

RESEARCH

Open Access



# In-TFK: a scalable traditional food knowledge platform, a new traditional food dataset, platform, and multiprocess inference service

Petrus Mursanto<sup>1\*</sup>, Ari Wibisono<sup>1</sup>, Puteri Khatya Fahira<sup>1</sup>, Zulia Putri Rahmadhani<sup>1</sup> and Hanif Arief Wisesa<sup>1</sup>

\*Correspondence:  
santo@cs.ui.ac.id

<sup>1</sup> Faculty of Computer Science,  
Universitas Indonesia, Kampus UI  
Depok 16424, Indonesia

## Abstract

Traditional Food Knowledge (TFK) is needed to define the acculturation of culture, society, and health in the context of food. TFK is essential for a human's cultural, economic, and health aspects. Variations of ethnicity, culture, and lifestyle affect the diversity of traditional Indonesian food. Recognition of food is needed to maintain the sustainability of traditional food. Nowadays, there are many food dataset collections, but there needs to be a dataset that specifically collects standard food datasets. Our main contributions to the TFK research field are professional food image data acquisition, innovative development of an automatic, scalable food recognition system, and multi-process inference service. There are 34 variations of traditional foods from all regions in Indonesia that were acquired in this dataset. The dataset comprises 1644 high-quality images captured by professional cameras and 1020 by a smartphone. Several deep learning models are implemented in food recognition systems. This system can accommodate the addition and reduction of food variations in the knowledge recognition system and is capable of multiple concurrent requests at a time. The current prototype incorporates traditional types of food from Indonesia. However, the food model can also be expanded to other countries traditional foods. The automatic recognition systems are evaluated using several deep-learning network models. The experiment results have shown that the AUROC score is 0.99, and the request success rate can be improved by 70% with a multiprocess inference service.

**Keywords:** Traditional food knowledge, Food dataset, Deep learning, Multiprocess inference service

## Introduction

Traditional food knowledge (TFK) is the interaction between human health, social culture, and the economy. TFK is commonly overlooked when talking about nutrition. However, it has a role in increasing food availability, enhancing the community's capacity building, and promoting biocultural diversity [1]. Some research has found that products from traditional food are healthy products. Baysal. A believes that ingredients from foods that go through natural processes for a long time (for example, braised meat, salted beef, grape, homemade macaroni, cracked wheat, yogurt, and pickles have

beneficial effects and nutrients for health. As an illustration, yogurt can fill the need for vitamins and calcium for bone development [2].

Although some traditional foods are unhealthy, nutritionists can try to understand the conventional food processing method. Therefore, nutritionists can provide sound nutritional advice to substitute harmful ingredients in traditional foods. Marie Douglas suggests that if we cannot understand the meaning of food for individuals, we cannot expect to be able to ask them to change their habits of consuming unhealthy food [3]. In addition to understanding the traditional way of processing food, nutritionists can benefit from conventional food by understanding the connection between culture, biodiversity, and food ecosystem in the daily life of a community. Vanhonacker researched 6 European states; the research has shown that the image of foreign consumers for local products is very positive. Traditional food knowledge will help document structured information on conventional food. This documentation can provide information about the types of food and the natural ingredients that traditional foods can offer. So that individuals have information about traditional foods that can be consumed in a healthy way.

Based on the Food and Agriculture Organization (FAO), Indonesia has 3,000 types of traditional foods that vary significantly between regions. It has the widest variety of traditional foods in Asia. Contrary to those facts, Indonesia has faced many problems related to food. Indonesia is classified as “generally food secure” on its main island of Java. However, some regions are categorized as “chronically food insecure,” such as the regions of Aceh and Nias and in the eastern part of Indonesia [4]. It is predicted that the number of poverty in Indonesia will reach 115 million people. Some regions in Indonesia that fall into this category include eastern Indonesia, East Java, Central Kalimantan, South Kalimantan, East Kalimantan, and West Kalimantan. West Sulawesi, Central Sulawesi, East Sulawesi, and Papua.

Some food shortages that occur in Indonesia happen because of inadequate access to food and road infrastructure for logistics in the eastern part of Indonesia (Papua and Maluku). In the last five years, the government has tried to make plans and make access to roads in Papua. Also, Female Literacy is quite significant in the regions of Nusa Tenggara Bara, East Nusa Tenggara, West Kalimantan, and South Sulawesi. Under-five mortality rate, 38 out of 1000 children under five years have died. This is due to the need for more water and nutrition availability. The infant mortality rate is also a concern; the ratio reaches 40 deaths among 1000 births—Indonesian stunting rates to 37% and 48% in remote areas. Underweight and micronutrient deficiency are also the leading causes as the Indonesian territory is categorized as a chronically food insecure region. In general, the territory of Indonesia can be sufficient for consumption and food production. However, sustainable food production is tough to be achieved. Over-exploitation of land and water resources, logging and making the land for oil palm plantations, and deforestation make the soil infertile, and nutrients in the soil are lost.

Based on these problems, we needed valid information about local products and traditional food in every region in Indonesia. A database of local products or traditional foods that can be accessed easily will benefit the local community. It will help build a healthier community and protect biodiversity, food ecosystems, and cultures well. The database can also serve as a geographical instrument to support the production of traditional food and local products used in some countries, which is required for logistics [5].

An ensemble approach between deep learning and reinforcement for short-term forecasting has been introduced by Liu [6].

Food security is the biggest challenge for developing countries. Identification of the food security situation is a must to be able to solve the problem. A combination of data from various aspects is needed so that the identification of food security problems can be reviewed and resolved. H Deléglise developed a food security prediction system based on multiple data sources, such as world bank data and GPS points (hospitals, schools, land use, and waters). Various meteorological data are also used to enrich data variations [7]. Explainable AI is a topic that is needed to provide relevant information about the predictive results of a system [8].

In this study, we contribute to creating a scalable platform for gathering traditional food knowledge. Some of the research methods that have been done are collecting Indonesian traditional food data from the largest islands in the Indonesian archipelago by using professional studio-lab standards. Food data is taken in a standard and professional studio lab to produce a common dataset, which is publicly open. The results of the food recognition model visualization can show the food's ethnicity, culture, and origin in a map visualization. Users can also use this platform to collect traditional food from various regions. For now, traditional Indonesian food is included as a prototype. Other users can form conventional foods from other countries. They can also add their training dataset to the existing models. So, the deep learning models of data are updated and enrich the knowledge model to recognize traditional food.

### **Related works**

Generally, most references related to food classification are intended for health purposes by monitoring food intake. In addition, there are cases where food classification is utilized for tourism. To our knowledge, only one reference classifies food from the Ingredients dataset to correlate ethnicities and countries. Regarding objectives, it can be concluded that food classification aims for health, tourist information, and ethical recognition [9]. However, until now, there has been no particular food dataset for traditional food and preserving cultural biodiversity. Therefore, this research proposes automatic, scalable food recognition and knowledge systems for traditional food knowledge. The current prototype can accommodate conventional types of food from Indonesia, but it is also possible that other traditional food can be further incorporated into the food database.

