Recipe Recommendation based on Food Ingredients Recognition using Deep Learning

Muhamad Ridhwan Mohamad Razali [1*], Hana Fakhira Almarzuki [2], Rabiahtul Adawiyah Bahar [3] and Nur Atiqah Sia Abdullah [4]

[1,2,3,4] School of Computing Sciences, College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, 40450 Shah Alam, Selangor Malaysia
[1*]mridwan84@uitm.edu.my, [2]hanafakhira@uitm.edu.my, [3]biaadawiyah.ra@gmail.com, [4]atiqah684@uitm.edu.my

Abstract. Every individual requires food for nutritional support, without which the body suffers from various diseases such as gastric issues and low blood sugar. While there are many recipe websites, finding a suitable recipe can be time-consuming, especially for those who lack cooking knowledge with available ingredients. This project addresses this challenge by developing a recipe recommendation model using a deep learning algorithm to suggest recipes based on recognized food ingredients. The model utilizes more than 2,000 images categorized into three food ingredient groups: rice, apple, and chicken. The data collection process includes thorough preprocessing, data cleaning, and annotation. The dataset is then augmented and split into training, validation, and testing sets in an 80-10-10 ratio. To build the food ingredient recognition model, the You Only Look Once (YOLO) technique, specifically YOLOv8, is employed. The system is designed as a web-based application, providing an accessible interface for users. After extensive training, validation, and testing, the model achieved an impressive accuracy of 98%, demonstrating its capability to accurately detect the specified ingredients. Future research aims to enhance the system's functionality by addressing current limitations. Planned improvements include integrating a comprehensive database for storing recipes, which would facilitate easier data insertion and the addition of nutritional information for each recipe. Furthermore, incorporating live-camera integration could enhance user engagement and practicality.

Keywords: Recipe Recommendation, YOLO, Food, Deep Learning, Image Recognition.

1 Introduction

According to a group of researchers, food is essential for survival and supplies important nutrition for the human body [1]. A recipe is a set of instructions for preparing a particular dish or meal. It typically includes a list of ingredients, along with their measurements and any specific instructions for their preparation or handling. The

N. A. S. Abdullah et al. (eds.), *Proceedings of the International Conference on Innovation & Entrepreneurship in Computing, Engineering & Science Education (InvENT 2024)*, Advances in Computer Science Research 117,
https://doi.org/10.2991/978-94-6463-589-8_52

recipe will also describe the steps involved in cooking the dish, including the specific cooking techniques, cooking times and temperatures, and any additional steps, such as seasoning or garnishing. A standard recipe is divided into three sections: title, ingredients, and cooking instructions, which may or may not correspond to the visual appearance of the cooked dish [2]. Recipes can be found in various formats, such as cookbooks, online articles, or even handwritten notes passed down through generations. The recipes guide anyone who wants to prepare a particular dish, and following the recipe can help ensure a successful outcome. People don't have time to stock up on their cooking ingredients. This situation can be difficult for them if they want to prepare their meals quickly because they only have limited ingredients. This can be proved by a journal that says many people often cook a dish with a cooking recipe on websites and magazines, but the recipes' listed ingredients sometimes cannot be prepared [3]. Due to this challenge, this paper aims to develop a mobile application that can recommend a recipe based on the ingredients.

The full utilization of advanced computer vision and machine learning approaches in the realm of food perception has not been fully addressed [4]. Hence, this paper makes a comparison of several deep-learning techniques to be implemented. Deep learning enhances computer predictive capabilities by leveraging big data and advanced algorithms, integrating machine learning with complex applications like image recognition, and autonomously training itself without human intervention [5][6]. It is essential to choose the most suitable technique to ensure the effectiveness of this project. Three approaches will be discussed for this study which are CNN-based ResNet50, Support Vector Machine (SVM), and You Only Look Once (YOLO). These are the most extensively used algorithms in food recognition, and one of them might be suitable for the project.

