International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Fine-Tuning InceptionV3 for Thai Cuisine Image Classification: A Mobile Deployment Perspective

Somsak Saksat[1], Nathaphon Boonnam[2], Siriwan Kajornkasirat[3]

Submitted: 12/02/2025 **Revised:** 12/03/2025 **Accepted:** 22/03/2025

Abstract: This work presents the development of a smartphone application that utilizes deep learning techniques for the automatic classification of Thai food images. Transfer learning and fine-tuning approaches were compared using the InceptionV3 model, initially trained on the ImageNet dataset and subsequently refined with a dataset consisting of 49 varieties of Thai cuisine images. Experimental results indicate that the fine-tuning model achieved superior performance, attaining an accuracy of 95.22% on the validation set, surpassing the transfer learning model, which achieved an accuracy of 85.43%. Additionally, the fine-tuning model exhibited a stable and consistent decrease in loss without significant overfitting, making it the preferred choice for application development. We converted this model to TensorFlow Lite to enable offline functionality on smartphones developed using Flutter. However, retrieving detailed nutritional information still requires an online database connection to ensure comprehensive nutrient data, including calories, protein, fat, and carbohydrates. This research demonstrates the potential of combining fine-tuning methods with mobile application development to promote mindful food consumption, reduce the risk of non-communicable diseases, and enhance quality of life in the digital era. Moreover, the application supports the United Nations Sustainable Development Goal 3: Good Health and Well-being by encouraging healthier lifestyle choices and contributing to improved health outcomes. Furthermore, it provides a valuable framework for the sustainable promotion of Thai food culture.

Keywords: *Thai Food Classification, InceptionV3, Transfer Learning, Fine-Tuning, TensorFlow Lite, Mobile app*

1. Introduction

Thai cuisine is classified as one of the world's cultural treasures due to its distinctive preparations and diverse menus [5], [6]. Its physical characteristics—vibrant colors, careful plating, and a variety of ingredients rooted in local wisdom—represent the geography, culture, and way of life of each region in Thailand. However, this diversity also makes it challenging to categorize Thai dishes based on images. Although cooking techniques and presentations may differ, many Thai dishes share similar ingredients, such as meats, vegetables, and herbs, which makes them visually identical. For instance, Pad Krapow, Boat Noodles, and Tom Yum Goong are popular dishes that often appear nearly similar from a computer's perspective [20].

This classification challenge is significant, considering that many Thai people lack access to easy-to-understand and reliable dietary information, especially when eating out or consuming single-plate meals without nutritional labels. According to the 2023 Thai Health Report published by the Thai Health Promotion Foundation (Thai Health, 2023, available at https://www.thaihealth.or.th/wp-content/uploads/2023/11/ Thai-Health-2023.pdf) and data from the Department of Health, Ministry of Public Health (Ministry of PublicHealth, 2023, available at https://nutrition2.anamai.moph.go.th/en/newsanamai/download/?did=221712&id=128340&reload=), Thailand, had more than 6.5 million people diagnosed with diabetes in 2023. Moreover, we recorded over 20 million people as overweight or obese. These conditions are strongly linked to non-communicable diseases (NCDs), including hypertension, cardiovascular disease, type 2 diabetes, and some types of cancer [7]–[10].

Artificial intelligence (AI) and deep learning have recently become key tools for building image classification models, particularly in areas such as food and agriculture that require high accuracy and real-time processing [1]–[3], [11]. Medina et al. [1] proposed a mobile application based on MobileNetV2 for detecting potato leaf diseases, achieving an accuracy of up to 98.7% even on offline Android devices. Tian and Sun [2] embedded a TensorFlow model in an AI_BIRDER Flutter-based app, which classifies bird images offline and displays the top 3 predictions along with their confidence scores. Rakesh et al. [3] compared MobileNetV2, ResNet152, and YOLOv8 in fish species classification, reporting YOLOv8 as the most accurate model, though its size is too large for lightweight mobile use.

From reviewing the literature, convolutional neural networks (CNNs), combined with transfer learning techniques—particularly fine-tuning InceptionV3—can achieve strong performance in food image classification tasks [12]–[15], [18], [19]. When these models are converted to TensorFlow Lite, they remain efficient enough for use on mobile devices [1], [2], [16]. Flutter also supports cross-platform app development that runs smoothly in environments with limited internet access [2], [16].