In terms of the dataset, to date, some of the latest references collect data by searching on search engines and taking data by using a smartphone or camera. The image resolution quality provided by the dataset also varies. In addition to the quality of the dataset, food image data licenses also vary; some of them are proprietary [10–12] and public [12–16]. To the best of our knowledge, in terms of the food dataset content, no reference specifically analyzes traditional food. All references focus on how much data can be retrieved and the divergence of data. The data acquisition process also does not consider the history and type of food, whether traditional or fast. In this research, we took the dataset independently with the professional and studio-lab standards. The foods are taken from five big islands in Indonesia, which consist of 38 types of food. We have obtained several states of the art ideas regarding food recognition based on the

existing literature study. In general, food recognition can automatically be classified into the automatic classification of food objects, classification of food objects with ingredients, analysis of food components in each country, and text processing for ingredients processing.

Measurement of food intake is essential to know the number of calories and nutrients consumed by the body. F. Zhu et al. implemented an application that recorded the amount of food and nutrition in one day. The approach analyzed food images and automatically measured food intake in the body. The author implements several segmentation approaches of eating sections, features for identification, and methods for estimating food portions. Nutritional information released by the system yields a 10% margin of correctness [17]. F. Zhu et al. improved the process for identifying and locating food in multiple forms. Identification and locating of food are carried out during the group eating event. Some concepts are combined; several parts of the segmented object are continuously divided into the same class object. This is done by utilizing local and global features. The resulting increase in accuracy is 30%. The author also mentions there are challenges to the illumination problem from the difference in the light of each data. We solve this problem by standardizing light when collecting food data [18].

Pouladzadeh et al. summarize several methods that other researchers have carried out to detect food and the number of calories in it using Vision-Based Measurement (VBM). Several algorithms used to recognize food are also mentioned in the paper, such as Convolutional Neural Network (CNN) and Handcrafted method (Feature Extraction + Machine Learning). The review also states several architecture and solutions for food identification [19]. In terms of the Handcrafted process, some authors use a variety of feature extraction methods to identify mobile devices, including SIFT Feature Extraction, Physical measurement, and Graph-based segmentation [20–23]. Other authors also identify using mobile and server infrastructure. Some of the techniques used are texture analysis, SIFT, Bag-of-Features (BoF), Color Histogram, Gabor Features, K-means Clustering, and Color-Texture Segmentation [24–27].

The framework for differentiating food classifications based on geolocalized settings has also been carried out by R. Xu et al. The strategy is geolocalized voting and a combination of several classifiers combined. Xu et al. collected data by visiting each restaurant by taking food pictures, dish tags, restaurant data, menus, and geolocation. From the results of experiments, using geolocation in research can increase accuracy by up to 30% [28].

Innovative ideas for using deep learning in food recognition are automatically proposed by P. Pandey et al. The author presented the concept of Foodnet; the foodie author uses a multi-layer CNN pipeline that is used to be able to use features from other deep networks to improve efficiency. Some traditional handcrafted methods and a combination of several CNN variations are done to see the best results. From the benchmark results using the ETH-Food-101 dataset, this model has an accuracy of 73.5 (Top-1), (94.4) (Top-5), and 97.6 (Top-10) [29]. The classification of food in the partially-labeled food dataset is done by B. Mandal et al. The author uses the Modification of a Generative Adversarial Network (GAN) to solve food recognition problems in the dataset. The experimental evaluation results using data with incomplete labels have better results than the state-of-the-art method. The experimental results of this model have an

accuracy of 75.34 (Top-1), (93.31) (Top-5), and 96.43 (Top-10) using the ETH-Food-101 dataset [30]. Hossain et al. classify fruits used in industrial applications. Two deep learning network methods evaluate the 6 CNN network layers and pre-trained Visual Geometry Group-16 (VGG-16). The results of the classification scores obtained were 99.49% [31].

In addition to image data, several authors also carried out the introduction of food ingredients in the form of text. Pugsee et al. proposed a method for conducting automated comment analysis. This is done on user comments to improve the quality of a recipe. The result of the accuracy obtained in this contribution is 70% accuracy [32]. Marta et al. used a copula-based clustering algorithm to analyze the EU country diets using data from the Food and Agriculture Organization (FAO). The data contains historical dietary records and food calorie data in Europe. The results indicated that central and eastern Europe have unhealthy nutritional habits [33]. Kim et al. provides insight or information about food ingredients and recipes in several countries. Each country has its ethnic group and different elements. Location factors will affect the classification of authentic ingredients in a country. Components can show some characterization of cultural food correlations in each area in the recipe [9]. Zhang reviewed some methods to evaluate the quality of food and argo-product b using spectroscopy and deep learning [34].

Ciocca et al. collected a food dataset from canteen plates arranged in several types. The dataset has 73 different food varieties and 1027 canteen tray images. The amount of food in all canteen trays is 361. Automatic food recognition is also implemented in the diet monitoring application. The accuracy results obtained are about 79% [35]. G. A., F. Alfarisy, et al. customized web crawlers to produce better relevance for food data retrieval. The web crawler proposed by the author focuses on crawling Indonesian food using uncomplicated classification techniques. The approach is based on the priority level of a link using URLs and text. The proposed method achieved a relevance of 81.75% from the experiments conducted, while the standard way achieved only 16% [36]. Text processing is also implemented and represented by the life cycle of a food object. The use of logic and graph-based knowledge are used to illustrate the transformation of cereal foods. This research aims to get the best query answers that prioritize the quality of the results of the query text [37].

Aside from automatic recognition and classification of food, several research also proposed automatic recognition of ingredients contained in the food image, which falls into the Multimodal category. W. Min et al. solve the correlation problem between the recipe and food image using the Multimodal multi-task deep belief network (M3TDBN) method. State the art presented by the author is the processing of food identification and retrieval of ingredients in a recipe. The author combines visual context and text retrieval to obtain a correlation between the two. The best results obtained were 0.866 based on Top K% for evaluation [38]. Zhu et al. also detected ingredients in the food dataset. The application is called Dietcam. Food ingredients are detected by combining texture verification and deformable part-based models. Categories are classified using multi-kernel SVM. The resulting classification accuracy is 60% [18]. Jingjing Chen et al. use deep architecture, which is simultaneously used for the recognition of recipes and food classification by looking at the relationship between recipes and food. The results

