Deep Learning-Based Smart System for Food Quality and Nutritional Content Detection in the Free Nutritious Meal Program

Fadiah Salsabiilah Zainal
School of Electrical Engineering
Telkom University
Bandung, Indonesia
hazellau@student.telkomuniversity.ac.id

Musda Muallimah Ramli
School of Electrical Engineering
Telkom University
Bandung, Indonesia
musdamaullimah@student.telkomuniversity.ac.id

Yukrim Haidar
School of Electrical Engineering
Telkom University
Bandung, Indonesia
yukrimhaidar@student.telkomuniversity.ac.id

Kris Sujatmoko
School of Electrical Engineering
Telkom University
Bandung, Indonesia
krissujatmoko@telkomuniversity.ac.id

Rita Magdalena
School of Electrical Engineering
Telkom University
Bandung, Indonesia
ritamagdalena@telkomuniversity.ac.id

***Abstract* — The availability of nutritious and safe food is a crucial factor in the Free Nutritious Meal (MBG) program. Currently, field food quality monitoring still relies heavily on manual organoleptic inspections, which are highly subjective and time-consuming. This research aims to design a Deep Learning-based Smart System as a software-only web application to detect consumption feasibility and estimate food nutritional content quickly and objectively. The proposed method integrates the YOLOv11 architecture to execute simultaneous object detection on 12 main food classes. The system validates food based on two main indicators: feasibility status (fresh or spoiled) and fulfillment of the minimum calorie standard. Macronutrient estimation is calculated dynamically referring to the Edible Portion (BDD) percentage mapped in a centralized database. Testing results show optimal AI model performance, achieving a mean Average Precision (mAP@0.5) of 95.3% in the validation phase and 91.5% in the testing data. The mathematical alignment of the nutritional calculation algorithm reached 100% accuracy with an average computational speed of under 3 seconds per scan. In conclusion, this smart system successfully provides an objective, accurate, and highly scalable food safety inspection solution without requiring additional physical hardware.**

Keywords* — *deep learning, nutrition, MBG, smart system, YOLOv11.

I. INTRODUCTION

The availability of nutritious and safe food is a crucial factor in supporting public health and the cognitive development of the younger generation through the Free Nutritious Meal (MBG) program [1]. The guarantee of safe, high-quality, and nutritious food availability is a national mandate strictly regulated in Law Number 18 of 2012 concerning Food [1]. However, current food quality monitoring in the field still heavily relies on conventional methods through manual organoleptic inspections. This method has major limitations as it is highly subjective, time-consuming, and not objectively measurable. The lack of adequate supervision risks failing the MBG program's objectives, considering that problematic or spoiled food products are still frequently found, which can negatively impact students' health [1].

To overcome these problems, artificial intelligence (AI) approaches have been widely explored as an alternative solution for food supervision [2]. Several previous studies have tested the use of biosensing instruments and gas sensor-based Electronic Nose (E-Nose) systems to detect spoiled food; however, this technology poses challenges due to the need for expensive physical hardware installation and its sensitivity to environmental conditions [3]. On the other hand, visual approaches based on Convolutional Neural Networks (CNN) have proven effective in detecting food spoilage from digital images with high accuracy [4]. Furthermore, Deep Learning models have also been successfully applied to identify food ingredients and estimate their nutritional content by mapping the detection results to a reference nutritional database [5]. This software-only visual solution offers a more practical, intelligent, and easily accessible inspection method in the field.

Based on the needs analysis and the limitations of conventional methods, this research aims to design and implement an efficient Deep Learning-based Smart System to support MBG operational activities. This system was developed using the lightweight and fast YOLOv11 object detection architecture to identify 12 main food classes that make up the MBG menu [6]. Specifically, the system is targeted to predict consumption feasibility (Fresh or Spoiled) while instantly calculating the estimation of four macronutrients (Calories, Protein, Fat, Carbohydrates) from visual image inputs. Through the implementation of this web-based application, it is hoped that field workers, kitchen staff, and school teachers will have an objective monitoring tool to ensure that the food served is truly safe and meets nutritional standards.

