Information Management and Business Review (ISSN 2220-3796)
Vol. 16, No. 4(S), pp. 72-82, 2024

Real-Time Vegetable Identification and Detection Implementing Yolo

*Mohd Rahmat Mohd Noordin, Mohamad Hafiz Khairuddin, Nurul Syuhaidah Binti Noor Asihin, Anis Amilah Shari
School of Computing Sciences, College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, Cawangan Melaka Kampus Jasin Melaka, Merlimau, Melaka, Malaysia
*mrahmat.noordin@uitm.edu.my, hafizk@uitm.edu.my, syuhaidah1401@gmail.com, anisamilah@uitm.edu.my
Corresponding Author: Mohd Rahmat Mohd Noordin

Abstract: Fruits and vegetables are essential sources of nutrition, crucial for preventing chronic diseases like diabetes, cancer, and cardiovascular issues. Maintaining a healthy diet, including the recommended servings of vegetables and fruits, is important for overall well-being. However, many individuals, particularly those unfamiliar with grocery shopping, struggle to distinguish between different types of vegetables due to their similar physical characteristics, leading to confusion and poor dietary choices. Therefore, this project aims to design and develop a real-time vegetable detection and identification system through a mobile application using deep learning and provide healthy recipes based on the identified vegetable. This project used the YOLOv5 object detection algorithm for vegetable detection, utilizing manually captured vegetable images from phones as the dataset. The collected data were organized into a unified folder. The Modified Waterfall model was adopted as the methodology, excluding the maintenance phase, encompassing requirement gathering, design, implementation, and testing. The testing phase demonstrated that the model met all the project's objectives and successfully identified and detected vegetables with a mAP of 95.7%. All functionality test cases and an accuracy test confirmed that the system effectively resolved the problem. The system was developed as a mobile application to enhance accessibility for target users. Future enhancements could include refining the detection model and incorporating a wider range of vegetables or including other ingredient datasets to broaden its scope.

Keywords: *Vegetable detection, Vegetable identification, YOLOv5, Mobile application*

1. Introduction

Fruits and vegetables are a constant and main source of nutrition for humans, animals and all other living creatures (Shri Ramswaroop Memorial University, 2019). It is also beneficial in health, especially for early prevention of several chronic diseases such as diabetes, certain cancers, and cardiovascular diseases. Having a healthy body has become an aspiration for many individuals, whether it entails maintaining good physical health or ensuring stable mental well-being. One of the ways to have a healthy body is people need to keep track of their daily dietary habits. Based on the knowledge that people have learned about the pyramid of food, fruit and vegetables are in the second level. What it means by level two is that people need to eat plenty of it, for example, three servings of vegetables and two servings of fruits in a day.

In today's society, people are increasingly recognizing the importance of consuming vegetables in daily life especially among groups of people either younger or older generations. It is a good thing when people are conscious of the importance of eating vegetables, but they also need to know there are various types of vegetables. Vegetables can be classified into biological groups including leafy green, cruciferous, marrow, root, edible plant stem, and allium. Nevertheless, some individuals still have difficulty distinguishing between certain vegetables because of their similarities in terms of size, shape and color.

An article from Malay Mail, Zi (2020) Reported that TESCO Malaysia has released a useful handbook aimed at husbands who are new to family shopping during the MCO. BBC News posted an article by BBC Monitoring, titled "Coronavirus: Malaysian men in shopping muddle amid lockdown.", in the article the journalist wrote that most husbands were having a hard time having to distinguish between bewildering varieties of vegetables, spices and herbs. The journalist also highlighted a tweet from a Twitter user who expressed he is feeling dizzy trying to differentiate which one is spinach and bok choy. The similarity in colors and shapes among various vegetables frequently leads to confusion among individuals when trying to differentiate between them, particularly for those who rarely cook and are not familiar with grocery shopping. This confusion is caused by the overlapping appearances of certain vegetables, making it difficult to discern one from another solely based on visual cues. As a result, individuals may struggle to accurately identify vegetables, leading to uncertainty and potential errors in grocery shopping and meal preparation.

According to Morol et al. (2022), the issue some people are encountering is that they sometimes choose ingredients without being familiar with their names or properties. Additionally, some people also may acquire items from the grocery store without knowing how to use them in recipes. When people do not have enough knowledge about the ingredients they buy, it might result in wasted food and poor dietary choices especially if the ingredients are vegetables.