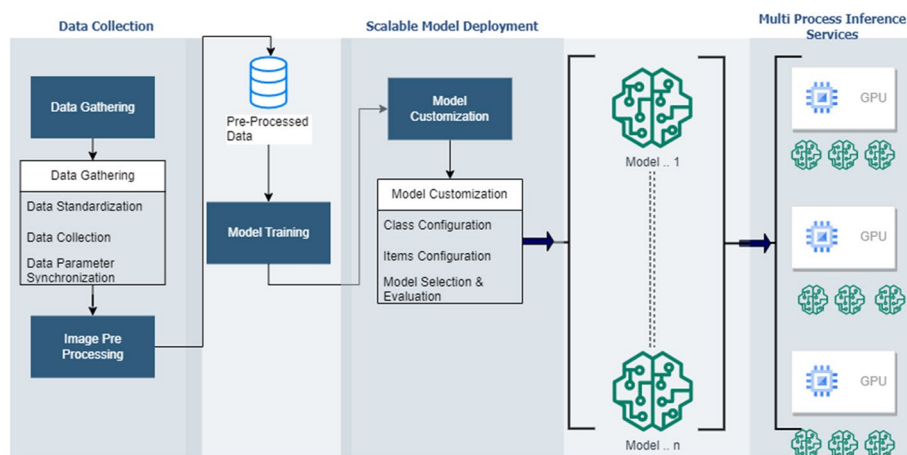
of the training features and semantic labels are applied to the zero-shot retrieval of the recipe. This paper demonstrates the possibility of introducing ingredients in the zero-shot problem for cooking recipe retrieval. The results obtained for top-10 hit rates are several 0.570 [39]. In addition to teaching food images and ingredients, J. Marin et al. contributed to collecting a dataset called Recipe1M. The author compiled a dataset of 800,000 images of food and a corpus of more than 1,000,000 for food recipes [40]. The dataset is called Recipe1M. Using this dataset; the authors automatically merge images with food recipes. Regularization, with the addition of high-level classification objectives, can increase performance retrieval compared to human capabilities [41].

In Indonesia, several research studies have been conducted on automatic food classification. The main purpose of these studies is mainly for tourism. Setyono et al. classifies 12 classes of traditional Indonesian foods from Jakarta. The dataset is collected by using a search on the search engine manually. The best classification was obtained using DenseNet169, which achieved 80.6% accuracy [42]. Food classification and ingredient labeling for tourist information was also carried out by Prasetya et al. The author used CNN to automatically display food labels and ingredients by CNN. The network used is inception; the total dataset is 400 data for training and 60 for testing data obtained through search engines. Classes from the traditional Javanese food dataset are six classes. The best accuracy results obtained are 82.5% [43]. Stanley et al. conducted a three-class classification of Indonesian food. The algorithm used in this research is the K-D tree and BPNN. The results obtained are 51% accuracy using the K-D tree algorithm [44].

From some references of authors contributing to food in Indonesia, the dataset is obtained from an image search on search engines or images captured by a smartphone camera. The amount of food variation is also focused on each region in Indonesia. In this paper, we propose data taken in a professional studio-lab standard with several variations representing five major islands in Indonesia, starting from the western part of Aceh to the east part of Papua. All focus conditions, aperture, ISO, and lumens are lighting standardized for each captured image. Our main contribution in this paper is the traditional food image data retrieval in a professional and innovative system of the automatic, scalable food recognition system. This system can accommodate the addition of food variations directly from the user. The current prototype can accommodate traditional types of food from Indonesia, but it can also be expanded to other countries.

## Methods

This paper proposes a scalable system that can adaptively recognize images of traditional Indonesian food. In conventional prediction systems, the user uploads a picture of the food, and the system will predict the type of food. In our proposed method, the user can also provide feedback to the model and develop the traditional food custom model. In addition, conventional prediction systems only provide requests to the backend and then receive the prediction results of the classification. Conventional prediction systems are weak; the GPU process will be overloaded when there are concurrent requests. We tried to solve this problem by creating multiple processes on each GPU. A detailed description of our proposed method is described in Fig. 1.



**Fig. 1** Proposed framework

**Table 1** Model Evaluation Result

Metrics/Models	DenseNet121	ResNet50	InceptionV3	NasnetMobile
Mean AUROC	0.999	0.993	0.997	0.999
Accuracy	0.992	0.920	0.900	0.969
Precision	0.993	0.952	0.926	0.965
Recall	0.993	0.937	0.906	0.968
F1-Score	0.993	0.929	0.898	0.962
Training time	3930	3789	3736	4360

**Data collection**

We also collect 34 types of traditional Indonesian food by formalizing the stages of data collection. This process standardizes the food images regarding lighting, angles of taking images regarding lighting, angles of taking parameters, and camera configurations. After the data collection, we perform data cleansing of the image. After the preprocessing, we tested several machine-learning models to classify the traditional food data. Some of the machine learning models that we have tested are Densenet121, Inception Restnet, Resnet50, and NasnetMobile [45].

**Scalable model deployment**

The modeling stages are carried out using several existing models, including DenseNet121, ResNet50, InceptionV3, and NasnetMobile. The measurement parameters are AUROC (Area Under Receiver Operating Characteristic), Precision, Recall, F1-Score, and training time. Training-test validation uses three-fold cross-validation to obtain a valid accuracy value. Based on the evaluation results in Table 1. The experiment was carried out using three folds. The results of fold-1–2–3 are averaged to obtain generalization results from the metrics. We then choose the best model that will be the best fit for the food recognition system. In a scalable food training system, the user can customize the model training and add new types of food. This module allows the user to enrich the model by adding the type of food.

### **Multiprocess inference services**

We proposed a multiprocess food inference module. This module will predict the type of food and the confidence score. Multiprocess food inference has a different mechanism compared to conventional systems. In a conventional prediction system, the front end will directly invoke the backend process of the model loaded in GPU RAM. However, this paper proposes a Multiprocess food inference with a load balancer on the backend server. The load balancer has a job divider and balancer for each process.

Finally, we combined all the components of the system. We also improved the prediction backend by creating multiple processes on multiple GPUs. This mechanism can accommodate the addition of concurrent users that make predictions at a time. We can accommodate more users than conventional prediction methods by using multiple inference processes on each GPU. In Fig. 1, the last stage depicts the process of each GPU. It will have multiple processes to perform inferencing from the endpoint (load balancer). The multiprocess inference process is carried out by one process and uses 2 or 3 processes for each GPU. We can maximize the process by up to nine processes. The multiprocess inference can increase the success rate of the prediction system by up to 70%.

### **Results and Discussions**

In this system, we propose a scalability model that can evolve with data input from the user. The user will have a model that can be updated according to the dataset the user owns. For example, at present, the model can predict 34 types of food. Should the user want to increase the type of food to be predicted by the system, the user can update the existing model so that the current food model can predict 35 types. Therefore, this system could not only load models on basic traditional food models but can also improve and grow the model along with the user's efforts to enrich the model. The infrastructure we use in this proposal is Intel(R) Core(TM) i7-6900 K CPU @ 3.20 GHz, 64 GB RAM, 2 TB HDD, and 3 × GeForce GTX 1080 (8118 MB). Apart from the scalability of the training model, we also added enhancements to the inference process. The inference process that we do is to initialize multiple processes on the GPU. So that the inference process also has a better scalability performance against the number of concurrent users compared to conventional prediction systems.

### **Dataset acquisition**

The dataset used as the training model is a traditional Indonesian food dataset. We got 34 types of Indonesian food from our previous research [45]. The dataset is taken formally by normalizing the camera settings and lighting conditions. Each food has several images of about 30–50 images of food from multiple angles inside a mini studio. The data capture resolution is 4000 × 2000 pixels. The professional camera uses an  $f/2.8$  24–70 mm lens, and the distance between the object and the camera is 70 cm. Our smartphone camera specifications are 12 MP  $f/1.5$ –2.5 26 mm, 12 MP  $f/2.4$  52 mm, and 16 MP  $f/2.2$  12 mm with an Exynos 9820 (8 nm)—EMEA/LATAM processor. The brightness of the lamp is set to 5,000 lumens. Based on this data collection



technique, we have 1,000 high-quality food images that will make the model easier to get a high level of accuracy.

### Model development

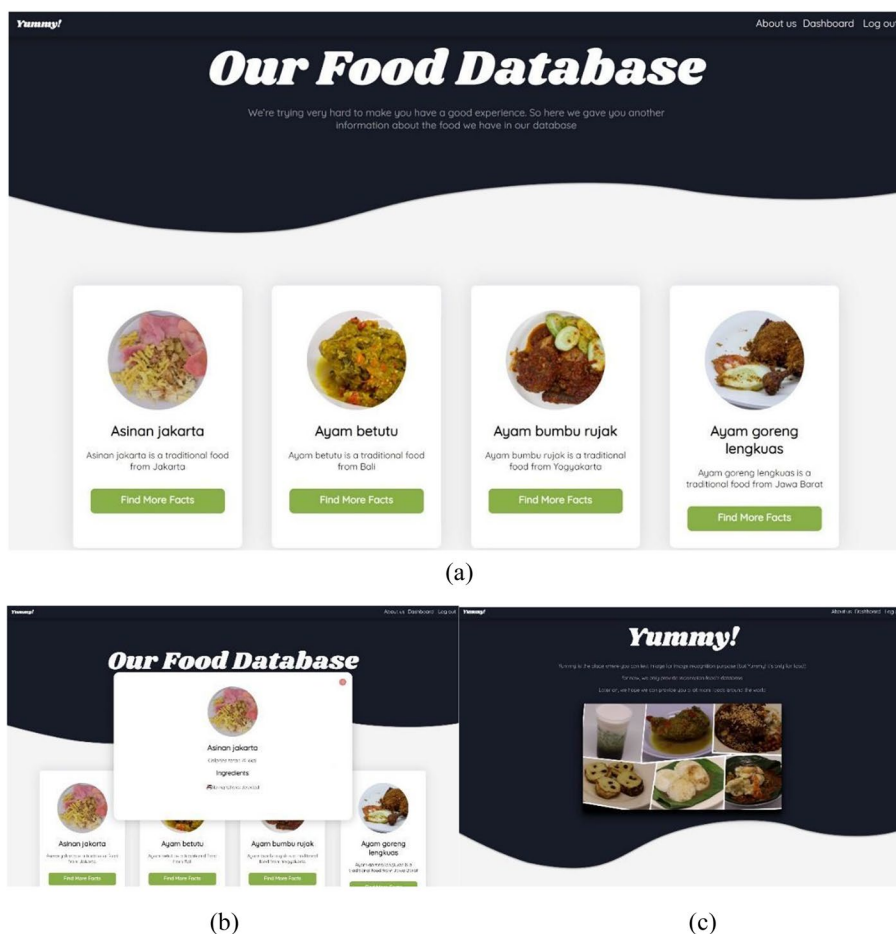
The modeling stages are carried out using several existing models, including DenseNet121, ResNet50, InceptionV3, and NasnetMobile. The measurement parameters are AUROC (Area Under Receiver Operating Characteristic), Precision, Recall, F1-Score, and training time. Training-test validation uses three-fold cross-validation to obtain a valid accuracy value. Based on the evaluation results in Table 1. The experiment was carried out using three folds. The results of fold-1–2–3 are averaged to obtain generalization results from the metrics. In Table 1, we can see that the best metric measurement value is obtained by the DenseNet121 network with AUROC (1.0), Accuracy (0.993), Precision (0.994), Recall (0.994), and F1-Score (0.994). The other network models also did not produce bad results. All metrics returned values greater than 0.89 for the ResNet50, NasNetMobile, and InceptionV3 models.

In Table 1 we can see the training speed for each model. Each model was tested with the same test parameters with 100 epochs. Nasnetmobile had the highest training time of 4361 s, Densenet 3963 s, ResNet50 3790 s, and InceptionV3 3793 s. DenseNet121 and NasnetMobile have a much longer training time of 5 and 10 min than ResNet50 and InceptionV3. However, both models (DenseNet121 and NasNet-Mobile) have AUROC values and accuracy parameters that are much better than ResNet50 and InceptionV3. Based on experiments on accuracy and training time (Table 1), we decided to use DenseNet121 as a model, considering its high AUROC (1.0), precision (0.994), recall (0.994), and F1-Score (0.993) performance values, and high training time and average training time (3.931 s).

### Scalable food training

This proposed system consists of several modules: the user management module and the food database module (Fig. 2a–c.), the food testing (inference) module (Fig. 4a–c), and the food model training module (Fig. 3a–e). This prototype system has a separate module between the front and back ends. In a scalable food training system, the user can customize the model training and add new types of food. This module allows the user to enrich the model by adding the type of food. For example, suppose the user wants to add a particular food from the Yogyakarta area, such as Javanese noodles. In that case, the user only needs to prepare several Javanese noodle image data for training on this system.

The training process is like the training process for making a new model. The average time for training with 100 epochs is 3800 s. To monitor the progress of the training carried out by our deep learning program, we apply it in an asynchronous process with the Redis implementation [46]. So that the user can leave the training page and the background process continues to run on the backend task. Users can use a scalable training model to enrich their knowledge of traditional food recognition systems.

