2025 IEEE 6th International Conference in Robotics and Manufacturing Automation (ROMA)
20-22 August 2025, Bangi, Selangor, Malaysia

2025 IEEE 6th International Conference in Robotics and Manufacturing Automation (ROMA) | 979-8-3315-9996-6/25/$31.00 ©2025 IEEE | DOI: 10.1109/ROMA66616.2025.11155482

Automated Food Nutrition Retrieval for Healthcare and Medical Applications using Image-Based Weight Prediction Models

Tang Jin Hang
Department of Electrical and Electronics Engineering
Universiti Teknologi PETRONAS
Bandar Seri Iskandar, Perak, Malaysia
tangjinhang2002@gmail.com

Azrina Abd Aziz
Department of Computer and Information Sciences
Universiti Teknologi PETRONAS
Bandar Seri Iskandar, Perak, Malaysia
azrina_aaziz@utp.edu.my

Kanimolli Arasu
Division of Nutritional and Dietetics
IMU University
Bukit Jalil, Kuala Lumpur, Malaysia
kanimolliarasu@imu.edu.my

Vergin Raja Sarobin M
School of Computer Science and Engineering
Vellore Institute of Technology
Chennai, India
verginraja.m@vit.ac.in

***Abstract*—Malaysia's diverse and rich cuisine is a fundamental part of Malaysia's culture. However, there are many Malaysia food that are high in calories and sugar contents, which can lead to health issues such as diabetes and obesity. Not only that, traditional calorie estimation methods such as manual food tracking are ineffective due to lack of a complete Malaysia food dataset and portion size estimation errors. To address the difficulties mentioned above, this study aims to propose a machine learning-based solution for automated food nutrition retrieval using image-based weight prediction models. First, data collection for Malaysia food images is performed across five food classes. Next, YOLOv11, which is an instance segmentation model is trained for food identification task and three gradient boosting algorithms, XGBoost, CatBoost and LightGBM are trained to identify best model for food weight prediction. Both food recognition model and weight prediction model are evaluated based on performance metrics such as precision and recall. Based on results obtained, food recognition model has achieved 99.20% precision and 99.50% recall while XGBoost has been chosen as best model to estimate food weight with MAE of 6.21g, RMSE of 8.00g and MAPE of 4.62%. After the food type and food weight are identified, they are provided as inputs to an application programming interface (API) that is created based on Malaysia food composition database to obtain nutritional information such as calories, fat and protein in healthcare applications.**

***Keywords*—Food Recognition, Food Weight Prediction, YOLOv11, XGBoost, CatBoost, LightGBM, Machine Learning, Image-based Food Analysis, Automated Food Nutrition Retrieval**

I. Introduction

Malaysia boasts a rich and diverse culinary food, with traditional dishes influenced by Malay, Chinese, Indians and other cultures. Some typical examples of Malaysian food include Nasi Lemak, Roti Canai, Cendol and etc [1]. These foods are popular among all Malaysians, but they are high in calories and complex in composition, which introduce health-related challenges to individuals in Malaysia [2]. Obesity, which is linked to several health issues such as heart disease, can result from the accumulation of body fat due to high-calories food. In Malaysia, traditional Malaysian cuisine relies heavily on coconut milk, deep-frying methods and thick sauces that contribute to raise of calorie densities of meals and population's risk of obesity [3]. As reported by [4,5], there are over half population (54.4%) of Malaysian adults that are facing issues related to overweight or obese. Between 2019 and 2024, the number of cases related to obesity has increased from 19.1% to 50.1% [4]. According to a research conducted by [6], there is approximately 1 in 5 adult Malaysians suffer from diabetes. Thus, the statistics shown above have highlighted the urgent need to emphasize dietary practices and lifestyle changes to minimize negative effects of high-calorie meals.

