Deep Learning for Meal Recognition and Calorie Estimation

Ahmad Nabil bin Ahmad Faris[1*], Nor Hafizah Abdul Razak[2], Mazura Mat Din[3]

[1,2,3] College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, Kedah Branch, Malaysia

Article Information

Received: 21-11-2024
Revised: 28-11-2024
Published: 05-12-2024

Keywords
Image processing, convolutional neural networks (CNN), YOLO, Calorie detection, Nutritional intake

****Correspondence Email:***

hafizah446@uitm.edu.my

Abstract

Accurate calorie estimates from foods are prerequisite for diet following and health monitoring. Manual calorie estimations according to age-old methods mostly tend to be inaccurate. This paper proposes the use of convolutional neural networks (CNNs) for precise identification from food images and prediction of meal calories to solve the concern. Therefore, the objective is to create a model capable of recognizing foodstuff besides estimating their caloric content. Developing a model that could correctly identify food ingredients and calculate their energy value through training and testing was important in this project. Our aim here was to verify the accuracy of the model using systematic reviewing means as well as an interface where it can be tested. A dataset of 1,337 high-quality images divided into 12 culinary classes cake, hamburger, noodles, spaghetti, pizza, chicken curry, croissant, French fries, fried chicken, roast chicken, lobster nasi goreng, and waffle was obtained from Roboflow Universe and used for this project. The selection of technique which is YOLO (You Only Look Once) model architecture and flow design because it proved to be highly efficient for real-time object recognition.

1. Introduction

Obesity is a crucial public health problem and is rising globally. According to The National Health and Morbidity Survey (NHMS) 2015, approximately 30.2% of elderly individuals in Malaysia were found to be obese. In addition to obesity, it is important to note the connection between obesity and various other diseases such as hypertension, diabetes, and hypercholesterolemia. (Hasani, 2020).

Existing solutions have been presented to the public but they seem to not care about the revolution that has been made. For example, calorie disclosure on restaurant menus is a popular strategy to address obesity. However, past research has found it doesn't reduce the calories people consume. The authors suggest this may be because when people choose foods, they don't consider how much mental effort it takes to process the calorie information. Furthermore, tracking diet apps can helps changing on eating habits, counting calories can be challenging. This difficulty may cause some people to give up on these programs. Traditional methods of keeping track of what you eat can be time-consuming and may not always be followed by users, resulting in less accurate and consistent information about your diet. (Tahir, 2021). This can make it tricky to determine portion sizes and calorie content. The main problem at hand is the need for accurate and efficient methods to measure daily calorie consumption, resulting in a healthy diet, and managing weight effectively (Shen,2019). Therefore, it is compulsory to do lots of studies to make sure these apps work well and are easy to use. This encourages healthier eating habits, leading to improved overall health.

The objective of this study is to develop a CNN model for image detection and meal calorie intake. conduct a comprehensive evaluation of the outcome's accuracy. To create a prototype system for users to view meal details and calories, helping them plan their meals easily. Transforming object recognition tasks using computer vision, such as CNNs, and YOLO, has led to faster and better methods of performing the tasks. Thus, the aim of this work is to develop and train a YOLO based model on various food items to identify it and predict the calorie content. The urge toward automated solutions keeps pushing the fitness-health applications. This systematic approach that outlines a procedure for data annotation, training of the model, and evaluation will work to address challenges of manual measuring of calories. Among the major task achievements include the development of an interactive interface for real-world applications and a very efficient pipeline in calorie estimation.

1.1 Literature Review

Numerous studies have explored the application of machine learning in dietary monitoring. Firstly, the study by Gustav Gath (2024) has done some result testing using YOLO technique. The evaluation of YOLOv7, YOLOv8, and YOLOv9 with respect to the design dataset made up of 3707 images from 42 food classes using metrics such as accuracy, recall, mean Average accuracy (mAP), and F1 score. YOLOv8 was found to be the best model among the three dues to its high accuracy, mAP at 0.5, and F1 score, making it very suitable for diet analysis and nutrition promotion. For further analysis of portion sizes and proportions to make healthy predictions based on models like the Swedish plate model, there is need for YOLO models for food-component recognition and bounding boxes even though they do not directly segregate healthiness without any specific context about the food. The work describes how YOLO models can add precision and power to nutritional evaluations that bring us closer to preventing diet-related diseases.

