Food Allergen Detection in Malaysian Food Using Convolutional Neural Networks

Pei Lin Khor* and Wong Jee Keen Raymond

Department of Electrical Engineering, Faculty of Engineering, Universiti of Malaya, Malaysia

*Corresponding Author

doi: https://doi.org/10.21467/proceedings.141.15

ABSTRACT

Food allergy is a rising, global epidemic. Some Malaysian cooking contains food-allergic-reaction-causing ingredients that may cause severe allergic reactions. A food allergen detection system in Malaysian food is proposed for tourists with food allergies who are unfamiliar with the wide variety of Malaysian dishes to prevent severe allergic reactions. This work focuses on three major food allergens, which include peanuts, cow's milk, and shellfish. A new Malaysian food image dataset was prepared, and transfer learning on the custom dataset was done via fine-tuning and feature extraction techniques. Comparisons on the ResNet50, InceptionV3, and VGG16 architectures are done based on the accuracy of each model on the testing data. The VGG16 architecture is concluded as the most suitable neural network model for food allergen detection in Malaysian food. The proposed classifier achieved an accuracy of 80.56% on the test samples. The final model is loaded into a Graphical User Interface (GUI) application to demonstrate the results of the Malaysian food classification model.

Keywords: Food allergen detection, Malaysian food, Transfer learning

1 Introduction

Food allergy is defined as an immunological mechanism of the human body to certain food [1]. It is a concern for the quality of life of food-allergic individuals and the food industry economy. This has led to the study of machine learning in food allergen detection via food recognition. In Malaysia, multi-race and multi-ethnicity give rise to various food habits, practices, cultures, and traditions. Malaysian cuisines pose a threat to food-allergy sufferers as it contains peanuts, shellfish, or milk. For instance, shrimp paste, known locally as *belacan*, is a common Malaysian cooking ingredient [2]. Unsuspecting foreigners may mistakenly consume food that could cause an allergic response.

In this research, a visual database with 5,284 images of Malaysian food from 36 different food types was created. Using convolutional neural network (CNN)-based approach, an application for food allergen identification in Malaysian food was developed. This project aims to contribute to the food-allergic community in detecting food allergens such as peanuts, milk, and shellfish through machine learning.

2 Related Work

Certain literature works have underlined that there is a difference even among food items of the same class. [3] stressed that two dishes of the same name may not be cooked using the same ingredients. For a dish under the same category, there is a high degree of variance and visual diversity in terms of shape, colour, texture, and each chef's preference. Thus, intraclass variation of food categories is a major drawback in food recognition applications [4]. Much research has been conducted on dietary monitoring applications such as portion tracking, calorie counting, and nutrition content approximation. Existing food-allergy-related applications aid the user in locating allergy-friendly restaurants and breaking down menu items that are safe to consume.

Proceedings DOI: 10.21467/proceedings.141; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-957605-4-1

Very few works highlight the problem of allergen detection in food since only a minority has mild to zero food allergies. Only the work by [5] explicitly targets mobile applications for food-allergic people. However, the setting is for food products in a German supermarket. This work contributed to the field of food recognition and on the topic of food allergens. A Malaysian food image dataset comprising 36 classes of food was compiled. A novel neural network, which has a high accuracy and precision in identifying the presence of food allergens in Malaysian food was developed.

3 Methodology

Comprising of 36 food classes, 5,284 photos of Malaysian food were collected using a Python script. The ingredients for all the food in the database were searched from online sources and summarized in a CSV file. The presence of peanut, cow's milk, or shellfish allergens in each food class was identified. Food that does not contain any of the three major allergens is given a 'Safe' label. Food that contains any one of the three allergens is rated as 'Unsafe' whereas food that may contain trace amounts of peanut, milk, or shellfish are labelled as 'Potentially unsafe'. As shown in Table 1, there are 21 food types categorized as 'Unsafe', 5 as 'Potentially Unsafe', while 10 classes are 'Safe'.