Manual food tracking or consulting food compositions tables are common methods of traditional calorie estimation techniques, which can be prone to errors and impractical for many users [7]. Conventional food tracking methods such as manual entry-based mobile applications and meal diaries would face several challenges. According to research by [8], it has identified major challenges including portion size estimating mistakes, inconsistent recordkeeping and user weariness. Furthermore, the vast variety of Malaysian cuisines may not be fully covered by the nutritional databases that are in use now, which could result in information gaps. For example, a research on traditional Malaysian kuih made with sticky rice has highlighted the need for updated and comprehensive nutrient composition data to improve the Malaysian Food Composition Database (MyFCD) [9]. The difficulties mentioned above underscore the necessity of a robust approach to dietary tracking and a specialized Malaysia food dataset to assist development of food identification and calorie estimation.

The purpose of this study is to collect Malaysia food images and corresponding weight datasets. Next, the collected images and weight datasets are used to develop machine-learning models for food image detection and food weight prediction.

The developed models are improved and evaluated throughout the study. Lastly, an application programming interface (API) is built for retrieving food nutritional contents from a Malaysia food composition database based on provided food type and food weight. In the end of the study, developed models are integrated with custom-build Malaysia food composition database API for automated retrieval of food nutritional information in healthcare and medical applications.

II. Literature Review

This section reviews the existing machine learning applications in food recognition and food weight prediction.

A. Machine Learning Applications in Food Recognition

Machine learning has transformed various industries such as food recognition by enabling precise classification of food images. Manual feature extraction and rule-based approaches used in traditional food recognition systems are frequently restricted by their inability to handle varied food datasets. In recent years, image processing has grown more effective with the introduction of deep learning, especially Convolutional Neural Networks (CNNs), which excel at solving difficult image-driven pattern recognition tasks such as food identification [10]. ML-driven food recognition is a useful tool for healthcare and wellness applications because it is important for nutrition tracking and dietary monitoring [11].

CNNs are one of the most popular architectures used in food identification due to their excellent ability to analyze picture properties such as colour, texture, and shape. CNNs are comprised of three different basic layers, which are convolutional layers, pooling layers and fully-connected layers. Firstly, the input layer contains pixel values of the input image. Next, the convolutional layers filter the input image to extract features such as edges, textures and patterns. The pooling layer reduces spatial dimensions of feature maps to increase computing efficiency while maintaining the most important characteristics. The fully-connected layers will map retrieved features to certain specified groups to perform classification.

YOLOv11 is a model introduced by Ultralytics that has a decent performance among all YOLO models [12]. In contrast to traditional CNNs, YOLOv11 uses a transformer-based backbone to capture long-range dependencies and enhance small object recognition. It also replaces non-maximum suppression (NMS) with a more effective algorithm that can speed up real-time inference speed. However, in order to achieve high accuracy for Malaysia food classification task, it requires a significant number of high-quality labelled images for transfer learning. Transfer learning method works by freezing early and central layers of model to maintain key information while retraining the latter layers for a specific task output [13]. By using this method, it helps to achieve solid performance with limited training data and decreases training time significantly.

B. Machine Learning Applications in Food Weight Prediction

Calorie estimation from food images can be relied on precisely estimating portion size given that calorie content is strongly correlated with food weight. Deep learning based method is one of the food weight prediction methods that can provide a flexible and scalable solution by directly estimating food weight. These deep learning models are capable of determining food weight by examining characteristics of food in the image such as shape, size, texture and density correlations.

It is suggested by [14] to combine deep-learning approach with other approaches to improve robustness of food weight prediction system. According to a study conducted by [15], it has combined Mask R-CNN and Multiple Linear Regression for food image classification and food weight prediction and has scored R-squared of 0.80. From the study above, it shows the feasibility to utilize capabilities of Mask R-CNN model to segment food mask from image to feed extracted features into a machine learning model to predict food weight. Among all machine learning algorithms, gradient boosting algorithms are studied to further integrate with deep-learning methods. Gradient boosting algorithms utilize ensemble learning methods such as bagging and boosting to achieve high accuracy and reduce bias error of model [16].