Recommender systems are tools that filter, and narrow information based on the material's content or a user's preferences or needs [7]. The application of recommendation algorithms allows quick decision-making in these complex data environments [8]. According to [9], recommendation systems, also known as recommendation engines, are widely used in the fields of artificial intelligence and machine learning. They also stated that recommendation systems based on user preferences are currently helping consumers make the best decision without tiring their cognitive resources [9].

2 Related Works

This section consists of related works in terms of deep learning approach, food recommendations system and web-based application.

2.1 Deep Learning Approach

Considering the three deep learning approaches, each approach offers advantages that could suit this project. However, it is crucial to thoroughly compare and select the most appropriate deep learning algorithms, as this choice will significantly impact the outcome of the project.

Table 1. Comparison between Deep Learning Approaches.

Deep Learning Approach	Purpose	Dataset	Result
Convolutional Neural Network (CNN) -based ResNet50 (Xiao et al., 2021)	The goal is to develop a CNN model for detecting straw-berry diseases through an image recognition method.	Two datasets detect straw-berry diseases in Taiwan's commercial production, identifying five major ones: crown rot, leaf blight, fruit rot, powdery mil-dew, and gray mold.	The ResNet50 model, trained for 20 epochs on 1306 feature imag-es, achieved 100% classification accuracy for leaf blight, 98% for gray mold, and 98% for pow-dery mildew cases.
Support Vector Machine (SVM) (Mia et al., 2019)	To develop a machine vision-based recognition of rare local fruits of Bangladesh that might benefit people in recognizing them.	The study includes 480 images of six rare local fruits: amla (amloki), sugar-apple (ata), bi-lombo, elephant apple (chalta), orboroi, and sapota (sopheda).	Earned a classification accuracy of 94.79%.
You Only Look Once (YOLO) (Shifat et al., 2021)	To design a system that will encourage individuals to avoid eating junk food to take good care of their health.	A new dataset of 10,000 data points from 20 junk food classifications was generated using the Google search engine.	98.05% accuracy rate.

The You Only Look Once (YOLO) technique is the best choice for developing a Recipe Recommendation based on food ingredients recognition due to its ability to train from start to finish, high accuracy rate of 98.05%, and faster detection speed of three days with image augmentation.

2.2 Food Recommendation System

Food recipe recommendation based on ingredients detection using deep learning was carried out by [10]. This study focuses on applying machine learning and deep learning techniques to the challenges of recognizing food elements and recommending recipes based on them. The researchers aim to detect food ingredients by achieving their highest classification rate and recommending worthy recipes. They used Kaggle's food ingredients dataset and a customized dataset with 32 categories. There are 9856 photos of 32 different ingredients in total. The images were divided into three sets: 70% were allocated for training, 20% for testing, and the remaining 10% were assigned for validation. After running 20 epochs using transfer learning, the proposed CNN-based ResNet50 model achieved a 99.71% accuracy for the training dataset. Moreover, 92.6% accuracy on the validation dataset, while the testing data achieved an accuracy of 94%. They also tried detection with VGG16 and MobileNetV2, but ResNet50 produced a 94% testing accuracy. They created a 2D matrix for the recipe suggestion algorithm after selecting 19 cooking recipes against 32 food elements. They also performed a linear search in the database to cross-check to match the recognized recognition with the recipes easily. However, the CNN model in this system the model is unable to categorize multiple objects simultaneously. It leads to the lack of ability to detect multiple food ingredients from an uploaded image.

2.3 Web-based Applications

This section focuses on web-based applications. The word web-based means the users can access the application by opening their web browser. They do not need to install other applications on their device. Based on [11], nowadays, the emphasis on proper and efficient design has become more crucial than the successful development or support of any web system. This change is primarily caused by the widespread adoption of web applications and the collaborative nature of sharing such applications globally [12].

According to [12]. rich web applications can be developed using modern HTML5 standards and JavaScript, enhancing functionality and user engagement. The web has become a full-fledged software delivery platform, replacing traditional methods like downloading and installing software. HTML5 and JavaScript can create web-based applications' structure and content, while image recognition algorithms can analyse ingredients and recommend recipes.