1 Faculty of Science and Industrial Technology, Prince of Songkla University, Surat Thani Campus, Surat Thani 84100, Thailand
ORCID ID : 0009-0006-3596-3474
2 Faculty of Science and Industrial Technology, Prince of Songkla University, Surat Thani Campus, Surat Thani 84100, Thailand
ORCID ID : 0000-0002-0625-4576
3 Faculty of Science and Industrial Technology, Prince of Songkla University, Surat Thani Campus, Surat Thani 84100, Thailand;
ORCID ID : 0000-0002-9229-8443
** Corresponding Author Email: nathaphon.b@psu.ac.th*

In this paper, we developed a Thai food image classification system using the InceptionV3 architecture. Both transfer learning and fine-tuning techniques were applied and compared. We then utilized the better-performing model to develop a mobile application that enables users to take a picture of food and receive database that provides real-time nutritional information. The integration of image classification technology with a smartphone app promotes informed food choices and supports the United Nations Sustainable Development Goal 3: Good Health and Well-being [17]. We aim for the app to be freely accessible, supporting the improvement of users' quality of life in the digital age [21], [22].

2. Materials And Methods

2.1. Data Collection

To create a reliable model for classifying Thai food, we generated a dataset to train the model by getting pictures from 49 different groups. There were 48 distinct varieties of well-known Thai food, and one category had pictures of things that weren't food. This last group was introduced to enable the model to distinguish between food and non-food content. The cuisine was presented in various ways in each session, including different angles, backgrounds, and lighting conditions. The images were collected using two primary methods. First, we utilized Google Image Search with carefully selected Thai and English keywords, such as "ต้มยำกุ้ง" and "Tom Yum Goong," to retrieve food images available on the internet. Second, we obtained additional photos from various food-related websites and open datasets. This included food blogs, culinary websites, and a widely recognized dataset known as THFOOD-50, which contains over 15,000 images spanning 50 categories of Thai dishes. The THFOOD-50 dataset is publicly available for research and educational purposes.

We manually checked each image after gathering the data. We removed anything blurry, had the wrong label, or had watermarks. We put them into folders by category so we could readily use them in training. The input supported by the InceptionV3 model requires all images to be resized to 299 x 299 pixels, and their pixel values are standardized to the range of 0 to 1 to enhance learning efficiency. The dataset was divided into three parts: a training set of 11,760 images (80% of the total data), a validation set of 1,470 images (10%) used to validate the performance of the trained model regarding overfitting, and a test set of 1,470 images (10%) used to evaluate the model's performance after formal training.

Therefore, we applied online image augmentation during each training epoch to increase data variability and prevent the model from overfitting to the original images. We randomly transformed the training images with the following augmentations: rotation within ±15 degrees, zooming by up to ±10%, horizontal and vertical shifts of up to ±10%, slight shear transformations with a shear intensity of 0.05, and random brightness adjustments ranging from 90% to 110% of the original brightness. Pixel values for the validation and test sets were normalized to the range of 0 to 1 without any augmentation, ensuring an accurate evaluation of the model's performance.

2.2. Model building and InceptionV3 training

Based on the architecture of the InceptionV3 model, one of the highly efficient Convolutional Neural Network (CNN) models, it has been pre-trained on the ImageNet dataset, which covers over 1.2 million images across 1,000 categories. This pre-training stage guides the development and training of the Thai food image classification model, enabling efficient extraction of deep photo features. Since the Thai food dataset differs from ImageNet, various structural changes were made, mainly to the classification layers, which were not suitable for the new problem. The original fully connected layers of ImageNet were excluded using the include_top=False parameter, allowing new classification layers explicitly designed for 49 varieties of Thai food to be added. Transfer learning and fine-tuning constitute two main approaches for model training. Each method employs a distinct set of models trained separately—not sequentially—to allow an honest and straightforward comparison.

The first technique, transfer learning, use the pre-trained InceptionV3 model without modifying the base model's weights, specifically setting base_model.trainable = False freezes the feature extraction layers, which we consider sufficiently capable of recognizing generic image features. In addition to the base model, a new classification head is introduced, comprising a GlobalAveragePooling2D layer that reduces the spatial dimensions to a small vector, which is then passed to the next layer. A Dense layer with 1024 units and ReLU activation learns specific features of Thai cuisine images. The final Dense layer uses a softmax activation to convert outputs into probabilities across all 49 classes. The model is compiled using the Adam optimizer with a learning rate of 0.001, suitable for training the new classification layers while keeping the base model frozen. Training is conducted for thirty epochs on the training and validation sets without early stopping, allowing the model to learn fully according to the set number of epochs.