II. THEORETICAL REVIEW

This section presents a comprehensive theoretical framework and literature review related to the main research variables. The discussion focuses on the fundamental concepts of food safety and balanced nutrition, visual image-based Deep Learning methodologies, and a comparative performance analysis of the You Only Look Once (YOLO) architecture against predecessor physical sensor solutions.

A. Food Safety and Nutrition

The availability and intake of nutritious food are fundamental factors in supporting public health, mindset formation, and the cognitive development of early childhood [2]. A lack of nutritional intake, such as iron, risks causing anemia, which directly impacts a decrease in productivity [7].

On the other hand, guaranteeing food safety remains a major challenge due to the risk of food spoilage and contamination. Therefore, a monitoring system capable of objectively selecting food feasibility is needed to prevent health problems and ensure compliance with balanced nutrition standards [8].

B. Deep Learning and Convolutional Neural Network

The Artificial Intelligence (AI) approach based on Deep Learning algorithms, particularly Convolutional Neural Networks (CNN), has been widely applied to identify spoilage and distinguish food freshness through visual images [4], [9]. CNN models can automatically extract visual image features, mimicking how the human brain processes information [9]. This architecture generally works through a convolutional layer to scan basic image patterns, a pooling layer to summarize key feature information, and a fully connected layer that acts as a decision-maker to generate probability scores [9].

C. You Only Look Once (YOLO) Architecture

To overcome the computational inefficiency of traditional CNNs (sliding window method) in real-time detection, the You Only Look Once (YOLO) architecture is widely used [6]. YOLO is a highly efficient one-stage detector because it can process all image pixels with just a single evaluation step [6]. The YOLOv11 architecture implemented in this system has three main organs: the Backbone to extract visual features, the Neck to combine features from various object scales, and the Head to draw bounding boxes while simultaneously predicting class types and freshness status (Fresh or Spoiled) [6].

D. Nutritional Estimation and Visual Advantages

Besides detecting physical quality, Deep Learning serves as a foundation for automatically estimating food nutritional content [5], [10]. Smart systems can be trained to recognize the type of dish from a photo and then extract nutritional composition attributes (calories, carbohydrates, protein, and fat) by mapping them to a standard reference nutritional database [8], [10]. This visual software-based approach offers a more practical, intelligent, and efficient solution compared to physical hardware instruments like the Electronic Nose (E-Nose) which are generally hindered by complex sample preparation, environmental sensitivity, and high equipment costs [3].

III. METHOD

This research employs an experimental approach by designing and testing a client-server web application-based Smart System. The research stages include data collection, system design, and performance testing of the detection algorithm.

A. Data Source and Acquisition

The system integrates a Next.js-based frontend, a backend providing API services using FastAPI, a YOLOv11-based artificial intelligence engine, and data storage using Cloud Storage. Image data captured by the user's device is transmitted over HTTPS to the backend, where it is processed by the deep learning engine before the result is returned to the client and archived in the cloud database, as illustrated in Fig. 1.

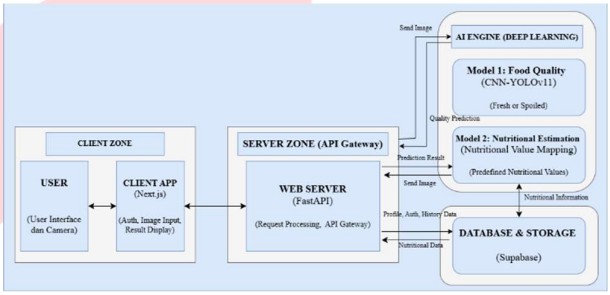

FIGURE 1
System System Block Architecture Diagram

The centralization of computation on the server side is intended to keep the client application responsive [11].

B. Dataset and Data Acquisition

The dataset consists of 2,400 self-collected food images, evenly balanced between the fresh and spoiled classes (1,200 images each), covering 12 primary food classes served in the MBG program (e.g., white rice, tofu, tempeh, chicken, fish, boiled egg, fried egg, carrot, stir-fried mustard greens, banana, apple, and orange). All images were uploaded, managed, and annotated (bounding-box labeling) using the Roboflow computer-vision platform to produce YOLO-format ground truth data. The dataset was then split into training (81.4%, 1,954 images), validation (9.3%, 223 images), and testing (9.3%, 223 images) subsets.