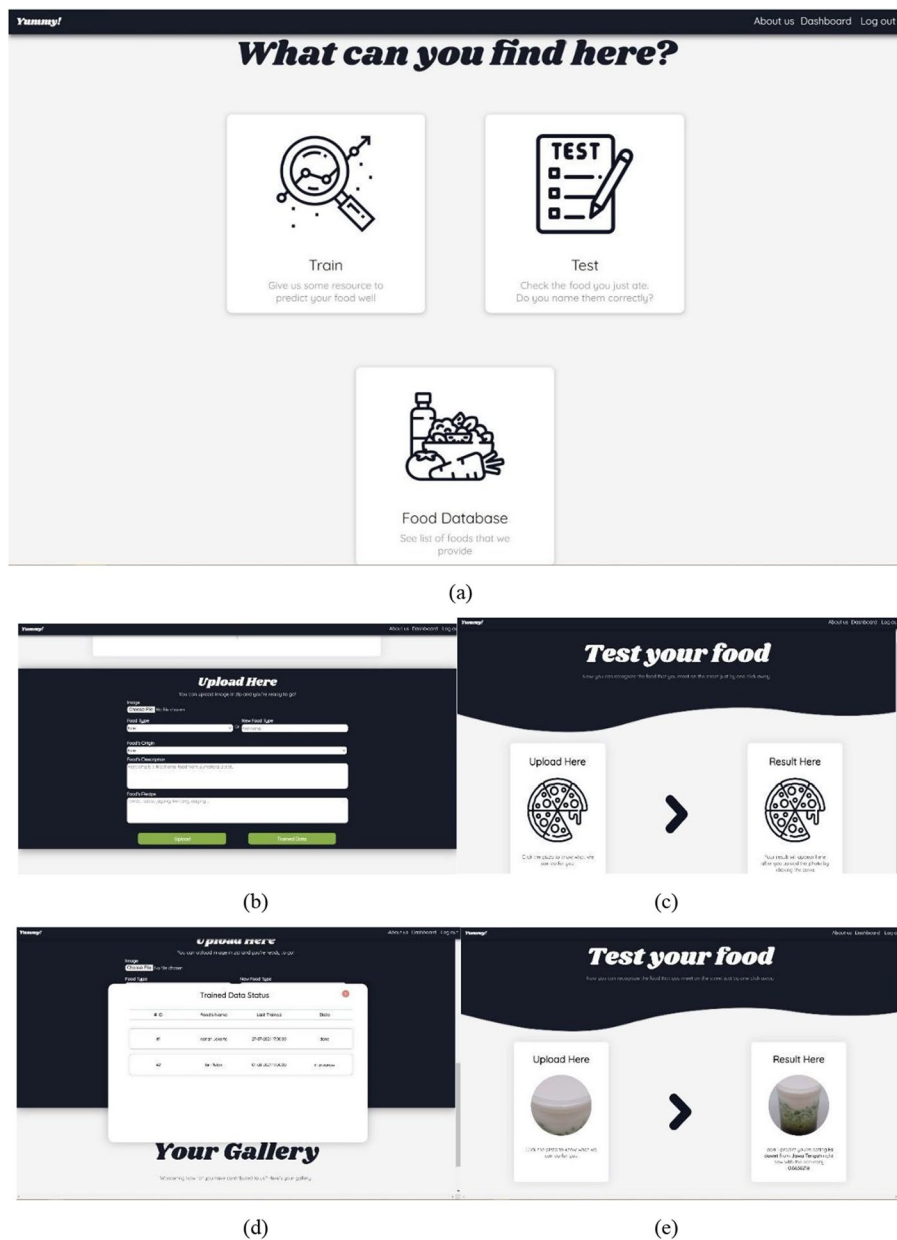


**Fig. 2** Conceptual interfaces **a** food database, **b** food details, **c** homepage

**Multiprocess inference services**

We proposed a multiprocess food inference module. This module will predict the type of food and the confidence score. The average time required for a user to perform inferring using the GPU is 0.3 s. Multiprocess food inference has a different mechanism compared to conventional systems. In a conventional prediction system, the front end will directly invoke the backend process of the model loaded in GPU RAM. However, this paper proposes a Multiprocess food inference with a load balancer on the backend server. The load balancer has a job divider and balancer for each process. The detailed process of the module is described in Fig. 1.

The average time to predict the image using the GPU is 0.3 s. If using the CPU, the average time required is 10 s. In this paper, we try to execute the backend process in parallel to maximize the performance of the GPU. The GPU we use is GTX 1080, which has 8118 MB of RAM. We have analyzed the GPU memory requirements to process one image; the image prediction system needs 2096–3096 MB of GPU global memory. Therefore, we proposed parallelizing the process in 1 GPU in this paper. We try to execute three processes in one GPU where the global memory allocation for each GPU is limited to 2400 MB. Using this method, we have 3 × more predictive abilities than using only one process. We tested multiprocess inference with multiple



**Fig. 3** Conceptual interface **a** features, **b** upload food image for training, **c** food inference (testing), **d** food inference result

scenarios. The first scenario uses 1 GPU ranging from one to three processes in one GPU. The maximum value of the process is three processes. The second scenario uses two GPUs wherein each GPU is first tested with one to three processes and continues until six processes. In the third scenario, we use three GPUs wherein one to three processes are tested for each GPU. We can then have nine processes running concurrently.

To evaluate the multi-inference scenario, we use the response correctness parameter. Where response correctness will be measured along with the increase in the number of concurrent users. In addition to response correctness, we also test the time it takes to

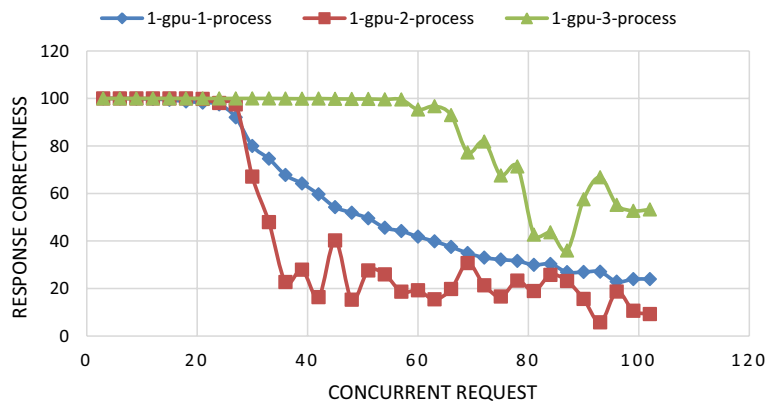


Fig. 4 Success rate % based on the concurrent request for 1 GPU

serve an increase in concurrent users. Increasing concurrent users tested from 1 to 120 users.

In all scenarios, we use image data with the same image size specification:  $4000 \times 2000$  pixels with a maximum size of 1 Mb. We tried to do a simulation with concurrent users ranging from 1 to 110 users for each scenario. We also use various number processes on each GPU. The multiprocessing inference testing metric that we use is the execution time

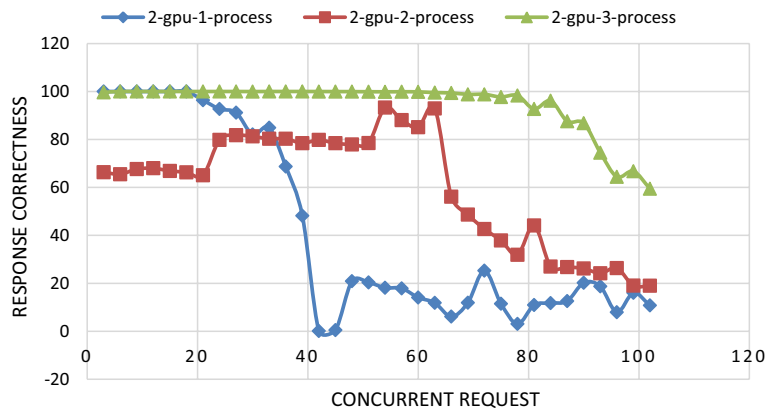


Fig. 5 Success rate % based on the concurrent request for 2 GPUs

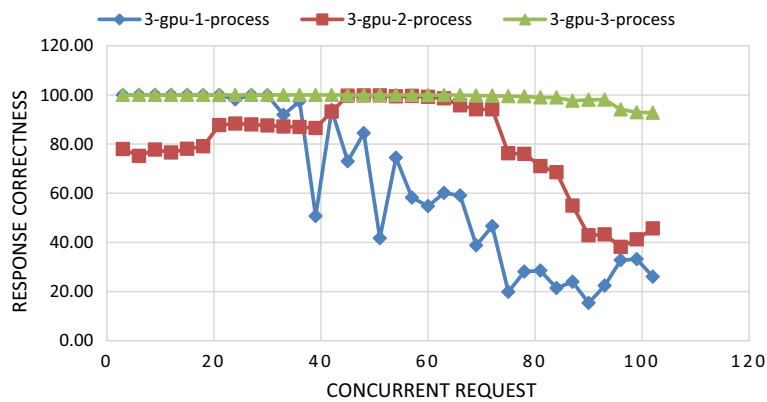


Fig. 6 Success rate % based on the concurrent request for 3 GPUs

and success rate of each scenario. The results of the success rate evaluation can be seen in Fig. 4. 1 GPU, Fig. 5. 2 GPUs, and Fig. 6. 3 GPUs.

In Fig. 4 it can be seen that with the use of 1 GPU 1 process (blue marker), there is a decrease in the success rate. After 30 concurrent requests, the success rate can only achieve 80%. The condition worsens when concurrent requests reach 60; 1-gpu-1-process and 1-gpu-2-process conditions drop the success rate to 40%. The use of 1-gpu-3-process in 1 GPU is quite promising, where a success rate can still achieve 90% with 60 concurrent requests. A higher success rate can be achieved by optimizing the number of processes on the GPU and managing memory usage.

In Fig. 5 we did a trial using 2 GPUs with scenarios 1 to 3 processes. It can be seen that the 2-gpu-1-process scenario has quite poor performance, almost the same as using 1-gpu-1-process and 1-gpu-2-process. The resulting success rate is only 20%, with 60 concurrent requests. However, we can see an improvement in the use of the 2-gpu-2-process; in this scenario, we can get a success rate of around 80% on 60 concurrent requests. The best result scenario for two GPUs is 2-gpu-3-process. At 80 concurrent requests, the success rate obtained is still above 90%.

The next scenario is a scenario using 3 GPUs. Generally, using 3 GPUs performs better than scenarios using 1 or 2 GPUs. Except for the 3-gpu-1-process scenario, the success rate is already at 60% on a concurrent request of 60. It is based on Fig. 6. We can see that the 3-gpu-3-process still has a success rate above 90% when there are 100 concurrent requests, and the 3-gpu-2-process can only have a success rate of 40% when there are 100 concurrent requests.

Based on several test scenarios. The use of multiprocessing inference can increase the success rate of the prediction in the recognition system. Users do not have to add more GPU hardware to achieve a higher success rate, but users can create multiprocessing in 1 GPU to maximize the utilization of the GPU. Based on the tests, we made 60 concurrent requests on 1 GPU, 1-gpu-1 process (20%), 1-gpu-2-process (50%), and 1-gpu-3-process (90%). From these results, we can conclude that an increase in the success rate of 30% can be obtained when adding one process, and adding a success rate of 70% can be obtained by adding two processes. Adding a GPU logically can certainly increase the success rate, but this increase can be maximized by performing multiple inference processes. Figure 7

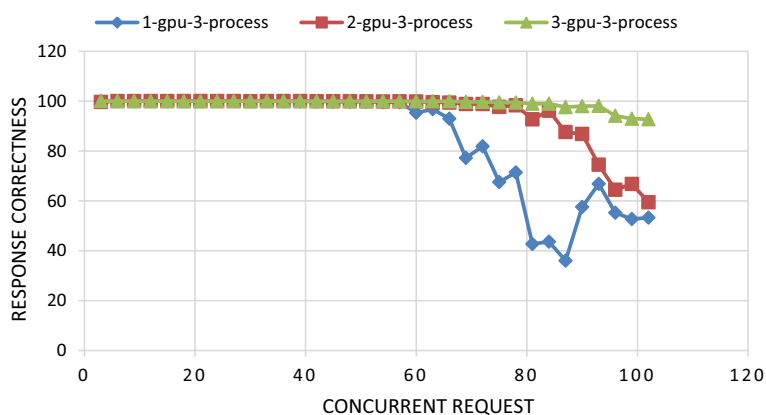


Fig. 7 Response correctness based on concurrent requests for 1, 2, 3 GPU with three processes