Rani Meree and Wisam Saad Aldin's (2024) work explores how artificial intelligence (AI), namely machine learning (ML) and deep learning (DL), might improve calorie tracking and food recognition. It highlights how computer vision, using cutting-edge methods like Artificial Neural Networks (ANNs), can properly recognize food items. Four versions of the YOLO, or You Only Look Once, object detection models YOLOv5, YOLOv6, YOLOv7, and YOLOv8 specifically smaller versions appropriate for web applications are evaluated in this study. The models are renowned for their speed and accuracy. Metrics like F1 score, precision, recall, and mean Average Precision (mAP) are used to evaluate performance, guaranteeing that accuracy and efficiency are balanced for efficient calorie management.

Next, another highlighted research made by Salaki Reynaldo Joshua and his team (2023) focuses on creating a model for diabetes patients. The project aims to make a smart plate system called “Health to Eat” aimed at assisting patients suffering from Type 2 Diabetes Mellitus to monitor their diet more effectively. This system employs a reliable technique based on the recognition, classification, and weighing of different foods to address the challenges encountered when categorizing food for diabetics. Findings indicated that it could differentiate between different types of foods and weigh them with reasonable accuracy. Perfect scores were given to particular meals like rice, soy sauce-braised quail eggs, spicy meat stew, and sun-dried radish. To incorporate these traits, innovative tools such as YOLOv5s’s photo-analyzing software, Chenbo’s weight measuring devices based on load cells or pressure sensors like HX711which converts this physical entity into an electrical signal for processing in computer systems were used besides it had IMX219-160 camera module. The calorie detection project chose YOLO because it has real-time detection capabilities that applications needing instant feedback require. Even more, YOLO is easier to create and integrate due to its simplicity.

2. Research Methods

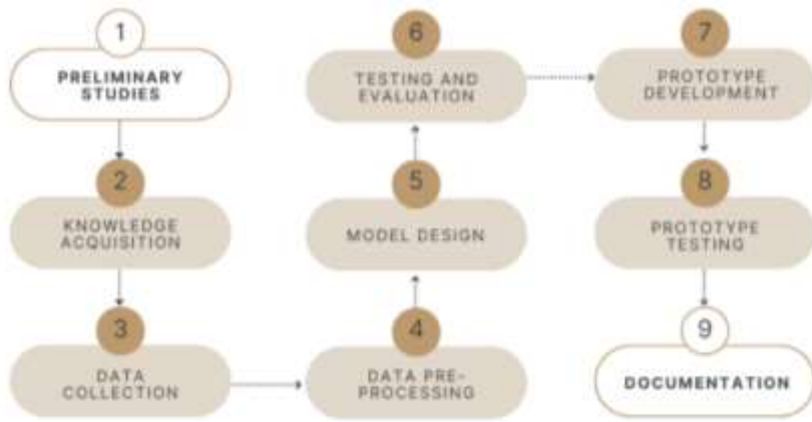

Fig. 1 Research phase of development process

To lay out the preliminary studies, there are now three main research questions with associated goals. These components specify the desired results and provide the study domain a distinct focus. They act as a road map, assisting the researcher in keeping track of developments and staying on course during the inquiry. The question included how can the calories in a meal be quickly and readily calculated using images, what is the methodology used to assess the validity of the study's findings and how user-friendly is the prototype in helping users view meal details and plan their meals. The objective to complete this research is to develop a CNN model for image detection and meal calorie intake, to conduct a comprehensive evaluation of the outcome's accuracy and to create a prototype system for users to view meal details and calories, helping them plan their meals easily.

Next, focusing on data acquisition, the data acquired from Rani Meree's (2024) research, 1,337 photos from 12 different food classes such as cake, hamburgers, and noodles were included in a bespoke food dataset that Roboflow provided. The dataset used for this study consists of 1,337 annotated images from Roboflow Universe, categorized into 12 classes which is cake, hamburger, noodles, spaghetti, pizza, chicken curry, croissant, french fries, fried chicken, roast chicken, fried rice, waffle

As for the data pre-processing there were 1,012 training photos, 256 validation images, and 100 testing images in the dataset. The model's ability to generalize well was ensured by the high quality and broad depiction of food items in the photos. Well-annotated pictures were essential for effective object detection model training, and balanced class distribution helped minimize bias.