There are three main gradient boosting algorithms utilized nowadays, which are XGBoost [17], CatBoost [18] and LightGBM [19]. XGBoost is a gradient boosting algorithm that reduces errors through iterative corrections and handles big datasets effectively. Next, CatBoost is designed and optimized to handle categorical data, which can effectively identify trends in different food varieties. Lastly, LightGBM is a lightweight and effective algorithm that excels in quick training and low memory utilization.

III. Methodology

The following section discusses the methodology of using two distinct models for food recognition and food weight prediction tasks respectively. The proposed methodology is illustrated in Fig. 1.

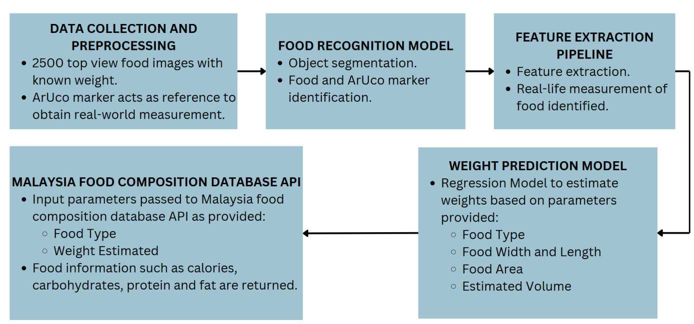

Fig. 1. Proposed methodology workflow

The proposed system begins with the creation of a comprehensive dataset consisting of 2500 Malaysia food images spread across 5 different food categories. Not only that, the actual weights of food will be measured using a kitchen scale and collected to train food weight prediction model later. Next, the food recognition model would be trained by the datasets collected to perform object detection to detect food in the images. After the food masks are obtained, the extracted food and marker from the images will be fed into a feature extraction pipeline to obtain important aspects such as food area. After that, a machine-learning based regression model will be trained using the features extracted from the images to perform food weight prediction. After the food weight is estimated, parameters such as food type and food weight will be passed as inputs to an application programming interface (API) that is created based on Malaysia food composition database to output nutritional information.

A. Data Collection and Preprocessing

The proposed 5 selected food classes for this study are banana, apple, fried noodles, fried rice and tangerine. Each food category will feature 50 weight variations with 10 images for each weight variant, resulting in a total of 2500 images. All images will be captured from a top-view perspective to ensure uniformity. At the same time, an ArUco marker [20] will be placed beside the food to act as a reference object to obtain real-world measurement. The actual weights of food will also be measured using a kitchen scale and collected to train food weight prediction model later. The sample food image of each food class is shown in Fig. 2.

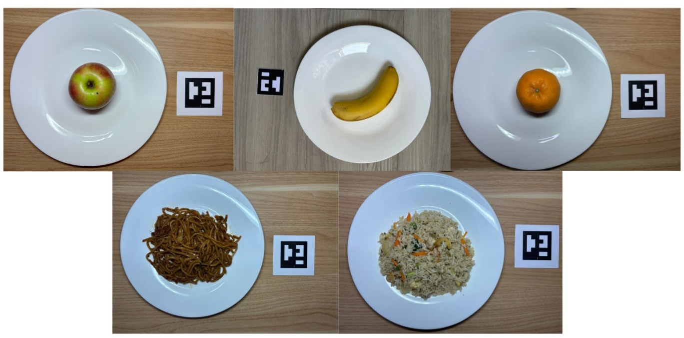

Fig. 2. Sample food image of each food class

Next, all collected food images are carefully preprocessed and labelled to maximize learning performance. Each raw image is stretched to a consistent 640 x 640-pixel size to help model processes inputs effectively while preserving the relative spatial relationships between food items and reference markers. Then, data augmentation is utilized to improve model robustness and avoid overfitting. Methods such as horizontal and vertical flipping are applied to double the range of food placement orientations.