From the Malaysian food dataset, images are split into train data, validation data, and test data. The training set allows the model to learn from the image data's features and patterns. The same data is fed to the model architecture at each epoch. The model continues to learn from the features during each forward pass, and update the weights assigned to each neuron connection. The validation set is used to validate the model's efficiency during model training. This process provides insight to tune the model's hyperparameters and configurations. The validation accuracy and validation loss tell whether the training is going in the right direction by giving the first test against new data. Model overfitting happens when the model is excellent in classifying the training samples but is unable to generalize and gives false classification on data that the model has not seen before [6]. Hence, the test set is used for model testing after the model completes its training and evaluation processes.

In general, the user uploads an image of a Malaysian food to the server which receives the raw data and performs data processing. The data was fed to the trained CNN model to predict the type of Malaysian food. The food allergen information in various Malaysian food types is stored in a CSV (comma-separated values) file. The allergen information is used to determine if the food is safe, unsafe, or potentially unsafe for consumption based on the type of food allergies the user has. Once the food is identified, the program will perform an analysis to determine if food allergens like peanuts, cow's milk, or shellfish are present in the food item. The type of foods identified by the model and any food allergens detected will be displayed to the user through a GUI (Graphical User Interface) application. The user can be a local Malaysian, a foreigner, a food-allergic person, or anyone who might find this application useful when consuming Malaysian food.

Table 1: *Allergen Rating for 36 Malaysian Food Classes*

Allergen Rating	Number of Classes	Food Names
Unsafe	21	Ais Kacang, Asam Pedas, Bak Chang, Char Keow Teow, Hokkien Mee (Dai Lok Mee), Ikan Bakar, Kangkung Belacan, Kek Batik, Lei Cha, Lotus Root Soup, Murukku, Nasi Kerabu, Nasi Lemak, Oyster Omelette, Penang Laksa, Pie Tee Kuih, Prawn Mee, Rempeyek, Rojak, Roti Canai, Sarawak Layer Cake

Proceedings DOI: 10.21467/proceedings.141
ISBN: 978-81-957605-4-1

Series: AIJR Proceedings
ISSN: 2582-3922

Potentially Unsafe	5	Ang Koo Kueh, Bak Kut Teh, Otak-Otak, Satay, Wan Tan Mee
Safe	10	Bubur Kacang Hijau, Bubur Pulut Hitam, Cendol, Keropok Lekor, Ketupat, Kuih Bangkit, Kuih Lapis, Mee Siput Muar, Putu Mayam, Putu Piring

3.1 Deep Learning Architectures

The ResNet50, InceptionV3, and VGG16 deep learning architectures are compared based on their performance on the Malaysian food dataset. These models are pre-trained using the ImageNet dataset. The most suitable architecture was selected, and transfer learning was applied to the pre-trained models. The pre-trained weights will be frozen or modified to cater to the new model to get better outcomes. Known for its nonlinearities and simple network, the ResNet (Residual Network) model builds pyramidal cells to skip connections or jump across layers [7]. The InceptionV3 network uses various kernel sizes in the same convolutional layer while the VGG and ResNet networks only use one kernel size in their convolution step. The VGG16 architecture records 92.7% test accuracy in the ImageNet dataset [8]. The convolution kernel size is 3×3 and 1 pixel in stride. 224×224 RGB (Red, Green, and Blue) images are set as the input to the neural network [7].

3.2 Image Augmentation and Activation Functions

To develop a neural network model that resembles the way humans identify food with just one look, a huge amount of data is required. To combat this issue, augmented images are introduced to the model during training at each epoch. Some of the techniques to be used to augment the food images are rotation, shift, flip, brightness, and zoom [9].

Activation functions are used to learn and approximate a continuous and complex relationship between variables, thus, adding non-linearity to the network [10]. In each hidden layer, a linear transformation is performed, and activation functions are applied. The output from the activation function is fed to the next layer and the process repeats. The activation functions used are ReLU (Rectified Linear Unit) and softmax. In ReLU, the neurons will be activated only when the result from the linear transformation is greater than zero. Contrarily, the softmax activation function is a more generalized form of the sigmoid function that is mainly used for multi-class classification [11].