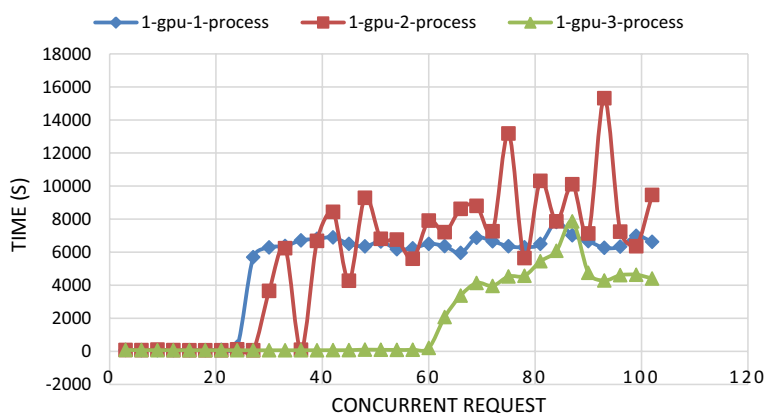


Fig. 8 Execution time (seconds) based on the concurrent request for 1 GPU

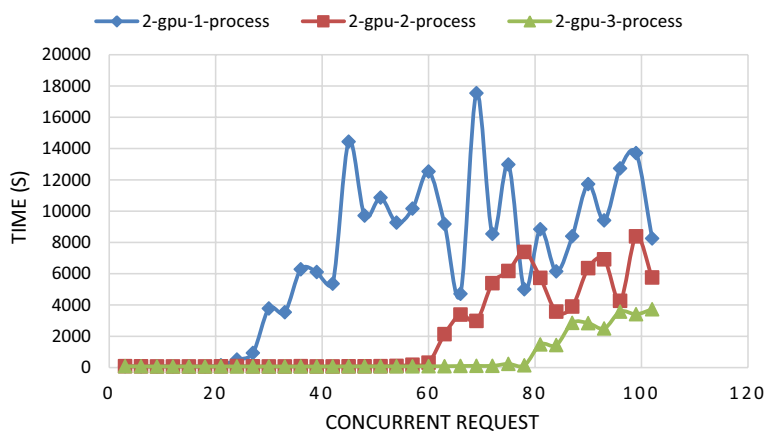
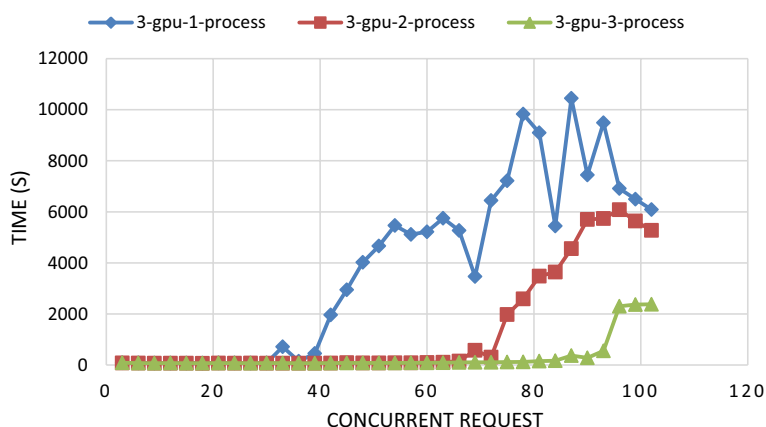


Fig. 9 Execution time (seconds) based on the concurrent request for 2 GPUs

is the ratio of the number of GPUs. On 90 concurrent requests, we can see that 1-gpu-3-process produces 50%, 2-gpu-3-process produces 75%, and 3-gpu-3-process produces 98%. We can see in Fig. 7 that if we use the 3-gpu-1-process on a concurrent request of 90, the success rate generated is 38%. It can be seen that by using the multiple inference process method, the success rate can increase up to 100% success rate.

We have also analyzed the execution time (second) of the test scenario. Based on Fig. 8., by increasing the number of processes at two GPUs, the time required to predict concurrent requests is 110, 2-gpu-1-process produces 14,000, 2-gpu-2-process takes 8000 s, and 2-gpu-3-process takes 4000 s. In general, multiprocessing inference increases the success rate of the prediction system and reduces the time execution of 1 GPU, 2 GPU, and 3 GPU. Based on Figs. 8, 9, and 10, the more processes in each GPU, the lower the execution time we will get. Except for scenario 1-gpu-2-processes in Fig. 8, we can see that using two processes in one GPU does not significantly reduce execution time. It has almost the same execution time compared to the 1-gpu-1-process. The success rate of scenario 1-gpu-2-processes cannot significantly improve the success rate. This condition happens because there are only a few processes in one GPU. Thus, those processes must accommodate a manageable concurrent request that invokes the GPU. Overall, the



**Fig. 10** Execution time (seconds) based on the concurrent request for 3 GPUs

multiprocess inference yields a higher success rate until 70% improvement, and the concept of a scalable model is proposed in this paper.

The limitation of this research is hardware management; GPU hardware management for inference has been done for one computer node with three GPUs in it. Using the hardware with GPU VRAM and more stream processors can increase the concurrency of users who make predictions using this system. The future works from this research are building a scalable training system scheduling from a multimode GPU and scheduling models with various parameters.

## Conclusions

In this paper, we propose a prediction system mechanism and scalable machine learning system to make predictions using traditional food data from Indonesia. The proposed system is scalable, where users can update machine learning models and train their data. Based on our evaluation, the best model performance is DenseNet121, with an AUROC value (of 0.99). In addition to the proposed scalable model, we also propose a multiprocess inference service. This method is used for accommodating the prediction system. Multiprocess inference services can accommodate concurrent users when making predictions. Based on our tests using 1–110 simultaneous users, multiple process inference services can increase the prediction success rate performance to 70%. With the combination of the two proposed mechanisms, namely a scalable knowledge model and multiple process inference services, those mechanisms can become a concept in preserving information on Indonesian culture. In the future, we will try to enrich the traditional food dataset from Indonesia and several other countries. We will also apply the calculation of the nutritional ingredients for each traditional food.

## Acknowledgements

We would like to express our gratitude for the grant received from Universitas Indonesia.

## Author contributions

PM: conceptualization, supervision, funding acquisition, writing—original draft. AW: conceptualization, software-hardware development, writing—original draft. Formal-analysis. PKF: software, data curation, investigation: software, data curation, investigation. HAW: formal analysis, writing original draft, writing—review and editing. All authors read and approved the final manuscript.

## Funding

Universitas Indonesia (2021).

**Availability of data and materials**

<https://www.dropbox.com/sh/88stvg9kraf17z8/AAC97ldMbHe7ksp8Cc8l08K2a?dl=0>, [https://www.dropbox.com/sh/a59su6gmzocbebw4/AAAB0r8hz3Z7nV703Df\\_P7Emda?dl=0](https://www.dropbox.com/sh/a59su6gmzocbebw4/AAAB0r8hz3Z7nV703Df_P7Emda?dl=0)<https://www.dropbox.com/s/wupw88x33v0cz03/stats-queue.xlsx?dl=0>.

**Declarations****Ethics approval and consent to participate**

Not Applicable.

**Consent for publication**

Not Applicable.

**Competing interests**

The authors declare no competing interests.