B. Food Recognition Model

For the food recognition model, YOLOv11 is used as the main object detection and segmentation model. This study is well-suited to the YOLOv11 architecture because of its improved ability to carry out both instance segmentation and object detection to precisely define the boundaries of food and classify food item within each image. Before the training starts, the data is split into 90:5:5 split ratio, allocating 90% of data for training, 5% for validation datasets and 5% for testing datasets. Next, the training process takes 100 epochs with early stopping to train the model to attain high accuracy without overfitting the dataset. For optimization algorithm, AdamW optimizer combined with enhanced weight decay regularization is used for optimization with an initial learning rate of 0.00167. After the training process has ended, model performance is evaluated based on two main performance metrics, which are Precision (P) and Recall (R) when executing bounding box and segmentation mask tasks.

C. Feature Extraction Pipeline

The following step creates a customized feature extraction pipeline to obtain dimensions measurements from output of food recognition model and serves as inputs for weight prediction model. This pipeline converts pixel-based image data into measures that may be used in the actual world by utilizing computer vision algorithms. The first stage in the pipeline is to detect ArUco marker in the image using OpenCV library [21]. By using the module, it has built-in functions that are able to determine marker's perimeter and pinpoint its corner with sub-pixel accuracy. The marker used in this study is 5x5cm with ID as 10.

Next, the pixels-per-centimeter of each image is calculated using detected ArUco markers since the real-life marker length is known to be 5cm. Thus, additional information of the food can be determined and extracted such as food type, food area, food width, food length, food colour, solidity, circularity, estimated volume and edge density. The sample output of feature extraction pipeline is illustrated in Fig. 3 by labelling the bounding box and segmentation mask of food. The dotted rectangle around the food denotes the bounding box whereas slanted black lines overlaid on the food represent the segmentation mask, which captures pixel-level of the food's exact area. Thus, this feature extraction pipeline closes the gap between image data and physical characteristics.

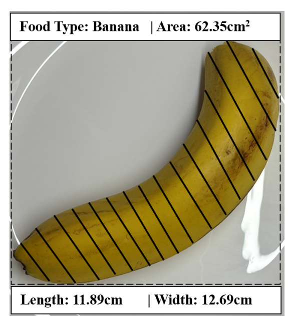

Fig. 3. Sample output of feature extraction pipeline

D. Weight Prediction Model

There will be three advanced gradient boosting models, which are XGBoost, CatBoost and LightGBM to be trained and evaluated to obtain the best model for weight prediction model. Before training the weight prediction model, the training data are obtained from previous feature extraction pipeline and arranged in a dataframe dataset format to simplify training process. Table I presents the sample structure of processed features dataset used for training of weight prediction model with food weight serves as the output variable.

TABLE I
SAMPLE STRUCTURE OF PROCESSED FEATURES DATASET

Feature	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
Food Type	Banana	Tangerine	Apple	Fried Rice	Fried Noodles
Pixels / cm	163.44	144.07	156.40	134.82	126.74
Food Area (cm^2)	62.35	42.78	56.49	187.50	187.64
Width (cm)	12.69	7.06	8.49	16.06	16.77
Length (cm)	11.88	7.22	8.33	16.38	17.92
Food Weight (g)	138.00	103.00	238.00	129.00	114.00

Next, the datasets are being split into 80:20 ratio, which 80% for training and 20% for testing purposes. In the training process, 3-fold cross-validation is applied to ensure robust evaluation and prevent overfitting. Not only that, model hyperparameters are optimized by utilizing cross-validation grid search as it helps to obtain key factors such as learning rate, number of estimators and tree depth. Cross-validation grid search will compare each key factor to select the best hyperparameters to obtain the best model performance in terms of prediction accuracy and generalization capabilities. After the training process, the performance of the weight prediction model using three different gradient boosting algorithms is evaluated using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE).

E. Malaysia Food Composition Database API

The final stage of the methodology workflow is the creation of a Malaysia food composition database application programming interface (API) system that connects food type and food weight with thorough nutritional insights. The API is created using FastAPI, which is a type of RESTful API that utilizes HTTP requests to access and use data [22]. It offers a reliable interface for obtaining Malaysia food composition data such as calories, protein and others by providing food type and food weight. Table II shows the nutritional profiles of each food class in Malaysia food composition database.