C. Deep Learning Model Design (YOLOv11)

The detection and classification model was developed using the YOLOv11 (You Only Look Once version 11) architecture, selected for its high computational efficiency without compromising accuracy [6]. The architecture consists of three main components operating in sequence: the Backbone, which extracts visual features (shape, color, texture) from the input image; the Neck, which fuses features across multiple scales so that both large and small food items are detected simultaneously; and the Head, which simultaneously predicts the bounding box, the food class, and the freshness status (Fresh/Spoiled) for each detected object, as shown in Fig. 2.

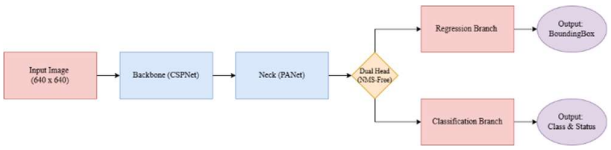

FIGURE 2
YOLOv11 Model Architecture

D. Nutrition Estimation Formula

Following object detection, macronutrient content is estimated through a lookup-table mapping to a standardized nutrition database. Total nutrients are calculated using (1).

$$N_{total} = \sum_{i=1}^{n} (w_i \times N_i) \quad (1)$$

where *N_total* is the total nutrient value (calories, protein, fat, or carbohydrate), *n* is the number of food items composing the portion, w_i is the estimated gross weight of ingredient *i* (based on Indonesian Household Measurement/URT standards [8], multiplied by its Edible Portion percentage), and N_i is the nutrient content per gram of ingredient *i* [10]. If an item is classified as *Spoiled*, its N_i is automatically set to zero as a food-safety mitigation measure.

E. Evaluation Metrics

Model performance was evaluated using object-detection and classification metrics [12], [13]. Detection correctness is determined using Intersection over Union (IoU) with a threshold of 0.5, calculated as in (2):

$$IoU = \frac{area(B_p \cap B_{gt})}{area(B_p \cup B_{gt})} \quad (2)$$

where B_p is the predicted bounding box and B_{gt} is the ground-truth bounding box. Based on the true positives (TP), false positives (FP), and false negatives (FN) obtained from the IoU threshold, the precision, recall, and F1-score are derived sequentially. Precision is defined as in (3):

$$Precision_i = \frac{TP_i}{TP_i + FP_i} \quad (3)$$

while the model's sensitivity or recall is formulated as in (4):

$$Recall_i = \frac{TP_i}{TP_i + FN_i} \quad (4)$$

To balance both precision and recall, the harmonic mean is computed using the F1-score formula as in (5):

$$F1_i = 2 \times \frac{Precision_i \times Recall_i}{Precision_i + Recall_i} \quad (5)$$

The overall multi-class object detection capability is comprehensively summarized using the mean Average Precision (mAP) metric, which has been established previously in the system design framework. Furthermore, freshness classification accuracy is evaluated using a confusion matrix. The overall accuracy rate is calculated as in (6):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

where TN represents True Negatives. Finally, system response time (latency) is measured as the difference between the response and request timestamps at the API gateway, formulated as in (7):

$$RT = t_{s(resp)} - t_{s(req)} \quad (7)$$

IV. RESULT AND DISCUSSION

This section presents the testing results of the implemented Deep Learning-based Smart System. The testing includes the performance evaluation of the object detection model (YOLOv11), the accuracy of the macronutrient estimation algorithm, response time (latency) testing, and the user interface implementation on the web application.

A. Detection and Classification Model Testing Results

The artificial intelligence modeling performance was evaluated using standard object detection metrics, namely Accuracy, Precision, Recall, F1-Score, and mean Average Precision (mAP) [12], [13]. In accordance with the initial data distribution, testing was conducted in two main stages: the validation stage and the final testing stage under real-world scenarios. The recapitulation of the metric testing results is presented in Table 1.

TABLE 1
Detection and Classification Test Results

Metric / Parameter	Validation Data	Testing Data
Number of Images	223 Images	223 Images
Accuracy (mAP@0.5)	95.3%	91.5%
Precision	94.1%	89.8%
Recall	90.0%	90.8%
F1-Score	92.0%	90.3%

Based on Table 1, the YOLOv11 model [6] demonstrates highly robust performance in simultaneously detecting and classifying 12 food classes. In the validation stage, the model successfully achieved an mAP@0.5 score of 95.3%, indicating a very high level of certainty in accurately predicting bounding boxes. When confronted with testing data (data previously unseen by the model), the mAP@0.5 value experienced a reasonable slight decrease to 91.5%. Despite the decrease, the Precision and Recall metric values remained stably balanced. This proves that the model did not experience overfitting and has excellent generalization capabilities when implemented to recognize food freshness conditions (fresh/spoiled) in the field [13].

In addition to numerical metric evaluation, the performance of the YOLOv11 algorithm was also visually validated through inference testing on lunch box images. The results of object localization detection and freshness classification are shown in Fig. 3.



FIGURE 3
YOLOv11 Inference Results on Food Images

In Fig. 3, it can be seen that the model is capable of performing bounding box regression with a very precise Intersection over Union (IoU) level against the actual objects. The model successfully recognizes various food classes simultaneously within a single image frame (such as rice, chicken, and vegetables) and labels each class with a classification accuracy level (freshness) represented by a high probability score above the threshold limit.

B. Nutritional Estimation Testing Results

Once the food is successfully detected and classified with a minimum confidence score of 50%, the system executes macronutrient calculations using the Black-Box Testing method. The testing was conducted by comparing the application's calculation output against manual calculation values referring to the Balanced Nutrition Guidelines standards [8]. Because this algorithm operates mathematically using a static variable in the form of the Edible Portion (BDD) percentage, the suitability level and accuracy of the nutritional estimation successfully achieved an absolute value of 100%. Furthermore, the system's safety mitigation functioned perfectly; whenever an object was detected with a "Spoiled" status, the system automatically canceled the calculation and converted all its nutritional content values to 0.

C. System Latency Testing Results

The operational effectiveness of the system in the field highly depends on the data processing speed. End-to-end latency testing was performed by calculating the time difference between the client sending an image request and receiving the response results from the API server [11]. The testing results show that the system is capable of performing the entire series of processes from image preprocessing, object detection, freshness classification, to nutritional estimation with an average response time of under 3 seconds per scan. This latency level is considered highly efficient and feasible for facilitating the massive and interactive inspection needs of the Free Nutritious Meal (MBG) program without significant delays.

D. System Interface Implementation

The integration results of the AI modeling and backend system are visualized through a software-only web application. The interface design is kept as concise as possible for easy operation by field supervisors. The main interface implementation results of the system are shown in Fig. 4 and Fig. 5.

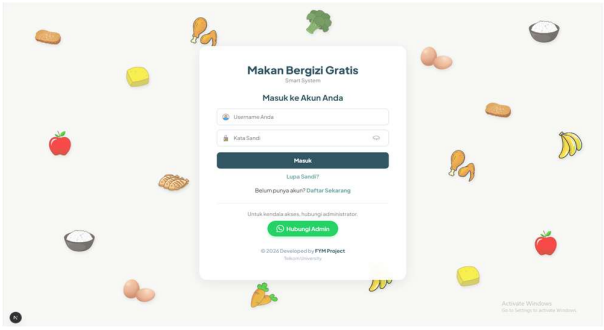

FIGURE 4
User Authentication (Login) Page

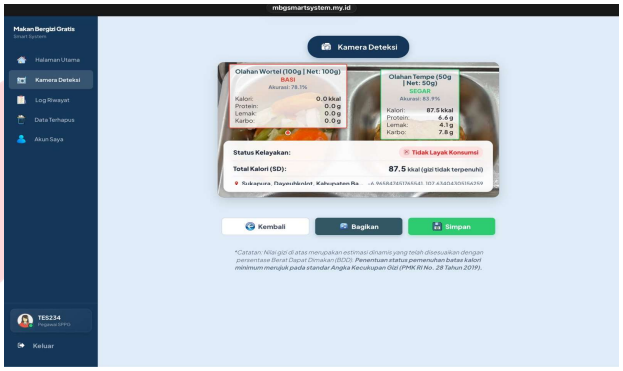

FIGURE 5
System Scanning and Nutrition Detection Results

Fig. 4 displays the authentication (login) page, which serves as the initial security gateway. This page is integrated with Row-Level Security (RLS) database protection to ensure that only authorized supervisors or personnel can access and submit inspection data to the central system. Furthermore, Fig. 5 represents the main functionality of the application. When the user scans a lunch box, the interface immediately highlights the food's location with a bounding box, displays the food name label along with its confidence level, and details the total nutritional estimation at the bottom of the screen.

V. CONCLUSION

This research successfully implemented a Deep Learning-based Smart System as a software-only web application to automate the feasibility and nutritional monitoring of the Free Nutritious Meal (MBG) program. The system, integrating the YOLOv11 architecture, proved robust with a mean Average Precision (mAP@0.5) of 95.3% on validation data and 91.5% on testing data. Furthermore, the nutritional calculation algorithm based on the Edible Portion (BDD) percentage achieved 100% mathematical accuracy, with a processing latency of under 3 seconds per scan. Although its field operation depends on internet connection stability and lighting conditions, this smart solution successfully provides an objective, fast, accurate, and highly scalable alternative for food safety inspection without requiring additional physical hardware.

REFERENCE

[1] Republik Indonesia, *Undang-Undang Republik Indonesia Nomor 18 Tahun 2012 tentang Pangan*. Jakarta: Badan Pemeriksa Keuangan RI, 2012.

[2] I. H. Sarker, "AI-based modeling: Techniques, applications and research issues towards automation, intelligent and smart systems," *SN Computer Science*, vol. 3, no. 2, Art. no. 158, 2022.

[3] W. N. F. S. W. Azman, K. N. F. K. Azir, and A. M. Khairuddin, "E-Nose: Spoiled food detection embedded device using machine learning for food safety application," in *ICOCI 2023*, CCIS 2001, Singapore: Springer Nature Singapore Pte Ltd, 2024, pp. 221-234.

[4] J. Nirmala Devi and V. R. Kiruthika, "Food spoilage alert system by deploying deep learning model," *International Journal of Health Sciences*, vol. 6, no. S1, pp. 8565-8575, 2022.

[5] Z. Shen, A. Shehzad, S. Chen, H. Sun, and J. Liu, "Machine learning based approach on food recognition and nutrition estimation," *Procedia Computer Science*, vol. 174, pp. 448-453, 2020.

[6] M. L. Ali and Z. Zhang, "The YOLO framework: A comprehensive review of evolution, applications, and benchmarks in object detection," *Computers*, vol. 13, no. 12, p. 336, 2024.

[7] N. Kristin, B. Jutomo, and D. L. A. Boeky, "Hubungan asupan zat gizi besi dengan kadar hemoglobin remaja putri," *SEHATRAKYAT (Jurnal Kesehatan Masyarakat)*, vol. 1, no. 3, pp. 189-195, 2022.

[8] Kementerian Kesehatan Republik Indonesia, Peraturan Menteri Kesehatan Republik Indonesia Nomor 41 Tahun 2014 tentang Pedoman Gizi Seimbang. Jakarta: Kementerian Kesehatan RI, 2014.

[9] R. K. Mega Lingam, G. S. Sree, G. M. Reddy, I. R. Sri Krishna, and L. U. Suriya, "Food spoilage detection using convolutional neural networks and K-means clustering," in *2019 3rd International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE)*, 2019, pp. 488-493.

[10] M. Al-Saffar and W. Baiee, "Nutrition information estimation from food photos using machine learning based on multiple datasets," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 5, pp. 2922-2929, 2022.

[11] A. Lawi, B. L. E. Panggabean, and T. Yoshida, "Evaluating GraphQL and REST API services performance in a massive and intensive accessible information system," *Computers*, vol. 10, no. 11, p. 138, 2021.

[12] R. Padilla, S. L. Netto, and E. A. B. da Silva, "A survey on performance metrics for object-detection algorithms," in *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 2020, pp. 237–242.

[13] A. Tharwat, "Classification assessment methods," *Applied Computing and Informatics*, vol. 17, no. 1, pp. 168–192, 2021.