The second technique, fine-tuning, enables the model to better adapt to the new dataset by unfreezing the base layers. Specifically, setting the base model.trainable = True allows the weights of the base layers to be updated. This adjustment enables the model to refine its features to suit Thai food images better. However, to prevent drastic changes that could erase previously learned ImageNet features, the learning rate is reduced to 0.00001. Apart from this learning rate change, the classification head's structure and compilation method remain the same as in transfer learning. The model is also trained for thirty epochs without early stopping to ensure a fair comparison. By training the models separately using these two techniques, we can independently evaluate the performance of transfer learning and fine-tuning. This enables us to select the most effective model for practical application in Thai food image classification within the developed mobile application.

2.3. Model Performance

In this study, two models were trained and evaluated, one using transfer learning and the other fine-tuning the InceptionV3 architecture. We tested both models on independent test datasets to assess their classification performance.

Model accuracy is used to measure the overall proportion of correctly classified instances [1]. However, accuracy alone is not sufficient to fully assess a model's performance, especially when dealing with imbalanced datasets. Therefore, additional metrics are used: precision, recall, and F1 score. Precision measures the proportion of true positive predictions among all positive predictions made by the model, reflecting the accuracy of the positive classification [2]. Recall measures the model's ability to identify all relevant instances within a class, indicating the model's sensitivity [3]. The F1 score provides the harmonic mean of precision and recall, providing a balanced metric that is particularly useful in situations where the class distribution is uneven [4].

In addition, a confusion matrix is generated to show the misclassification pattern and identify the food categories that are most frequently confused by the model. As expected, foods with similar characteristics show higher confusion rates. This highlights the challenges inherent in fine-grained food classification [5].

We tracked the training progress by constructing accuracy and loss curves over 30 epochs. The fine-tuned model demonstrated more stable convergence and superior performance compared to the transfer learning model, exhibiting smoother improvements and fewer irregular fluctuations in the validation metrics [6]. These results suggest that fine-tuning yields a more robust model, which is suitable for practical applications in mobile food recognition.

2.4. Model TensorFlow Lite conversion

After reviewing the results of the two models from the previous stage, we selected the one that performed better for use on mobile devices. We then switched to TensorFlow Lite (TFLite), a version optimized for devices with limited resources, such as smartphones and tablets. This conversion method makes the model smaller and simpler, which speeds up processing, uses less memory, and maintains high accuracy in identifying Thai cuisine photographs. The TFLite model can also operate offline, allowing users to obtain results immediately. Plus, the image data doesn't have to be transferred to the server, which makes it safer and more private for users. Therefore, the step of converting to TFLite is crucial for utilizing AI models in mobile apps.

2.5. Application Development and Integration

Once the model was in TensorFlow Lite form, incorporating it into the mobile app was a simple second step. Designed with Flutter, a cross-platform framework that allows deployment on both Android and iOS from a single Dart script, the app features a simple, user-friendly layout designed for ease of use. Users could select fresh images straight from the program's camera function or photos from the gallery. This capability was developed using Flutter's basic widgets, along with additional packages such as image_picker. Trained on a dataset of 49 Thai cuisine categories, the enhanced InceptionV3 model, which we converted to TensorFlow Lite, was inserted into the app via the TensorFlow Lite Flutter plugin. Resizing, scaling, and normalizing input images helps them to fit the expected input form of the model before classification. Once processed, the system produces the top three expected food categories, along with their corresponding confidence ratings.

After classification, the app links to a Firebase database to retrieve matching nutritional information, which is then displayed alongside the food picture, predicted labels, and confidence ratings. Since the dietary data is collected online, an internet connection is required to access the most current information. Figure 1 aggregates the flow from image input to model inference to provide nutritional information on the user interface and highlight the entire system architecture.

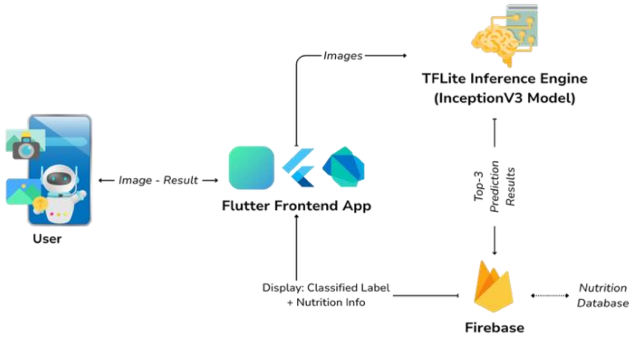


Fig. 1. System Architecture Diagram

3. Results

We aim to compare the performance of two image processing techniques—transfer learning and fine-tuning—using the InceptionV3 architecture. Pre-trained on the ImageNet dataset, which contains over one million images spanning more than 1,000 categories, InceptionV3 is a large Convolutional Neural Network (CNN) capable of extracting highly generalizable features.