TABLE II
NUTRITIONAL PROFILES OF EACH FOOD CLASS

Food Type	Weight (g)	Calories (cal)	Protein (g)	Fat (g)
Apple	115.0	64.0	0.2	0.5
Banana	62.0	65.0	0.6	0.2
Tangerine	165.0	73.0	1.8	0.5
Fried Rice	481.0	911.0	27.5	47.4
Fried Noodles	200.0	346.0	9.0	13.0

There are two key functional endpoints created for this study, which are named as “/data” and “/nutrient”. The first endpoint is a GET-type endpoint that allows users to gain information from the database. The next endpoint is named as “/nutrient” and it will return user about nutritional information of a specific food item. It will first accept two input parameters which are food name and food weight. Next, it will perform dynamic nutritional calculations based on input weight and search food name in the database. Nutrients information such as calories, protein and fat are calculated and returned as JSON output using a linear scaling approach as they are directly proportional to food weight.

IV. RESULTS AND ANALYSIS

This section presents and analyses the performance of food recognition model and weight prediction model. It also displays the integration of two developed models with Malaysia food composition database API to create a web-based application for users.

A. Food Recognition Model

After the model is trained with food images, the performance of the model is evaluated based on precision and recall. Precision measures how accurate the model’s positive predictions are by computing the ratio of true positives to the total of true positives and false positives. Recall calculates the proportion of true positives to total of true positives and false negatives. Table III shows the performance metrics of food recognition model executing bounding box and segmentation mask tasks.

TABLE III
PERFORMANCE METRICS OF FOOD RECOGNITION MODEL: PRECISION AND RECALL

Class	Bounding Box		Segmentation Mask	
	Precision	Recall	Precision	Recall
All	99.20%	99.50%	99.20%	99.50%
Apple	99.10%	99.50%	99.10%	99.50%
Banana	99.90%	99.50%	99.90%	99.50%
Tangerine	99.20%	99.50%	99.20%	99.50%
Fried Rice	99.10%	99.50%	99.10%	99.50%
Fried Noodles	99.20%	99.50%	99.20%	99.50%

It has shown that the precision of bounding box and segmentation mask is 99.20% for all classes, with individual class precisions of 99.10% for apple, 99.90% for banana, 99.20% for tangerine, 99.10% for fried rice and 99.20% for fried noodles. The high precision of model has been validated that it can classify food types consistently without recognizing background objects as food incorrectly. Next, the model has demonstrated 99.50% of recall for bounding box and segmentation mask, which shows the model’s ability to accurately identify food type with minimal false negatives. The fine-tuned YOLOv11 model has displayed high precision and recall as it utilizes past information from other pre-trained food datasets, which contributes its great generalization capabilities to the current food dataset. Thus, the results have proved that the proposed model using YOLOv11 for food detection task is highly effective, especially for clearly defined food classes such as apples and bananas.

B. Weight Prediction Model

There are three weight prediction models that are trained with features extracted, which are XGBoost, CatBoost and

LightGBM. The performance of each model is evaluated based on Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE). MAE calculates average absolute differences between predicted weights and actual weights whereas RMSE is the square root of Mean Squared Error (MSE). Then, MAPE calculates the error as an average percentage of actual weight values. Table IV has shown the performance metrics of three distinct weight prediction models.

TABLE IV
PERFORMANCE METRICS OF WEIGHT PREDICTION MODELS: MAE, RMSE AND MAPE

Model	MAE (g)	RMSE (g)	MAPE (%)
XGBoost	**6.21**	**8.00**	**4.62**
CatBoost	7.07	10.90	5.22
LightGBM	9.67	18.63	7.05

XGBoost has outperformed other models, which has achieved the lowest MAE (6.21g), RMSE (8.00g) and MAPE (4.62%). This has proved that XGBoost offers the most precise weight prediction, with the smallest prediction error of 4.62% and an average deviation of 6.21g between predicted weight and actual weight. Next, CatBoost followed closely with slightly higher MAE (7.07g), RMSE (10.90g) and MAPE (5.22%). However, LightGBM has shown the worst performance among all, which has the highest MAE (9.67g), RMSE (18.63g) and MAPE (7.05%).