4 Results and Discussion

4.1 Transfer Learning

Training a neural network from scratch is computationally expensive. Transfer learning makes use of the robust and discriminative filters learned by advanced networks to identify images that it was not trained on [12]. Pre-trained models are used to lower the errors and the time needed to develop a working and efficient model. Weights in the pre-trained layers are utilized to adapt to the new classification of Malaysian food. In general, transfer learning can be classified into two types: via fine-tuning and via feature extraction.

4.2 Transfer Learning via Fine-Tuning

Transfer learning through fine-tuning modifies the pre-trained model by removing the fully connected (FC) layers at the top. New FC layers connect to the body of the initial CNN architecture. The modified model is then fine-tuned to the Malaysian food dataset by changing the number of epochs, batch size, number of neurons, number of dense layers, dropout rates, optimizer algorithm, learning rate, and unfreezing some of the convolutional layers of the original network. The performance of each of the pre-trained models, ResNet50, InceptionV3, and VGG16 was recorded and analyzed.

The 3 best testing accuracy for each pre-trained model under different modifications were compared and shown in Figure 1. Both InceptionV3 and VGG16 models outperform ResNet50 for all three modifications. It was found that the performance of VGG16 is better than InceptionV3 in this fine-tuning technique.

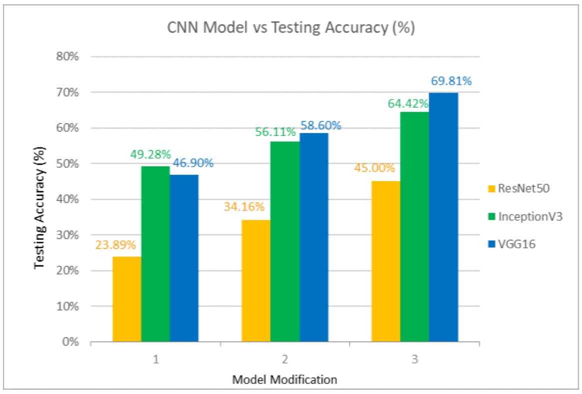

Figure 1: *Accuracy Performance of Transfer Learning via Fine-Tuning*

4.3 Transfer Learning via Feature Extraction

Transfer learning through feature extraction deals with the pre-trained network as an arbitrary feature extractor. The image sample goes through the network and stops at a pre-specified layer. The outputs from that layer are used as the image features [12]. The features from the specified layer are extracted as NumPy arrays and loaded as feature vectors with a specified classifier or neural network model to classify based on the extracted features.

Custom layers were initialized to perform image classification from the extracted features. Due to the poor performance of ResNet50 in the previous technique, only InceptionV3 and VGG16 models were studied using this technique. The 3 best testing accuracy for both pre-trained models during three different custom layers were compared and shown in Figure 2.

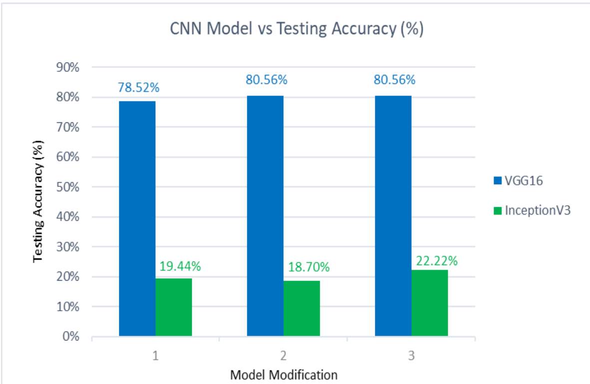

Figure 2: *Accuracy Performance of Transfer Learning via Feature Extraction*

4.4 Graphical User Interface (GUI) Application

It was discovered that the best method is using transfer learning via feature extraction on the pre-trained VGG16 network. The finalized model achieved a testing accuracy of 80.56% on data that it has not seen before. A Graphical User Interface (GUI) application to demonstrate the results of Malaysian food classification was developed by using PyQt in Qt Designer. The image uploaded to the GUI will pass through the VGG16 network for feature extraction. Subsequently, the feature vectors are passed into the

Proceedings DOI: 10.21467/proceedings.141
ISBN: 978-81-957605-4-1

Series: AIJR Proceedings
ISSN: 2582-3922

finalized model for image prediction. The model performs image classification and returns the predicted outcome.

