Ingredients estimation and recommendation of Thai-foods
Lawankorn Soimart¹,*, Pakpoom Mookdarsanit²
¹Department of Business Computer, Faculty of Management Science, Chandrakasem Rajabhat University, Bangkok, 10900, Thailand
²Department of Computer Science, Faculty of Science, Chandrakasem Rajabhat University, Bangkok, 10900, Thailand
*Corresponding Author: lawankorn.s@gmail.com

Abstract

Thai-foods are one of the most delicious tastes in the world. Many visitors around the world are fascinated in Thai-foods. In 2011, Thailand had the foods more than other countries from the list of “World’s 50 Most Delicious Foods” such as Tom-yam-goong (4th ranking), Pad-thai (5th ranking), Som-tam (6th ranking), Mas-sa-man (10th ranking), Green-curry (19th ranking), Thai-fried-rice (24th ranking) and Moo-nam-tok (36th ranking), respectively. However, some visitors cannot recognize these foods’ name. What’s more, the ingredients of food are difficult to be considered by the visitors’ eyes. To that end, this paper firstly introduces a novel ingredients estimation of Thai-foods as a Thai-food guidance of visitors that opens a new gate for Thai-foods with image processing and artificial intelligence. For the measurement, our system provided the high correctness of estimation in the criterions of accuracy, recall and precision which were summarized in term of their average values as 0.74, 0.64 and 0.78 respectively.

Keywords: Food recognition; Deep learning; Ingredients estimation of foods

©2017 Sakon Nakhon Rajabhat University reserved

1. Introduction

Thai-foods are one of the most delicious tastes in the world. Such in 2011, the 7 Thai-foods appeared on the list of the "World’s 50 Most Delicious Foods (Readers' Pick)"— a worldwide voting of 35,000 people by CNN [1]. Thai-foods appeared more than other nations from the list. From Fig. 1, there were Tom-yam-goong (4th ranking), Pad-thai (5th ranking), Som-tam (6th ranking), Mas-sa-man (10th ranking), Green-curry (19th ranking), Thai-fried-rice (24th ranking) and Moo-nam-tok (36th ranking), respectively.
Thai-government has designed the policy for supporting Thailand to be one of the world’s top cuisine and tourism. The total of 9,881,091 visitors around the world came to Thailand in 2015 [2] which made the country’s income about 2.30 trillion baht of Thai-GDP [3]. However, many visitors cannot recognize these foods’ names with their ingredients. Some visitors do not eat some ingredients such as beef, pork or glutamate, etc. Some of them may be in diet. Hence, they do not buy the foods because of unknown ingredients which make the country lose off another income. Since Image processing and Artificial intelligence in computer science are available in other researches such as gender/age estimation of an image [4, 5], face detection [6, 7], place recognition [8 – 11], remote sensing [12 – 16], image forensics [17, 18] and biomedical images [19], etc. Likewise, it is possible to estimate the ingredients of the foods using only an image. This paper firstly introduces a novel method to estimate the ingredients of Thai-foods by image matching with the list of Thai-food images collected in the large-scale computer-brain [20]. For the computer-brain creation, all Thai-food images with their textual recommendations such as names, ingredients and other properties were crawled from the accessible data source [21, 22] over the internet [23, 24]. Then, all features of each image were extracted in term of a vector. After that, all similar images were grouped together. Finally, all images with their recommendations were modeled as the large-scale computer-brain by Deep learning (DP) [25 – 27]. For the ingredients estimation, a visitor can input an unknown Thai-food image to the system. The system will estimate the ingredients from an input image which considers from the most similar images from the computer-brain.

The organization of this paper consists of section 2 “Preliminary” as basic knowledge explanation. The system architecture is crucially described in section 3 and the measurement in section 4, respectively. The future work and our research direction are mentioned in section 5.
2. Preliminary

Feature Extraction

Feature extraction is a mathematical representation of the important features within the image [28] in term of vector. Features can be texture, color and shape [29] which are unique from other images [30]. Basically, important features are dependent on the high contrast position such as edges or corners on the image which are robust for illumination, scale, rotation, and shadowing. All features of an image are mathematically represented in term of a vector. Later, these vectors are input to the process for grouping of similar images.