Deep learning techniques play a crucial role in vegetable identification and detection. This may not only help people when buying vegetables but also for production, such as automating processes in agriculture. While these techniques have indeed offered a better result when identifying and detecting a vegetable, it is also important to acknowledge that challenges still exist when using deep learning techniques. Based on Shri Ramswaroop Memorial University et al. (2019), the author wrote in the journal that vegetable detection and classification are a challenging objective in daily production and use, and the complexity increases when other parameters such as shape, size, and color are taken into consideration. As the complexity increases, it may lead to inaccurate results. Given that this challenge persists in the realm of deep learning, there is a need to develop a better solution by employing a more promising algorithm such as YOLO.

The goal of the project is to develop a mobile application that will help the targeted user identify the vegetable in real time. This project is mainly focusing on leafy green vegetables. By using this mobile application, users will have a reliable tool for quickly and accurately identifying leafy green vegetables, it can also enhance their overall grocery shopping experience. The importance of the project lies in the fact that the system can offer to help in identifying vegetables based on real-time and cease the confusion about vegetables. In terms of the usefulness of the proposed system, the system offers real-time vegetable detection and healthy recipes for people who are having a hard time deciding what kind of healthy dish to cook while going on a groceries shopping. Other than the importance and usefulness of the proposed system, it also is a benefit because it may lead to improved health outcomes, and make meal planning and preparation more convenient.

2. Literature Review

Vegetables play an important role in promoting overall health and well-being. For some people that want to lose weight, they need to know the right way of healthy eating plate. Vegetables are a variety of foods that provide nourishment of essential vitamins and minerals to the body. The various categories of vegetables have been mentioned in the previous chapter. The variety of colors from different arrays of vegetables is evidence of the numerous phytochemicals present in vegetables. These phytochemicals have been shown to have a positive impact on human health, including reducing the risk of chronic diseases such as cancer, heart disease, and diabetes (Ülger et al., 2018).

Vegetables are edible when raw or cooked and serve an important role in human nutrition because it has low fat and carbohydrates but is high in vitamins, minerals, and dietary fiber. Nutrition experts advised people to consume more fruit and vegetables and also recommended five or more portions a day (Ebabhi & Adebayo, 2022). Vegetables offer plenty of health benefits and a rich array of essential nutrients that contribute to overall well-being. The nutritional value of vegetables varies across different types, but collectively, they provide essential nutrients such as vitamin A, vitamin C, potassium, folate, and dietary fiber. These nutrients are crucial for maintaining a healthy immune system, promoting proper digestion, and supporting cardiovascular health.

Even though everyone knows that vegetables are important to their body, the challenges and complexities surrounding their proper identification of vegetables are still lacking. The challenges faced in 2020 due to the COVID-19 pandemic prompted the Malaysian government to implement a Movement Control Order (MCO), restricting grocery shopping to only the head of the household. TESCO Malaysia responded with a handbook for husbands navigating this new responsibility. However, reports from Malay Mail and BBC News highlighted the difficulties men faced in distinguishing vegetables during shopping, leading to humorous anecdotes and potential food waste. The issue extends beyond the pandemic, as some people often lack knowledge about the ingredients they purchase, such as when they do not know how to make recipes using those ingredients (Morol et al., 2022). This problem might result in wasted food and suboptimal dietary choices.

There are various types of vegetables available in this world. As mentioned in the previous chapter, vegetables can be categorized into five different types, which are leafy green, cruciferous, marrow, root, edible plant stem, and allium. This project will be focused on leafy green vegetables that are available in Malaysia. The top five leafy green vegetables commonly recognized by people are Kangkung, Sawi, Bayam, Kailan, and Bok Choy. According to Kumar et al. (2020), Green leafy vegetables occupy an important place among food crops as these provide adequate amounts of vitamins and minerals for humans. They are rich sources of vitamins like beta-carotene, ascorbic acid, riboflavin, folic acid and minerals like calcium, iron, phosphorous, etc. Table 1 lists some of the leafy green vegetables available in Malaysia.

Table 1: Green Leafy Vegetable

Name	Image	Nutritional Value
1. Water Spinach (Kangkung)		• Vitamin A • Vitamin C • Vitamin K • Calcium • Iron
2. Spinach (Bayam)		• Vitamin A • Vitamin C • Calcium • Iron • Folic acid • Protein • Fiber
3. Chinese Broccoli (Kailan)		• Vitamin A • Vitamin C • Vitamin K • Calcium • Folate

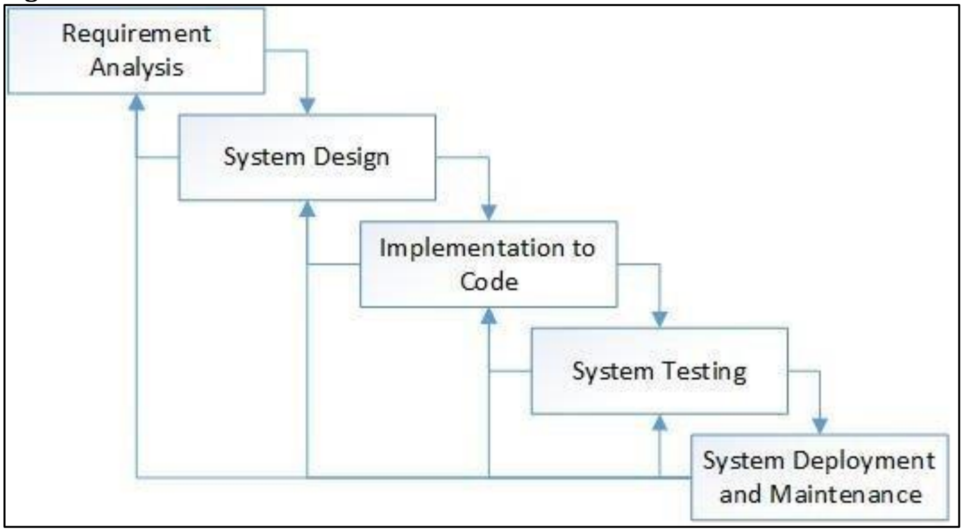
Chinese Mustard Greens (Sawi)		• Vitamin A • Vitamin C • Calcium • Iron • Fiber • Folic acid
. Chinese Cabbage (Bok Choy)		• Vitamin A • Vitamin C • Vitamin K • Iron • Calcium • Potassium

You Only Look Once (YOLO) is a viral and widely used algorithm, and it is famous for its object detection characteristics. (Jiang et al., 2021). YOLO is a one-stage object detector that uses a specific CNN network and bounding boxes to predict and locate a specific object in each region of feature maps (Sumit et al., 2020). Its ability to process in real-time makes it ideal for applications that need immediate analysis, including security surveillance and autonomous vehicle technology. Additionally, YOLO's integrated method simplifies the classification and localization process by removing the necessity for distinct stages, thereby boosting efficiency. A series of YOLO versions have been released since the first version of YOLO was released and developed in 2016 through a research paper titled "You Only Look Once: Unified, Real-Time Object Detection" by Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. Joseph Redmon and Ali Farhadi only introduced the version of YOLO until YOLOv3. YOLOv4 was introduced by Alexey Bochkovskiy in 2020 and the version has been improvised to address the limitation of the previous version. As of 2023, YOLO has progressed to its seventh version, referred to as YOLOv7. YOLO breaks through the max-speed limit of CNN and realizes an excellent balance of speed and accuracy. YOLOv2 outperforms state-of-the-art methods like Faster R-CNN with ResNet in terms of 76.8 mAP at 67 FPS and 78.6 mAP at 40 FPS. (Du, 2018).

3. Methodology

This project will use the Modified Waterfall Model as the project methodology that will help to provide a structured framework that facilitates the efficient implementation of project tasks. The Modified Waterfall Model is an iterative and phase project management methodology that builds upon the traditional Waterfall Model by introducing elements of flexibility and adaptability. The traditional waterfall model is a sequential approach to project methodology, where a project flows through a series of steps or phases. In contrast, the Modified Waterfall Model allows the phases to overlap and allows changes to be made in the design phase (Nugroho & Izzah, 2017). The phases in the Modified Waterfall Model are similar to the traditional design which includes, requirements gathering, design, implementation, testing, and maintenance. Figure 1 shows the phases of the Modified Waterfall Model.