The predicted and actual labels of the food class are displayed in Figure 3. For food items that contain peanuts, cow's milk, or shellfish in their ingredients, the information is shown in the interface. All 540 images in the test set are passed into the GUI application. A summary of the outcomes was recorded in Table 2, showing the number of correctly and incorrectly predicted images based on the food allergen rating, 'Safe', 'Potentially Unsafe', and 'Unsafe'.

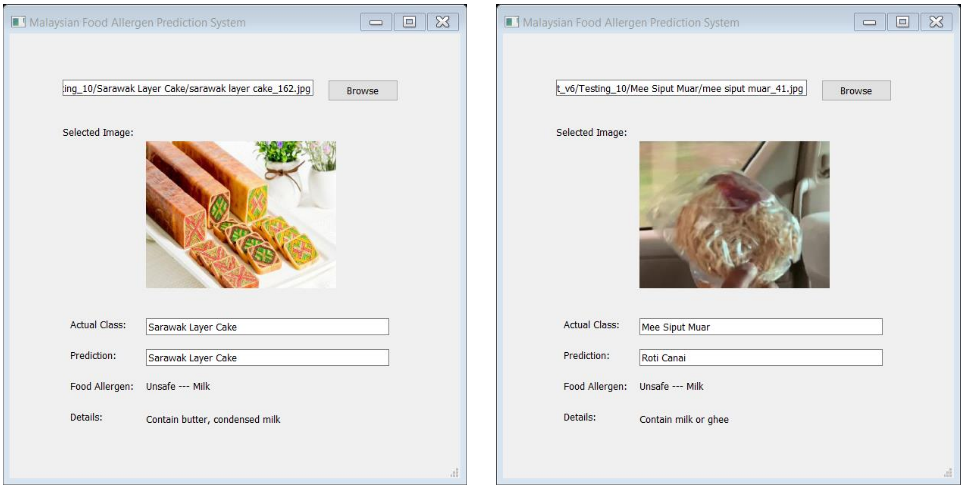

Figure 3: *Classification of Food Item Displayed in GUI*

Table 2: *Classification Accuracy for 3 Types of Malaysian Food Allergen Rating*

Allergen Rating	**Predicted Result**		**Total Images**	**Accuracy (%)**
	Correct	**Incorrect**		
Safe	119	23	142	83.80
Potentially Unsafe	67	12	79	84.81
Unsafe	249	70	319	78.06

5 Conclusions

This research project aims to implement a food allergen detection system in Malaysian food using convolutional neural networks (CNNs) for people with food allergies. The targeted food allergens are peanuts, cow's milk, and shellfish. A new Malaysian food image dataset was compiled. Stratified splitting was utilized to segregate the images of each food classes into a balanced distribution. The pre-trained models that were studied are ResNet50, InceptionV3, and VGG16 architectures. The food types were labelled 'Unsafe', 'Potentially Unsafe', or 'Safe' to indicate the presence of the targeted allergens. The performance for all three models through fine-tuning showed varied testing accuracies. Further enhancement was done by applying transfer learning via feature extraction. For this work, the most suitable neural network model for food allergen detection in Malaysian food is the VGG16 network. The proposed classifier recorded an accuracy of 80.56% in classifying 36 types of Malaysian food. The allergen information is displayed in the developed GUI application such that the user can avoid consuming it if he or she has a food allergy to this ingredient.

6 Declarations

6.1 Study Limitations

The dataset size and number of classes are possible limitations. A larger dataset and more classes could help in the model's generalization capability and reduce overfitting.

