

Topics

Image Recognition Specific Object Recognition Food Product Judgment

Special Articles on AI Supporting a Prosperous and Diverse Society

A Food Product Judgment System Supporting Food Diversity —Enabling People Who Have Food and Drink Prohibitions to Select Foods Simply with an App—

Service Innovation Department Seiya Kojima^{†1} Fatina Putri^{†2}

Certain food or drinks or ways of consuming foods are restricted by some cultures or religions. This is known as having food and drink prohibitions. There are many examples of foods whose consumption is prohibited by certain religions, for instance under the Islamic categorization of foods as Halal or Haram^{*1}, there are many things that Muslims are not permitted to consume such as pork, pork-derived products or alcohol. In addition to religious or cultural reasons, there are also many vegetarians all over the world whose diets entail partial or full avoidance of animal-based foods for health reasons and so forth.

As more and more visitors from overseas are expected to visit Japan with the 2020 Olympics as a trigger, there will be many visitors with such food and drink prohibitions. This means it will be necessary to handle an unprecedented diversity of foods, for example handling greater numbers of food products labeled with Halal certification.

Conventionally, when Muslims or vegetarians who

cannot read Japanese buy some food at a Japanese convenience store or a supermarket, they have to pick up each item and check the ingredients written in Japanese using a translation app, or check if the product is okay to eat by taking a photo of it and sending it to a friend for confirmation using a social networking service, etc. before purchasing the item. This inconvenience has even resulted in cases of travelers to Japan bringing foods from their own country to consume during their stay.

To address this issue, NTT DOCOMO has developed a “food product judgment system” for purchasing foods in convenience stores and supermarkets that enables people who have food and drink prohibitions to determine whether they can consume a product just by photographing shelved merchandise using their smartphone, etc. before they make purchase [1].

This system consists of two functions.

- The first is a food product recognition function that uses DOCOMO’s “shelved merchandise

©2019 NTT DOCOMO, INC.

Copies of articles may be reproduced only for personal, noncommercial use, provided that the name NTT DOCOMO Technical Journal, the name(s) of the author(s), the title and date of the article appear in the copies.

†1 Currently, Solution Service Department

†2 Served at the Service Innovation Department until the end of June 2019.

image recognition engine” [2]. This image recognition engine enables identification of various items on display from an image captured of the shelved merchandise.

- The second is a function to judge food products for people who have food and drink prohibitions. By combining data on the ingredients in a product with information about food and drink prohibitions (food product judgment logic), the system determines whether the product can be consumed by people who have food and drink prohibitions such as Muslims or vegetarians.

With these two functions, the system enables people who have food and drink prohibitions to determine whether they can consume a product just by photographing the shelved merchandise using

their smartphone, etc. (**Figure 1**). Thus, the system reduces the bother of purchasing foods because it eliminates the need for the users to hold a product and translate what is written on its packaging to decrypt its ingredients.

This article describes the details of the food product judgment system we developed.

1) Food Product Recognition Function Enabled by Image Recognition Technology

The food product judgment system is achieved with two image recognition technologies.

(1) Object detection technology using deep learning^{*2}

The first technology is object detection technology that uses deep learning to analyze the position information of products from an image of shelved merchandise (**Figure 2**). An orange frame is displayed for the results of object detection, and the upper left and