Next, focusing on model design the performance of food identification and calorie estimate will necessitate the training of models based on YOLOv8, YOLOv9, and YOLOv10. At the beginning of the training, the learning rate, which mostly begins from 0.001, and all the other parameters are fixed although they can be altered depending on the performance of the model. Given the dataset size and the existing technology, batch size set to 8 limits the number of images processed at once. The models will be trained for 30 epochs and the training will proceed until there is no further improvement. In order to improve the performance of object detection and recognition systems in particular on food products, advanced metrics that are meant for loss, have been integrated into the designs, training and use of YOLOv8, YOLOv9 and YOLOv10. The implementation of management and enhancement of data designs into the training system of all the three versions will make the training quicker and efficient enough to allow a sufficient performance comparison.

Later on, the next step is to conduct the testing and evaluation phases where a number of important measures will be used to evaluate the model's performance. An average precision which measures precision over classes through a single metric that reflects how accurately items are detected is mean average precision (mAP). Precision is based on the ratio of properly recognized food items to all detected items, which determines the accuracy of the model in classifying food items. Recall measures the model's ability to determine all relevant items by calculating the ratio of correct food item recognition to all food items that are present. Finally, the F1-Score provides an honest evaluation of the precision of the model by virtue of the fact that it is a harmonic mean of precision and recall. Taken together, these metrics promise a thorough assessment of the model's performance. The performance of the three models namely YOLOv8, YOLOv9 and YOLOv10 will also be assessed over various key performance indicators which include precision, recall, F1 score and mean average precisions (mAP). In particular, these metrics will assess how well the models are capable of detecting food and estimating the calories present.

The next phases are prototype development where based on the prototype architecture shown in the Figure 1, the system begins with the input of food photos being taken and readied for analysis as per the prototype architecture, as represented in the image. The image goes through a data preparation pipeline, including normalization, scaling to 640 by 640 pixels, and augmentations such cropping, 90-degree rotations, horizontal and vertical flips, and bit little random movements. These preprocessing processes help improve the resilience of the model by ensuring that the dataset used is a varied one and standardized.

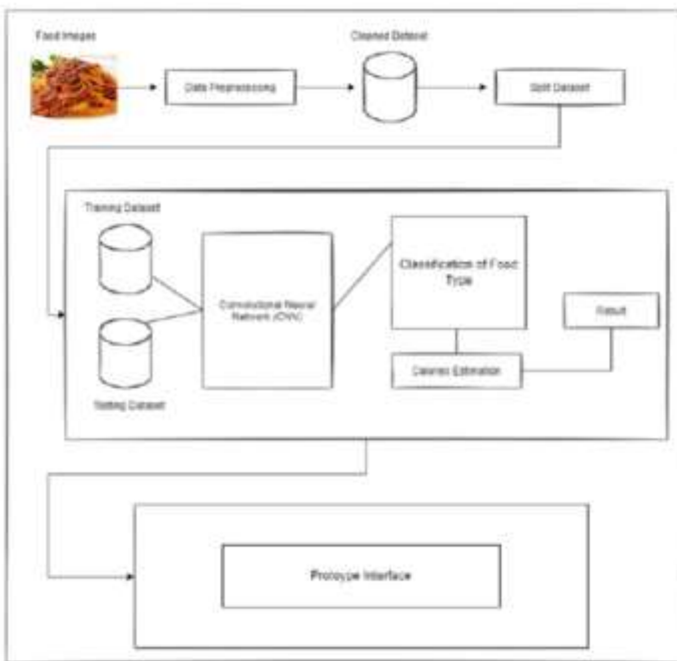

Fig. 2 Architecture Design

The dataset is divided into three subsets: 10% for the test set, 20% for the validation set, and 70% for the training set. This approach ensures that the model is evaluated, validated, and fitted with different datasets, thus providing a comprehensive review of its operation It takes the training and testing stage to apply Google Colab's YOLO (You Only Look Once) for effective and real-time item recognition and categorization. The subsequent step is to install PyTorch, Torchvision, and Ultralytics to prepare the environment. After the setup is complete, the dataset should be obtained from Roboflow. Prior to starting the training process, it is essential to modify the data.yaml file to specify the correct paths for the training, testing, and validation directories according to their local machine locations. The model is augmented by introducing a calorie estimation module after training and testing, which will predict calorie values on the basis of the set nutritional data. The interfacing prototype is thus able to communicate the findings in a very clear real-time manner using display categories of food type and calorie estimates for consumption.

After the development phases finish, the prototype testing phases will be conducted. First of all, the accuracy of the food classification by the system will be evaluated. This step requires the actual comparisons between the true food types and the one predicted by the system so as to compute the values of F1 score, recall, accuracy, and precision of the classification. This will help in ensuring that the system is not misidentifying different categories of food substances. Next, the precision of the estimation of calories will be evaluated. This involves comparing the real caloric content of the food with the caloric value expected based on the weight the user entered and the type of food anticipated. The aim is to give an assessment of calorie content accurately after confirming that the formula applied, and weight entered is correct.

3. Result and Discussion

The YOLOv8, YOLOv9, and YOLOv10 were all developed and trained using the same customized food dataset and parameters. All of the images were resized to 640x640 and automatically oriented as part of the pre-processing stages. There were three outputs per training example sourced from data augmentation strategies, consisting of cropping, with a zoom range of 0% -30%. Rotational angles from -15° to +15°, horizontal and vertical flips, and rotations of 90° clockwise, counter-clockwise, and upside down constituted such techniques. In this manner, the models were trained at 30 epochs and an 8-batch size, while the image size was 640, which could assure similarity and maximum efficiency in each version. For evaluation of YOLOv8s, YOLOv9s, and YOLOv10s on food image detection, assessing their performance in terms of accuracy, speed, and resource utilization during result testing for different version of each YOLO is on Table 1.

Table 1. Evaluation of YOLOv8s, YOLOv9s, and YOLOv10s on food image detection

Metric	**YOLOv8**	**YOLOv9**	**YOLOv10**
mAP50	0.88	0.89	0.91
mAP50-95	0.72	0.74	0.76
Precision (B)	0.86	0.87	0.88
Recall (B)	0.75	0.76	0.78
F1 Score	0.80	0.81	0.83
Training time(minutes)	36	85	42

This approach could be further developed into a web-based application using Flask and python post successful implementation of the food categorization and calorie estimation system. It would enable users to upload pictures of food and classify them into specific categories and get an estimate of the calorie value through a responsive online interface.

The calorie estimation process integrates the formula as shown equation 1:

$$\frac{calories\ per\ 100g}{100g} \times weight = calories$$

This could be realized with user-inputs or weight estimation integrated with food type classification for accurate calorie computation. The system, using credible nutritional reference, will be complete for accuracy and usability.

An extensive range of applications for this approach is possible. It could be used in dietary monitoring, enabling people to diet for weight loss or improve health through accurate calorie tracking and balanced meal planning. Healthcare institutions may use this technology, allowing physicians and dietitians to create personalized meal plans for their patients according to the nutritional needs of individual patients. It also has pedagogic importance: it provides a convenient means to teach nutrition science through interactive graphical representations and practical calorie computation exercises.

4. Conclusions

In conclusion, the proposed food classification and calorie estimation system is a showcase for the potential integration of web-based applications with deep learning to facilitate dietary regulation. The system can accurately identify types of foods and estimate the number of calories consumed by using YOLO for effective image recognition combined with calorie computation according to established nutritional data. The solution assesses for an extensible platform via easy real-time interaction through a simple web interface that utilizes Flask and Python. It creates pathways toward advanced living solutions in potential areas of hospital nutrition management to restaurant transparency to personal health tracking.

5. References

Aldin, W. S., & Meree, R. (2024). AUTOMATED FOOD RECOGNITION AND CALORIE MANAGEMENT SYSTEM: A MACHINE LEARNING APPROACH FOR HEALTHY DIETARY MANAGEMENT.

Ariaratnam, S., Rodzlan Hasani, W. S., Krishnapillai, A. D., Abd Hamid, H. A., Jane Ling, M. Y., Ho, B. K., ... & Mohd Yusoff, M. F. (2020). Prevalence of obesity and its associated risk factors among the elderly in Malaysia: Findings from The National Health and Morbidity Survey (NHMS) 2015. Plos one, 15(9), e0238566.

Chrintz-Gath, G. (2024). Deep Learning for Dietary Assessment: A Study on YOLO Models and the Swedish Plate Model.

Joshua, S. R., Shin, S., Lee, J. H., & Kim, S. K. (2023). Health to eat: a smart plate with food recognition, classification, and weight measurement for type-2 diabetic mellitus patients' nutrition control. Sensors, 23(3), 1656.

Shen, Z., Shehzad, A., Chen, S., Sun, H., & Liu, J. (2020). Machine learning-based approach on food recognition and nutrition estimation. Procedia Computer Science, 174, 448-453.

Tahir, G. A., & Loo, C. K. (2021, December). A comprehensive survey of image-based food recognition and volume estimation methods for dietary assessment. Healthcare, 9(12), 1676