Fig. 4 has presented three scatter plots that compares predicted vs actual weight for the three models. Firstly, XGBoost has displayed its points are closely situated along the red line across all weight range, which indicates its consistency of predicting accurate values except for two weights that exceed 500g. Next, CatBoost has shown similar pattern to XGBoost, with points aligning well to the red line. However, it has a larger spread in the higher weight range, which suggests that it is less accurate when predicting greater weight compared to XGBoost. Then, LightGBM has shown the worst performance as it often overpredicts weights in 200g to 350g range and underpredicts weights in 450g to 500g range.

Overall, it has validated that XGBoost is the most suitable model for food weight prediction as it shows the lowest errors across all performance metrics and predicts more accurate weights across all weight range compared to CatBoost and LightGBM. XGBoost has demonstrated a better performance for food weight prediction due to several factors. First, XGBoost uses both L1 (Lasso) and L2 (Ridge) regularization techniques to prevent overfitting. Due to this feature, it has enabled the model to be more generalizable to smaller datasets and achieve better performance and results. Not only that, XGBoost implements depth-first approach for tree pruning, which allows the model to discover complex patterns more effectively.

C. Malaysia Food Composition Database API

The final stage of the study involves developing a Malaysia food composition database API that provides nutritional information. The self-created API is then integrated with food recognition model and weight prediction model that are previously trained to obtain food contents. Fig. 5 displays the user interface of web-based application created to upload food image and obtain food contents.

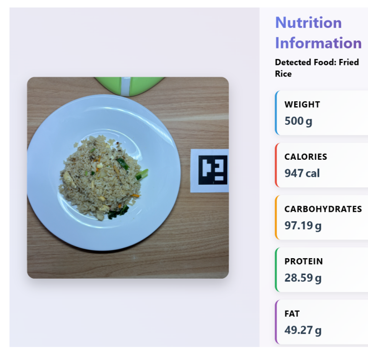

Fig. 5. Web-based application user interface

From the web-based application, it allows users to identify food type and predicts food weight accurately based on a top-view image. Next, crucial parameters such as food type and food weight are passed to the API created to calculate nutrients information using a linear scaling approach. Standardised nutrient values per 100g of a specified food will be scaled proportionally to the projected food weight provided by user. Finally,

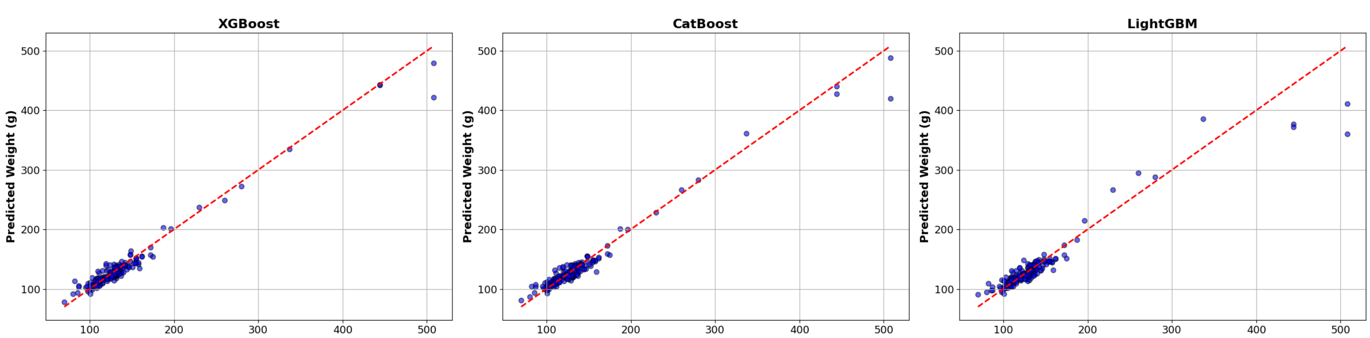

Fig. 4. Predicted vs actual weight of XGBoost, CatBoost and LightGBM

nutrients information including calories, carbohydrates, protein and fats are displayed to the user.