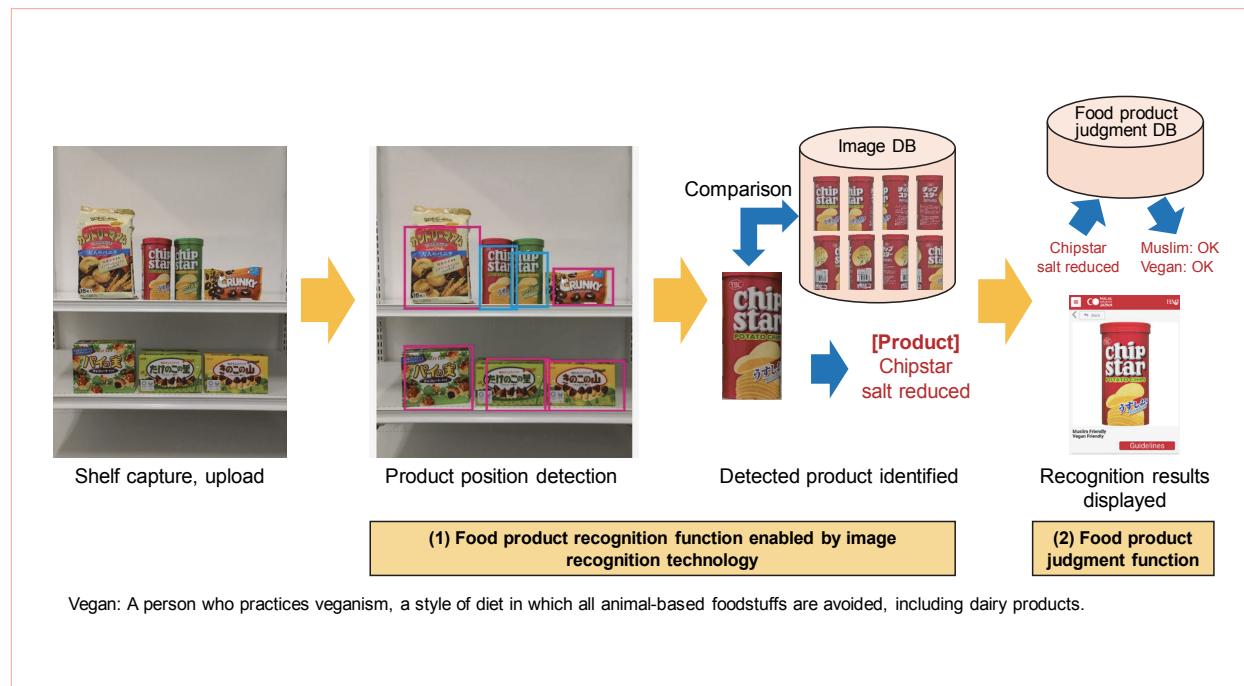


Figure 1 Food product judgment system recognition flow

*1 Halal, Haram: Items that are allowed under Islam law are referred to as Halal, while items that are not allowed are referred to as Haram. Usually applied to determine whether dishes or ingredients can be consumed by Muslims.

*2 Deep learning: A method of machine learning (see *3) using a multilayered neural network.

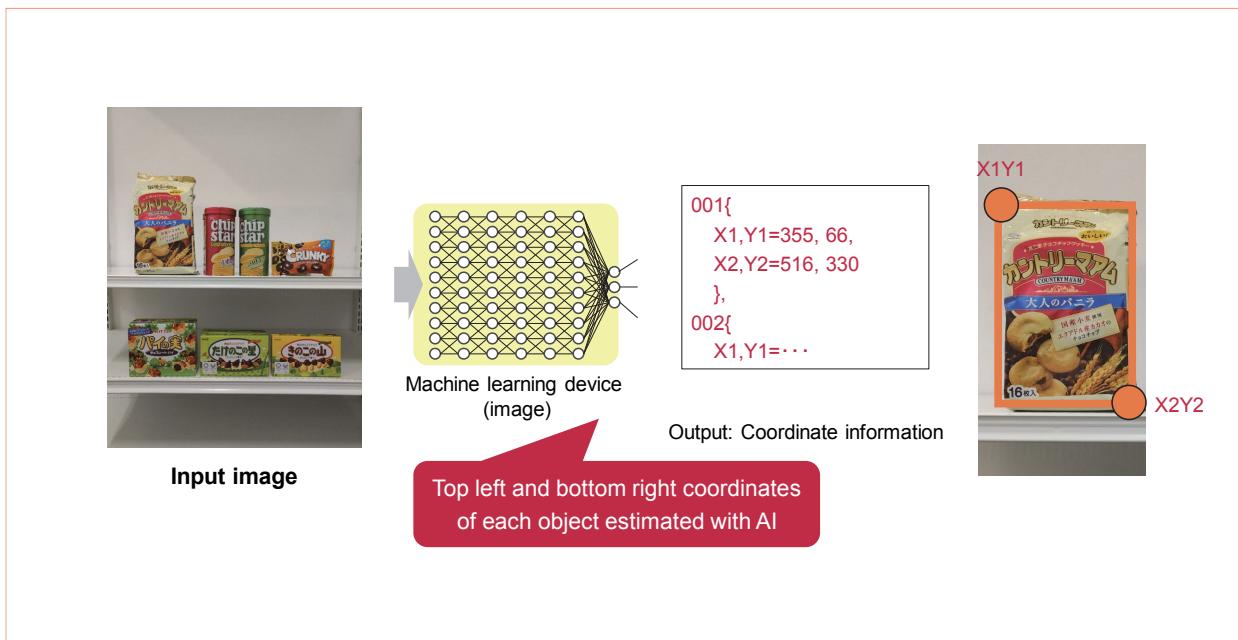


Figure 2 Object detection technology

bottom right coordinates of the frame surrounding the product are estimated. Because DOCOMO's object detection technology uses deep learning, the object detection engine must undergo machine learning^{*3} in advance to detect the desired objects. Hundreds and thousands of images of product displays in various actual stores and annotation data^{*4} have been prepared for this deep learning. Teaching this data to the system enables it to detect products with a high degree of accuracy even when products are crammed into tiny display spaces. Please refer to reference [3] for details of the object detection technology algorithm.

(2) Specific object recognition technology using local feature values^{*5}

Second is specific object recognition technology that uses local feature values to identify

products from partial images in the merchandise area detected as described above (Figure 3). As the target image, the merchandise area is input and compared to large amounts of product image data preregistered in an image database. This identifies products in the merchandise area by determining whether the input image is similar to any preregistered images. Images of products captured from various angles are preregistered in the database to make it possible to compute the similarity of the input image to the preregistered images. However, this could be impractical because the level of similarity is computed by comparing with all of the large number of images in the database, which could take several tens of seconds or more for one image of shelved merchandise. We addressed this issue with our specific object

*3 Machine learning: Technology that enables computers to acquire knowledge, decision criteria or behaviors, etc. from data in ways similar to how humans acquire these things from perception and experience.

*4 Annotation data: In this article, refers to metadata indicating what is in an image.

*5 Local feature values: Extracted from data, values (numbers) that characterize the data. In this article, "feature values" refers specifically to image feature values, which are characteristic points (corners) extracted from the image and the surrounding distribution of brightness.

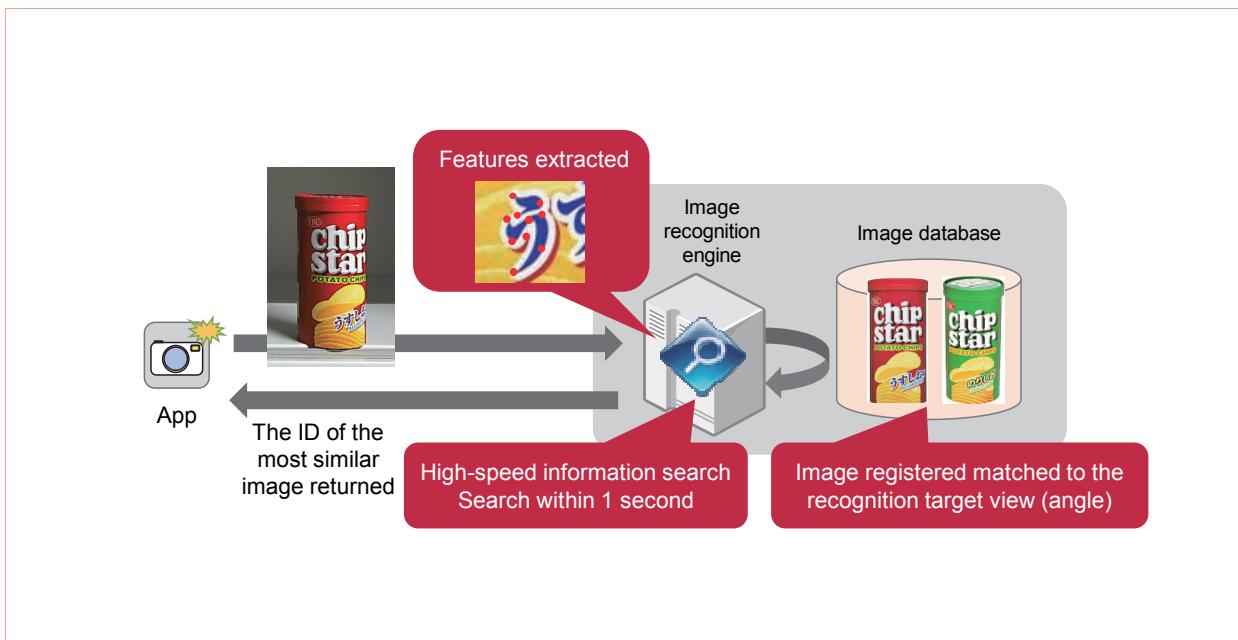


Figure 3 Specific object detection technology

recognition technology. With this technology, high-speed, high accuracy recognition of not only front view images but also images from different angles enabled by an algorithm we developed within one second from the several million preregistered images captured from various angles in the large database. Please refer to reference [4] for details of the specific object detection algorithm.