Deep Learning

Deep learning is the latest calculator-brain in artificial intelligence; was officially proposed by Deng, et al. in 2014 [25]. Deep learning is a well-known approach in AlphaGo that was implemented by Google [26]. AlphaGo also was validated by providing the Go-matches between AlphaGo and some famous Go-players. Especially, AlphaGo [27] won Fun Hui in 2015 and Lee Sedol in 2016, respectively. Traditionally, the infrastructure of deep learning is based on the higher number of hidden nodes in Multi-layer perceptron which is one of a supervised learning. The supervised learning consists of a training set and test set [31]. Training set is a model which is constructed from analysis of huge independent and dependent variables [32]. Test set is a set of independent variables which input to the model for determining the results [33] (or named dependent variables).

3. Methodology of ingredients estimation and recommendation of Thai-foods

Our ingredients estimation of Thai-foods opens a new gate about the relation between Thai-food patterns and image processing with artificial intelligence which is a new field for researching as shown in Fig. 2. In this section, we crucially explain our system: RGB-to-grayscale conversion, Feature extraction and Deep learning: training, and testing, respectively.
Input a food image

3.1 RGB-to-grayscale conversion

3.2 Feature extraction

3.3 Deep learning: Training part

3.4 Deep learning: Testing part

Train or test?

Computer Brain

3.1 RGB-to-grayscale conversion

3.2 Feature extraction

3.3 Deep learning: Training part

3.4 Deep learning: Testing part

Fig. 2 System architecture of “ingredients estimation and recommendation of Thai-foods”

RGB-to-grayscale Conversion

Either an unknown Thai-food image or Thai-food image with its information is input to the system as shown in Fig. 3. Firstly, the image must be normalized the (color) dimensions. Gray-scale has the value between 0 – 255.

Fig. 3 An unknown Thai-food image (in case of Tom-yam-goong)
where Gray is the gray-scale value of a pixel within the image, R is the red-axis within the image, G is the green-axis within the image and B is the blue-axis within the image.

All pixels within the image are compressed from RGB to gray-scale using the equation (1) as shown in Fig. 4.

**Feature Extraction**

Before training or testing part, the gray-scale image must be in the form of a mathematical vector representation that automatically and recursively computed by the algorithm of computer (according to these 6 steps). The output of this procedure is a vector that represents all important features such as texture, color and shape of the image as shown in Fig. 5.

**Step 1**: Calculate the high-contrasts through the image in x and y axis. The high-contrast is a derivative distribution of intensity, by equation (2) and the transpose of (2) as (3).

\[
D_y = \sum_{i}^{n_{\text{pixels}}} [-1 \ 0 \ 1], \quad (2)
\]

\[
D_v = \sum_{i}^{n_{\text{pixels}}} [-1 \ 0 \ 1]^{T}, \quad (3)
\]

where $D_y$ and $D_v$ is gradient value in horizontal and vertical axis.

**Step 2**: Separate the image into 4 (or more) blocks as shown in Fig. 6.
Step 3: Separate each block into 4 (or more) cells as shown in Fig. 7.

Step 4: Calculate the magnitude of gradient in each block by (4)

$$M_g(x, y) = \sqrt{D_h(x, y) \cos \theta + D_v(x, y) \sin \theta}$$

where $M_g$ is the magnitude of a gradient at the position $(x, y)$

Step 5: Calculate the derivative of gradients by (5)

$$Q_g(x, y) = \frac{\partial D_h(x, y)}{\partial D_h(x, y)}$$

Step 6: Extract the important features into a vector in (6) which parallel considers the step 4-5.

$$Features = \sum_1^M M_g(x, y) \ast Q_g(x, y)$$
**Deep learning: training part**

Only the Thai-food image with its information (name, ingredients and other textual recommendation) is input to this procedure. All important features of an image are trained in a large number of hidden nodes in the deep learning which uses the least-squared error. Target value is the name, ingredients and textual recommendation of Thai-food images to construct a computer-brain as visualized in Fig. 8. For the mathematical definition, the parameters and target classes as \( \{(x_{tk}, t_{tk}), \ldots, (x_{tk}, t_{tk})\} \) where \( x_{tk} = (x_{tk}, \ldots, x_{tk}) \), \( t_{tk} = [t_{tk}, t_{tk}] \), \( k = 1, 2 \).