Table 2. Table mapping for Similar Systems.

Systems	Number
Food recipe recommendation based on ingredients detection using deep learning. [10]	1
Real-time mobile recipe recommendation system using food ingredient recognition. [13]	2
Design of Low-Cost Object Identification Module for Culinary Applications. [14]	3
Vision Based Intelligent Recipe Recommendation System. [15]	4

Table 3. Comparison between Similar Systems.

Systems	1	2	3	4
Deep Learning Approaches	CNN-based Resnet50 to recognize food ingredients and a linear search algorithm in the database to recommend cooking recipes.	For recognition, they built bag-of features by extracting SURF and color histogram from multiple images and used linear kernel SVM at	You Only Look Once (YOLO) v2	You Only Look Once (YOLO) v7

		one-vs-all as a classifier.		
Accuracy Results	Achieved 99.71% accuracy for the training dataset. And 92.6% accuracy on the validation dataset. Achieved a testing accuracy of 94%.	Achieved the 83.93% recognition rate within the top six candidates	93.77% accuracy result	At first, achieved classification accuracy of more than 81.6%. By combining more vegetable classes, it provided a wide range of ingredients that could be found in a recipe and improved the model's accuracy to an excellent 88.2% on average.
Allow users to use a camera for recognition	No	Yes	No	Yes

3 Methodology

This paper develops a web-based application using waterfall model. The waterfall model is a software development process that follows a sequential and systematic approach. It is named the "waterfall" model due to its similarity to a cascading waterfall, where progress moves steadily downwards, without any backward movement. Each phase relies on the completion of the previous step. Below is the figure of the waterfall model process:

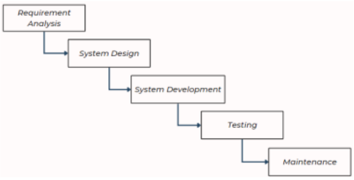

Fig. 1. Waterfall Model of Software Development Life Cycle.

The figure above shows that the model starts with requirement analysis and continues with system design, development, testing and maintenance phases.

3.1 Requirements Analysis

The datasets have been collected from Kaggle, like Fruit360. Other than that, FoodCamMobi also provides a dataset named UECFOOD-256. There are more than 2,000 images that can be gained. For Fruit360, this project used apples, and for UECFOOD-256, this project used rice. Meanwhile, chicken image datasets will be collected from Roboflow. Besides, the food recipes will be gathered from the recipe websites Yummly, Allrecipes, and any other resources related to the ingredients.

3.2 Design

The figure below shows a system architecture that shows an overview of the implementation of the YOLO approach.

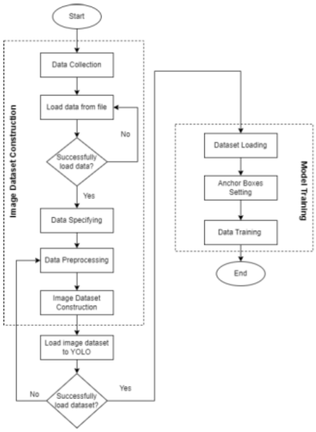

Fig. 2. System Architecture

The dataset construction starts with data collection, and the data will be loaded into one folder. After that, the data needs to be specified to the respective category, which has been mentioned in the Data Collection phase. Then, data preprocessing will resize the images to construct the dataset. Next, the dataset must load into YOLO, and if the process is successful, the model training section can start. However, if it is not successfully loaded, it continues to do the data preprocessing phase. In the model training section, after the dataset has been loaded, YOLO will do the anchor box setting before data training and testing.

3.3 Development

The You Only Look Once (YOLO) algorithm has been used. Besides, this project also uses open-source frameworks, which are Streamlit, along with the Python language. Three platforms are being used in this project which are Visual Studio Code, Google Colab, and Roboflow. Meanwhile, there are three languages to implement the web-based application in this project, which are Hypertext Markup Language (HTML), and Cascading Style Sheet (CSS).

3.4 Testing

The process to identify the model's accuracy starts with validating the custom model using the command shown in Figure 5.1 below.

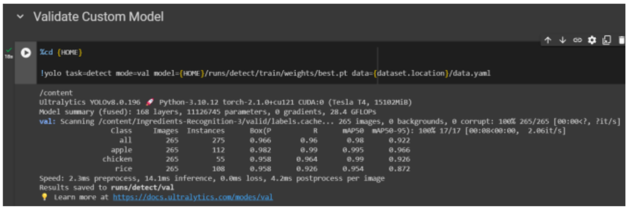

Fig. 3. YOLO Model Validation

As shown in Figure 5.1, the command validates the YOLO model using a validation set. It uses the YOLO toolkit with parameters specifying the detection task, validation mode, the path to the model weights file, and the location of the data configuration file. By executing this command containing "mode=val" the model's performance results are evaluated using the validation dataset. As a result, after running this command, the detection accuracy result on the validation dataset is 98% mAP50. The results are saved in the runs/detect/val path file. The figure below shows one compilation of the results from the validation model process for this project.

Fig. 4. Results of the Validation Model Process

Figure 5.2 displays the results of the validation model process that contains each ingredient's bounding box, label, and accuracy for detected objects in the images.

4 Results and Discussion

This paper develops a web-based application. This section contains a description of developing a web-based application.

4.1 Model Evaluation

After the image dataset is ready, the model can be trained, and for this project, the model is being trained in Google Colab. Firstly, the Ultralytics library must be installed to use the class from the library. The YOLO (You Only Look Once) class can be accessed from the library to train the model. Besides, it also imports the 'display' and 'Image' modules from IPython, which display images in the notebook. This can be shown in the figure below.

Fig. 5. Environment Setup for Ultralytics and YOLO Object Detection

4.2 User Interface

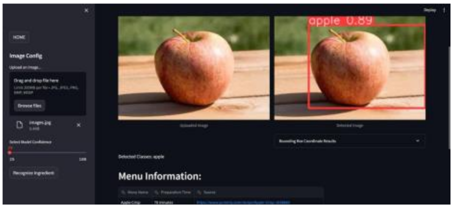

Fig. 6. User Interface for Output page

As shown in Fig 4.12, next to the uploaded image, the system displays the detected image containing their bounding boxes with accuracy and class, whether rice, apple, or chicken. Then, menu information will be displayed depending on the detected class, and a list of menus will be shown in a table format. The table has three columns, which are menu names, preparation times, and sources. When the user clicks on the sources, it will bring users to the recipe that matches the recognized ingredient.

4.3 Discussion

Limitation in terms of number of data Even though the system was successfully developed, it also has some limitations that need to be encountered. Firstly, the fewer ingredients make the system less attractive because there are only rice, apple, and chicken that the model can detect. Besides, the image file types might affect their recognition process. This is because the system can only detect the mentioned type.

5 Conclusion

Nutrition is important for the body. Without it, the body suffers from various diseases. While there are many recipe websites, finding a suitable recipe can be time-consuming, especially for those who lack cooking knowledge with available ingredients. This project addresses this challenge by developing a recipe recommendation model using a deep learning algorithm to suggest recipes based on recognized food ingredients. The model utilizes more than 2,000 images categorized into three food ingredient groups: rice, apple, and chicken. The data collection process includes thorough preprocessing, data cleaning, and annotation. The dataset is then augmented and split into training, validation, and testing sets in an 80-10-10 ratio. To build the food ingredient recognition model, the You Only Look Once (YOLO) technique, specifically YOLOv8, is employed. The system is designed as a web-based application, providing an accessible interface for users. After extensive training, validation, and testing, the model achieved an impressive accuracy of 98%, demonstrating its capability to accurately detect the specified ingredients. Future research aims to enhance the system's functionality by addressing current limitations. Planned improvements include integrating a comprehensive database for storing recipes, which would facilitate easier data insertion and the addition of nutritional information for each recipe. Furthermore, incorporating live-camera integration could enhance user engagement and practicality. These future enhancements are geared towards offering a more user-oriented cooking experience and promoting healthier eating habits.

Acknowledgments. A special appreciation for College of Computing, Informatics and Mathematics, Universiti Teknologi MARA for supporting the publication of this paper.

References

1. Jahan I, Karmakar P, Hossain MM, Jahan N, & Islam MZ. (2020). Fast Food Consumption and its Impact on Health. https://www.researchgate.net/publication/346220090
2. Zhu, B., Ngo, C.-W., Chen, J., & Hao, Y. (2019). R 2 GAN: Cross-modal Recipe Retrieval with Generative Adversarial Network. https://ink.library.smu.edu.sg/cgi/viewcontent.cgi?article=7459&context=sis_research
3. Shinde, Mr. P., Sayyed, Mr. A., & Bhosale, Prof. R. (2019). Recipe Recommendation Based on Ingredients using Machine Learning. IJARCCE, 8(3), 73–77. https://doi.org/10.17148/ijarcce.2019.8313
4. Min, W., Jiang, S., Liu, L., Rui, Y., & Jain, R. (2019). A survey on food computing. ACM Computing Surveys, 52(5). https://doi.org/10.1145/3329168
5. Khan, M., Jan, B., Farman, H., Ahmad, J., Farman, H., & Jan, Z. (2019). Deep Learning: Convergence to Big Data Analytics. http://www.springer.com/series/10028
6. Purwono, P., Ma'arif, A., Rahmaniar, W., Fathurrahman, H. I. K., Frisky, A. Z. K., & Haq, Q. M. ul. (2023). Understanding of Convolutional Neural Network (CNN): A Review. International Journal of Robotics and Control Systems, 2(4), 739–748. https://doi.org/10.31763/ijrcs.v2i4.888
7. Chavan, P., Thoms, B., & Isaacs, J. (2021). A recommender system for healthy food choices: Building a Hybrid Model for Recipe Recommendations using Big Data Sets. Proceedings of the Annual Hawaii International Conference on System Sciences, 2020-January, 3774–3783. https://doi.org/10.24251/hicss.2021.458
8. Chhipa, S., Berwal, V., Hirapure, T., & Banerjee, S. (2022). Recipe Recommendation System Using TF-IDF. ITM Web of Conferences, 44, 02006. https://doi.org/10.1051/itmconf/20224402006
9. Wayesa, F., Leranso, M., Asefa, G., & Kedir, A. (2023). Pattern-based hybrid book recommendation system using semantic relationships. Scientific Reports, 13(1), 3693. https://doi.org/10.1038/s41598-023-30987-0
10. Shafaat, M., Rokon, J., Morol, K., Hasan, I. B., Saif, A. M., Khan, R. H., & Hussain Khan, R. (2022).

11. Singh, A., & Gupta, A. (2022). Comparison of Available Web-System Design Methodologies. International Journal of Advances in Engineering and Management (IJAEM), 4, 473. https://doi.org/10.35629/5252-0401473477
12. Azad, B. A., Laperdrix, P., & Nikiforakis, N. (2019). Less is More: Quantifying the Security Benefits of Debloating Web Applications. https://www.usenix.org/conference/usenixsecurity19/presentation/azad
13. Maruyama, T., Kawano, Y., & Yanai, K. (2012). Real-time mobile recipe recommendation system using food ingredient recognition. https://doi.org/10.1145/2390821.2390830
14. Pranav Vijay Chakilam, B., Revanth, Muppirala, V., Anilet Bala, A., & Maik, V. (2021). Design of Low-Cost Object Identification Module for Culinary Applications. Journal of Physics: Conference Series, 1964(6). https://doi.org/10.1088/1742-6596/1964/6/062088
15. Jadhav, T., Khatik, V., Navghane, A., Rajpurohit, M., & Kalshetti, R. (2023). Vision Based Intelligent Recipe Recommendation System. In Chem. Bull (Vol. 2023). https://www.eurchembull.com/uploads/paper/7575c510aabd16f306d20d5db8fa6894.pdf