Figure1: Modified Waterfall Model



(Source: Nugroho & Izzah, 2017)

The main change is that the phases in this advanced model are permitted to overlap. Since the phases are allowed to overlap, a lot of flexibility has been introduced. Several tasks can function concurrently, which ensures that the defects or errors are removed in the development stage which can save the any overhead cost of making changes before implementation. In case there are any errors introduced because of the changes made, rectifying them is also easy. This helps to reduce any oversight issues. Table 2 shows the summary of each phase in the Modified Waterfall Model. In this project, the maintenance phase is not implemented because the mobile application being developed is a proof-of-concept prototype and does not require maintenance.

Table 2: Summary of Modified Waterfall Model

Phase	Activities	Deliverable
Requirement Analysis	Define information about vegetables. Determine the problem statements and objectives Compile the evidence of previous related works Design the appropriate use case diagram for the system Design the expected user interfaces of the mobile application	Background of study Problem statement Objectives Project scope Significance Use Case diagram User Interface
Design	Design the appropriate flowchart of the system Select a suitable technique to execute in the project Plan the project timeline using the Gantt chart	System Flowchart System Algorithm
Implementation	Define the hardware and software requirement Build the mobile application	Implementation of the vegetable identification mobile application
Testing	Execute system testing Review all the functionality to ensure the system works as intended	Test case

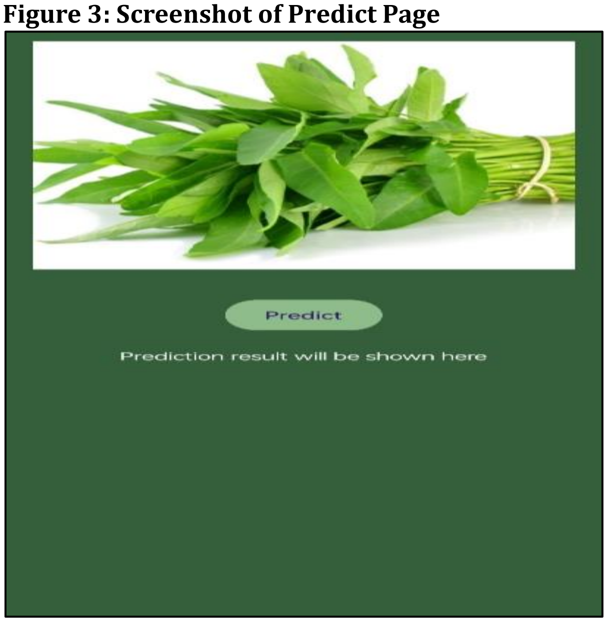
4. Results and Discussion

For the implementation of the system, Android Studio was used. Android Studio is a specialized integrated development environment (IDE) for Android that offers coding, testing, and debugging tools, making it crucial for creating and launching Android systems. Below in Figure 2 is a screenshot of the Home Page. When users first open the application, this is the first page they will see, designed to provide immediate and easy access to the primary functions of the app. The layout is user-friendly, with three prominently displayed buttons: How To Use, Take Photo, and Choose Photo.

Figure 2: Screenshot of Home Page

Once input is received from the user, it will be shown on this page by the system. When the user selects the 'Predict' button, the system will examine the input and forecast the type of vegetable. Figure 3 shows the screenshot of the Predict Page.

Figure 3: Screenshot of Predict Page



When the user clicks on the 'Predict' button, the prediction result of the input will be displayed. If the input is valid and the model correctly predicts the vegetable, a 'Recipe' button will become visible. This prediction is made using the TensorFlow Lite file. The TensorFlow Lite file, named 'best-fp16.tflite', along with 'customlabels.txt', is loaded in the assets directory. Figure 4 shows a screenshot of the Result Page, while Figure 4 shows a code snippet of the model and label being loaded in the code.

Figure Error! No text of specified style in document.**: Screenshot of Result Page**

The recipes shown are determined by the predicted vegetable. For instance, if the predicted outcome is Water Spinach, a compilation of recipes with Water Spinach as an ingredient will be shown. All recipes are sourced from the internet and compiled into a JSON file the code uses to display the relevant recipes. Figure 5 shows the screenshot of the Recipe Page.

Figure 1: **Screenshot of Recipe Page**

Table 3 presents the functionality testing of the system, indicating whether the expected status is a failure or success based on the outputs produced.

Table 3: Results of Functionality Testing

Test Case	Expected Output	Pass/Fail
View Home Page	When the user clicks on the mobile application's icon it will immediately show the home page. On the home page, there are two buttons. Users can choose whether to take a photo which is a real-time image or upload a photo such as a still image from the gallery.	Pass
View Detection Page	The vegetable will be detected on this page.	Pass
View Result Page	After the user clicks on the identify button on the detection page, it will go to this page to show the result of the vegetable detected.	Pass
View Recipe Page	This page will show up when the user decides to scroll down to see the suggested healthy recipe.	Pass

Testing for accuracy in deep learning is an integral part of the model development process, as it conveys how well the model would generalize to new, unseen data. Normally, one would split up the data into a training and a validation set. While training shall be on the former, checking performance should be on the latter. This is where a validation set comes in handy: estimating how well it is going to perform on real-world data by assessing the accuracy.

Figure 6: Result of YOLOv5 model training

```
Validating runs/train/yolov5s_results/weights/best.pt...
Fusing layers...
custom_YOLOv5s summary: 182 layers, 7257306 parameters, 0 gradients
                 Class     Images  Instances          P          R      mAP50   mAP50-95: 100% 12/12 [00:08<00:00,  1.36it/s]
                   all        376        375      0.969      0.948      0.957      0.904
                 Bayam        376         83      0.957      0.813      0.848      0.792
              Bok-Choy        376         64      0.995          1      0.995      0.973
                Kailan        376         67      0.949       0.97      0.962      0.916
              Kangkung        376         92      0.978      0.973      0.986      0.855
                  Sawi        376         69      0.967      0.986      0.993      0.983
Results saved to runs/train/yolov5s_results
CPU times: user 49.2 s, sys: 5.85 s, total: 55.1 s
Wall time: 1h 20min 6s
```

Based on the above Figure 6, shows the result of training a YOLOv5 model for 100 epochs, providing a detailed summary of the training and validation process. Results include detailed class-wise performance metrics that are valuable to understanding how well the model works on each category in the dataset. Key metrics presented are Precision, Recall, mean average precision at an IoU threshold of 0.50, and mean average precision averaged over IoU thresholds from 0.50 to 0.95 (mAP50-95). These metrics are critical in evaluating the accuracy and robustness of the model.

Results showed good performance of the YOLOv5 model in correctly detecting and classifying the different kinds of vegetables in the dataset. The results, along with the weights from this training, were saved in the runs/train/exp directory and could be used for further analysis and execution in live applications.

The most important metric for assessing the performance of object detection models like YOLO is mean average precision since it provides a very fine-grained measure of how well a model can identify and classify different objects within images by combining precision and recall: precision explains how many of the predicted positive cases are positive. In contrast, recall defines the capability of identifying all the positive cases. The mean average precision (mAP) integrates both these factors to give a comprehensive appraisal of the model's performance.

To calculate the map, start by gathering predicted bounding boxes with confidence scores and ground truth boxes for each class. Sort the predictions in descending order by confidence, then determine true positives (TP) and false positives (FP) using Intersection over Union (IoU), considering predictions with IoU above a threshold of 0.5 as TP. The formula for calculating precision and recall is shown in Equation 1 and Equation 2.

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

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

After getting the precision and the recall values at various confidence thresholds, the next step is to calculate the average precision (AP) for each class. The formula is based on the Equation 5.3.

$$AP = \sum_n (Recall_n - Recall_{n-1}). Precision_n \quad (3)$$

Finally, calculate mAP by averaging the AP values across all classes using Equation 5.4, where N is the number of classes, and AP_i Is the average precision for class? i.