Using these two technologies, the position and the identity of the product are recognized.

2) Food Product Judgment Function

Food product judgment is enabled for products in the photograph by referencing recognition results from the image recognition engine with the database. Judgment information based on information about product ingredients is preregistered in the database to determine whether products are Muslim

or vegetarian-friendly. The food product judgment logic has been achieved through collaboration with FOOD DIVERSITY Inc., a company in Japan taking initiatives regarding food and drink prohibitions particularly for Muslims and vegetarians. Products are judged to be Muslim or vegetarian-friendly based on information about the primary ingredients^{*6} of products.

3) Application Provision

A trial offering is underway of a food product judgment service incorporating the food product judgment system with the aforementioned two functions in the “Halal Gourmet Japan^{*7}” restaurant search app provided by FOOD DIVERSITY Inc. for Muslims and vegetarians (Figure 4) [5]. This app displays food products that are Muslim or vegetarian-friendly in the image captured of shelved merchandise in different colors – Muslim-friendly products are displayed in a red frame, while Muslim and

^{*6} Primary ingredients: The ingredients that directly comprise a final product. With foods products in particular, these are the ingredients listed on the product label.

^{*7} Halal Gourmet Japan: A smartphone application designed for Muslims that enables search of restaurant information, etc., and is operated by FOOD DIVERSITY Inc.

vegetarian-friendly products are displayed in a blue frame. Products that are not Muslim and vegetarian-friendly or are unregistered are displayed with a white frame. Users can tap the product in the colored frame to display details about it. “Muslim-friendly” is displayed for Muslim-friendly products.

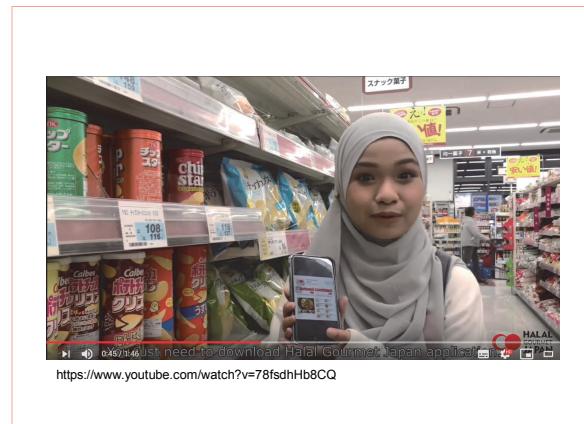


Figure 4 Food product judgment service trial offering

In addition to judgments about whether products are Muslim or vegetarian-friendly, the details screen also clarifies ingredient names and informs whether the products are Muslim or vegetarian-friendly under the different judgments of different people*8 in a way that is easy to understand (Figure 5).

This article has described the two functions of the food product judgment system we developed, and introduced a trial offering of a food product judgment service using those functions.

The NTT DOCOMO image recognition technology used with this system is capable of identifying food products on display just by capturing an image of the shelved merchandise, and thus in addition to Muslims and vegetarians, it could be used to provide services to people with various other food issues just by mapping information to various

