# Exploring ResNet-18 Estimation Design through Multiple Implementation Iterations and Techniques in Legacy Databases

Nuntachai Thongpance [1], Pareena Dangyai [2], Kittipan Roongprasert [3*], Anantasak Wongkamhang [4],
Ratchanee Saosuwan [5], Rawiphon Chotikunnan [6], Pariwat Imura [7], Anuchit Nirapai [8],
Phichitphon Chotikunnan [9], Manas Sangworasil [10], Anuchart Srisiriwat [11]

[1, 3, 4, 5, 6, 7, 8, 9, 10] College of Biomedical Engineering, Rangsit University, Pathum Thani, Thailand
[2] Saipanyarangsit School, Pathum Thani, Thailand
[11] Department of Electrical Engineering, Pathumwan Institute of Technology, Bangkok, Thailand
Email: [1] nuntachai.t@rsu.ac.th, [2] pareenaginnie@gmail.com, [3] kittipan.r@rsu.ac.th, [4] anantasak.w@rsu.ac.th,
[5] ratchanee.sa@rsu.ac.th, [6] rawiphon.c@rsu.ac.th, [7] pariwat.i@rsu.ac.th, [8] anuchit.ni@rsu.ac.th,
[9] phichitphon.c@rsu.ac.th, [10] manas.s@rsu.ac.th, [11] anuchart@pit.ac.th
*Corresponding Author

*Abstract*—In a rapidly evolving landscape where automated systems and database applications are increasingly crucial, there is a pressing need for precise and efficient object recognition methods. This study contributes to this burgeoning field by examining the ResNet-18 architecture, a proven deep learning model, in the context of fruit image classification. The research employs an elaborate experimental setup featuring a diverse fruit dataset that includes Rambutan, Mango, Santol, Mangosteen, and Guava. The efficacy of single versus multiple ResNet-18 models is compared, shedding light on their relative classification accuracy. A unique aspect of this study is the establishment of a 90% decision threshold, introduced to mitigate the risk of incorrect classification. Our statistical analysis reveals a significant performance advantage of multiple ResNet-18 models over single models, with an average improvement margin of 15%. This finding substantiates the study's central hypothesis. The implemented 90% decision threshold is determined to play a pivotal role in augmenting the system's overall accuracy by minimizing false positives. However, it's worth noting that the increased computational complexity associated with deploying multiple models necessitates further scrutiny. In sum, this study provides a nuanced evaluation of single and multiple ResNet-18 models in the realm of fruit image classification, emphasizing their utility in practical, real-world applications. The research opens avenues for future exploration by refining these methodologies and investigating their applicability to broader object recognition tasks.

*Keywords*—*ResNet-18; Multiple ResNet-18; Nutritional Food.*

## I. INTRODUCTION

Nutritional food systems are a cornerstone of human health, offering insights into daily nutrient requirements [1]-[2]. Fruits, an essential part of these systems, are a rich source of dietary fiber, vitamins, and minerals, playing a significant role in overall human well-being [3]-[6]. Studies have highlighted the anticancer properties of fruits, emphasizing the importance of specific nutrients like beta-carotene, vitamin E, and vitamin C. Research by Thailand's Ministry of Public Health has particularly identified fruits that are high in these nutrients, underscoring their potential role in cancer prevention strategies [6].

Concurrent with this, the rise of web applications and nutritional databases has significantly impacted how health-related information is disseminated. These platforms often employ sophisticated databases capable of handling image or video data [7]-[11]. In this context, Convolutional Neural Networks (CNNs) have become indispensable tools for accurate image classification [8]-[9]. Various algorithms for image processing have been seamlessly integrated into a myriad of fields, including agriculture [12], medical diagnostics [13], and dental imaging diagnostics [14]. Within the sphere of more complex systems, ResNet models have emerged as versatile deep neural networks, capable of addressing a wide array of tasks, ranging from disease detection to advanced image processing [15]-[35]. Specifically, ResNet models like ResNet-18 [15], [36], ResNet-34 [37], ResNet-50 [38], and ResNet-101 [39] have gained popularity due to their ease of use.

Visual technology finds extensive application in robotics within various modern industries. Moreover, its application in robotics extends to facial recognition [40]-[46], public identity verification, and mask detection [47]. In the agricultural sector, visual technology aids in fruit detection using faster R-CNN systems, estimates fruit quantity and ripeness, evaluates crop harvesting, and identifies plant diseases [48]-[56]. Beyond agriculture, visual technology is employed for object detection in diverse fields, including traffic management [57]-[65], passenger detection in vehicles, and geographical surveys [66]-[68]. It also plays a pivotal role in medical research [69]-[72], assisting in the diagnosis of COVID-19 [73], segmentation of lung regions in chest X-rays [74], histopathological image analysis [75], and classification of brain tissues [76]-[79].

This study aims to address the existing research gap by developing and evaluating multiple ResNet-18 models for the specific purpose of fruit image categorization at the web application level. The research enhances the functionality of the ResNet-18 neural network system through the design and comparative evaluation of single and multiple ResNet-18 models for image grouping and categorization. Overall

accuracy is improved by employing a voting mechanism. Additionally, the research introduces the concept of prediction thresholds to manage "Indecision" outcomes, thereby enhancing the robustness of fruit type categorization. The research presented in this case study serves as a practical example of fruit type detection with broader implications in the development of nutritionally valuable fruit databases. This work makes a significant contribution to the nutritional assessment of fruits within a healthcare organization that is currently in the process of development at the university level. Additionally, it extends the potential applications of object recognition techniques and automated systems in the future.

## II. RESEARCH METHOD

### A. Research Overview

To provide a comprehensive overview of the methodology, this section is structured into several subsections, including system configuration, dataset characteristics, and evaluation methods. A flowchart illustrating the entire research methodology is presented in Fig. 1 to facilitate understanding. Based on the flowchart, it outlines the process of utilizing image classification. The initial step involves checking if the ResNet18 model exists within the program. If it does, the image classification process begins. If not, the program proceeds to create the ResNet18 model. This involves inputting the dataset and defining the learning system conditions for ResNet18. Subsequently, the ResNet18 model is trained until the system's algorithm is generated. Following this, input images for testing purposes are fed into the system. The algorithm designed for the ResNet18 system processes the data and predicts the system's output. The results of the predictions are then displayed, marking the completion of the program's operation.
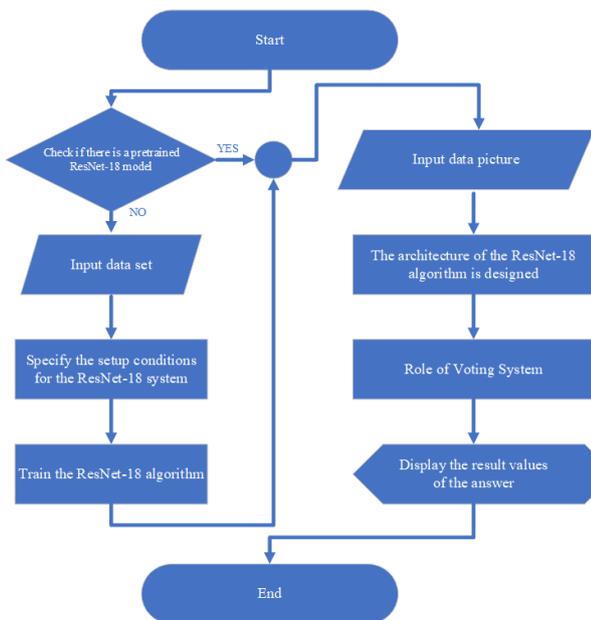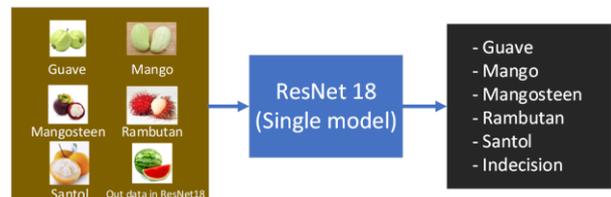


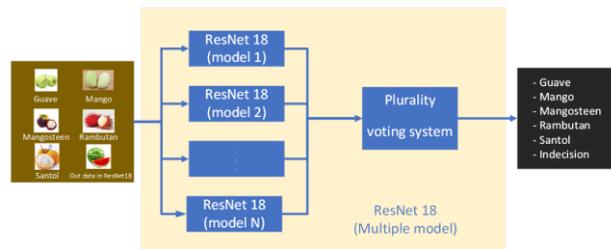Fig. 1. Block diagram of the single ResNet-18 model

### B. System Configuration and Model Setup

The study explores two unique configurations of ResNet-18 models for image classification tasks. The first

configuration is illustrated in Fig. 2 (a), which displays a block diagram of a standalone ResNet-18 model adhering to the framework's original architecture. This standard setup is primarily used for traditional image classification tasks and serves as the baseline for comparison. In contrast, Fig. 2 (b) shows the alternative configuration, featuring multiple ResNet-18 models working in unison. This novel arrangement promotes collaborative decision-making during the classification of a set of images. The array of interconnected ResNet-18 models enhances the system's accuracy in predicting image classification. The key difference between the configurations lies in their structural composition. The single ResNet-18 model retains the architecture historically used in image classification tasks within the ResNet-18 framework. On the other hand, the multiple ResNet-18 model presents an advanced system in which various instances of ResNet-18 work in synergy. This collective effort significantly boosts the system's ability to deliver accurate predictions across an array of image classifications.