Received: 18 February 2022 Accepted: 3 April 2023

Published online: 14 April 2023

**References**

- Hancock T. The mandala of health: a model of the human ecosystem. *J Family Commun Health*. 1985;8(3):1–10.
- Baysal A. Geleneksel Gıdaların, Üzerine Etkileri.İl:Geleneksel Gıdalar Sempozyumu 27–29 Mayıs 2009. pp.5–6.
- Douglas M. *Implicit meanings: selected essays in anthropology*. New York: Routledge; 1999.
- UN. World food program, executive brief: Indonesia food security assessment and classification, 2007.
- Birinci Y. Yöresel Ürünler için Yeni Açılımlar: Corafi aretler, GEME'den Bakı,Sayı:36, Ankara,s. 2008. pp. 85–86
- Liu H, et al. A new hybrid ensemble deep reinforcement learning model for wind speed short-term forecasting. *Energy*. 2020;202:117794.
- Deléglise H, et al. Food security prediction from heterogeneous data combining machine and deep learning methods. *Expert Syst Appl*. 2022;190:116189.
- Van der Velden BHM, et al. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Med Image Anal*. 2022;79:102470.
- Kim K, Chung C. Tell me what you eat, and i will tell you where you come from a data science approach for global recipe data on the web. *IEEE Access*. 2016;4:8199–211.
- Joutou T, Yanai K. A food image recognition system with multiple kernel learning. In: Proc. 16th IEEE Int. Conf. Image Process. 2009. p. 285–288.
- Mariappan A, et al. Personal dietary assessment using mobile devices. In: IS&T/SPIE Electron. Image., International Society for Optics and Photonics. 2009;7246:72460Z–72460Z.
- Hoashi H, Joutou T, Yanai K. Image recognition of 85 food categories by feature fusion. In: Proc. IEEE Int. Symp. Multimedia. 2010. p. 296–301.
- Ciocca G, Napoletano P, Schettini R. Food recognition and leftover estimation for daily diet monitoring. In: Proc. New Trends Image Anal. Process. Workshops, 2015;9281:334–341.
- Chen M-Y, et al. Automatic Chinese food identification and quantity estimation. In: Proc. SIGGRAPH Asia Tech. Briefs, 2012. p. 29.
- Bossard L, Guillaumin M, Van Gool L. Food-101—mining discriminative components with random forests. In: Proc. Comput. Vis. 2014. p. 446–461.
- Chen M, Dhirga K, Wu W, Yang L, Sukthankar R, Yang J. Pfd: Pittsburgh fast-food image dataset. In: Proc. 16th IEEE Int. Conf. Image Process. 2009. p. 289–292.
- Zhu F, et al. The use of mobile devices in aiding dietary assessment and evaluation. *IEEE J Sel Top Signal Process*. 2010;4(4):756–66.
- Zhu F, Bosch M, Khanna N, Boushey CJ, Delp EJ. Multiple hypotheses image segmentation and classification with application to dietary assessment. *IEEE J Biomed Health Inform*. 2015;19(1):377–88.
- Pouladzadeh P, Shirmohammadi S, Yassine A. You are what you eat: so measure what you eat! *IEEE Instrum Meas Mag*. 2016;19(1):9–15.
- Yang J, Wu W. Fast food recognition from videos of eating for calorie estimation. In: Proc. IEEE Intl. Conf. on Multimedia and Expo. 2009. p. 1210–1213.
- Rebro SM, Patterson R, Kristal A, Cheney C. The effect of keeping food records on eating patterns. *J Amer Dietetic Assoc*. 1998;98:1163–5.
- Takeda F. Dish extraction method with neural network for food intake measuring system on medical use. In: Computational Intelligence for Meas. Syst. and Applications. 2003. p. 56–59
- Wang Y, He Y, Zhu F, Boushey C, Delp E. The use of temporal information in food image analysis. *New Trends Image Anal Process ICIAP 2015 Workshops*. 2015;9281:317–25.
- Pouladzadeh P, Shirmohammadi S, Almaghrabi R. Measuring calorie and nutrition from food image. *IEEE Trans Instrum Meas*. 2014;63(8):1947–56.
- Joutou T, Yanai K. A food image recognition system with multiple kernel learning. In: Proc. 16th IEEE Int. Conf. on Image Processing (ICIP). 2009. p. 285–288.
- Hoashi H, Joutou T, Yanai K. Image recognition of 85 food categories by feature fusion. In: Proc. 2010 IEEE Int. Symp. on Multimedia (ISM). 2010. p. 296–301.



27. Pouladzadeh P, Shirmohammadi S, Yassine A. Using graph cut segmentation for food calorie measurement. In: Proc. IEEE Int. Symp. on Medical Meas. and Applications. 2014. p. 1–6.
28. Xu R, Herranz L, Jiang S, Wang S, Song X, Jain R. Geolocalized modeling for dish recognition. *IEEE Trans Multimed.* 2015;17(8):1187–99.
29. Pandey P, Deepthi A, Mandal B, Puan NB. FoodNet: recognizing foods using ensemble of deep networks. *IEEE Signal Process Lett.* 2017;24(12):1758–62.
30. Mandal B, Puan NB, Verma A. Deep convolutional generative adversarial network-based food recognition using partially labeled data. *IEEE Sens Lett.* 2019;3(2):1–4.
31. Hossain MS, Al-Hammadi M, Muhammad G. Automatic fruit classification using deep learning for industrial applications. *IEEE Trans Industr Inf.* 2019;15(2):1027–34.
32. Pugsee P, Niyomvanich M. Suggestion analysis for food recipe improvement. In: 2015 2nd International Conference on Advanced Informatics: Concepts, Theory, and Applications (ICAICTA), Chonburi, 2015, p. 1–5.
33. Di Lascio FML, Disegna M. A copula-based clustering algorithm to analyze EU country diets. *Knowl-Based Syst.* 2017;132:72–84.
34. Zhang X, et al. Food and agro-product quality evaluation based on spectroscopy and deep learning: a review. *Trends Food Sci Technol.* 2021;112:431–41.
35. Ciocca G, Napoletano P, Schettini R. Food recognition: a new dataset, experiments, and results. *IEEE J Biomed Health Inform.* 2017;21(3):588–98.
36. Alfariy GAF, Bachtiar FA. Focused web crawler for Indonesian recipes. In: 2017 International Conference on Sustainable Information Engineering and Technology (SIET), Malang, 2017, p. 196–202.
37. Thomopoulos R, Bourguet J-R, Cuq B, Ndiaye A. Answering queries that may have resulted in the future: a case study in food science. *Knowl-Based Syst.* 2010;23(5):491–5.
38. Min W, Jiang S, Sang J, Wang H, Liu X, Herranz L. Being a supercook: joint food attributes and multimodal content modeling for recipe retrieval and exploration. *IEEE Trans Multimed.* 2017;19(5):1100–13.
39. Chen J, Ngo C-W. Deep-based ingredient recognition for cooking recipe retrieval. In: Proceedings of the 24th ACM international conference on Multimedia, p. 32–41.
40. Marin J, et al. "Recipe1M+": a dataset for learning cross-modal embeddings for cooking recipes and food images. *IEEE Trans Pattern Anal Mach Intell.* 2019. <https://doi.org/10.1109/TPAMI.2019.2927476>.
41. Salvador A, et al. Learning cross-modal embeddings for cooking recipes and food images. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, p. 3068–3076.
42. Setyono NFP, Cahyati D, Fanany MI. Betawi traditional food image detection using ResNet and DenseNet. In: 2018 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Yogyakarta. 2018. p. 441–445.
43. Prasetya RP, Bachtiar FA. Indonesian food items labeling for tourism information using Convolution Neural Network. In: 2017 International Conference on Sustainable Information Engineering and Technology (SIET), Malang. 2017. p. 327–331.
44. Giovany S, Putra A, Hariawan AS, Wulandhari LA. Machine learning and SIFT approach for Indonesian food image recognition. *Proced Computer Sci.* 2017;116:612–20.
45. Wibisono A, Wisesa HA, Rahmadhani ZP, Fahira PK, Mursanto P, Jatmiko W. Traditional food knowledge of Indonesia: a new high-quality food dataset and automatic recognition system. *Journal of Big Data.* 2020;7(1):1–19.
46. Redis | The Real-time data platform. <https://redis.com/>. Accessed 8 June 2021.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---