6.2 Acknowledgements

The authors would like to thank the lab personnel in the Department of Electrical Engineering, Universiti Malaya for their assistance in this work.

6.3 Competing Interests

The authors declare that the research was conducted without the presence of any commercial or financial relationships that could be construed as a potential conflict of interest.

6.4 Publisher's Note

AIJR remains neutral with regard to jurisdiction claims in published maps and institutional affiliations.

References

[1] A. Cianferoni and J. M. Spergel, "Food Allergy: Review, Classification and Diagnosis," *Allergology International,* vol. 58, no. 4, pp. 457-466, 2009, doi: https://doi.org/10.2332/allergolint.09-RAI-0138.

[2] N. Din, B. Rashid, K. Ramli, N. A. Mansor, and D. Abdullah, *Revelation of Children and Adult's Food Allergen in Malaysia*. 2019.

[3] J. Chen, B. Zhu, C. W. Ngo, T. S. Chua, and Y. G. Jiang, "A Study of Multi-Task and Region-Wise Deep Learning for Food Ingredient Recognition," *IEEE Transactions on Image Processing,* vol. 30, pp. 1514-1526, 2021, doi: 10.1109/TIP.2020.3045639.

[4] H. He, F. Kong, and J. Tan, "DietCam: Multiview Food Recognition Using a Multikernel SVM," *IEEE Journal of Biomedical and Health Informatics,* vol. 20, no. 3, pp. 848-855, 2016, doi: 10.1109/JBHI.2015.2419251.

[5] K. Gassner, B. Vollmer, M. Prehn, M. Fiedler, and S. Ssmoller, "Smart food: mobile guidance for food-allergic people," in *Seventh IEEE International Conference on E-Commerce Technology (CEC'05)*, July 19-22 2005, pp. 531-534, doi: 10.1109/ICECT.2005.85.

[6] X. Shi, V. De-Silva, Y. Aslan, E. Ekmekcioglu, and A. Kondoz, "Evaluating the Learning Procedure of CNNs through a Sequence of Prognostic Tests Utilising Information Theoretical Measures," *Entropy*, vol. 24, no. 1, p. 67, 30 December 2021, doi: 10.3390/e24010067.

[7] P. Verma, V. Tripathi, and B. Pant, "Comparison of different optimizers implemented on the deep learning architectures for COVID-19 classification," *Materials Today: Proceedings,* vol. 46, pp. 11098-11102, 2021/01/01/ 2021, doi: https://doi.org/10.1016/j.matpr.2021.02.244.

[8] P. Thakur, A. Chug, and A. P. Singh, "Plant Disease detection of Bell Pepper Plant Using Transfer Learning over different Models," in *2021 8th International Conference on Signal Processing and Integrated Networks (SPIN)*, 26-27 August 2021, pp. 384-389, doi: 10.1109/SPIN52536.2021.9565945.

[9] A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in *2018 International Interdisciplinary PhD Workshop (IIPhDW)*, 9-12 May 2018, pp. 117-122, doi: 10.1109/IIPHDW.2018.8388338.

[10] W. Hao, W. Yizhou, L. Yaqin, and S. Zhili, "The Role of Activation Function in CNN," in *2020 2nd International Conference on Information Technology and Computer Application (ITCA)*, 18-20 December 2020, pp. 429-432, doi: 10.1109/ITCA52113.2020.00096.

[11] M. Pavlova, "Comparison of Activation Functions in Convolution Neural Network," in *2020 28th National Conference with International Participation (TELECOM)*, 29-30 October 2020, pp. 65-67, doi: 10.1109/TELECOM50385.2020.9299559.

[12] L. Shao, F. Zhu, and X. Li, "Transfer Learning for Visual Categorization: A Survey," *IEEE Transactions on Neural Networks and Learning Systems,* vol. 26, no. 5, pp. 1019-1034, 2015, doi: 10.1109/TNNLS.2014.2330900.