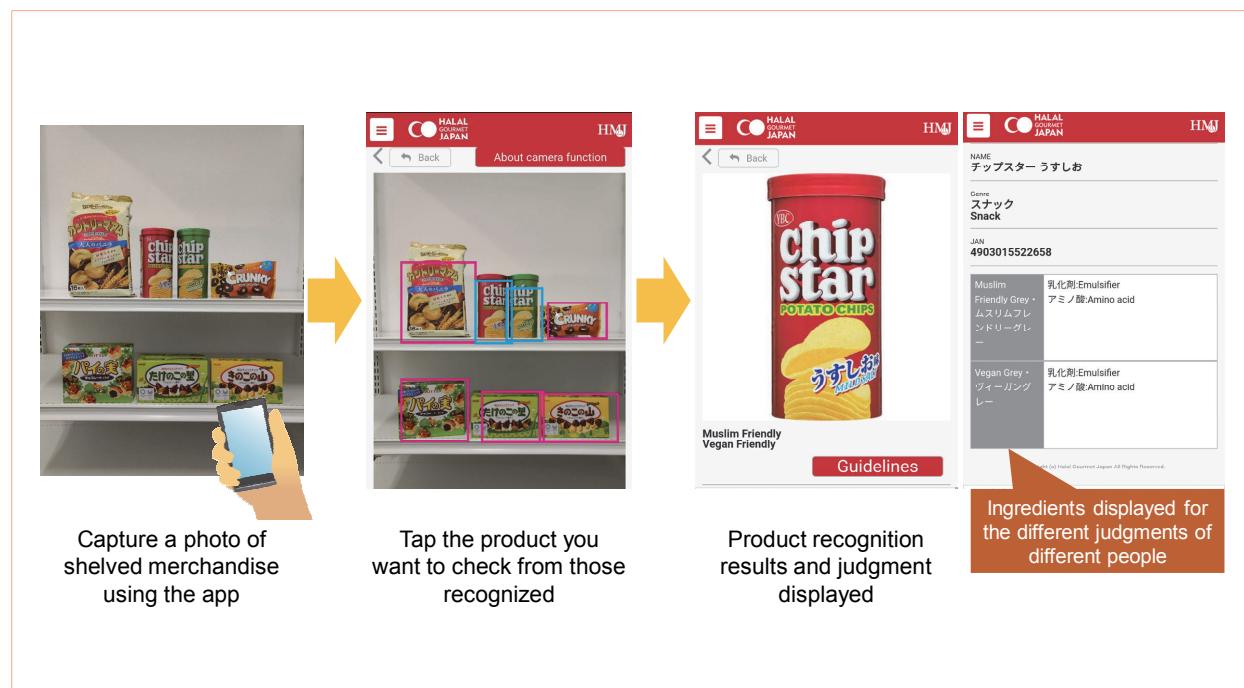


Figure 5 “Halal Gourmet Japan” food product judgment service usage image

*8 Different judgments of different people: In this system, this refers to such things as fresh cream. Because Muslim judgments are based on a promise between oneself and Allah, ultimately decisions about whether a food product can be consumed are up to the individual, and standards are not uniform.

products. For example, information about allergens^{*9} could be added so that people with various allergies could determine whether food products are consumable, or information about low-protein or low sugar foods or even information for picky eaters, etc. could be added to enable judgments. We plan to make judgment of these possible in the future.

Since only a limited number and types of products are currently registered, we will expand the numbers and types of food products that can be identified by this system by partnering with food providers and retailers.

This will enable the system to respond smoothly to various dietary restrictions. We greatly expect that this system will help to get the increasing numbers of visitors to Japan to recognize that when it comes to food, Japan is a safe place where people can feel confident about the things they eat.

REFERENCES

- [1] NTT DOCOMO Press Release: "(Announcement)

NTT DOCOMO Develops a "Food Product Judgment System" that Lets Muslims and Vegetarians Determine with an App what Food Products They Can Eat – Just by taking a picture of shelved merchandise with a smartphone –," Sep. 2018 (In Japanese).

https://www.nttdocomo.co.jp/info/news_release/2018/09/26_01.html

- [2] NTT DOCOMO Press Release: "DOCOMO Launches AI Engine for Fast, Accurate Shelf Analysis – Recognizes shelf allocation by analyzing photos of shelved merchandise –," Mar. 2018.

https://www.nttdocomo.co.jp/english/info/media_center/pr/2018/0316_00.html

- [3] H. Akatsuka et al: "A Retail Shelving Analysis Solution Using Image Recognition – Recognizes Shelving Allocation and Quantifies Inventory by Analyzing Photos of Shelved Merchandise," NTT DOCOMO Technical Journal, Vol.20, No.2, pp.22–31, Nov. 2018.

- [4] H. Akatsuka et al: "High-speed, Large-scale Image Recognition and API," NTT DOCOMO Technical Journal, Vol.17, No.1, pp.10–17, Jul. 2015.

- [5] HALAL MEDIA JAPAN: "Just Take a Picture" and You Will Know Which Products are "Muslim-Friendly"."

<https://www.youtube.com/watch?v=78fsdhHb8CQ>

^{*9} Allergen: A substance which can cause an allergic reaction. Among foodstuffs, allergens are defined as specific ingredients in food items such as shrimp, crab or wheat that can cause a high incidence or severity of symptoms of allergic reactions.