Recognition of Corona Virus Disease (COVID19) using Deep Learning Network," *Int. J. Electr. Comput. Eng.*, vol. 11, no. 1, pp. 365–374, 2021, doi: 10.11591/ijece.v11i1.
- [58] M. S. Diab, M. A. Elhosseini, M. S. El-Sayed, and H. A. Ali, "Brain Strategy Algorithm for Multiple Object Tracking Based on Merging Semantic Attributes and Appearance Features," *Sensors*, vol. 21, p. 7604, May 2021, doi: 10.3390/s21227604.
- [59] A. R. Luaibi, T. M. Salman, and A. H. Miry, "Detection of citrus leaf diseases using a deep learning technique," *Int. J. Electr. Comput. Eng.*, vol. 11, no. 2, pp. 1719–1727, 2021, doi: 10.11591/ijece.v11i2.
- [60] Y. Xu and R. Goodacre, "On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning," *J. Anal. Test.*, vol. 2, no. 3, pp. 249–262, 2018, doi: 10.1007/s41664-018-0068-2.
- [61] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Comput. Electron. Agric.*, vol. 145, pp. 311–318, 2018, doi: <https://doi.org/10.1016/j.compag.2018.01.009>.
- [62] B. Jabir and N. Fali, "Deep learning-based decision support system for weeds detection in wheat fields," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 1, pp. 816–825, May 2022, doi: 10.11591/ijece.v12i1.
- [63] M. G. Selvaraj, A. Vergara, H. Ruiz, N. Safari, S. Elayabalan, W. Ocimati, and G. Blomme, "AI-powered banana diseases and pest detection," *Plant Methods*, vol. 15, no. 1, p. 92, 2019, doi: 10.1186/s13007-019-0475-z.
- [64] K. Bresilla, G. D. Perulli, A. Boini, B. Morandi, L. Corelli Grappadelli, and L. Manfrini, "Single-Shot Convolution Neural Networks for Real-Time Fruit Detection Within the Tree," *Front. Plant Sci.*, vol. 10, 2019, doi: 10.3389/fpls.2019.00611.

BIOGRAPHIES OF AUTHORS






Ahmad Shukri Firdhaus Kamaruzaman    is a Master Student at the Faculty of Electrical Engineering, Universiti Teknologi MARA Cawangan Pulau Pinang (UiTM CPP). He obtained his Bachelor's Degree in Electrical & Electronics Engineering from Universiti Teknologi MARA Cawangan Pulau Pinang (UiTM CPP) in 2018. His research interests are in the fields of artificial intelligence and deep learning. He can be contacted at email: 2021459494@student.uitm.edu.my.






Adi Izhar Che Ani    is a Senior Lecturer at the Centre for Electrical Engineering, Universiti Teknologi MARA, Cawangan Pulau Pinang (UiTM CPP), with a Master's Degree in Engineering from Universiti Malaya Malaysia (2012). He obtained his Bachelor's Degree in Electrical & Electronics Engineering from the University of Miyazaki (Japan) in 2007. His research interests are the fields of artificial intelligence. He can be contacted at email: adiizhar@uitm.edu.my.






Mohammad Afiq Hamdani Mohammad Farid    was born in Kedah, Malaysia in 1999. He is a final-year student with a Bachelor's Degree in Electrical and Electronic Engineering from Universiti Teknologi MARA Cawangan Pulau Pinang (UiTM CPP). His research interests include deep learning and computer vision. He can be contacted at email: 2018297868@student.uitm.edu.my.






Siti Juliana Abu Bakar    received her B.Eng degree (Hons.) in electronic engineering from UTeM, Malaysia in 2019 and Master ESDE (Electronic System Design Engineering) from USM, Malaysia, in the year 2015. Completed her Ph.D. in the field of Automation & Control systems from Universiti Sains Malaysia, Engineering Campus in 2020. She is currently working as a senior lecturer at the Centre for Electrical Engineering Studies, Universiti Teknologi MARA Campus Pulau Pinang Malaysia. Prior to working as a senior lecturer at UiTM, she worked at Intel Product (M) Sdn.Bhd more than 5+ years as a Validation Engineer. She can be contacted at email: sitijuliana@uitm.edu.my.






Mohd Ikmal Fitri Maruzuki    received his B.Eng. in Computer Engineering from Ehime University (Japan) in 2004 and his Master's degree in Communication and Computer Engineering from Universiti Kebangsaan Malaysia (UKM) in 2010. He is currently employed as a Lecturer at the Faculty of Electrical Engineering in Universiti Teknologi MARA Cawangan Pulau Pinang (UiTMCPP) Malaysia. His research interests include on machine learning, image processing, and embedded systems. He can be contacted at email: ikmalf@uitm.edu.my.



Samsul Setumin    received B.Eng. degree (Hons.) in electronic engineering from the University of Surrey, in 2006, and M.Eng. degree in electrical-electronic and telecommunication from the Universiti Teknologi Malaysia in 2009. He obtained his Ph.D. degree from Universiti Sains Malaysia in 2019 in the imaging field. Since 2010, he has been a lecturer at Universiti Teknologi MARA, Malaysia. He was a Test Engineer with Agilent Technologies (M) Sdn. Bhd., and Intel Microelectronics (M) Sdn. Bhd., for one year. His research interests include computer vision, image processing, pattern recognition, and embedded system design. He can be contacted at email: samsuls@uitm.edu.my.



Mokh. Sholihul Hadi    received his B.S. degree in Electrical Engineering, from Brawijaya University Indonesia in 2004. He received his M.S and Ph.D. degrees in Electronics and Applied Physics from the Tokyo Institute of Technology, Japan, in 2010 and 2014. Since 2009, he has worked as Associate Professor at the Department of Electrical Engineering, Faculty of Engineering, State University of Malang Indonesia. His current research interest is embedded IoT systems, smart devices, robotics, semiconductor devices, and nano electronics. He can be contacted at email: mokh.sholihul.ft@um.ac.id.