(a) Block diagram of the single ResNet-18 model



(b) Block diagram of multiple ResNet-18 models

Fig. 2. Block diagram of the single and multiple ResNet-18 model

### C. Dataset and Illustrative Scenarios

The research utilizes a comprehensive dataset comprising images of five unique fruit types, namely, Guava, Mango, Mangosteen, Rambutan, and Santol. These images are visually presented across Fig. 3 to Fig. 7 and were purposefully selected through a randomized methodology to ensure a diverse array of conditions. For instance, the dataset includes images of fruits on trees, peeled fruits, and fruits presented in various dishes. The dataset contains different quantities of images for each fruit type - Guava has 65, Mango holds 92, Mangosteen includes 90, Rambutan shows 73, and Santol features 102.

This dataset serves a dual purpose. First, it is designed for seamless integration into the ResNet-18 system to enhance its predictive analysis capabilities. Second, it aims to contribute to the development of an application that provides nutritional information about different fruit categories, using a structured database framework for storing and managing the image data. Significantly, this study not only focuses on the machine learning and classification aspects but also aims to

compile relevant nutritional information related to the featured fruits.



Fig. 3. The Guava image dataset serves as the foundational learning dataset for training the ResNet-18 system



Fig. 4. The Mango image dataset serves as the foundational learning dataset for training the ResNet-18 system

### D. Testing Dataset and Evaluation Metrics

To evaluate the performance of both individual and collaborative ResNet-18 configurations, a specialized testing dataset is deployed, as illustrated in Fig. 8 to Fig. 12. The dataset includes 10 images for each type of fruit under investigation, Guava, Mango, Mangosteen, Rambutan, and Santol. These images were carefully selected to cover a broad spectrum of conditions, such as the natural appearance of the fruit, fruit on trees, peeled state, and arranged presentations.

Additionally, Fig. 13 features a dataset of 10 images that are not part of the ResNet-18 training data, designated solely for system testing. By excluding these test images from the training set, the study ensures an unbiased evaluation of the classification capabilities of the ResNet-18 model.

The testing dataset serves two primary functions in this research. Firstly, it quantifies the accuracy of the ResNet-18 configuration in classifying fruit types. Secondly, it gauges the models' robustness by subjecting them to an assortment of conditions in which fruit images may be encountered.

Through the careful selection of test images and their deliberate exclusion from the training set, the study aims to conduct a thorough performance assessment, offering crucial insights into the system's operational viability and reliability under real-world conditions.



Fig. 5. The Mangosteen image dataset serves as the foundational learning dataset for training the ResNet-18 system



Fig. 6. The Rambutan image dataset serves as the foundational learning dataset for training the ResNet-18 system

Fig. 7. The Santol image dataset serves as the foundational learning dataset for training the ResNet-18 system



Fig. 8. An image of a guava for testing the ResNet-18 model



Fig. 9. An image of a Mango for testing the ResNet-18 model



Fig. 10. An image of a Mangosteen for testing the ResNet-18 model



Fig. 11. An image of a Rambutan for testing the ResNet-18 model



Fig. 12. An image of a Santol for testing the ResNet-18 model



Fig. 13. An image of outdate for testing the ResNet-18 model

## III.　ResNet-18 Method

### A. Single ResNet-18 Model

In this research, the ResNet-18 architecture was employed. A learning framework designed to improve the training efficiency of deep residual networks. The challenge of vanishing gradients can emerge when dealing with a substantial number of layers. However, ResNet effectively tackles this issue by incorporating shortcut techniques that allow the network to bypass certain layers. The specific implementation details of the ResNet-18 model can be found in [80], which establishes a consistent learning configuration for the ResNet-18 system in this research. Additionally, the configuration of the Add-On explorer is visually depicted in Fig. 14.
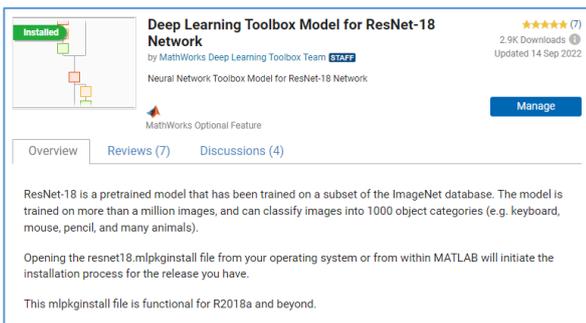


Fig. 14. Add-On explorer of the deep learning toolbox model for ResNet-18 network

In the current research, data from Fig. 3 to Fig. 7 serve as the foundation for machine learning with the specific aim of pattern recognition in images. Eight models were meticulously configured for this purpose, with their detailed specifications documented in Table I, Table II, and Table III. The choices for model parameters, including the number of training epochs and oversampling techniques, were rigorously determined to strike a balance between computational efficiency and predictive accuracy.

TABLE I. Data of Model Training in ResNet-18

| No | Model | Detail |
|----|-------|--------|
| 1 | M 1 | setup epoch loop 15 |
| 2 | M 2 | setup epoch loop 20 |
| 3 | M 3 | setup epoch loop 25 |
| 4 | M 4 | setup epoch loop 20 with out guave data |
| 5 | M 5 | setup epoch loop 20 with out mango data |
| 6 | M 6 | setup epoch loop 20 with out mangosteen data |
| 7 | M 7 | setup epoch loop 20 with out rambutan data |
| 8 | M 8 | setup epoch loop 20 with out santol data |

TABLE II. Datas of Detail in ResNet-18 Training (M1-M3)

| | Model | | |
|---|---|---|---|
| | M 1 | M 2 | M 3 |
| **Results** | | | |
| Accuracy | 0.9756 | 0.9024 | 0.9756 |
| **Training Cycle** | | | |
| Epoch | 15 | 20 | 25 |
| Iterations per epoch | 6 | 6 | 6 |
| Iteration | 90 | 120 | 150 |
| Over sampled | 82 | 82 | 82 |

Models M1 through M8 were trained using specialized datasets and parameters. For example, Model M1 utilized a dataset consisting of five sets and was trained for 15 epochs,

resulting in 90 iterations and achieving an accuracy rate of 0.9756. In contrast, Model M2 was trained on a similar dataset but underwent 20 epochs, totaling 120 iterations, and reached an accuracy of 0.9024. Model M3 used the same dataset size but increased the training epoch to 25, equivalent to 150 iterations, yielding an accuracy of 0.9756.

Models M4 through M8 focused on more specific dataset configurations. Model M4, which excluded the guava dataset, was trained for 20 epochs and achieved a perfect accuracy of 1. Models M5 and M6, which omitted the mango and mangosteen datasets, respectively, underwent 20 epochs, amounting to 120 iterations, and reached accuracies of 0.9375 and 0.9688. Similarly, Model M7, which left out the rambutan dataset, was trained for 20 epochs, achieving an accuracy of 0.9462. Lastly, Model M8, trained without the santol dataset, completed 20 epochs, and reached an accuracy of 0.9677.

TABLE III. Datas of Detail in ResNet-18 Training (M4-M8)

| | Model | | | | |
|---|---|---|---|---|---|
| | M 4 | M 5 | M 6 | M 7 | M 8 |
| **Results** | | | | | |
| Accuracy | 1.000 | 0.9375 | 0.9688 | 0.9462 | 0.9677 |
| **Training Cycle** | | | | | |
| Epoch | 20 | 20 | 20 | 20 | 20 |
| Iterations per epoch | 5 | 5 | 5 | 5 | 4 |
| Iteration | 100 | 100 | 100 | 100 | 80 |
| Over sampled | 82 | 82 | 82 | 82 | 74 |

### B. Multiple ResNet-18 model

The innovative approach behind the multiple ResNet-18 model system is based on leveraging distinct Single ResNet-18 models, each individually trained on unique datasets. These models adhere to the architecture of the original Single ResNet-18 and are organized in a parallel configuration, as visually illustrated in Fig. 2 (b).

The architecture of the multiple ResNet-18 model system in this study comprises three main composite models. These are named "S Model 2", "S Model 3", and "S Model 4". Specifically, S Model 2 integrates models M1, M2, and M3. S Model 3 combines models M2, M4, M5, M6, M7, and M8. Meanwhile, "S Model 4" is a comprehensive assembly that includes models M1, M2, M3, M4, M5, M6, M7, and M8. Once predictions are generated from each of these individual models, a voting mechanism is implemented to determine a collective decision for the final system output, thereby enhancing the overall data processing efficiency.

It's also important to note that there is a stand-alone Single ResNet-18 model known as "S Model 1", which corresponds to model M2 as detailed in Table IV. This configuration adds another layer of complexity and adaptability to the overall system.

TABLE IV. Datas of Design Single ResNet-18 Model to Multiple ResNet-18

| Model | M 1 | M 2 | M 3 | M 4 | M 5 | M 6 | M 7 | M 8 |
|-------|-----|-----|-----|-----|-----|-----|-----|-----|
| S Model 1 | | / | | | | | | |
| S Model 2 | / | / | / | | | | | |
| S Model 3 | | / | | / | / | / | / | / |
| S Model 4 | / | / | / | / | / | / | / | / |

## IV. VOTING SYSTEM

In this research, a specialized scoring system serves as an integral component to facilitate decision-making processes. This system operates on the principle of tallying votes, thereby assisting in the selection of one or more options from a set of available choices. The framework for the scoring system is developed based on the method of vote counting, as well as the number of choices available for selection. One of the key innovations in this study is the incorporation of a voting mechanism within multiple ResNet-18 models, aimed at enhancing classification performance. This feature holds particular relevance for web-based applications, as it offers a more nuanced approach to data categorization.

In the specific context of this study, the voting mechanism integrated into the system with multiple ResNet-18 models employs a single-choice plurality voting system, also known as the First-past-the-post system. To understand how this works, consider that there are three options for selection, namely, A, B, and C. The option that receives the most votes becomes the final output, as defined by the criteria in Equation 1. Efficiency and quick decision-making are two essential qualities of this selected voting methodology, making it particularly suitable for real-time web applications. A thorough discussion on the role and advantages of this voting system serves to highlight its significance in enhancing the overall performance and reliability of the multiple ResNet-18 models used in this study.

$$Winner = argmax(Votes(A), Votes(B), Votes(C)) \quad (1)$$

Where Votes(A), Votes(B), and Votes(C) represent the number of scores received from voters for each respective option A, B, C.

After incorporating (1) into the data outlined in Table 4, the formula that emerges for the plurality voting system is specified below.

In this context, $Winner_{S\_Model\_1}$ in (2) represents the outcome from "S Model 1". $Winner_{S\_Model\_2}$ in (3) signifies the result of "S Model 2". $Winner_{S\_Model\_3}$ in (4) indicates the result of "S Model 3", and $Winner_{S\_Model\_4}$ in (5) corresponds to the result of "S Model 4".

$$Winner_{S\_Model\_2} = argmax(Votes(M2)) \quad (2)$$

$$Winner_{S_{Model_3}} = argmax(Votes(M1), Votes(M2), Votes(M3)) \quad (3)$$

$$Winner_{S\_Model\_4} = argmax(Votes(M2), Votes(M3), Votes(M4), Votes(M5), Votes(M6), Votes(M7), Votes(M8)) \quad (4)$$

$$Winner_{S_{Mode5}} = argmax(Votes(M1), Votes(M2), Votes(M3), Votes(M4), Votes(M5), Votes(M6), Votes(M7), Votes(M8)) \quad (5)$$

## V. EXPERIMENTAL RESULTS

This section aims to elucidate the findings from two interconnected but distinct experiments, both designed to rigorously evaluate the ResNet-18 model's performance in fruit type categorization. The overarching objective extends beyond measuring predictive accuracy to include an examination of different prediction criteria and the utility of a voting system when using multiple ResNet-18 models.

The first experiment focuses on classifying fruit types based on the highest prediction scores generated by a single ResNet-18 model. Six rounds of testing were conducted for this purpose, each involving 10 images assessed with eight different ResNet-18 configurations.

In contrast, the second experiment shifts the focus on the role of prediction thresholds in accurate fruit categorization. Here, a prediction score exceeding 90% is considered necessary for confident classification. Any score below this threshold is labeled as "Indecision," offering an alternative evaluation metric compared to the first experiment.

Additionally, a voting system is implemented in both experiments to reach collective decisions when using multiple ResNet-18 models. Under this system, if the maximum voting score surpasses 2, the outcome is labeled as "Indecision," indicating either an inconclusive result or a degree of model uncertainty.

By adopting these multi-faceted testing paradigms, a comprehensive understanding of the performance of both single and multiple ResNet-18 models under various conditions is achieved. This prepares the ground for subsequent, in-depth analyses and discussion, the key conclusions of which are captured in the summary of the results section.

### A. Testing ResNet-18 by Selecting Maximum Prediction Scores for Fruit Type Categorization

This section presents the outcomes of the initial experiment, which involves evaluating the ResNet-18 model by selecting the highest prediction scores to categorize different fruit types. The experiment was conducted on six separate occasions, with each trial consisting of 10 images for each of the 8 models that were designed as outlined in Table IV. The results of the experiment were classified into three distinct categories as correctness, incorrectness, and indecision.

Referring to the data presented in Table V, an image of a guava, as shown in Fig. 8, was employed to assess the system. The prediction with the highest score was taken as the model's response. Notably, the system's predictions were consistently congruent across all models, and the image marked as "Incorrectness" was uniformly identified as Image 3.

Similarly, the findings from Table VI pertain to the examination of the system using an image of a mango, as illustrated in Fig. 9. The model's response was determined by the highest predicted score. S Model 1 accurately predicted 5 images while incorrectly predicting 5 others. In contrast, S Model 2 achieved correctness for over 2 images compared to S Model 1. Despite S Model 3 correctly predicting 4 images, it deliberately refrained from giving answers, thereby leading to 2 images being categorized as "Indecision" and 4 images as "Incorrectness". Conversely, S

Model 4 correctly predicted 5 images, with 3 images marked as "Incorrectness" and 2 images as "Indecision".

TABLE V. DATAS OF GUAVE BY SELECTING MAXIMUM PREDICTION SCORES FOR FRUIT TYPE CATEGORIZATION

| Picture | S Model 1 | S Model 2 | S Model 3 | S Model 4 |
|---|---|---|---|---|
| Guave 01 | Guave | Guave | Guave | Guave |
| Guave 02 | Guave | Guave | Guave | Guave |
| Guave 03 | Mango | Mango | Mango | Mango |
| Guave 04 | Guave | Guave | Guave | Guave |
| Guave 05 | Guave | Guave | Guave | Guave |
| Guave 06 | Guave | Guave | Guave | Guave |
| Guave 07 | Guave | Guave | Guave | Guave |
| Guave 08 | Guave | Guave | Guave | Guave |
| Guave 09 | Guave | Guave | Guave | Guave |
| Guave 10 | Guave | Guave | Guave | Guave |
| **Correctness** | 9 | 9 | 9 | 9 |
| **Incorrectness** | 1 | 1 | 1 | 1 |
| **Indecision** | 0 | 0 | 0 | 0 |

TABLE VI. DATAS OF MANGO BY SELECTING MAXIMUM PREDICTION SCORES FOR FRUIT TYPE CATEGORIZATION

| Picture | S Model 1 | S Model 2 | S Model 3 | S Model 4 |
|---|---|---|---|---|
| Mango 01 | Guave | Mango | Guave | Indecision |
| Mango 02 | Guave | Mango | Indecision | Mango |
| Mango 03 | Mango | Mango | Mango | Mango |
| Mango 04 | Mango | Mango | Indecision | Indecision |
| Mango 05 | Mango | Mango | Mango | Mango |
| Mango 06 | Guave | Guave | Guave | Guave |
| Mango 07 | Guave | Guave | Guave | Guave |
| Mango 08 | Mango | Mango | Mango | Mango |
| Mango 09 | Guave | Guave | Guave | Guave |
| Mango 10 | Mango | Mango | Mango | Mango |
| **Correctness** | 5 | 7 | 4 | 5 |
| **Incorrectness** | 5 | 3 | 4 | 3 |
| **Indecision** | 0 | 0 | 2 | 2 |

Concurrently, the findings showcased in Table VII demonstrate an experiment employing an image of a mangosteen, as shown in Fig. 10, to assess the system's performance. The model's response was determined by the highest predicted score. Remarkably, S Model 1 exhibited precise predictions for 9 images while making 1 incorrect prediction. Similarly, the remaining models yielded similar results. Conversely, S Model 3 achieved accurate predictions for 9 images, with 1 image categorized as "Indecision".

TABLE VII. DATAS OF MANGOSTEEN BY SELECTING MAXIMUM PREDICTION SCORES FOR FRUIT TYPE CATEGORIZATION

| Picture | S Model 1 | S Model 2 | S Model 3 | S Model 4 |
|---|---|---|---|---|
| Mangosteen 01 | Mangosteen | Mangosteen | Mangosteen | Mangosteen |
| Mangosteen 02 | Mangosteen | Mangosteen | Mangosteen | Mangosteen |
| Mangosteen 03 | Mangosteen | Mangosteen | Mangosteen | Mangosteen |
| Mangosteen 04 | Mangosteen | Mangosteen | Mangosteen | Mangosteen |
| Mangosteen 05 | Mangosteen | Mangosteen | Mangosteen | Mangosteen |
| Mangosteen 06 | Mangosteen | Mangosteen | Mangosteen | Mangosteen |
| Mangosteen 07 | Mangosteen | Mangosteen | Mangosteen | Mangosteen |
| Mangosteen 08 | Mangosteen | Mangosteen | Mangosteen | Mangosteen |
| Mangosteen 09 | Rambutan | Rambutan | Indecision | Rambutan |
| Mangosteen 10 | Mangosteen | Mangosteen | Mangosteen | Mangosteen |
| **Correctness** | 9 | 9 | 9 | 9 |
| **Incorrectness** | 1 | 1 | 0 | 1 |
| **Indecision** | 0 | 0 | 1 | 0 |

Drawing insights from the findings presented in Table VIII, an evaluation was conducted using an image of a Rambutan, as shown in Fig. 11, to test the system. The model's answer was determined by selecting the highest predicted score. The results revealed consistent predictions across all models, with each model achieving a remarkable 100% accuracy in predicting the outcomes

TABLE VIII. DATAS OF RAMBUTAN BY SELECTING MAXIMUM PREDICTION SCORES FOR FRUIT TYPE CATEGORIZATION

| Picture | S Model 1 | S Model 2 | S Model 3 | S Model 4 |
|---|---|---|---|---|
| Rambutan 01 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 02 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 03 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 04 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 05 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 06 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 07 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 08 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 09 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 10 | Rambutan | Rambutan | Rambutan | Rambutan |
| **Correctness** | 10 | 10 | 10 | 10 |
| **Incorrectness** | 0 | 0 | 0 | 0 |
| **Indecision** | 0 | 0 | 0 | 0 |

In the context of the experimental findings presented in Table IX, an evaluation was carried out using an image of a mangosteen, as shown in Fig. 12, to assess the performance of the system. The highest predicted score was employed as the model's response. The results revealed that S Model 1 accurately predicted the correct classification for 9 images and had 1 incorrect prediction. Similarly, S Model 3 achieved accurate predictions for 9 images, with 1 image categorized as "Indecision". Conversely, the remaining models correctly predicted the classification for all 10 images.

TABLE IX. DATAS OF SANTOL BY SELECTING MAXIMUM PREDICTION SCORES FOR FRUIT TYPE CATEGORIZATION

| Picture | S Model 1 | S Model 2 | S Model 3 | S Model 4 |
|---|---|---|---|---|
| Santol 01 | Santol | Santol | Santol | Santol |
| Santol 02 | Santol | Santol | Santol | Santol |
| Santol 03 | Santol | Santol | Santol | Santol |
| Santol 04 | Santol | Santol | Santol | Santol |
| Santol 05 | Santol | Santol | Santol | Santol |
| Santol 06 | Santol | Santol | Santol | Santol |
| Santol 07 | Santol | Santol | Santol | Santol |
| Santol 08 | Mango | Santol | Indecision | Santol |
| Santol 09 | Santol | Santol | Santol | Santol |
| Santol 10 | Santol | Santol | Santol | Santol |
| **Correctness** | 9 | 10 | 9 | 10 |
| **Incorrectness** | 1 | 0 | 0 | 0 |
| **Indecision** | 0 | 0 | 1 | 0 |

The results presented in Table X showcase an experiment involving an image that was not part of the learning dataset, as shown in Fig. 13. The primary goal of this experiment was to assess the system's response when confronted with an unfamiliar image, expecting to receive an "Indecision" response from the system. This outcome aligns with the system's correctness prediction within this specific experiment. The experiment underscores the system's inability to categorize images absent from its training dataset, while still generating responses based on the knowledge it has acquired.

TABLE X. DATAS OF OUTDATE BY SELECTING MAXIMUM PREDICTION SCORES FOR FRUIT TYPE CATEGORIZATION

| Picture | S Model 1 | S Model 2 | S Model 3 | S Model 4 |
|---|---|---|---|---|
| Outdate 01 | Guave | Guave | Guave | Guave |
| Outdate 02 | Santol | Santol | Santol | Santol |
| Outdate 03 | Rambutan | Indecision | Mangosteen | Mangosteen |
| Outdate 04 | Santol | Santol | Santol | Santol |
| Outdate 05 | Rambutan | Mangosteen | Santol | Mangosteen |
| Outdate 06 | Rambutan | Mango | Rambutan | Mango |
| Outdate 07 | Santol | Santol | Santol | Santol |
| Outdate 08 | Santol | Guave | Santol | Santol |
| Outdate 09 | Guave | Guave | Guave | Guave |
| Outdate 10 | Mango | Mango | Mango | Mango |
| **Correctness** | 0 | 1 | 1 | 0 |
| **Incorrectness** | 10 | 9 | 9 | 10 |

*B. Testing ResNet-18 by Setting Prediction Thresholds for Fruit Type Categorization*

The second series of experiments focuses on assessing ResNet-18 by defining prediction thresholds for the categorization of fruit types. An accurate classification of a fruit type requires a prediction score exceeding 90%, otherwise, the outcome is categorized as "Indecision". This approach introduces a distinction from the first experiment.

The results presented in Table XI depict an experiment involving an image of guava, as illustrated in Fig. 8, to evaluate the system. The highest predicted score determined the model's answer. The findings revealed that S Model 1 correctly predicted the classification for 6 images, made an incorrect prediction for 1 image, and indicated "Indecision" for 3 images. Conversely, S Model 2 achieved accurate classification for 8 images and labeled 2 images as "Indecision". Meanwhile, both S Model 3 and S Model 4 achieved correct classification for 9 images, with 1 incorrect prediction.

TABLE XI. DATAS OF GUAVE BY SETTING PREDICTION THRESHOLDS FOR FRUIT TYPE CATEGORIZATION

| Picture | S Model 1 | S Model 2 | S Model 3 | S Model 4 |
|---|---|---|---|---|
| Guave 01 | Guave | Guave | Guave | Guave |
| Guave 02 | Guave | Guave | Guave | Guave |
| Guave 03 | Mango | Indecision | Mango | Mango |
| Guave 04 | Guave | Guave | Guave | Guave |
| Guave 05 | Indecision | Indecision | Guave | Guave |
| Guave 06 | Guave | Guave | Guave | Guave |
| Guave 07 | Indecision | Guave | Guave | Guave |
| Guave 08 | Guave | Guave | Guave | Guave |
| Guave 09 | Guave | Guave | Guave | Guave |
| Guave 10 | Indecision | Guave | Guave | Guave |
| **Correctness** | 6 | 8 | 9 | 9 |
| **Incorrectness** | 1 | 0 | 1 | 1 |
| **Indecision** | 3 | 2 | 0 | 0 |

Simultaneously, the outcomes in Table XII showcase an experiment involving an image of Mango, illustrated in Fig. 9, to assess the system. The highest predicted score was employed as the model's response. Notably, S Model 1 accurately predicted the correctness for 2 images, made 2 incorrect predictions, and indicated "Indecision" for 6 images. Conversely, S Model 2 achieved correctness for 2 images and identified 8 images as "Indecision". Despite S Model 3 predicting correctness for 5 images, it deliberately chose not to provide an answer for 1 image, categorizing them as "Indecision," leading to 4 images being labeled as "Incorrectness". In contrast, S Model 4 achieved accurate

classification for 5 images, with 3 images marked as "Incorrectness" and 2 images as "Indecision".

TABLE XII. DATAS OF MANGO BY SETTING PREDICTION THRESHOLDS FOR FRUIT TYPE CATEGORIZATION

| Picture | S Model 1 | S Model 2 | S Model 3 | S Model 4 |
|---|---|---|---|---|
| Mango 01 | Indecision | Indecision | Guave | Guave |
| Mango 02 | Guave | Indecision | Guave | Indecision |
| Mango 03 | Indecision | Indecision | Mango | Mango |
| Mango 04 | Indecision | Indecision | Indecision | Indecision |
| Mango 05 | Indecision | Indecision | Mango | Mango |
| Mango 06 | Indecision | Indecision | Mango | Mango |
| Mango 07 | Indecision | Indecision | Guave | Guave |
| Mango 08 | Mango | Mango | Mango | Mango |
| Mango 09 | Guave | Indecision | Guave | Guave |
| Mango 10 | Mango | Mango | Mango | Mango |
| **Correctness** | 2 | 2 | 5 | 5 |
| **Incorrectness** | 2 | 0 | 4 | 3 |
| **Indecision** | 6 | 8 | 1 | 2 |

Continuing with the findings presented in Table XIII, an experimental trial was conducted using an image of Mangosteen, showcased in Fig. 10, for system evaluation. The highest predicted score was employed as the model's response. It was observed that both S Model 1 and S Model 2 accurately predicted the correctness for 7 images and indicated "Indecision" for 3 images. Similarly, the remaining models demonstrated comparable results. In contrast, both S Model 3 and S Model 4 accurately predicted the correctness for all 10 images.

TABLE XIII. DATAS OF MANGOSTEEN BY SETTING PREDICTION THRESHOLDS FOR FRUIT TYPE CATEGORIZATION

| Picture | S Model 1 | S Model 2 | S Model 3 | S Model 4 |
|---|---|---|---|---|
| Mangosteen 01 | Mangosteen | Mangosteen | Mangosteen | Mangosteen |
| Mangosteen 02 | Mangosteen | Mangosteen | Mangosteen | Mangosteen |
| Mangosteen 03 | Indecision | Indecision | Mangosteen | Mangosteen |
| Mangosteen 04 | Mangosteen | Mangosteen | Mangosteen | Mangosteen |
| Mangosteen 05 | Mangosteen | Mangosteen | Mangosteen | Mangosteen |
| Mangosteen 06 | Indecision | Indecision | Mangosteen | Mangosteen |
| Mangosteen 07 | Mangosteen | Mangosteen | Mangosteen | Mangosteen |
| Mangosteen 08 | Mangosteen | Mangosteen | Mangosteen | Mangosteen |
| Mangosteen 09 | Indecision | Indecision | Mangosteen | Mangosteen |
| Mangosteen 10 | Mangosteen | Mangosteen | Mangosteen | Mangosteen |
| **Correctness** | 7 | 7 | 10 | 10 |
| **Incorrectness** | 0 | 0 | 0 | 0 |
| **Indecision** | 3 | 3 | 0 | 0 |

In relation to the experimental results showcased in Table XIV, an experiment involving an image of Rambutan, as depicted in Fig. 11, was conducted to assess the system's performance. The highest predicted score was utilized as the model's response. Significantly, the predictions made by the system remained uniform across all models, with each model accurately forecasting correctness for 9 images and identifying 1 image as "Incorrectness".

In the context of the outcomes presented in Table XV, a test was carried out utilizing an image of Santol, as shown in Fig. 12, to evaluate the system's functionality. The highest predicted score was employed as the model's output. The outcomes disclosed that S Model 1 adeptly predicted the correctness of 8 images while classifying 2 images as "Indecision". Similarly, the other models demonstrated consistent patterns, correctly predicting 9 images and designating 1 image as "Indecision".

TABLE XIV. DATAS OF RAMBUTAN BY SETTING PREDICTION THRESHOLDS FOR FRUIT TYPE CATEGORIZATION

| Picture | S Model 1 | S Model 2 | S Model 3 | S Model 4 |
|---|---|---|---|---|
| Rambutan 01 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 02 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 03 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 04 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 05 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 06 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 07 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 08 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 09 | Rambutan | Rambutan | Rambutan | Rambutan |
| Rambutan 10 | Indecision | Indecision | Indecision | Indecision |
| Correctness | 9 | 9 | 9 | 9 |
| Incorrectness | 0 | 0 | 0 | 0 |
| Indecision | 1 | 1 | 1 | 1 |

TABLE XV. DATAS OF SANTOL BY SETTING PREDICTION THRESHOLDS FOR FRUIT TYPE CATEGORIZATION

| Picture | S Model 1 | S Model 2 | S Model 3 | S Model 4 |
|---|---|---|---|---|
| Santol 01 | Santol | Santol | Santol | Santol |
| Santol 02 | Santol | Santol | Santol | Santol |
| Santol 03 | Indecision | Santol | Santol | Santol |
| Santol 04 | Santol | Santol | Santol | Santol |
| Santol 05 | Santol | Santol | Santol | Santol |
| Santol 06 | Santol | Santol | Santol | Santol |
| Santol 07 | Santol | Santol | Santol | Santol |
| Santol 08 | Indecision | Indecision | Indecision | Indecision |
| Santol 09 | Santol | Santol | Santol | Santol |
| Santol 10 | Santol | Santol | Santol | Santol |
| Correctness | 8 | 9 | 9 | 9 |
| Incorrectness | 0 | 0 | 0 | 0 |
| Indecision | 2 | 1 | 1 | 1 |

The outcomes presented in Table XVI showcase an experiment involving an image not included in the learning dataset, as illustrated in Fig. 13. The primary objective was to elicit an "Indecision" response from the system, based on a predefined prediction threshold of over 90% for correctness. This experiment highlighted those systems utilizing a prediction threshold exceeding 90% could effectively distinguish responses. On average, correctness was accurately predicted for 7 images, while 3 images were categorized as "Incorrectness". Notably, S Model 4 deviated from this trend, with 4 images classified as "Incorrectness".

TABLE XVI. DATAS OF OUTDATE BY SETTING PREDICTION THRESHOLDS FOR FRUIT TYPE CATEGORIZATION

| Picture | S Model 1 | S Model 2 | S Model 3 | S Model 4 |
|---|---|---|---|---|
| Outdate 01 | Guave | Guave | Guave | Guave |
| Outdate 02 | Santol | Santol | Santol | Santol |
| Outdate 03 | Indecision | Indecision | Indecision | Indecision |
| Outdate 04 | Santol | Santol | Santol | Santol |
| Outdate 05 | Indecision | Indecision | Indecision | Indecision |
| Outdate 06 | Indecision | Indecision | Indecision | Indecision |
| Outdate 07 | Indecision | Indecision | Indecision | Santol |
| Outdate 08 | Indecision | Indecision | Indecision | Indecision |
| Outdate 09 | Indecision | Indecision | Indecision | Indecision |
| Outdate 10 | Indecision | Indecision | Indecision | Indecision |
| Correctness | 7 | 7 | 7 | 6 |
| Incorrectness | 3 | 3 | 3 | 4 |

## C. Summary of the Results

The research introduces compelling evidence that systems equipped with multiple ResNet-18 models deliver more consistent and often superior prediction outcomes

compared to those running on a single ResNet-18 model. This distinction becomes crucial when one examines the system's behavior with predefined answer thresholds, as detailed in Table XVII to Table XIX. When the system operates under these constraints, it shows a discernable preference for abstaining from answering or displaying "Indecision" rather than offering incorrect responses. This behavioral nuance contributes to higher correctness values but also indicates the limitations of such an approach, especially in scenarios requiring absolute answers.

In contrast, the system grapples with a specific set of challenges when it faces data not originally included in its learning process. This is particularly evident in Table 18 and Table XX, where the absence of decision boundaries leads to difficulties in distinguishing unlearned data, causing a noticeable decline in the rate of correct predictions. These observations provide a counterbalance to the potential advantages of deploying multiple ResNet-18 models for categorization tasks. Where this study differs from existing research is in its nuanced comparison of the effectiveness of multiple models versus single models, a fresh angle that brings new considerations to the table. The overarching implication suggests that multiple model systems, while more reliable in many aspects, still need fine-tuning to better handle unlearned or outlier data. This duality of strengths and limitations highlights the need for further research, aiming to refine the system's ability to make robust predictions across varied datasets.

TABLE XVII. SUMMARY DATAS OF FRUIT BY SELECTING MAXIMUM PREDICTION SCORES FOR FRUIT TYPE CATEGORIZATION

| Picture | S Model 1 | S Model 2 | S Model 3 | S Model 4 |
|---|---|---|---|---|
| Correctness | 84 | 90 | 82 | 86 |
| Incorrectness | 16 | 10 | 10 | 10 |
| Indecision | 0 | 0 | 8 | 4 |

TABLE XVIII. SUMMARY DATAS OF OUTDATE BY SELECTING MAXIMUM PREDICTION SCORES FOR FRUIT TYPE CATEGORIZATION

| Picture | S Model 1 | S Model 2 | S Model 3 | S Model 4 |
|---|---|---|---|---|
| Correctness | 0 | 10 | 10 | 0 |
| Incorrectness | 100 | 90 | 90 | 100 |

TABLE XIX. SUMMARY DATAS OF FRUIT BY SETTING PREDICTION THRESHOLDS FOR FRUIT TYPE CATEGORIZATION

| Picture | S Model 1 | S Model 2 | S Model 3 | S Model 4 |
|---|---|---|---|---|
| Correctness | 64 | 70 | 84 | 84 |
| Incorrectness | 6 | 0 | 10 | 8 |
| Indecision | 30 | 30 | 6 | 8 |

TABLE XX. SUMMARY DATAS OF OUTDATE BY SETTING PREDICTION THRESHOLDS FOR FRUIT TYPE CATEGORIZATION

| Picture | S Model 1 | S Model 2 | S Model 3 | S Model 4 |
|---|---|---|---|---|
| Correctness | 70 | 70 | 70 | 60 |
| Incorrectness | 30 | 30 | 30 | 40 |

## VI. CONCLUSION

In summary, this study examines how single and multiple ResNet-18 models perform in classifying fruit images. Single ResNet-18 models are effective when there is limited storage or a short learning period. This provides important information about the strengths and weaknesses of object recognition methods using ResNet-18. Using multiple

ResNet-18 models results in better accuracy and more reliable system performance. This is relevant for industries that need automated systems and large databases. However, it's also important to note the computational costs and challenges in deploying multiple models. These issues point to areas that could benefit from future research. Looking ahead, there's an opportunity to improve the classification between closely related categories. This could be valuable for both practical applications and academic study, as it would make image classification systems more effective. To conclude, this study offers valuable insights for both practical use and academic research in image classification, especially using ResNet-18 models. As industries adopt these technologies, the findings from this study could guide future improvements in automated systems and databases.

### REFERENCES

[1] S. Çakmakçı and R. Çakmakçı, "Quality and nutritional parameters of food in agri-food production systems," *Foods*, vol. 12, no. 2, p. 351, 2023, doi: 10.3390/foods12020351.

[2] A. S. Roy *et al.*, "Food systems determinants of nutritional health and wellbeing in urban informal settlements: A scoping review in LMICs," *Social Science & Medicine*, vol. 115804, 2023, doi: 10.1016/j.socscimed.2023.115804.

[3] H. M. Habib *et al.*, "Palm Fruit (Phoenix dactylifera L.) Pollen Extract Inhibits Cancer Cell and Enzyme Activities and DNA and Protein Damage," *Nutrients*, vol. 15, no. 11, p. 2614, 2023, doi: 10.3390/nu15112614.

[4] M. Y. Wani *et al.*, "The phenolic components extracted from mulberry fruits as bioactive compounds against cancer: A review," *Phytotherapy Research*, vol. 37, no. 3, pp. 1136-1152, 2023, doi: 10.1002/ptr.7713.

[5] L. Zhao *et al.*, "Specific botanical groups of fruit and vegetable consumption and liver cancer and chronic liver disease mortality: a prospective cohort study," *The American Journal of Clinical Nutrition*, vol. 117, no. 2, pp. 278-285, 2023, doi: 10.1016/j.ajcnut.2022.12.004.

[6] H. Huang *et al.*, "Edible and cation-free kiwi fruit derived vesicles mediated EGFR-targeted siRNA delivery to inhibit multidrug resistant lung cancer," *Journal of Nanobiotechnology*, vol. 21, no. 1, pp. 1-14, 2023, doi: 10.1186/s12951-023-01766-w.

[7] R. Muwardi, M. Yunita, H. Ghifarsyam, and H. Juliyanto, "Optimize Image Processing Algorithm on ARM Cortex-A72 and A53," *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, vol. 8, no. 3, pp. 399-409, 2022, doi: 10.26555/jiteki.v8i3.24457.

[8] D. Sharma and N. Agrawal, "Development of Modified CNN Algorithm for Agriculture Product: A Research Review," *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, vol. 8, no. 1, pp. 167-174, 2022, doi: 10.26555/jiteki.v8i1.23722.

[9] R. Bello and C. Oluigbo, "Deep Learning-Based SOLO Architecture for Re-Identification of Single Persons by Locations," *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, vol. 8, no. 4, pp. 599-609, 2022, doi: 10.26555/jiteki.v8i4.25059.

[10] B. Suprapto, A. Wahyudin, H. Hikmarika, and S. Dwijayanti, "The Detection System of Helipad for Unmanned Aerial Vehicle Landing Using YOLO Algorithm," *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, vol. 7, no. 2, pp. 193-206, 2021, doi: 10.26555/jiteki.v7i2.20684.

[11] R. Bello, C. Oluigbo, and O. Moradeyo, "Motorcycling-Net: A Segmentation Approach for Detecting Motorcycling Near Misses," *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, vol. 9, no. 1, pp. 96-106, 2023, doi: 10.26555/jiteki.v9i1.25614.

[12] A. Kamilaris and F. X. Prenafeta-Boldú, "A review of the use of convolutional neural networks in agriculture," *The Journal of Agricultural Science*, vol. 156, no. 3, pp. 312-322, Mar. 2018, doi: 10.1017/S0021859618000436.

[13] R. H. Abiyev and M. K. S. Ma'aitaH, "Deep convolutional neural networks for chest diseases detection," *Journal of Healthcare Engineering*, vol. 2018, pp. 1-16, Nov. 2018, doi: 10.1155/2018/4168538.

[14] F. Schwendicke, T. Golla, M. Dreher, and J. Krois, "Convolutional neural networks for dental image diagnostics: A scoping review," *Journal of Dentistry*, vol. 91, pp. 103226, Mar. 2019, doi: 10.1016/j.jdent.2019.103226.

[15] G. N. Nguyen, N. H. Le Viet, M. Elhoseny, K. Shankar, B. B. Gupta, and A. A. Abd El-Latif, "Secure blockchain enabled Cyber–physical systems in healthcare using deep belief network with ResNet model," *Journal of parallel and distributed computing*, vol. 153, pp. 150-160, 2021, doi: 10.1016/j.jpdc.2021.03.011.

[16] C. D. Vo, D. A. Dang, and P. H. Le, "Development of Multi-Robotic Arm System for Sorting System Using Computer Vision," *Journal of Robotics and Control (JRC)*, vol. 3, no. 5, pp. 690-698, 2022, doi: 10.18196/jrc.v3i5.15661.

[17] A. N. Hidayah, S. A. Radzi, N. A. Razak, W. H. M. Saad, Y. C. Wong, and A. A. Naja, "Disease Detection of Solanaceous Crops Using Deep Learning for Robot Vision," *Journal of Robotics and Control (JRC)*, vol. 3, no. 6, pp. 790-799, 2022, doi: 10.18196/jrc.v3i6.15948.

[18] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye, "Object Detection in 20 Years: A Survey," in *Proceedings of the IEEE*, vol. 111, no. 3, pp. 257-276, March 2023, doi: 10.1109/JPROC.2023.3238524.

[19] P. Rosyady and R. Sumiharto, "Highway Visual Tracking System using Thresholding and Hough Transform," *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, vol. 4, no. 2, pp. 93-99, 2019, doi: 10.26555/jiteki.v4i2.12016.

[20] T. Agrawal and P. Choudhary, "Segmentation and classification on chest radiography: a systematic survey," *The Visual Computer*, vol. 39, no. 3, pp. 875-913, 2023, doi: 10.1007/s00371-021-02352-7.

[21] S. Saha Roy, S. Roy, P. Mukherjee, and A. Halder Roy, "An automated liver tumour segmentation and classification model by deep learning based approaches," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 11, no. 3, pp. 638–650, 2023, doi: 10.1080/21681163.2022.2099300.

[22] M. Xu, K. Huang, and X. Qi, "A Regional-Attentive Multi-Task Learning Framework for Breast Ultrasound Image Segmentation and Classification," *IEEE Access*, vol. 11, pp. 5377–5392, 2023, doi: 10.1109/ACCESS.2023.3236693.

[23] A. Iqbal and M. Sharif, "BTS-ST: Swin transformer network for segmentation and classification of multimodality breast cancer images," *Knowledge-Based Systems*, vol. 267, p. 110393, 2023, doi: 10.1016/j.knosys.2023.110393.

[24] M. K. Hasan, M. A. Ahamad, C. H. Yap, and G. Yang, "A survey, review, and future trends of skin lesion segmentation and classification," *Computers in Biology and Medicine*, p. 106624, 2023, doi: 10.1016/j.compbiomed.2023.106624.

[25] T. Dang and H. Tran, "A Secured, Multilevel Face Recognition based on Head Pose Estimation, MTCNN and FaceNet," *Journal of Robotics and Control (JRC)*, vol. 4, no. 4, pp. 431–437, 2023, doi: 10.18196/jrc.v4i4.18780.

[26] M. I. Rusydi *et al.*, "Autonomous Movement Control of Coaxial Mobile Robot based on Aspect Ratio of Human Face for Public Relation Activity Using Stereo Thermal Camera," *Journal of Robotics and Control (JRC)*, vol. 3, no. 3, pp. 361-373, 2022, doi: 10.18196/jrc.v3i3.14750.

[27] X. Zheng, Q. Lei, R. Yao, Y. Gong, and Q. Yin, "Image segmentation based on adaptive K-means algorithm," *EURASIP Journal on Image and Video Processing*, vol. 2018, no. 1, pp. 1-10, 2018, doi: 10.1186/s13640-018-0309-3.

[28] R. Srikanth and K. Bikshalu, "Multilevel thresholding image segmentation based on energy curve with harmony Search Algorithm," *Ain Shams Engineering Journal*, vol. 12, no. 1, pp. 1-20, 2021, doi: 10.1016/j.asej.2020.09.003.

[29] G. Aletti, A. Benfenati, and G. Naldi, "A semiautomatic multi-label color image segmentation coupling Dirichlet problem and colour distances," *Journal of Imaging*, vol. 7, no. 10, pp. 208, 2021, doi: 10.3390/jimaging7100208.

[30] C. Suguna and S. P. Balamurugan, "Computer Aided Diagnosis for Cervical Cancer Screening using Monarch Butterfly Optimization with Deep Learning Model," in *2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pp. 1059-1064, January 2023, doi: 10.1109/ICSSIT55814.2023.10060959.

[31] G. Sun, X. Jia, and T. Geng, "Plant diseases recognition based on image processing technology," *Journal of Electrical and Computer Engineering*, vol. 2018, 2018, doi: 10.1155/2018/6070129.

[32] A. Mohan and S. Poobal, "Crack detection using image processing: A critical review and analysis," *Alexandria Engineering Journal*, vol. 57, no. 2, pp. 787-798, 2018, doi: 10.1016/j.aej.2017.01.020.

[33] B. Li and Y. He, "An improved ResNet based on the adjustable shortcut connections," *IEEE Access*, vol. 6, pp. 18967-18974, 2018, doi: 10.1109/ACCESS.2018.2814605.

[34] D. Marmanis, K. Schindler, J. D. Wegner, S. Galliani, M. Datcu, and U. Stilla, "Classification with an edge: Improving semantic image segmentation with boundary detection," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 135, pp. 158-172, 2018, doi: 10.48550/arXiv.1612.01337.

[35] J. Duan, T. Shi, H. Zhou, J. Xuan, and S. Wang, "A novel ResNet-based model structure and its applications in machine health monitoring," *Journal of Vibration and Control*, vol. 27, no. 9-10, pp. 1036-1050, 2021, doi: 10.1177/107754632093650.

[36] H. Yu, H. Sun, J. Tao, C. Qin, D. Xiao, Y. Jin, and C. Liu, "A multi-stage data augmentation and AD-ResNet-based method for EPB utilization factor prediction," *Automation in Construction*, vol. 147, p. 104734, 2023, doi: 10.1016/j.autcon.2022.104734.

[37] S. Ayyachamy, V. Alex, M. Khened, and G. Krishnamurthi, "Medical image retrieval using Resnet-18," in *Medical Imaging 2019: Imaging Informatics for Healthcare, Research, and Applications*, vol. 10954, pp. 233-241, 2019, doi: 10.1117/12.2515588.

[38] M. Gao, D. Qi, H. Mu, and J. Chen, "A transfer residual neural network based on ResNet-34 for detection of wood knot defects," *Forests*, vol. 12, no. 2, p. 212, 2021, doi: 10.3390/f12020212.

[39] L. Wen, X. Li, and L. Gao, "A transfer convolutional neural network for fault diagnosis based on ResNet-50," *Neural Computing and Applications*, vol. 32, pp. 6111-6124, 2020, doi: 10.1007/s00521-019-04097-w.

[40] B. Yu, L. Yang, and F. Chen, "Semantic segmentation for high spatial resolution remote sensing images based on convolution neural network and pyramid pooling module," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 9, pp. 3252-3261, Sep. 2018, doi: 10.1109/JSTARS.2018.2860989.

[41] F. Tabassum, M. I. Islam, R. T. Khan, and M. R. Amin, "Human face recognition with combination of DWT and machine learning," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 3, pp. 546-556, 2022, doi: 10.1016/j.jksuci.2020.02.002.

[42] M. Sajjad *et al*., "Raspberry Pi assisted face recognition framework for enhanced law-enforcement services in smart cities," *Future Generation Computer Systems*, vol. 108, pp. 995-1007, 2020, doi: 10.1016/j.future.2017.11.013.

[43] L. Li, X. Mu, S. Li, and H. Peng, "A review of face recognition technology," *IEEE Access*, vol. 8, pp. 139110-139120, 2020, doi: 10.1109/ACCESS.2020.3011028.

[44] H. Ling, J. Wu, L. Wu, J. Huang, J. Chen, and P. Li, "Self residual attention network for deep face recognition," *IEEE Access*, vol. 7, pp. 55159-55168, 2019, doi: 10.1109/ACCESS.2019.2913205.

[45] A. Jha, "Classroom attendance system using facial recognition system," *The International Journal of Mathematics, Science, Technology, and Management*, vol. 2, no. 3, pp. 4-7, 2007, doi: 10.1051/itmconf/20203202001.

[46] R. I. Bendjillali, M. Beladgham, K. Merit, and A. Taleb-Ahmed, "Illumination-robust face recognition based on deep convolutional neural networks architectures," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 18, no. 2, pp. 1015-1027, 2020, doi: 10.11591/ijeecs.v18.i2.pp1015-1027.

[47] M. Z. Khan, S. Harous, S. U. Hassan, M. U. G. Khan, R. Iqbal, and S. Mumtaz, "Deep unified model for face recognition based on convolution neural network and edge computing," *IEEE Access*, vol. 7, pp. 72622-72633, 2019, doi: 10.1109/ACCESS.2019.2918275.

[48] A. Alzu'bi, F. Albalas, T. Al-Hadhrami, L. B. Younis, and A. Bashayreh, "Masked face recognition using deep learning: A review,"

[49] Y. Tang *et al*., "Recognition and localization methods for vision-based fruit picking robots: A review," *Frontiers in Plant Science*, vol. 11, p. 510, 2020, doi: 10.3389/fpls.2020.00510.

[50] H. A. Williams *et al*., "Robotic kiwifruit harvesting using machine vision, convolutional neural networks, and robotic arms," *Biosystems Engineering*, vol. 181, pp. 140-156, 2019, doi: 10.1016/j.biosystemseng.2019.03.007.

[51] S. D. Kumar *et al*., "A microcontroller based machine vision approach for tomato grading and sorting using SVM classifier," *Microprocessors and Microsystems*, vol. 76, p. 103090, 2020, doi: 10.1016/j.micpro.2020.103090.

[52] M. Halstead, C. McCool, S. Denman, T. Perez, and C. Fookes, "Fruit quantity and ripeness estimation using a robotic vision system," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 2995-3002, 2018, doi: 10.1109/LRA.2018.2849514.

[53] S. Wan and S. Goudos, "Faster R-CNN for multi-class fruit detection using a robotic vision system," *Computer Networks*, vol. 168, p. 107036, 2020, doi: 10.1016/j.comnet.2019.107036.

[54] M. T. Habib *et al*., "Machine vision based papaya disease recognition," *Journal of King Saud University-Computer and Information Sciences*, vol. 32, no. 3, pp. 300-309, 2020, doi: 10.1016/j.jksuci.2018.06.006.

[55] J. J. Zhuang *et al*., "Detection of orchard citrus fruits using a monocular machine vision-based method for automatic fruit picking applications," *Computers and Electronics in Agriculture*, vol. 152, pp. 64-73, 2018, doi: 10.1016/j.compag.2018.07.004.

[56] A. Gongal, M. Karkee, and S. Amatya, "Apple fruit size estimation using a 3D machine vision system," *Information Processing in Agriculture*, vol. 5, no. 4, pp. 498-503, 2018, doi: 10.1016/j.inpa.2018.06.002.

[57] Z. Wang, J. Underwood, and K. B. Walsh, "Machine vision assessment of mango orchard flowering," *Computers and Electronics in Agriculture*, vol. 151, pp. 501-511, 2018, doi: 10.1016/j.compag.2018.06.040.

[58] X. Zhao, P. Sun, Z. Xu, H. Min and H. Yu, "Fusion of 3D LIDAR and Camera Data for Object Detection in Autonomous Vehicle Applications," *IEEE Sensors Journal*, vol. 20, no. 9, pp. 4901-4913, 2020, doi: 10.1109/JSEN.2020.2966034.

[59] W. Zhao, W. Ma, L. Jiao, P. Chen, S. Yang and B. Hou, "Multi-Scale Image Block-Level F-CNN for Remote Sensing Images Object Detection," *IEEE Access*, vol. 7, pp. 43607-43621, 2019, doi: 10.1109/ACCESS.2019.2908016.

[60] T. Zhou, D. P. Fan, M. M. Cheng, J. Shen and L. Shao, "RGB-D Salient Object Detection: A Survey," *Computational Visual Media*, vol. 7, pp. 37-69, 2021, doi: 10.48550/arXiv.2008.00230.

[61] M. Shirpour, N. Khairdoost, M. A. Bauer and S. S. Beauchemin, "Traffic Object Detection and Recognition Based on the Attentional Visual Field of Drivers," in *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 1, pp. 594-604, Jan. 2023, doi: 10.1109/TIV.2021.3133849.

[62] S. S. A. Zaidi, M. S. Ansari, A. Aslam, N. Kanwal, M. Asghar and B. Lee, "A Survey of Modern Deep Learning Based Object Detection Models," *Digital Signal Processing*, vol. 115, p. 103514, 2022, doi: 10.48550/arXiv.2104.11892.

[63] A. Kuznetsova *et al*., "The Open Images Dataset V4: Unified Image Classification, Object Detection, and Visual Relationship Detection at Scale," *International Journal of Computer Vision*, vol. 128, no. 7, pp. 1956-1981, 2020, doi: 10.48550/arXiv.1811.00982.

[64] X. Chen, H. Li, Q. Wu, K. N. Ngan and L. Xu, "High-Quality R-CNN Object Detection Using Multi-Path Detection Calibration Network," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 2, pp. 715-727, 2021, doi: 10.1109/TCSVT.2020.2987465.

[65] S. Zhang, Y. Wu, C. Men and X. Li, "Tiny YOLO Optimization Oriented Bus Passenger Object Detection," *Chinese Journal of Electronics*, vol. 29, no. 1, pp. 132-138, 2020, doi: 10.1049/cje.2019.11.002.

[66] A. Suhail, M. Jayabalan and V. Thiruchelvam, "Convolutional Neural Network Based Object Detection: A Review," *Journal of Critical Reviews*, vol. 7, no. 11, pp. 786-792, 2020, doi: 10.48550/arXiv.1905.01614.

Electronics, vol. 10, no. 21, p. 2666, 2021, doi: 10.3390/electronics10212666.

[67] M. D. Hossain and D. Chen, "Segmentation for Object-Based Image Analysis (OBIA): A review of algorithms and challenges from remote sensing perspective," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 150, pp. 115-134, 2019, doi: 10.1016/j.isprsjprs.2019.02.009.

[68] H. Costa, G. M. Foody, and D. S. Boyd, "Supervised methods of image segmentation accuracy assessment in land cover mapping," *Remote Sensing of Environment*, vol. 205, pp. 338-351, 2018, doi: 10.1016/j.rse.2017.11.024.

[69] Y. Xiao, L. Daniel, and M. Gashinova, "Image segmentation and region classification in automotive high-resolution radar imagery," *IEEE Sensors Journal*, vol. 21, no. 5, pp. 6698-6711, 2020, doi: 10.1109/JSEN.2020.3043586.

[70] M. Z. Alom, C. Yakopcic, M. Hasan, T. M. Taha, and V. K. Asari, "Recurrent residual U-Net for medical image segmentation," *Journal of Medical Imaging*, vol. 6, no. 1, pp. 014006-014006, 2019, doi: 10.1117/1.JMI.6.1.014006.

[71] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, "U-net and its variants for medical image segmentation: A review of theory and applications," *IEEE Access*, vol. 9, pp. 82031-82057, 2021, doi: 10.1109/ACCESS.2021.3086020.

[72] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, "Deep learning techniques for medical image segmentation: achievements and challenges," *Journal of digital imaging*, vol. 32, pp. 582-596, 2019, doi: 10.1007/s10278-019-00227-x.

[73] A. Sinha and J. Dolz, "Multi-scale self-guided attention for medical image segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 1, pp. 121-130, 2021, doi: 10.48550/arXiv.1906.02849.

[74] F. Shi, J. Wang, J. Shi, Z. Wu, Q. Wang, Z. Tang, et al., "Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19," *IEEE Reviews in Biomedical Engineering*, vol. 14, pp. 4-15, 2021, doi: 10.1109/RBME.2020.2987975.

[75] F. Munawar, S. Azmat, T. Iqbal, C. Grönlund, and H. Ali, "Segmentation of lungs in chest X-ray image using generative adversarial networks," *IEEE Access*, vol. 8, pp. 153535-153545, 2020, doi: 10.1109/ACCESS.2020.3017915.

[76] S. Wang, D. M. Yang, R. Rong, X. Zhan, and G. Xiao, "Pathology image analysis using segmentation deep learning algorithms," *The American Journal of Pathology*, vol. 189, no. 9, pp. 1686-1698, 2019, doi: 10.1016/j.ajpath.2019.05.007.

[77] S. M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami, and M. K. Khan, "Medical image analysis using convolutional neural networks: a review," *Journal of Medical Systems*, vol. 42, pp. 1-13, 2018, doi: 10.1007/s10916-018-1088-1.

[78] Y. Lu *et al*., "Differentiate Xp11.2 Translocation Renal Cell Carcinoma from Computed Tomography Images and Clinical Data with ResNet-18 CNN and XGBoost," *CMES-Computer Modeling in Engineering & Sciences*, vol. 136, no. 1, 2023, doi: 10.32604/cmes.2023.024909.

[79] V. Narayan *et al*., "Enhance-Net: An Approach to Boost the Performance of Deep Learning Model Based on Real-Time Medical Images," *Journal of Sensors*, 2023, doi: 10.1155/2023/8276738.

[80] S. A. El-Feshawy *et al*., "IoT framework for brain tumor detection based on optimized modified ResNet-18 (OMRES)," *The Journal of Supercomputing*, vol. 79, no. 1, pp. 1081-1110, 2023, doi: 10.1007/s11227-022-04678-y.

[81] P. Chotikunnan *et al*., "Evaluation of Single and Dual image Object Detection through Image Segmentation Using ResNet-18 in Robotic Vision Applications," *Journal of Robotics and Control (JRC)*, vol. 4, no. 3, pp. 263-277, 2023, doi: 10.18196/jrc.v4i3.17932.