In the transfer learning approach, the pre-trained InceptionV3 model is utilized without modifying the convolutional base (i.e., the weights are frozen), and only the classifier head is retrained using our dataset. This method significantly reduces training time and requires less data to achieve the same results. However, its major limitation lies in the model's inability to adapt to domain-specific features, especially if those features differ significantly from the patterns seen in the original ImageNet dataset.

Conversely, fine-tuning involves unfreezing some or all of the convolutional base layers to allow further learning from new data. This approach enables the model to refine feature extraction for the specific domain—in this case, Thai cuisine—by adjusting weights from low-level (edges, textures) to high-level features (shapes and components of food items). In this study, we fine-tuned the entire model to achieve more profound and more relevant learning.

Both models were trained using 299×299 pixel images—the default input size for InceptionV3—and optimized using the Adam algorithm. This gradient-based method adaptively adjusts the learning rate for each parameter. We used a learning rate of 1e-3 for transfer learning, which is suitable for training dense layers with significant weight updates. Fine-tuning employed a lower learning rate of 1e-5 to mitigate the risk of catastrophic forgetting, where the model loses the general knowledge gained from pre-training.

We conducted training over 30 epochs, a choice guided by both literature and practical evaluation. Prior studies have shown that training image classification models using architectures like InceptionV3 typically achieves stable convergence and optimal performance within 25 to 35 epochs, beyond which improvements tend to plateau or lead to overfitting [1], [14], [15]. In our pilot runs, validation accuracy began to stabilize around the 28th epoch, which reinforced our decision to cap the training at 30 epochs, thereby striking a balance between accuracy and efficiency.

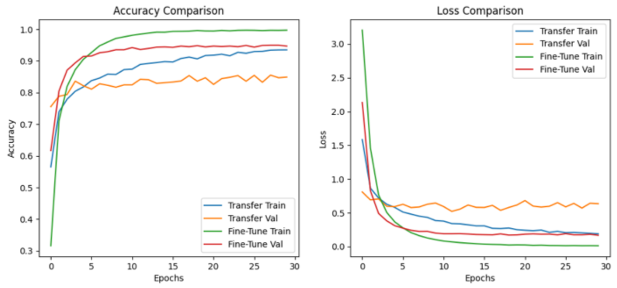


Fig. 2. Accuracy and Loss Comparisons

The fine-tuning model achieved a validation accuracy of 95.22%, significantly outperforming the transfer learning model's 85.43%, as shown in Figure 2. This result indicates that fine-tuning enables the model to capture the distinct features of Thai food images more effectively, resulting in superior generalization to unseen images. Toward the final epochs, the loss of the fine-tuned model decreased steadily and plateaued, suggesting stable learning. The training accuracy reached 99.98%, and the low training loss indicates strong learning capacity without signs of overfitting, as the validation metrics remained consistent.

Retaining the pre-trained convolutional base from ImageNet and training just the classifier head, the transfer learning model attained 92.31% training accuracy and 85.43% validation accuracy. Although this method provides efficiency in terms of training time and data requirements, the results suggest that it struggles to generalize effectively to Thai cuisine features. Table 1 reveals explicitly that this model achieved a macro-averaged precision, recall, and F1-score of 0.85. Class 23 (0.43) had the lowest recall, while the same class (0.55) had the lowest F1-score. These principles suggest challenges in differentiating meals with comparable aesthetic qualities. Although Class 48 achieved perfect accuracy (1.00), general performance among the other classes was variable. On the other hand, the finely adjusted model—which allows updates to the entire network, including the convolutional base—achieved a training accuracy of 99.98% and a significantly higher validation accuracy of 95.22%. It maintained macro and weighted average scores of 0.96 throughout precision, recall, and F1-score, as shown in Table 1.