**Step 1:** Identify the deep learning, the learning rate \( \eta \) as a small value convergence to 0, Least-error value as \( E \) (error threshold), and iteration of learning as \( \hat{L} \) (maximum epoch) by defining a counter in finding and the counter for all inputs.

**Step 2:** Randomize the weight \( W \) value and bias \( b \) value.

**Step 3:** Input the training set as a series of \( \{x_{tk}, t_{tk}\} \), input to the deep learning and also calculate the output and error of target class.

**Step 4:** Update the weight value and the bias value of backward again from the output layer to the first hidden layer. Calculate the derivative of error. Update the new weight value and the bias value in the output layer, the second hidden layer and the first hidden layer.

**Step 5:** If \( k < K; k = k + 1 \) and go to step 3, else ; go to the step 6

**Step 6:** Calculate average square of error from (7).

\[
MSE = \frac{1}{K} \sum_{k=1}^{K} \left\| e_k - t_k \right\|^2
\]  
(7)

In case of \( MSE > E \) and no. of iterations \( m < \hat{L} \) then identify \( k = 1, m = m + 1 \) and start a new iteration of training at step 3

In case of \( MSE \leq E \) or \( m \geq \hat{L} \), then the training is finished.

---

**Fig. 8** A vector (with its name, ingredients and recommendation) trained to the computer brain by deep learning

<table>
<thead>
<tr>
<th>0.25</th>
<th>0.84</th>
<th>0.53</th>
<th>0.34</th>
<th>…</th>
</tr>
</thead>
</table>

Name: Tom-yam-goong
Ingredients: Stock, Fish-sauce, Shrimp, Garlic, crushed, ...
Recommendation: the most famous of Thai soups and is popular not only in ...

Computer Brain

- Pad-thai
- Green-curry
- Som-tam
The most famous of Thai soups and is popular not only in Thailand but in Thai restaurants worldwide. It is a clear, sour soup flavored with fragrant lemon grass, fresh galangal root and kaffir lime leaf. This potent herbal mixture is well known for its medicinal properties.

**Fig. 9** Name, ingredients and textual recommendation tagged to the unknown image

Deep learning: testing part

The unknown Thai-food image is input to testing part for estimating its name, ingredients and textual-recommendation. Many images from the computer brain are used to compute the number of feature similarities with the unknown Thai-food. Deep learning automatically considers the highest similar number of features to estimate the name, ingredients and textual-recommendation. The unknown Thai-food image is finally tagged its information as shown in Fig. 9.

4. Experimental Measurement

In this experiment, all Thai-food images with their information were randomly crawled from the internet (for 458 images). The 80% of images were input to construct the computer-brain. The target classes were Tom-yam-goong, Pad-thai, Som-tam, Mas-sa-man, Green-curry, Thai-fried-rice, and Moo-nam-tok. For testing, we used another 20% of images; and to compute these criterions: Precision, Recall and Accuracy.

\[
\text{Precision} = \frac{TP}{TP + FP} \quad (8)
\]

\[
\text{Recall} = \frac{TP}{TP + FN} \quad (9)
\]

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)
\]

where \(TP\) = True Positive, \(FP\) = False Positive, \(TN\) = True Negative and \(FN\) = False Negative, respectively
The RGB to gray scale conversion, feature extraction and deep learning (either training or testing) were executed, respectively. This system was developed by Matlab. We used Precision, Recall and Accuracy in (8), (9) and (10) as criterions for the measurement of our system. The quantitative measurements from the names of unknown foods were calculated as the average of (8), (9) and (10) as shown in Table 1.

<table>
<thead>
<tr>
<th>Table 1 Precision, Recall and Accuracy criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>0.71</td>
</tr>
<tr>
<td>0.74</td>
</tr>
<tr>
<td>0.65</td>
</tr>
<tr>
<td>.</td>
</tr>
<tr>
<td>.</td>
</tr>
<tr>
<td>.</td>
</tr>
<tr>
<td>0.74</td>
</tr>
</tbody>
</table>

5. Conclusion

This paper introduces a novel ingredients estimation of Thai-foods which opens a new gate about the relation between Thai-food patterns and image processing with artificial intelligence for Thai-foods guidance of visitors. The system was categorized into RGB-to-grayscale conversion, Feature extraction and Deep learning (either training or testing). The system provides high correctness in criterions of accuracy, recall and precision which were measured from the food’s name. In every minute, there are more than 100,000 uploaded-images around the world. