UDC 004.9
doi:10.31799/1684-8853-2024-2-9-20
EDN: XSSHJI

Articles

Deep learning-based IoT system for fruit and vegetable recognition

***Pham Cong Thang*[a]**, *PhD, Lecturer, orcid.org/0000-0002-6428-102X, pcthang@dut.udn.vn*
***Nguyen Van Dung*[a]**, *Student, orcid.org/0009-0004-1282-184X*
***Le Tuan Nguyen Khoi*[a]**, *Student, orcid.org/0009-0001-1191-2502*
***Pham Duy Tin*[a]**, *Student, orcid.org/0009-0002-2314-1503*
***Tran Thi Thu Thao*[b]**, *M. Sc., Lecturer, orcid.org/0000-0001-7705-2405*
[a]*The University of Danang — University of Science and Technology, 54 Nguyen Luong Bang St., Da Nang 550000, Vietnam*
[b]*The University of Danang — University of Economics, 71 Ngu Hanh Son St., Da Nang 550000, Vietnam*

Introduction: *In the context of the increasing importance of vegetables and fruits as essential nutritional sources, the demand for advanced computer vision-based fruit recognition technology within the supply chain has surged. This technology is critical at various stages, including harvesting, grading, and quality control.* ***Purpose:*** *To develop advanced fruit recognition systems through the integration of IoT devices, such as cameras and sensors, with deep learning algorithms.* ***Results:*** *This research utilizes convolutional neural networks to enhance fruit recognition capabilities by capturing intricate image features. We feed images of vegetables and fruits into pre-trained deep learning models to extract their deep features. Among these models, ResNet152V2 is notable for its robustness against noise and distortions, which makes it suitable for real-world applications. Its scalability allows it to handle larger datasets and more complex tasks, consistently achieving high accuracy in recognizing fruits and vegetables, even under challenging conditions. The dataset comprises around 31,000 images, spanning a diverse array of fruits and vegetables. The training process utilizes techniques like GlobalAveragePooling2D, fully connected layers, dropout for overfitting prevention, and softmax activation, culminating in an impressive accuracy of 98.01% for ResNet152V2 after 20 epochs. This surpasses the performance of a basic convolutional neural network model, which achieves 88.3%. Notably, when deployed on mobile platforms and Raspberry Pi 4, identification times are recorded at 4.3 seconds and 2.25 seconds, respectively. Concurrently, we develop both application software and an Internet of Things hardware system to monitor and enhance the fruit and vegetable recognition process.* ***Practical relevance:*** *The research addressed the fruit and vegetable recognition challenge by employing the ResNet152V2 deep learning model along with a dataset sourced from online and field channels. It achieved high accuracy, swift time series prediction, and the development of cost-effective, durable Internet of Things application hardware.*

Keywords *— fruit and vegetable recognition, deep learning, machine learning, ResNet152V2.*

For citation: Pham C. T., Nguyen V. D., Le T. N. K., Pham D. T., Tran T. T. T. Deep learning-based IoT system for fruit and vegetable recognition. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 2, pp. 9–20. doi:10.31799/1684-8853-2024-2-9-20, EDN: XSSHJI

Introduction

Vegetables and fruits are important food sources for humans, providing essential nutrients such as vitamins, minerals, fiber, and antioxidants, which can positively impact human health [1–3]. In contemporary times, there has been a notable surge in the demand for fruit consumption, necessitating the adoption of modern production and processing techniques. Consequently, fruit recognition technology has emerged as a pivotal component within the fruit supply chain, finding applications at various stages. During the harvesting phase, it aids farmers in pinpointing the optimal time for harvest, thereby ensuring the superior quality of fruits. Moreover, during the grading phase, fruit recognition facilitates the categorization of fruits based on factors such as type, size, and ripeness, thus streamlining the transportation and storage processes. Finally, during the quality control phase, this technology plays a crucial role in identifying surface defects on fruits, thereby upholding food safety standards. The integration of deep learning techniques and robotic systems to automate agricultural processes has garnered considerable interest [4]. Fruits often thrive in complex environments fraught with uncertainty, making a robust fruit vision detection system imperative for smart agriculture and automated harvesting. The primary attributes of such a system include image sensors and fruit image data. Typically, fruit vision detection systems undergo five stages: image acquisition, preprocessing, feature extraction, segmentation, and recognition. Although still under development, these systems have the potential to revolutionize the agricultural industry. As the technology continues to improve, fruit vision detection systems are likely to become more widely used in agricultural applications [5, 6].

Additionally, the Internet of Things (IoT) is revolutionizing the detection and recognition of vegetables and fruits by integrating advanced sensors and data analytics to enhance accuracy and efficiency in agriculture. IoT technologies enable real-time monitoring of fruit crops, assessing factors such as ripeness, size, and health directly from the field. This data is invaluable for optimizing harvest times and improving crop management. IoT systems facilitate automated harvesting by guiding robotic systems to pick ripe fruits, thereby reducing labor costs and increasing operational efficiency. With the help of IoT, farmers can also detect early signs of disease and pest infestation, allowing for timely interventions to protect their crops. As IoT continues to evolve, its integration with artificial intelligence promises even greater advancements in automated fruit recognition and quality assessment, potentially transforming traditional farming practices into highly efficient, data-driven operations [7–9].

This research explores the utilization of convolutional neural networks (CNNs) to advance fruit recognition capabilities by acquiring intricate image features. We feed images of vegetables and fruits into pre-trained deep learning models to extract their deep features. Its scalability allows for handling larger datasets and more complex tasks, consistently achieving high accuracy in fruit and vegetable recognition, even under challenging conditions. Moreover, we evaluate the efficacy of deep features derived from various pre-trained deep learning models across different architectures for fruit and vegetable detection. Concurrently, we are developing both application software and an IoT hardware system to monitor and refine the fruit and vegetable recognition process.

Related work

The detection and recognition of fruits and vegetables is a multifaceted endeavor that necessitates the integration of diverse methodologies [10]. Fruit and vegetable recognition methodologies can broadly be categorized into two groups: feature-based recognition methods and machine learning-based recognition methods. Feature-based methods identify fruits and vegetables using geometric, color, or texture features [11–16]. In contrast, machine and deep learning-based methods employ models trained on datasets of labeled fruit images. Recent years have seen the proposal of various machine learning-based recognition methods [17, 18], including advanced deep learning techniques such as YOLO [19, 20], Single Shot Multibox Detector [21, 22], AlexNet [23, 24], VGG [25, 26], MobileNet [27, 28], ResNet [29, 30], and R-CNN [31, 32]. Although fruit vision detection systems are still under development, they have the potential to revolutionize the agricultural industry, with their usage likely to expand as the technology improves [33, 34].

Additionally, the integration of IoT and deep learning technologies has led to the development of advanced fruit recognition systems, marking a significant breakthrough in fruit processing and monitoring. At the core of this system lies the synergy between IoT devices, such as cameras and sensors, and deep learning algorithms [35, 36]. These devices capture real-time data from orchards or processing lines, providing a continuous flow of information. This data, which includes images, size measurements, color profiles, and other relevant features, is processed by the deep learning models. Trained on extensive datasets containing labeled fruit images and metadata, these algorithms learn intricate patterns and associations crucial for accurate fruit recognition and classification. Through iterative processes, the models are fine-tuned and optimized to achieve high levels of accuracy and efficiency.

The system's real-time capabilities enable instant recognition and classification of various fruit types, facilitating tasks such as sorting, quality control, and inventory management [37, 38]. Moreover, its scalability allows it to handle increasing volumes of data and adapt to evolving requirements in the fruit processing industry. By automating labor-intensive tasks and minimizing errors, this technology enhances operational efficiency while reducing costs. Furthermore, by ensuring consistent quality standards across different batches of fruits, it minimizes wastage and maximizes market value. Leveraging IoT capabilities, the system provides comprehensive traceability of fruits throughout the supply chain, addressing food safety concerns and ensuring compliance with regulatory standards [39, 40]. Stakeholders can remotely access and monitor the system's performance, receiving real-time updates on fruit processing activities, quality assessments, and inventory levels. Additionally, the system's predictive analytics capabilities offer valuable insights and forecasts regarding fruit yields, market demand, and supply chain optimization. Seamlessly integrating with existing enterprise resource planning systems, it facilitates data exchange and synchronization, streamlining business processes. A user-friendly interface allows operators to interact with the system, visualize key metrics, and configure parameters according to specific requirements. Through ongoing feedback loops and data-driven insights, the system undergoes continuous improvement, enhancing its accuracy, efficiency, and adaptability over time. The fruit recognition system utilizing IoT and deep learning technologies not only revolutionizes the fruit processing industry but also enhances productivity, quality assurance, and market competitiveness [41–46].

Materials and methods

Currently, the challenges in achieving high accuracy and speed in fruit recognition stem from several factors. These include the lack of quality assurance in fruit datasets, difficulties in detecting various types of fruit, limitations in the size and clarity of the target detection frame, and the use of lightweight fruit detection models. Collecting high-quality fruit datasets is a critically important task. These datasets act as foundational elements in training deep learning models, significantly influencing their post-training accuracy in fruit recognition. In the realm of deep learning-based fruit recognition, datasets must meet two essential criteria: sufficiency and diversity. Sufficiency means having an ample amount of data, while diversity refers to the variation in types of fruits and vegetables represented. Typically, fruit datasets are sourced from outdoor environments or online repositories. Moreover, it's crucial to ensure a proportional distribution of fruit and vegetable data during collection. Depending on the research focus, researchers select datasets that align with their specific requirements. Different datasets exhibit significant variations in the quantity, quality, and categories of images, as well as the types of fruits and vegetables included.

In this study, we utilized the Kaggle dataset, known for its popularity and extensive use. This dataset includes a total of 40 categories and a substantial number of images. However, it suffers from drawbacks such as single-image backgrounds, insufficient diversity, and category imbalances. Despite these limitations, the Kaggle dataset remains a valuable resource for research and development in fruit recognition tasks.

The dataset

The dataset for this study was aggregated from two sources: Fruit and vegetables (https://www.kaggle.com/datasets/shadikfaysal/fruit-and-vegetables-ssm) and Vegetable images (https://www.kaggle.com/datasets/misrakahmed/vegetable-image-dataset). It comprises a total of 31,000 images, which are divided into three subsets in an 8:1:1 ratio: the training dataset, the validation dataset, and the test dataset. Specifically, the training dataset includes 24,000 images, the validation dataset contains 3,500 images, and the test dataset also consists of 3,500 images (Fig. 1).

Since the data is compiled from various sources and the images are captured under unrealistic conditions, the pictures in reality will not always meet criteria such as angle, brightness, and size. Therefore, it is necessary to process the input data so that training the model can yield more accurate results reflective of real-world conditions.

Resize image: The data is compiled from various sources and varies in size. For model training, it's crucial that the input data adhere to a specific condition: the images must have matrix data of identical dimensions. In this case, we have chosen an image size of 224×224 for training the model. Using this size can reduce memory usage and increase computational efficiency, while still retaining essential information.

Image augmentation: In reality, images are always subject to varying lighting conditions and angles, so training should be supported by robust image data processing.

Rotation range: Randomly rotating images at various angles diversifies the data and helps the model generalize beyond the original dataset. This enhancement supports training across various angles and shooting techniques, improving the model's robustness.

Zoom range: Random zoom enables the creation of different versions of the same image, reducing the training model's dependence on a fixed size. This adjustment enhances both the accuracy and robustness of the model, more accurately reflecting real-world size variations.

Width shift range: In real-life images, objects may not always be centrally located but can appear in various positions. By randomly shifting the image horizontally, each shift creates a slightly offset version of the original, which helps the model better adapt to practical situations.

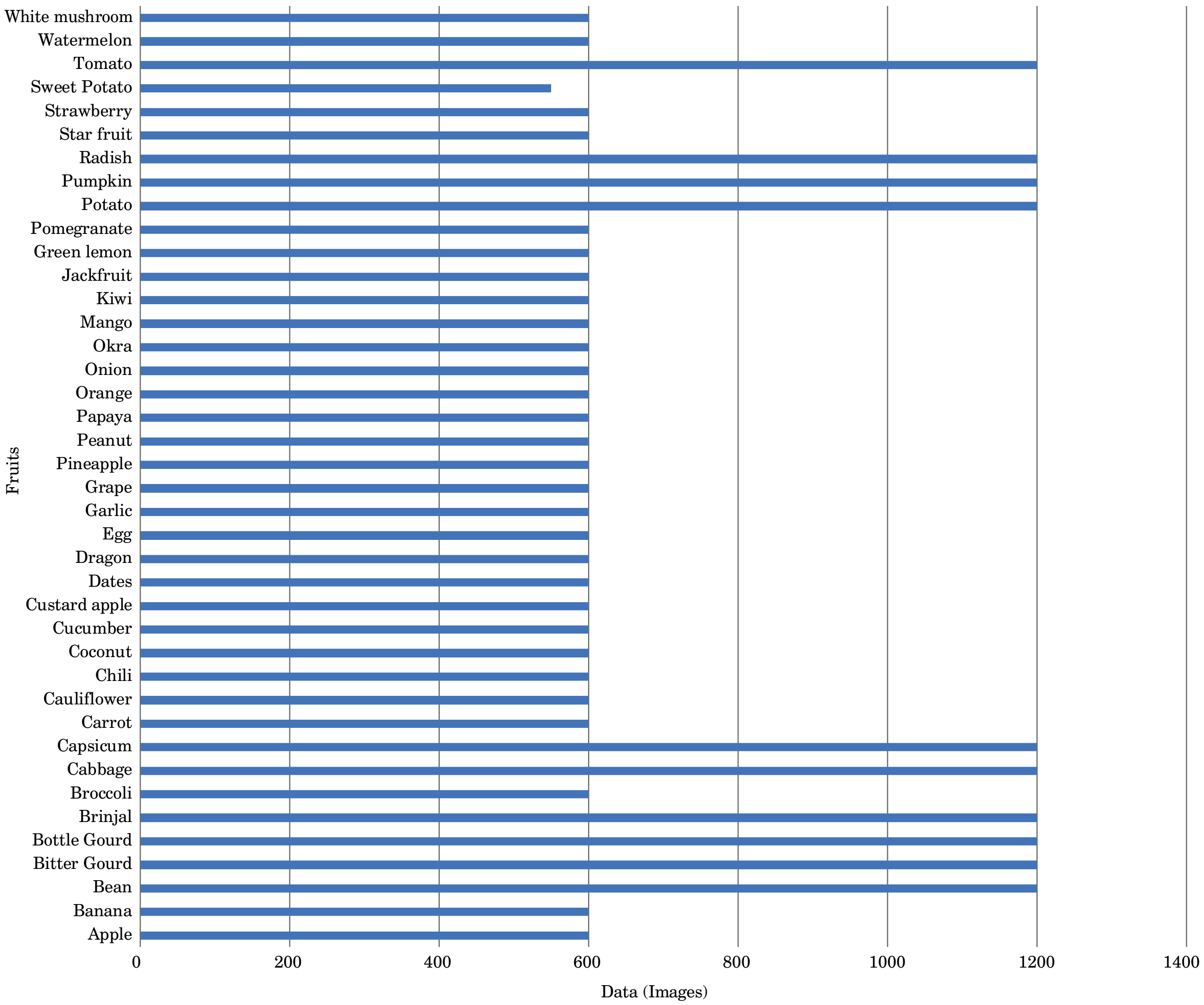
Brightness range: Randomly adjusting the brightness in images makes the data more representative of real-world conditions, thereby helping the model to become less dependent on specific lighting conditions.

Rescale: Rescaling pixel values from a range of 0–255 down to 0–1 stabilizes and enhances the efficiency of neural network training. This reduction helps prevent gradient saturation and explosion, stabilizes the backpropagation process, and accelerates the model's learning speed.

Deep learning model

In this work, we chose ResNet152V2 for data training [47]. ResNet152V2, an enhanced version of the original ResNet152, is a deep CNN architecture that features 152 layers (Fig. 2). It incorporates various improvements over its predecessor to enhance training convergence, reduce overfitting, and boost overall performance. Below is a detailed overview of the ResNet152V2 architecture.

Input layer: Accepts RGB images that are typically resized to a fixed dimension, such as 224×224 pixels.



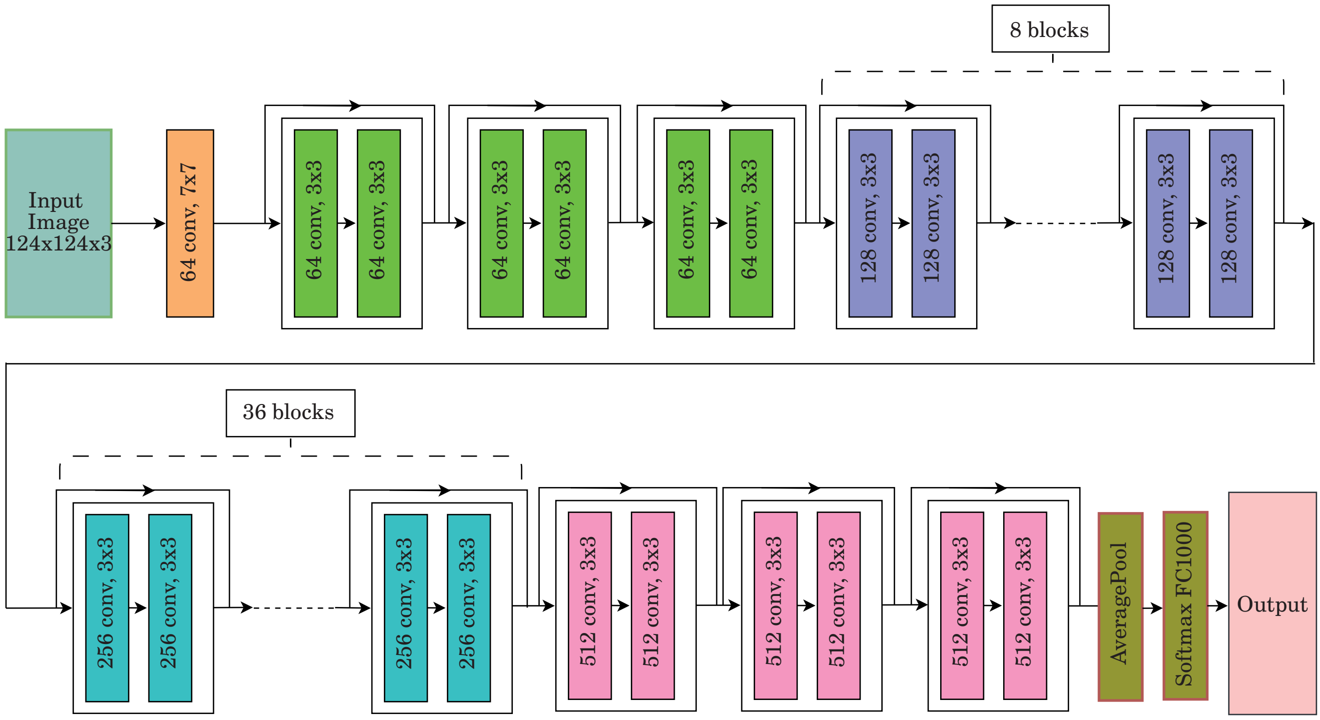
■ ***Fig. 1.*** Number of images for each fruit and vegetable

Initial convolutional layer: The input image undergoes an initial convolution; typically, this involves a 7×7 convolution with 64 filters, applied with a stride of 2. This step is followed by batch normalization and ReLU activation. After the initial convolution, a max-pooling layer with a 3×3 kernel and a stride of 2 is used to reduce the spatial dimensions of the feature maps. This sequence of steps marks the initial stage after the input layer, referred to as stage 1.

Residual blocks: ResNet152V2 consists of multiple residual blocks, each containing a stack of convolutional layers with skip connections [47]. These skip connections enable gradients to flow directly through the network, preventing them from vanishing. The residual blocks are grouped into stages, each containing a different number of blocks. The blocks within each stage have similar architectures but vary in the number of filters. In ResNet152V2, there are four stages, referred to as stages 2 to 5. The distribution of residual blocks across these stages is as follows: stage 2 includes 3 blocks with 64 filters; stage 3 has 8 blocks with 128 filters; stage 4 comprises 36 blocks with 256 filters; and stage 5 includes 3 blocks with 512 filters.

Global average pooling layer: After the final residual stage, global average pooling is applied to aggregate spatial information across the entire feature map. This operation effectively reduces the spatial dimensions to a single vector for each feature map, simplifying the output for further processing.

Fully connected layer: A fully connected layer, typically equipped with softmax activation, is added to neural network architectures for classification tasks. The number of neurons in this output layer corresponds directly to the number of classes in the



■ ***Fig. 2.*** Architecture of ResNet152V2

classification problem, providing a probability distribution across all possible outcomes.

Output layer: The output layer produces the final class predictions based on the softmax probabilities generated by the fully connected layer. This ensures that the output represents the likelihood of each class, allowing for the determination of the most probable category.

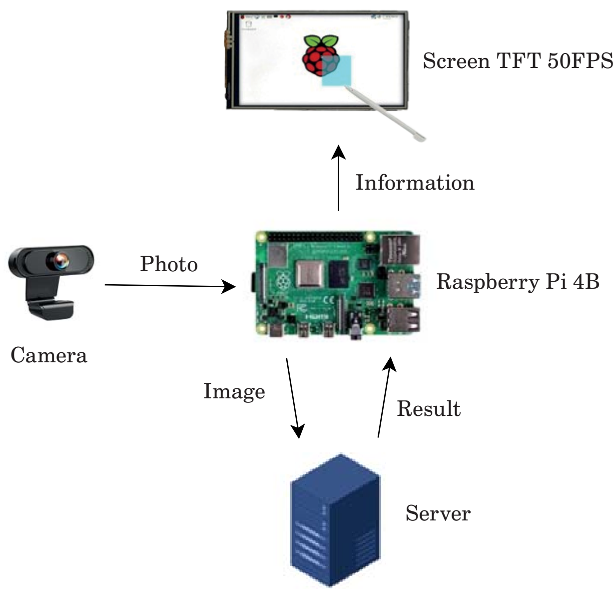
Overall, ResNet152V2's architecture enables the training of very deep neural networks while addressing issues such as vanishing gradients. It has demonstrated state-of-the-art performance on various image classification benchmarks and is widely used in research and applications that require high-performance computer vision models.

System design

We have developed a system to empirically evaluate the performance of a deep learning model in recognizing fruits and vegetables. The system consists of three main components: hardware, software, and an intermediary server, all of which are interconnected. A diagram of the system is depicted in Fig. 3.

Hardware: This setup includes a Raspberry Pi 4B board for processing, complemented by a screen and a camera. These components serve dual purposes: they not only facilitate the display of the user interface but also capture images for the recognition of vegetables and fruits. Furthermore, the hardware is designed to interact directly with the server, which enables data reception and streamlines the identification process.

Software: The software component is a mobile application developed using the React Native framework (JavaScript) [48]. The advantages of using React Native include the ability to develop applications for both iOS and Android from a single source code, significantly reducing development effort and costs. Additionally, applications built with React

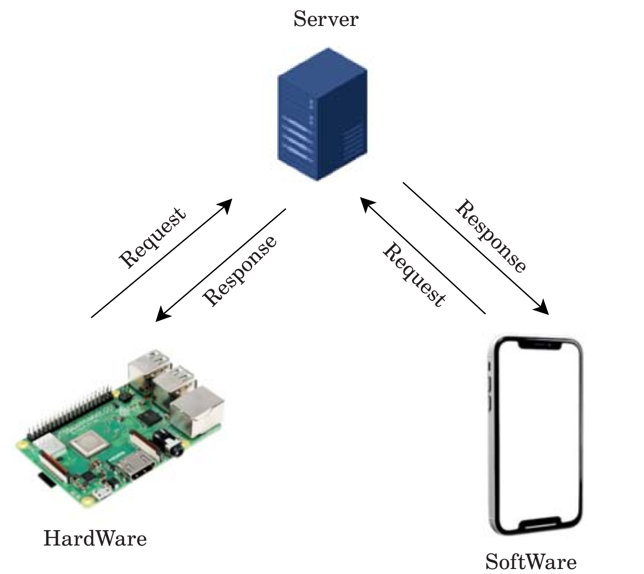


■ ***Fig. 3.*** The hardware circuit diagram

Native offer performance close to native apps due to optimizations and the ability to incorporate native components. The large and active React Native community provides extensive resources and support, aiding developers in promptly resolving issues. The framework's hot reloading feature allows developers to immediately see the effects of their changes, thereby streamlining the development and testing process (Fig. 4).

The software is directly linked to the server for data retrieval and for recognizing vegetables and fruits. Users can log into the system to perform fruit recognition, save recognized items or related dishes to their favorites list, and access their recognition history. Furthermore, our mobile application enables users to identify fruits and vegetables anywhere by capturing photos. Upon capturing an image, it is sent to the server for identification, and the server returns detailed information about the type of fruit or vegetable to the client. Each type of fruit or vegetable in the application is associated with a menu of dishes tailored to that specific product. This feature is designed to assist users in making informed dietary choices based on the identified fruits and vegetables.

Server: Our system utilizes two servers to optimize functionality. The first server, developed with NodeJS (JavaScript), supports user functionalities such as registration, login, storing recognition history, and maintaining records of recognized fruits/vegetables and preferred dishes. The second server, built with Flask (Python), focuses on the recognition of fruits and vegetables. For the mobile application component, we chose React Native (JavaScript) due to its numerous advantages over other programming languages. However, it is important to note that, despite its many benefits, we have encountered certain limitations during the development process.



■ ***Fig. 4.*** System diagram

In this study, we showcase an efficient application architecture that employs two servers simultaneously: one using NodeJS for database management and user connectivity, and another using Flask to host the recognition model and manage data prediction tasks. The integration of NodeJS and Flask offers significant advantages in developing web applications, particularly when addressing complex and diverse requirements. This dual-server setup allows for robust, scalable solutions that can efficiently handle both backend data management and real-time data processing tasks, as follows:

Performance and concurrent processing: Employing NodeJS for the database management server leverages its non-blocking architecture, optimizing performance and enabling concurrent processing when interacting with data from multiple sources. This capability is crucial, especially when prompt feedback from the database is essential for efficiently responding to user requests.

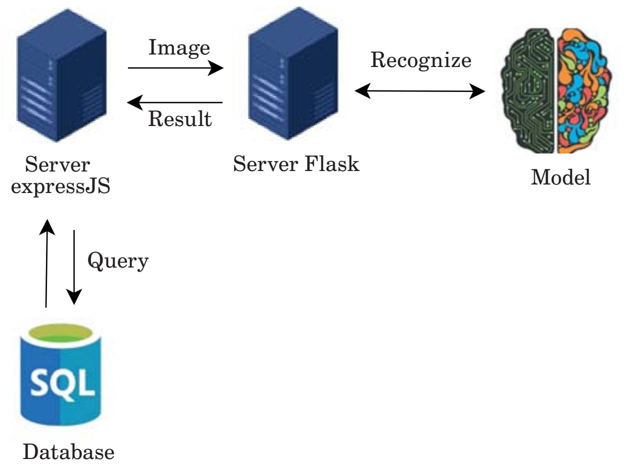
Work separation and flexibility in development: Using NodeJS for database management and Flask for hosting the model facilitates a natural separation of tasks. This division allows the development team to focus on database-related logic and data processing in NodeJS, while concurrently developing and maintaining the machine learning model in Flask.

Unified JavaScript and efficient interaction: Uniformity in using the JavaScript language across the NodeJS server and the browser simplifies data transmission and interaction between application components. This consistency enhances the efficiency of interactions between the database management server and the model-hosting server.

Fast feedback and real-time application: This architecture is ideal for applications requiring fast feedback and real-time processing. NodeJS, with its non-blocking architecture, excels in efficiently handling these requirements. Meanwhile, Flask's lightweight and flexible nature makes it convenient for deploying and maintaining the machine learning model.

Security and efficient resource management: The combination of NodeJS and Flask offers advantages in resource management and security. NodeJS, with its capability for efficient concurrent processing, helps alleviate pressure on system resources. Simultaneously, Flask's lightweight nature contributes to maintaining a high level of security.

In conclusion, the simultaneous use of NodeJS and Flask servers in our application architecture provides a robust solution for developing web applications with enhanced performance, clear task separation, and efficient interaction. This architecture is particularly advantageous for applications re-



■ ***Fig. 5.*** Server system architecture diagram

quiring real-time processing and aims to maintain a secure and resource-efficient system. The intermediary server architecture diagram is illustrated in Fig. 5.

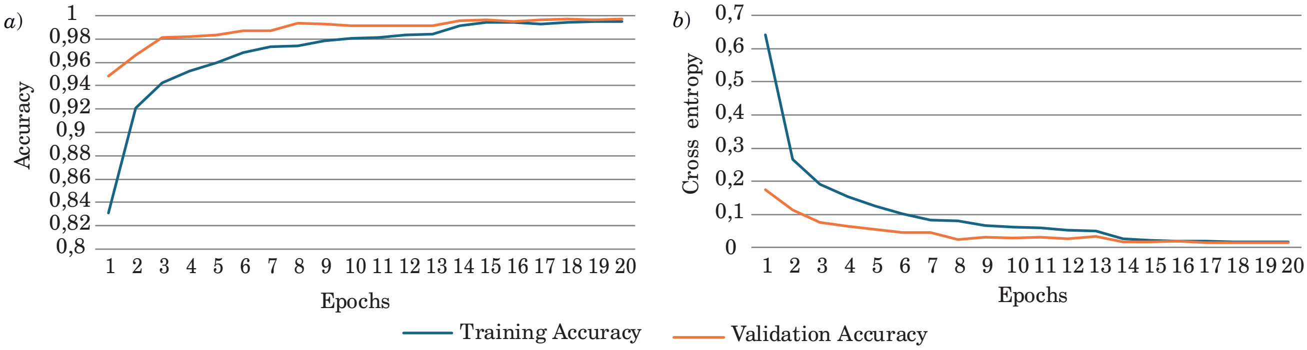
Results

Based on the designed system depicted in Fig. 3, the hardware configuration includes a Raspberry Pi 4B equipped with a 64-bit quad-core processor [49], two HDMI ports, two USB 2.0 ports, two USB 3.0 ports, one audio jack, and a MicroSD card slot. Additionally, it features a 5MP camera capable of capturing images at resolutions of 1080p and 720p, providing an image resolution of 2592×1944 pixels for 1080p and 1280×720 pixels for 720p. Complementing this is a 3.5-inch TFT screen with a resolution of 320×480 and a refresh rate of 50 FPS, integrated into the setup. The Raspberry Pi 4B connects to both the screen and the camera. Image data captured by the camera is transmitted to the Raspberry Pi 4B and displayed on the screen interface. Access to this interface is facilitated through a website hosted on a NodeJS server. Images are transmitted from the Raspberry Pi to the server for recognition, with the results subsequently displayed on the 50 FPS TFT screen. This setup enables real-time image recognition and display, making it suitable for various applications in scientific research and development (Fig. 6). Figure 7 shows the interface of the designed mobile application. The source code of the designed system can be accessed here (https://github.com/pacotha/DL_Based-IoT.git).

The deep learning model ResNet152V2 undergoes a training process consisting of approximately 20 epochs (Fig. 8, *a*, *b*), with a dropout rate of 0.2. This dropout rate leads to the disabling of neurons, resulting in the inevitable loss of information pertaining to each sample. Consequently, subsequent layers must rely on incomplete representations to formulate predictions. As a result of this loss, the training loss tends to increase, posing artificial challenges for the network in providing accurate answers. However, during the validation phase, all units remain accessible, allowing the network to leverage its complete computational capacity. This

■ ***Fig. 6.*** Hardware simulation of the system

■ ***Fig. 7.*** The interface of the designed mobile application



■ ***Fig. 8.*** Accuracy (*a*) and loss (*b*) after 20 epochs

■ ***Table 1.*** Comparison of optimizers and accuracy

Optimizer	Accuracy, %
Adam	95.3
RMSPROP	98.01

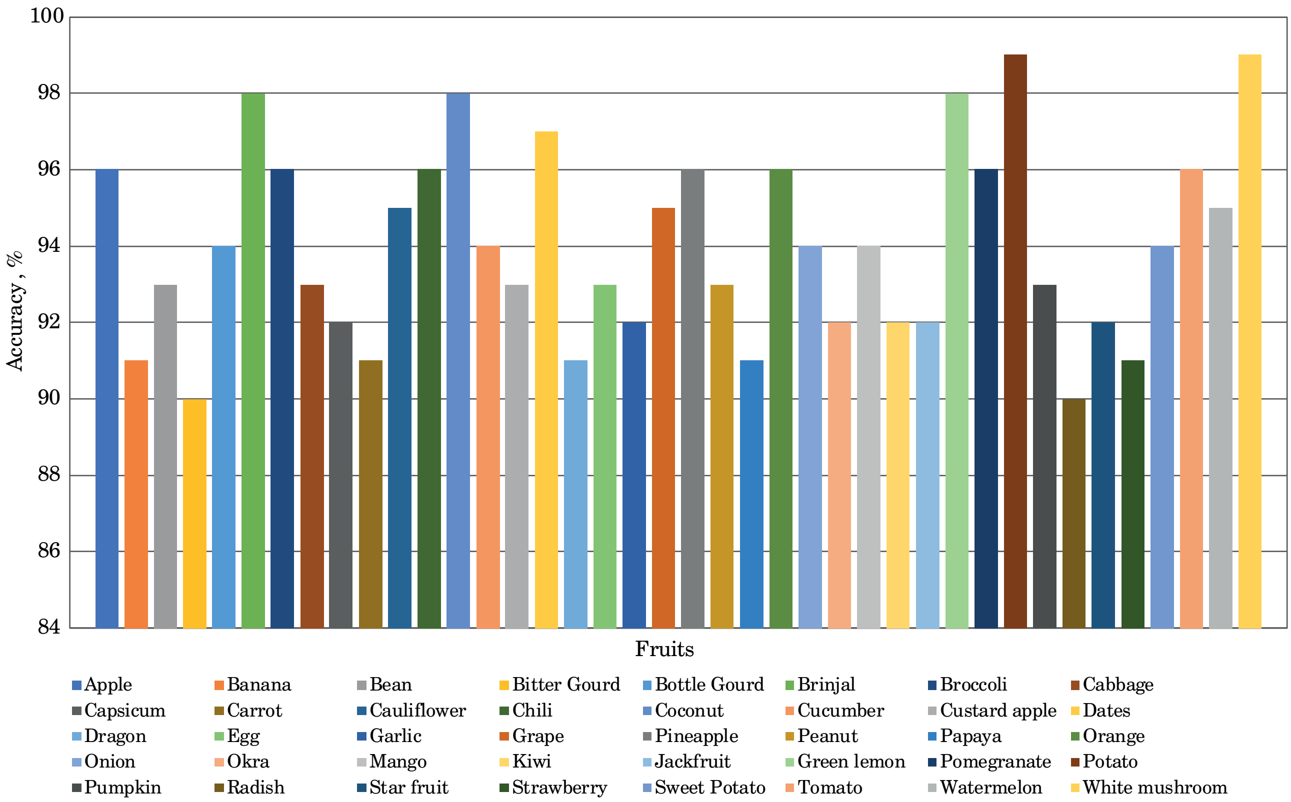
■ ***Table 2.*** Accuracy between CNN and ResNet152V2

Model	Accuracy, %
CNN	88.3
ResNet152V2	98.01

■ ***Table 3.*** Comparison of time to identify and produce results on two platforms

Platform	Time, s
Mobile	4.3
Raspberry Pi 4	2.25

enables the network to potentially exhibit superior performance during validation compared to training, as it can fully exploit its computational prowess without the hindrance of disabled neurons. Although the advanced 'Adam' optimizer may perform better on many tasks, it yields accuracy and



■ ***Fig. 9.*** Prediction accuracy of each fruit and vegetable

■ ***Fig. 10.*** Fruit and vegetable recognition result interface

recognition results that are not as close to reality for the model trained on our dataset (Table 1). Given the simplicity of the problem and the manageable size of the dataset, the 'rmsprop' optimizer yields better results and more appropriate recognition outcomes. We use the Categorical Crossentropy loss function, which is well-suited for multiclass classification problems. This approach is particularly apt for our large variety of fruits and vegetables, as it effectively prioritizes improving recognition rates for those categories with lower accuracy, and is compatible with the Softmax function [50].

During the training process, we employ callback functions such as Reduce Learning Rate, Model Checkpoint, and Early Stopping. These functions monitor the model's performance and halt training when they detect minimal or no improvement in accuracy across iterations, ensuring that only the best results are retained.

After a training process spanning approximately 20 epochs, the model achieved an accuracy of up to 98.01%. The prediction accuracy for each fruit and vegetable is illustrated in Fig. 9.

We also utilized a basic CNN model for identifying fruits and vegetables, although it yielded lower accuracy compared to ResNet152V2, as shown in Table 2. All models were trained on a P100 GPU, with the training process taking over three hours to complete. The time required by the model to identify and produce results on two different platforms is detailed in Table 3.

We also conducted extensive surveys on various types of fruits available in the market using the systems we designed. The results indicate that the majority of fruits and vegetables in the database set can be successfully identified, providing accurate results from various perspectives. Figure 10 illustrates the recognition results from the designed system.

Conclusions

Ensuring the high quality of fruit and vegetable datasets is paramount, especially for detecting real-world environmental conditions such as small targets within frames, low-light situations, blurriness, and obstructions. The ability to detect a wide variety of fruits and vegetables is also crucial in today's context. These aspects require careful attention when constructing and developing deep learning models for fruit and vegetable identification. Both the quality of the dataset and the architecture of the utilized model play pivotal roles, significantly impacting the performance of fruit and vegetable recognition systems.

In this study, we employed the ResNet152V2 model for fruit and vegetable recognition tasks, leveraging deep learning-based features. The preprocessing phase involved resizing, zooming, rotating, and shifting fruit and vegetable images to enrich and enhance the dataset's realism. Fruit and vegetable recognition achieved a peak accuracy of 98.01% across various field conditions. Furthermore, we developed a mobile application system featuring an intuitive, aesthetically pleasing, and fully functional interface. This system is supported by a stable and high-speed server infrastructure. Additionally, we engineered cost-effective, durable, compact, and portable IoT application hardware. Continuous improvements have been made to the system's database to enhance operational efficiency, accelerate data processing, and elevate accuracy levels.

Acknowledgments

Pham Cong Thang (corresponding author) would like to thank his colleagues at the Faculty of Information Technology, DUT, for their valuable feedback. The authors appreciate and are grateful to the reviewers and editors for their insightful comments and suggestions.

References

1. Aular J., Natale W. Mineral nutrition and fruit quality of some tropical fruit: guava, mango, banana, and papaya. *Revista Brasileira de Fruticultura*, 2013, no. 35, pp. 1214–1231. doi:10.1590/S0100-29452013000400033
2. Kumoro A. C., Alhanif M., Wardhani, D. A. Critical review on tropical fruits seeds as prospective sources of nutritional and bioactive compounds for functional foods development: A case of Indonesian exotic fruits. *International Journal of Food Science*, 2020, vol. 2020, pp. 1–15. doi:10.1155/2020/4051475
3. Hsouna A. B., Sadaka C., Mekinić I. G., Garzoli S., Švarc-Gajić J., Rodrigues F., Morais S., Moreira M. M., Ferreira E., Spigno G., Brezo-Borjan T., Akacha B. B., Saad R. B., Delerue-Matos C., Mnif W. The chemical variability, nutraceutical value, and food-industry and cosmetic applications of citrus plants: A critical review. *Antioxidants*, 2023, vol. 12, iss. 2, pp. 1–37. doi:10.3390/antiox12020481
4. Xiao F., Wang H., Xu Y., Zhang R. Fruit detection and recognition based on deep learning for automatic harvesting: An overview and review. *Agronomy*, 2023, vol. 13, iss. 6, pp. 1–32. doi:10.3390/agronomy13061625
5. Ukwuoma C. C., Zhiguang Q., Heyat Md B. B., Ali L., Almaspoor Z., Monday H. N. Recent advancements in fruit detection and classification using deep learning techniques: A critical review. *Mathematical Problems in Engineering*, 2022, vol. 2022, pp. 1–29. doi:10.1155/2022/9210947
6. Ismail N., Malik O. A. Real-time visual inspection system for grading fruits using computer vision and deep learning techniques. *Information Processing in Agriculture*, 2022, vol. 9, iss. 1, pp. 24–37. doi:10.1016/j.inpa.2021.01.005
7. Ray P., Pradhan S., Sharma R., Rasaily A., Swaraj A., Pradhan A. IoT based fruit quality measurement system. *Intern. Conf. on Green Engineering and Technologies (IC-GET)*, 2016, pp. 1–5. doi:10.1109/GET.2016.7916620
8. Gawas A., et al. E-fresh: Computer vision and IoT framework for fruit freshness detection. *Intern. Conf. on Advances in Computing, Communication, and Control (ICAC3)*, 2021, pp. 1–6. doi:10.1109/ICAC353642.2021.9697306
9. Behera S. K., Sethy P. K., Sahoo S. K., Panigrahi S., and Rajpoot S. C. On-tree fruit monitoring system using IoT and image analysis. *Concurrent Engineering*, 2021, vol. 29, iss. 1, pp. 6–15. doi:10.1177/1063293X20988395
10. Gupta S., Tripathi A. K. Fruit and vegetable disease detection and classification: Recent trends, challenges, and future opportunities. *Engineering Applications of Artificial Intelligence*, 2024, no. 133, pp. 1–30. doi:10.1016/j.engappai.2024.108260
11. Jana S., Basak S., Parekh R. Automatic fruit recognition from natural images using color and texture features. *Devices for Integrated Circuit (DevIC)*, 2017, pp. 620–624. doi:10.1109/DEVIC.2017.8074025
12. Lv J., Zhao D. A., Wei J., Ding S. Recognition of apple fruit in natural environment. *Optik*, 2016, vol. 127, iss. 3, pp. 1354–1362. doi:10.1016/j.ijleo.2015.10.177
13. Seng W. C., Mirisaee S. H. A new method for fruits recognition system. *Intern. Conf. on Electrical Engineering and Informatics*, 2009, pp. 130–134. doi:10.1109/ICEEI.2009.5254804
14. Shakil R. Addressing agricultural challenges: An identification of best feature selection technique for dragon fruit disease recognition. *Array*, 2023, vol. 20, 100326, pp. 1–9. doi:10.1016/j.array.2023.100326
15. Wu G., Li B., Zhu Q., Huang M., Guo Y. Using color and 3D geometry features to segment fruit point cloud and improve fruit recognition accuracy. *Computers and Electronics in Agriculture*, 2020, vol. 174, 105475, pp. 1–8. doi:10.1016/j.compag.2020.105475
16. Rachmawati E., Supriana I., Khodra M. L., Firdaus F. Integrating semantic features in fruit recognition based on perceptual color and semantic templat. *Information Processing in Agriculture*, 2022, vol. 9, iss. 2, pp. 316–334. doi:10.1016/j.inpa.2021.02.004
17. Zaki N., Singh H., Krishnan A., Alnaqbi A., Alneyadi S., Alnaqbi S., Alhindaassi S., Alam M., Eldin A. K. Transfer learning and explainable artificial intelligence enhance the classification of date fruit varieties. *International Conf. on Innovations in Information Technology (IIT)*, 2023, pp. 222–227. doi:10.1109/IIT59782.2023.10366495
18. Gill H. S., Murugesan G., Mehbodniya A., Sajja G. S., Gupta G., Bhatt A. Fruit type classification using deep learning and feature fusion. *Computers and Electronics in Agriculture*, 2020, vol. 211, 107990, pp. 1–6. doi:10.1016/j.compag.2023.107990
19. Chen J., Liu H., Zhang Y., Zhang D., Ouyang H., Chen X. A multiscale lightweight and efficient model based on YOLOv7: Applied to citrus orchard. *Plants*, 2022, vol. 11, 3260, pp. 1–17. doi:10.3390/plants11233260
20. Bai Y., Yu J., Yang S., Ning J. An improved YOLO algorithm for detecting flowers and fruits on strawberry seedlings. *Biosystems Engineering*, 2024, vol. 237, pp. 1–12. doi:10.1016/j.biosystemseng.2023.11.008
21. Qiang J., Liu W., Li X., Guan P., Du Y., Liu B., Xiao G. Detection of citrus pests in double backbone network based on single shot multibox detector. *Computers and Electronics in Agriculture*, 2023, vol. 212, 108158, pp. 1–11. doi:10.1016/j.compag.2023.108158
22. Liu W., et al. SSD: Single shot MultiBox detector. Computer vision — ECCV. *Springer International Publishing*, 2016, pp. 21–37. doi:10.1007/978-3-319-46448-0_2
23. Al-Hami M., Pietron M., Casas R., Wielgosz M. Methodologies of compressing a stable performance convolutional neural networks in image classification. *Neural Processing Letters*, 2020, vol. 51, pp. 105–127. doi:10.1007/s11063-019-10076-y

24. Jiang B., He J., Yang S., Fu H., Li T., Song H., He D. Fusion of machine vision technology and AlexNet-CNNs deep learning network for the detection of postharvest apple pesticide residues. *Artificial Intelligence in Agriculture*, 2019, vol. 1, pp. 1–8. doi:10.1016/j.aiia.2019.02.001
25. Li Z. Vegetable recognition and classification based on improved VGG deep learning network model. *International Journal of Computational Intelligence Systems*, 2020, vol. 13, iss. 1, pp. 559–564. doi:10.2991/ijcis.d.200425.001
26. Yang H., Ni J., Gao J., Han Z., Luan T. A novel method for peanut variety identification and classification by improved VGG16. *Scientific Reports*, 2021, vol. 11, pp. 1–17. doi:10.1038/s41598-021-95240-y
27. Xiang Q., Wang X., Li R., Zhang G., Lai J., Hu Q. Fruit image classification based on MobileNetV2 with transfer learning technique. *Intern. Conf. on Computer Science and Application Engineering*, 2019, pp. 1–7. doi:10.1145/3331453.3361658
28. Jain P., Chawla P., Masud M., Mahajan S., Pandit A. K. Automated identification algorithm using CNN for computer vision in smart refrigerators. *Computers, Materials and Continua*, 2022, vol. 71, no. 2, pp. 3337–3353. doi:10.32604/cmc.2022.023053
29. Yang Y., Wang L., Huang M., Zhu Q., Wang R. Polarization imaging based bruise detection of nectarine by using ResNet-18 and ghost bottleneck. *Postharvest Biology and Technology*, 2022, vol. 189, 111916, pp. 1–11. doi:10.1016/j.postharvbio.2022.111916
30. Buyukarikan B., Ulker E. Classification of physiological disorders in apples fruit using a hybrid model based on convolutional neural network and machine learning methods. *Neural Computing and Applications*, 2022, vol. 34, pp. 16973–16988. doi:10.1007/s00521-022-07350-x
31. Yao N., Ni F., Wu M., Wang H., Li G., Sung W.-K. Deep learning-based segmentation of peach diseases using convolutional neural network. *Frontiers in Plant Science*, 2022, vol. 13, pp. 1–14. doi:10.3389/fpls.2022.876357
32. Min W. Vision-based fruit recognition via multi-scale attention CNN. *Computers and Electronics in Agriculture*, 2023, vol. 210, 107911, pp. 1–11. doi:10.1016/j.compag.2023.107911
33. Hadipour-Rokni R., Askari Asli-Ardeh E., Jahanbakhshi A., Esmaili paeen-Afrakoti I., Sabzi S. Intelligent detection of citrus fruit pests using machine vision system and convolutional neural network through transfer learning technique. *Computers in Biology and Medicine*, 2023, vol. 155, 1–8. doi:10.1016/j.compbiomed.2023.106611
34. Azadnia R., Fouladi S., Jahanbakhshi A. Intelligent detection and waste control of hawthorn fruit based on ripening level using machine vision system and deep learning techniques. *Results in Engineering*, 2023, vol. 17, 100891, pp. 1–13. doi:10.1016/j.rineng.2023.100891
35. Rajak P., Ganguly A., Adhikary S., Bhattacharya S. Internet of things and smart sensors in agriculture: Scopes and challenges. *Journal of Agriculture and Food Research*, 2023, vol. 14, pp. 1–13. doi:10.1016/j.jafr.2023.100776
36. Kasera R. K., Gour S., Acharjee T. A comprehensive survey on IoT and AI based applications in different pre-harvest, during-harvest and post-harvest activities of smart agriculture. *Computers and Electronics in Agriculture*, 2024, vol. 216, pp. 1–24. doi:10.1016/j.compag.2023.108522
37. Wason R., Choudhary P., Tomar A., Arora D. A novel, low-cost, smart IoT based framework for fruit and vegetable quality detection during transit in India. *International Journal of Information Technology*, 2023, vol. 15, pp. 1509–1519. doi:10.1007/s41870-023-01177-y
38. Nirale P., Madankar M. Design of an IoT based ensemble machine learning model for fruit classification and quality detection. *Intern. Conf. on Emerging Trends in Engineering and Technology — Signal and Information Processing*, 2022, pp. 1–6. doi:10.1109/ICETET-SIP-2254415.2022.9791718
39. Rebelo R. M. L., Pereira S. C. F., Queiroz M. M. The interplay between the Internet of things and supply chain management: Challenges and opportunities based on a systematic literature review. *Benchmarking: An International Journal*, 2022, vol. 29, iss. 2, pp. 683–711. doi:10.1108/BIJ-02-2021-0085
40. Taj S., Imran A. S., Kastrati Z., Daudpota S. M., Memon R. A., Ahmed J. IoT-based supply chain management: A systematic literature review. *Internet of Things*, 2023, vol. 24, 100982, pp. 1–25. doi:10.1016/j.iot.2023.100982
41. Mishra S., Khatri S. K., Johri P. IoT based automated quality assessment for fruits and vegetables using infrared. *Intern. Conf. on Information Systems and Computer Networks (ISCON)*, 2019, pp. 134–138. doi:10.1109/ISCON47742.2019.9036165
42. Putra B. T. W., Indracahyana K. S., Fanshuri B. A. Development of a handheld IoT-based fruit harvester to support agrotourism. *Microprocessors and Microsystems*, 2022, vol. 91, pp. 1–7. doi:10.1016/j.micpro.2022.104550
43. Kavitha R., Shanthi T., Maithili P., Roopika J., Naveen Kumar K., Saravanakumar K. Deep learning and Internet of Things based detection of diseases and prediction of pesticides in fruits. *Intern. Conf. on Trends in Electronics and Informatics (ICOEI)*, 2023, pp. 1133–1140. doi:10.1109/ICOEI56765.2023.10125946
44. Nasir I. M., Bibi A., Shah J. H., Khan M. A., Sharif M., Iqbal K., Nam Y., Kadry S. Deep learning-based classification of fruit diseases: An application for precision agriculture. *Computers, Materials & Continua*, 2021, vol. 66, no. 2, pp. 1949–1962. doi:10.32604/cmc.2020.012945
45. Hayajneh A. M., Batayneh S., Alzoubi E., Alwedyan M. TinyML olive fruit variety classification by means of

convolutional neural networks on IoT Edge devices. *AgriEngineering*, 2023, vol. 5, iss. 4, pp. 2266–2283. doi:10.3390/agriengineering5040139

46. Augustin A., Kiliroor C. C. (2024) IoT-based pesticide detection in fruits and vegetables using hyperspectral imaging and deep learning. *Cognitive Computing and Cyber Physical Systems, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, Springer, Cham, 2023, vol. 536, pp. 74–83. doi:10.1007/978-3-031-48888-7_6

47. He K., Zhang X., Ren S., Sun J. *Identity Mappings in Deep Residual Networks*. In: *Computer Vision — ECCV 2016. ECCV 2016*. Lecture Notes in Computer Science, 2016, vol. 9908, pp. 630–645. doi:10.1007/978-3-319-46493-0_38

48. *React Native*. Available at: https://reactnative.dev (accessed 30 January 2024).

49. *Raspberry Pi 4*. Available at: https://www.raspberrypi.org/products/raspberry-pi-4-model-b (accessed 30 January 2024).

50. Murphy K. P. *Probabilistic Machine Learning: An Introduction. Adaptive Computation and Machine Learning series*. The MIT Press, 2020. 864 p.

УДК 004.9
doi:10.31799/1684-8853-2024-2-9-20
EDN: XSSHJI

Система интернета вещей на основе глубокого обучения для распознавания фруктов и овощей

Фам Конг Тханг[а], PhD, преподаватель, orcid.org/0000-0002-6428-102X, pcthang@dut.udn.vn
Нгуен Ван Зунг[а], студент, orcid.org/0009-0004-1282-184X
Ле Туан Нгуен Хой[а], студент, orcid.org/0009-0001-1191-2502
Фам Зуи Тин[а], студент, orcid.org/0009-0002-2314-1503
Чан Тхи Тху Тхао[б], магистр, преподаватель, orcid.org/0000-0001-7705-2405
[а]Университет Дананга — Университет науки и техники, Нгуен Лыонг Банг, 54, Дананг, 550000, Вьетнам
[б]Университет Дананга — Университет экономики, Нгу Хань Шон, 71, Дананг, 550000, Вьетнам

Введение: в контексте растущей важности овощей и фруктов как значимых источников питания возрастает спрос на передовую технологию их распознавания на основе компьютерного зрения в цепочке поставок. Эта технология имеет решающее значение на различных этапах, включая сбор урожая, сортировку и контроль качества. **Цель:** разработать передовые системы распознавания фруктов путем интеграции устройств интернета вещей, таких как камеры и датчики, с алгоритмами глубокого обучения. **Результаты:** использованы сверточные нейронные сети для усовершенствования возможностей распознавания фруктов путем извлечения сложных особенностей изображений, для чего изображения овощей и фруктов загружались в предварительно обученные модели глубокого обучения. Среди этих моделей ResNet152V2 отличается устойчивостью к шуму и искажениям, что делает ее подходящей для реальных приложений. Ее масштабируемость позволяет ей обрабатывать большие наборы данных и более сложные задачи, постоянно достигая высокой точности распознавания изображений даже в сложных условиях. Набор данных, полученный из онлайн-каналов и полевых каналов, включал около 31 000 изображений фруктов и овощей. В процессе обучения использовались такие методы, как GlobalAveragePooling2D, полностью связанные слои, отсев для предотвращения переобучения и активация softmax, что привело к впечатляющей точности 98,01 % для ResNet152V2 после 20 эпох по сравнению с производительностью базовой модели сверточной нейронной сети 88,3 %. Примечательно, что при развертывании на мобильных платформах и Raspberry Pi 4 время идентификации составило 4,3 и 2,25 с соответственно. Одновременно разработано прикладное программное обеспечение и аппаратная система интернета вещей для мониторинга и улучшения процесса распознавания фруктов и овощей. **Практическая значимость:** решена проблема распознавания фруктов и овощей с помощью модели глубокого обучения ResNet152V2 и набора данных. Достигнуты высокая точность и быстрое прогнозирование, разработано экономичное и надежное оборудование для приложений интернета вещей.

Ключевые слова — распознавание фруктов и овощей, глубокое обучение, машинное обучение, ResNet152V2.

Для цитирования: Pham C. T., Nguyen V. D., Le T. N. K., Pham D. T., Tran T. T. T. Deep learning-based IoT system for fruit and vegetable recognition. *Информационно-управляющие системы*, 2024, № 2, с. 9–20. doi:10.31799/1684-8853-2024-2-9-20, EDN: XSSHJI

For citation: Pham C. T., Nguyen V. D., Le T. N. K., Pham D. T., Tran T. T. T. Deep learning-based IoT system for fruit and vegetable recognition. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2024, no. 2, pp. 9–20. doi:10.31799/1684-8853-2024-2-9-20, EDN: XSSHJI