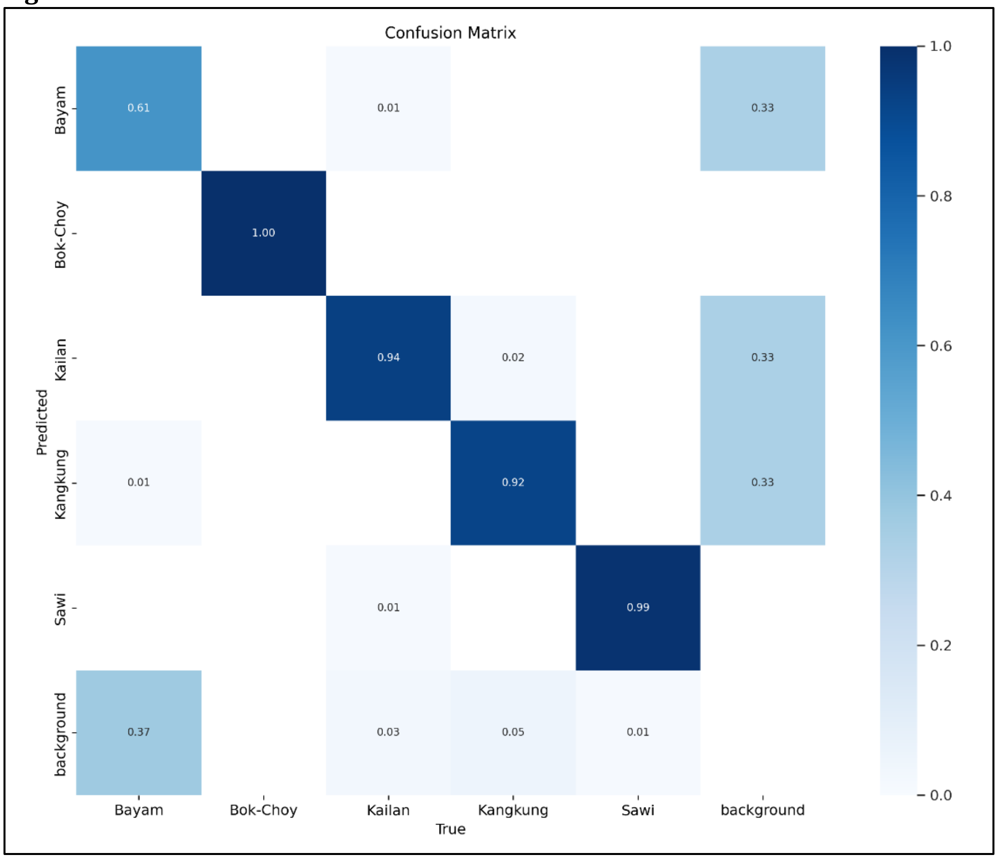
$$mAP = \frac{1}{N} \sum_{i=1}^{N} AP_i \quad (4)$$

Table 4 shows the result of the accuracy testing of the model. The following table shows some key performance metrics describing the performance of the YOLOv5 classification model on a given dataset, showing very strong performance in almost all measures. Given the precision is 96.9%, this means that the positive predictions of the model are highly correct; with a recall of 94.8%, it is also good at identifying nearly all the actual positive instances. A balanced measure between the two measures mentioned above is the F1-Score, which is at 95.8%, thus confirming the model for reliability and robustness in practical applications. Additionally, the mAP of 95.7% is indicative of very high accuracy, both across classes and recall levels, hence further verifying overall performance. These results collectively suggest that the model is well-optimized and effective for the given dataset, making it suitable for tasks requiring precise and comprehensive classification.

Table 4: The result of the accuracy testing

Metric	Result
Precision	96.9%
Recall	94.8%
F1-Score	95.8%
Mean Average Precision	95.7%

Figure 7: Confusion matrix



According to Figure 7, the confusion matrix represents the performance of the project YOLOv5 classification model implemented on a custom dataset containing six classes: Bayam, Bok-Choy, Kailan, Kangkung, Sawi, and Background. Each cell in the matrix refers to the proportion of correct actual class instances (True) predicted as a specific class (Predicted). Cells in the main diagonal indicate correct classifications, whereas the cells outside of the main diagonal represent misclassifications.

The model shows perfect classification for Bok-Choy and near-perfect classification for Sawi. It also performs well for Kangkung and Kailan, with high accuracy rates. However, there are notable misclassifications for Bayam, with 33% being misclassified as Background and 1% as Bok-Choy. Similarly, 37% of Background instances were misclassified as Bayam. Other minor misclassifications include some confusion between Kailan and Kangkung. These results suggest that while the model is effective for certain classes, improvements are needed for Bayam and Background, possibly through increased training data diversity and enhanced feature extraction. Overall, the confusion matrix highlights the model's strengths and weaknesses, guiding future improvements to improve its reliability.

5. Conclusion and Recommendations

The developed mobile application utilizing the YOLO algorithm effectively identifies and recognizes vegetable types in real-time input with high performance. It accurately displays the five vegetable types: Bayam, Bok Choy, Kailan, Kangkung, and Sawi simultaneously, achieving a high mean average precision (mAP) of 95.7%. This application has met the three objectives outlined in Chapter 1 by leveraging the YOLO algorithm. The dataset used for this project was obtained by taking a picture of the vegetables using a smartphone camera. The development of the YOLOv5 model was conducted on Google Colab while the mobile application was developed by using Android Studio.

Recommendations

Based on the functionality and limitations of the mobile application, several ideas and potential enhancements can be proposed to expand the utility of the project in the future. The recommendations for further enhancements include:

- Collect more vegetable images to increase the dataset size for the training process to improve the performance of the model.
- Expand the capabilities of the mobile application to recognize a wider variety of vegetables beyond the current five types.
- Explore and compare different model architectures to potentially enhance performance and gain insights into the most effective models for the dataset.

Acknowledgment: The authors express their gratitude and thanks to UiTM Cawangan Melaka for the assistance and support of this article.

References

Ebabhi, A., & Adebayo, R. (2022). Nutritional Values of Vegetables. In E. Yildirim & M. Ekinci (Eds.), Vegetable Crops. IntechOpen. https://doi.org/10.5772/intechopen.101090

Jiang, P., Ergu, D., Liu, F., Cai, Y., & Ma, B. (2021). A Review of Yolo Algorithm Developments. *Procedia Computer Science*, *199*, 1066–1073. https://doi.org/10.1016/j.procs.2022.01.135

Kumar, D., Kumar, S., Gochar, C. S., Rampur, M., Saharanpur, M., Pradesh, U., Corresponding, I., & Shekhar, C. (2020). Nutritional components in green leafy vegetables: A review. J Pharmacogn Phytochem 2020;9(5), 2498-2502.

Morol, Md. K., Rokon, Md. S. J., Hasan, I. B., Saif, A. M., Khan, R. H., & Das, S. S. (2022). Food Recipe Recommendation Based on Ingredients Detection Using Deep Learning. Proceedings of the 2nd International Conference on Computing Advancements, 191–198. https://doi.org/10.1145/3542954.3542983

Nugroho, B. A., & Izzah, A. (2017). Implementation of Google's Technology in Android Mobile App "Kediri City Fasum and Fasos Information System". *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, *3*(1), 45-56. https://doi.org/10.22219/kinetik.v3i1.543

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2015). *You Only Look Once: Unified, Real-Time Object Detection*. http://arxiv.org/abs/1506.02640

Shri Ramswaroop Memorial University, Institute of Electrical and Electronics Engineers, & Institute of Electrical and Electronics Engineers. Uttar Pradesh Section. (2019). 2019 International Conference on Cutting-edge Technologies in Engineering (ICon-CuTE) : Shri Ramswaroop Memorial University, Barabanki, India, Nov. 14-16, 2019.

Sumit, S. S., Watada, J., Roy, A., & Rambli, D. R. A. (2020). In object detection deep learning methods, YOLO shows supremum to Mask R-CNN. *Journal of Physics: Conference Series*, *1529*(4). https://doi.org/10.1088/1742-6596/1529/4/042086

Ülger, T. G., Songur, A. N., Çırak, O., & Pınar Çakıroğlu, F. (2018). Role of Vegetables in Human Nutrition and Disease Prevention. In Md. Asaduzzaman & T. Asao (Eds.), Vegetables. IntechOpen.

Zi, T. M. (2020). Covid-19: Tesco Malaysia 'rescues' clueless husbands with an illustrated guide to grocery shopping. https://ww.Malaymail.Com/News/Life/2020/03/25/Covid-19-Tesco-Malaysia-Rescues-Clueless-Husbands-with-Illustrated-Guide-to/1850117.