Table 1. Transfer Learning and Fine-tuning classification report

Key Points	Transfer Learning	Fine-Tuning
Overall Accuracy	0.85	0.96
Class with Highest Precision	Class 48: 1.00	Classes 16, 17, 34, 43: 1.00
Class with Lowest Precision	Class 27: 0.66	Class 22: 0.83
Class with Highest Recall	Class 48: 1.00	Multiple classes such as 0, 3, 14, 16, 20, 24, 34, 43
Class with Lowest Recall	Class 23: 0.43	Class 31: 0.83
Class with Highest F1-Score	Class 48: 1.00	Classes 3, 24, 34, 43: 1.00
Class with Lowest F1-Score	Class 23: 0.55	Class 31: 0.89
Macro Avg: Precision/Recall/F1	0.86 / 0.85 / 0.85	0.96 / 0.96 / 0.96
Weighted Avg: Precision/Recall/F1	0.86 / 0.85 / 0.85	0.96 / 0.96 / 0.96

With flawless F1-scores in Classes 3, 24, 34, and 43, and no class dropping below a 0.83 recall or 0.89 F1-score, the model demonstrated great performance in several categories, showing considerable generalization and class-wise consistency.

Furthermore, as shown in Appendices A and B, we calculate the prediction distribution by class, providing additional data to support our analysis. Out of 30 photos in each class, the refined model consistently produced 25–30 accurate predictions, thereby emphasizing strong true positive (TP) rates. The transfer learning model revealed reduced TP values as well as increased false positives, while some classes had as few as 13 accurate predictions. This implies that the transfer learning model often confuses visually identical dishes, such as those with overlapping ingredients or presentation techniques.

Building upon these promising results, we integrated the fine-tuned InceptionV3 model into a mobile application. The workflow begins on the main screen of the application, which displays a dashboard showing the remaining daily nutrient content and a list of the user's last three recorded menu items. When the user wants to record a new meal, there is a "Nutrilens" option to utilize the AI feature for automatic food image analysis. When entering the Nutrilens page, users can select a food image in two ways: by taking a new photo or selecting an image from the gallery. After preprocessing, including image scaling and normalization, the resultant picture is fed into the model for classification. A screen showing the model's processing state appears throughout this period. The system displays the classified results—top-3 labels with a confidence score—when processing is complete. After that, the application links to Firebase to retrieve the nutritional information that matches the top-ranked result and displays it on the results screen. The user can then register the food data in the system using the meal recording screen.

The development and practical application of the fine-tuned InceptionV3 model demonstrate that it can be efficiently deployed on mobile devices via TensorFlow Lite. This provides low-latency inference and seamless connectivity to external databases, reflecting the AI model's readiness for real-world use cases, such as daily health and nutrition tracking.

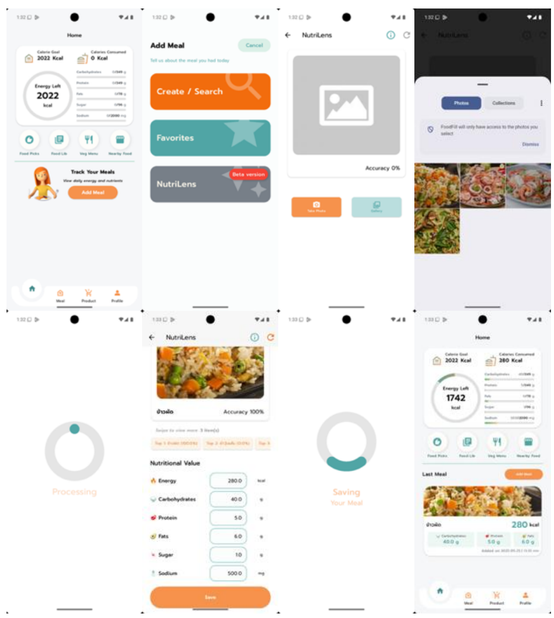


Fig. 3. Screen of Thai Food Classification and Logging using Nutrilens Feature

4. Discussion

Experimental results in classifying Thai food images using the

InceptionV3 architecture revealed that the fine-tuning technique yielded significantly better results than transfer learning. Transfer learning involves using a model trained on the ImageNet dataset while retaining the weights of the base layers and training only the model's head with our data. This method has the advantage of saving time and using less data. However, the main limitation is that the model struggles to adapt to the specific characteristics of Thai food images, particularly when these characteristics differ from those in ImageNet. For example, dishes with similar ingredients, such as chili or spices, can confuse the model, resulting in lower accuracy.

Moreover, fine-tuning enables the model to partially or fully "relearn" from Thai food image data, allowing it to grasp specific characteristics such as color, texture, and shape of each dish more deeply. This resulted in a validation accuracy of up to 95.22% with stable performance and minimal overfitting. The confusion matrix further confirms that fine-tuning improves classification accuracy across all classes. At the same time, transfer learning showed lower recall and more incorrect predictions, particularly among classes with similar visual features. Additionally, data augmentation techniques—such as rotation, zooming, brightness adjustment, and image distortion—play a crucial role in enhancing the diversity of training data, enabling the model to learn image variations under different conditions better and thereby improving its ability to generalize to new, unseen images.

However, this study has certain limitations. The dataset size, although sufficient for training, could be expanded further to improve the model's robustness. The choice of input image resolution (299×299 pixels) was based on the default requirement for InceptionV3; experimenting with different image sizes might yield additional insights into model performance. Moreover, the training duration of 30 epochs, while adequate to reach convergence in this study, might be optimized using early stopping or learning rate scheduling techniques to reduce training time and computational costs.

Based on all the experimental results, fine-tuning the InceptionV3 model is more suitable for Thai food image classification because it can effectively adapt to the specific characteristics of the data, resulting in significantly higher overall accuracy.

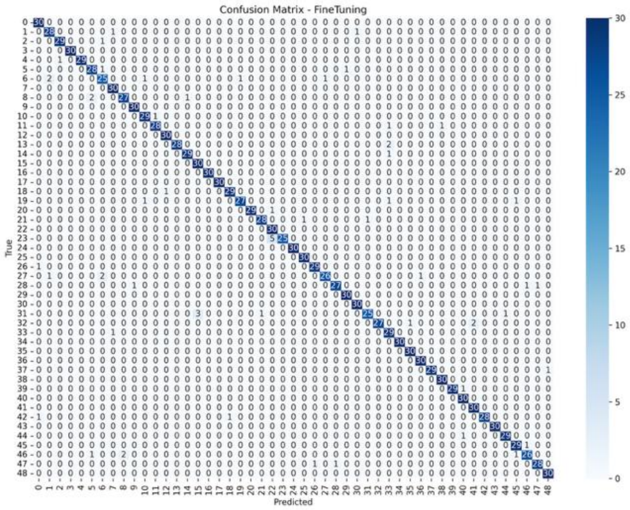
5. Conclusion

The comprehensive research in this work clearly demonstrates, from both numerical results (accuracy, loss) and qualitative results (confusion matrix), that fine-tuning on the InceptionV3 architecture outperforms transfer learning in the context of multi-class image classification, exhibiting comparable characteristics. This demonstrates how fine-tuning facilitates the learning of more complex and specialized features. Although it offers advantages in terms of speed and simplicity of use, transfer learning lacks sufficient accuracy in the context of challenging photo categorization tasks with many classes. Therefore, fine-tuning should be chosen as the primary approach when there is sufficient data and the necessary resources to support model training costs.

For future research and development, we can further enhance the model by incorporating other advanced architectures such as EfficientNet, DenseNet, or ResNet. Comparative studies involving these architectures, along with both transfer learning and fine-tuning approaches, could provide deeper insights. Moreover, exploring automatic data augmentation techniques, such as AutoAugment or RandAugment, could reduce the risk of overfitting and improve generalization. Additionally, optimizing hyperparameters such as learning rate, batch size, and epochs using learning rate scheduling or discriminative learning rates can enhance the model's efficiency.

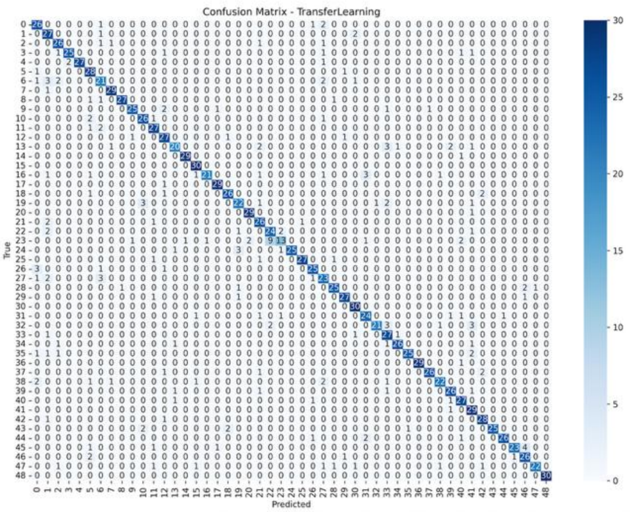
5.1. Appendix

Appendix A: Confusion Matrix – Fine-Tuned InceptionV3 Model This external view displays a detailed Confusion Matrix of the InceptionV3 model, providing a detailed examination of each class of supplements. The table shows the True Positives and the concentrations between classes. The model exhibits high predictive accuracy within the class, yielding 25-30 True Positives (TPs) from 30 images per class, excellent potency concentrations, and integrated food images.



Appendix B: Confusion Matrix – Transfer Learning InceptionV3 Model

This appendix presents the Confusion Matrix of the InceptionV3 model using Transfer Learning without deep tuning. It reveals a reduced number of True Positives (TP) compared to the Fine-Tuning model and an increase in False Positives in some classes, especially those with similar food characteristics. This makes it easy for the model to confuse images of food with similar ingredients or arrangements.



Appendix C: Table of Menu Code, Thai Name, and English Name for Thai Food Classification

This appendix compiles a table of menu codes, along with their

Thai and English names, for each Thai food used in image data classification. This information helps readers clearly understand and verify the accuracy of each class and is also an essential reference for interpreting the model's results in this research.

Code No.	Thai Name	English Name
00	แกงเขียวหวานไก่	Chicken Green Curry
01	แกงเทโพ	Pork Curry with Morning Glory
02	แกงเลียง	Spicy mixed vegetable soup
03	แกงจืดเต้าหู้หมูสับ	Pork Chopped Tofu Soup
04	แกงจืดมะระยัดไส้	Stuffed Bitter Gourd Broth
05	แกงมัสมั่นไก่	Chicken Mussaman Curry
06	แกงส้มกุ้ง	Sour Soup
07	ไก่ผัดเม็ดมะม่วงหิมพานต์	Stir Fried Chicken with Chestnuts
08	ไข่เจียว	Omelet
09	ไข่ดาว	Fried egg
10	ไข่พะโล้	Egg and Pork in Sweet Brown Sauce
11	ไข่ลูกเขย	Egg with Tamarind Sauce
12	กล้วยบวชชี	Banana in coconut milk
13	ก๋วยเตี๋ยวคั่วไก่	Stir Fried Rice Noodles with Chicken
14	กะหล่ำปลีผัดน้ำปลา	Fried Cabbage with Fish Sauce
15	กุ้งแม่น้ำเผา	Grilled River Prawn
16	กุ้งอบวุ้นเส้น	Baked Prawns With Vermicelli
17	ขนมครก	Coconut rice pancake
18	ข้าวเหนียวมะม่วง	Mango Sticky Rice
19	ข้าวขาหมู	Thai Pork Leg Stew
20	ข้าวคลุกกะปิ	Shrimp Paste Fried Rice
21	ข้าวซอย	Curried Noodle Soup with Chicken
22	ข้าวผัด	Fried rice
23	ข้าวผัดกุ้ง	Shrimp Fried Rice
24	ข้าวมันไก่	Steamed capon in flavored rice
25	ข้าวหมกไก่	Thai Chicken Biryani
26	ต้มข่าไก่	Thai Chicken Coconut Soup
27	ต้มยำกุ้ง	River prawn spicy soup
28	ทอดมัน	Fried fish-paste balls
29	ปอเปี๊ยะทอด	Deep fried spring roll
30	ผักบุ้งไฟแดง	Stir-Fried Chinese Morning Glory
31	ผัดไท	Fried noodle Thai style with prawns
32	ผัดกะเพรา	Stir fried Thai basil with minced pork
33	ผัดซีอิ๊วเส้นใหญ่	Fried Noodle in Soy Sauce
34	ผัดฟักทองใส่ไข่	Stir-fried Pumpkin with Eggs
35	ผัดมะเขือยาวหมูสับ	Stir-Fried Eggplant with Soybean Paste Sauce
36	ผัดหอยลาย	Stir Fried Clams with Roasted Chili Paste
37	ฝอยทอง	Golden Egg Yolk Threads
38	พะแนงไก่	Chicken Panang
39	ยำถั่วพู	Thai Wing Beans Salad
40	ยำวุ้นเส้น	Spicy Glass Noodle Salad
41	ลาบหมู	Spicy minced pork salad
42	สังขยาฟักทอง	Egg custard in pumpkin
43	สาคูไส้หมู	Tapioca Balls with Pork Filling
44	ส้มตำ	Green Papaya Salad
45	หมูปิ้ง	Thai-Style Grilled Pork Skewers
46	หมูสะเต๊ะ	Pork Satay with Peanut Sauce
47	ห่อหมก	Steamed Fish with Curry paste
48	unknown	unknown

6. References And Footnotes

6.1. References

[1] D. Medina, I. M. Cabrera, R. A. de la Cruz, L. Guerra Arzuaga, S. Cuello Portal, and M. Bianchini, "A mobile app for detecting potato crop diseases," J. Imaging, vol. 10, no. 2, p. 47, 2024, doi: 10.3390/jimaging10020047.

[2] C. Tian and Y. Sun, "AI_BIRDER: An intelligent mobile application to automate bird classification using artificial intelligence and deep learning," Comput. Sci. Inf. Technol., 2022.

[3] R. G. Rakesh et al., "Fish classification with machine learning: Enhancing accuracy and efficiency," Int. J. Comput. Digit. Syst., 2024.

[4] P. Chotwanvirat, A. Prachansuwan, P. Sridonpai, and W. Kriengsinyos, "Automated artificial intelligence–based Thai food dietary assessment system," Curr. Dev. Nutr., vol. 8, p. 102154, 2024.

[5] A. Phonrat and S. Pholphirul, "Culinary culture and Thai food exports," J. Asian Dev., vol. 6, no. 1, pp. 45–59, 2020.

[6] S. Sirirak, "Thai food culture: Identity and globalization," Thai J. Cult. Stud., vol. 12, no. 2, pp. 10–18, 2022.

[7] Thai Health Promotion Foundation, "Thai health report 2023." [Online]. Available: https://www.thaihealth.or.th

[8] Ministry of Public Health Thailand, "Obesity and diabetes data 2023." [Online]. Available: https://nutrition2.anamai.moph.go.th

[9] World Health Organization, "Noncommunicable diseases," 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases

[10] B. M. Popkin, S. Kim, and N. Slining, "The nutrition transition and its health implications in lower- and middle-income countries," Public Health Nutr., vol. 10, no. 1, pp. 56–65, 2021.

[11] H. Fang et al., "Food recognition using CNNs with weakly supervised learning," IEEE Trans. Image Process., vol. 29, pp. 685–698, 2020.

[12] S. Kawano and K. Yanai, "Food image recognition with deep convolutional features," in Proc. IEEE Int. Conf. Image Process. (ICIP), 2014, pp. 1894–1898.

[13] X. Liu et al., "DeepFood: Deep learning-based food image recognition for dietary assessment," in Proc. IEEE Int. Conf. Smart Homes Health Telemat. (ICOST), 2016.

[14] M. Hassannejad et al., "Food image classification using deep convolutional networks," in Proc. Int. Conf. Biomed. Health Inform. (ICBHI), 2016.

[15] C. Szegedy et al., "Rethinking the inception architecture for computer vision," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 2818–2826.

[16] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in Proc. Int. Conf. Mach. Learn. (ICML), 2019, pp. 6105–6114.

[17] United Nations, "Sustainable development goal 3: Ensure healthy lives and promote well-being for all," 2023. [Online]. Available: https://sdgs.un.org/goals/goal3

[18] Z. Yang et al., "Chinese food recognition using deep neural networks," Neurocomputing, vol. 372, pp. 187–197, 2020.

[19] T. Mezgec and B. Koroušić Seljak, "Nutrinet: A convolutional neural network for food and drink image recognition," Nutrients, vol. 9, no. 7, p. 657, 2017.

[20] K. Tanprasert and C. Boonbrahm, "Thai food image recognition using convolutional neural networks," in Proc. Int. Conf. Signal Process. Inf. Secur. (ICSPIS), 2021, pp. 1–6.

[21] Y. W. Teo and K. F. Lim, "Nutrition knowledge and food choice among Malaysian adults," Asia Pac. J. Clin. Nutr., vol. 27, no. 2, pp. 390–398, 2018.
[22] C. A. Monteiro et al., "Ultra-processed foods: What they are and how to identify them," Public Health Nutr., vol. 22, no. 5, pp. 936–941, 2019.
[23] THFOOD-50 Dataset. Fine-grained Thai food image classification dataset. [Online]. Available: https://drive.google.com/ (access restricted for educational/non-commercial use)

Author contributions

Somsak Saksat: Contributed significantly to the theoretical development, system and experimental design, prototype implementation, and the analysis and interpretation of the results. Also responsible for drafting the manuscript and revising it for intellectual content. Approved the final version of the article for publication.
Nathaphon Boonnam*: Contributed to the theoretical framework, system design, and data analysis. Participated in discussions regarding experimental validation. Approved the final version of the article for publication.
Siriwan Kajornkasirat: Provided critical revision of the manuscript for important intellectual content. Contributed to reviewing and editing the final draft. Approved the final version of the article for publication.