Thus, it has indicated that the Malaysia food composition database API is created and integrated successfully with food recognition and weight prediction models. The API is self-developed and remains in a prototype stage, as it relies on a private dataset of Malaysia food composition that is not yet intended for public release. Next, it results in a creation of a web-based application to allow users to upload food image and obtain nutritional information. The web-based application has enabled users to make well-informed food choices to aid in nutrition tracking and health monitoring.

V. Conclusion and Future Works

This study has successfully implemented an automated food nutrition retrieval system for healthcare and medical applications using image-based weight prediction models. For the food recognition model, YOLOv11 model has been trained successfully to detect food classes accurately with precision of 99.20% and recall of 99.50% for bounding boxes and mask segmentation. For the food weight prediction model, XGBoost has been trained successfully to estimate food weight with MAE (6.21g), RMSE (8.00g) and MAPE (4.62%). In between the food identification model and weight prediction model, there is a feature extraction pipeline to obtain useful measurement information from the image such as food length, food width and food area. Next, a Malaysia food composition database API is built to obtain nutritional information easily based on input information such as food type and food weight. In the end, the trained models and API is integrated to create a web-based application for users to assist in fitness tracking and personalized diet planning. Thus, this study has proved that automated food nutrition retrieval system is feasible to serve as a foundation for upcoming developments in AI-driven healthcare monitoring applications. In the near future, the focus will be incorporating a broader range of higher-quality food classes and improving feature extraction pipeline to increase features extracted for better performance.

Acknowledgment

The authors acknowledge the support of this research by the national Collaborative Research Fund (CRF) between Universiti Teknologi PETRONAS (UTP) and International Medical University (IMU) under the cost centre 015MD0-211.

References

[1] Food'n Road, "Malaysian Food Guide — The Main Typical Dishes Of Malaysia," *Food'n Road*, Sep. 23, 2019. [Online]. Available: https://foodandroad.com/malaysian-food/.

[2] Z. N. Zainal Arifen, S. Shahar, K. Trieu, H. Abdul Majid, M. F. Md Noh, and H. Haron, "Individual and total sugar contents of street foods in Malaysia – Should we be concerned?," *Food Chemistry*, vol. 450, p. 139288, Aug. 2024. DOI:10.1016/j.foodchem.2024.139288.

[3] Vitamode, "Savoring Malaysian Delights With Health In Mind," *Vitamode*, Aug. 15, 2023. [Online]. Available: https://www.vitamode.com.my/post/savoring-malaysian-delights-with-health-in-mind.

[4] M. V. Ang, "Health Ministry: 54.4% Of Malaysian Adults Are Either Overweight Or Obese," *SAYS*, Jun. 06, 2024. [Online]. Available: https://says.com/my/news/health-ministry-half-of-malaysian-adults-are-overweight.

[5] C. T. Chong, W. K. Lai, S. Mohd Sallehuddin, and S. S. Ganapathy, "Prevalence of overweight and its associated factors among Malaysian adults: Findings from a nationally representative survey," *PLOS ONE*, vol. 18, no. 8, p. e0283270, Aug. 2023. DOI:10.1371/journal.pone.0283270.

[6] World Health Organization, "The annual health-care cost of cardiovascular diseases, diabetes and cancer in Malaysia exceeds RM 9.65 billion," *WHO*, Aug. 9, 2022. [Online]. Available: www.who.int.

[7] X. Li, A. Yin, H. Y. Choi, V. Chan, M. Allman-Farinelli, and J. Chen, "Evaluating the quality and comparative validity of manual food logging and artificial intelligence-enabled food image recognition in apps for nutrition care," *Nutrients*, vol. 16, no. 15, p. 2573, Aug. 2024. DOI:10.3390/nu16152573.

[8] X. Zhao, X. Xu, X. Li, X. He, Y. Yang, and S. Zhu, "Emerging trends of technology-based dietary assessment: A perspective study," *European Journal of Clinical Nutrition*, vol. 75, no. 4, pp. 582–587, Oct. 2020. DOI:10.1038/s41430-020-00779-0.

[9] A. Mahmood, L. Y. Mei, M. F. M. Noh, and H. M. Yusof, "Nutrient composition of five selected glutinous rice-based traditional Malaysian kuih," *Malaysian Applied Biology*, vol. 47, no. 4, pp. 71–77, Oct. 2018.

[10] X. Zhao, L. Wang, Y. Zhang, X. Han, M. Deveci, and M. Parmar, "A review of convolutional neural networks in computer vision," *Artificial Intelligence Review*, vol. 57, no. 4, Mar. 2024. DOI:10.1007/s10462-024-10721-6.

[11] S. Jha, I. K. Garewal, L. Aathisaya, L. Alphonso, and B. Aher, "FoodMO: A food nutrient analysis application using optical character recognition and machine learning," *Lecture Notes in Networks and Systems*, pp. 589–600, Jan. 2025. DOI:10.1007/978-981-97-8526-1_47.

[12] N. Dwivedi, "YOLOv11: The next leap in real-time object detection," *Analytics Vidhya*, Oct. 24, 2024. [Online]. Available: https://www.analyticsvidhya.com/blog/2024/10/yolov11/.

[13] P. Sharma, "Transfer learning: Understanding transfer learning for deep learning," *Analytics Vidhya*, Oct. 30, 2021. [Online]. Available: https://www.analyticsvidhya.com/blog/2021/10/understanding-transfer-learning-for-deep-learning/.

[14] F. P. W. Lo, Y. Sun, J. Qiu, and B. Lo, "Image-based food classification and volume estimation for dietary assessment: A review," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 7, pp. 1926–1939, Jul. 2020. DOI:10.1109/JBHI.2020.2987943.

[15] N. K. Aditama and R. Munir, "Indonesian street food calorie estimation using Mask R-CNN and multiple linear regression," in *Proc. 2nd Int. Conf. Power, Control and Computing Technologies (ICPC2T)*, Mar. 2022. DOI:10.1109/ICPC2T53885.2022.9776804.

[16] A. Bonnet, "What is ensemble learning?," *Encord*, Nov. 24, 2023. [Online]. Available: https://encord.com/blog/what-is-ensemble-learning/.

[17] L. Xu, S. Wen, H. Huang, Y. Tang, Y. Wang, and C. Pan, "Corrosion failure prediction in natural gas pipelines using an interpretable XGBoost model: Insights and applications," *Energy*, p. 136157, Apr. 2025. DOI:10.1016/j.energy.2025.136157.

[18] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "CatBoost: Unbiased boosting with categorical features," *arXiv:1706.09516*, Jan. 2019. [Online]. Available: https://arxiv.org/abs/1706.09516.

[19] T.-K. Yu, I.-C. Chang, S.-D. Chen, H.-L. Chen, and T.-Y. Yu, "Predicting potential soil and groundwater contamination risks from gas stations using three machine learning models (XGBoost, LightGBM, and Random Forest)," *Process Safety and Environmental Protection*, vol. 199, p. 107249, Jul. 2025. DOI:10.1016/j.psep.2025.107249.

[20] T. Tocci, L. Capponi, and G. Rossi, "ArUco marker-based displacement measurement technique: Uncertainty analysis," *Engineering Research Express*, vol. 3, no. 3, p. 035032, Aug. 2021. DOI:10.1088/2631-8695/ac1fc7.

[21] OpenCV, "OpenCV library," *opencv.org*, 2019. [Online]. Available: https://opencv.org/.

[22] FastAPI, "FastAPI," *fastapi.tiangolo.com*, 2023. [Online]. Available: https://fastapi.tiangolo.com/.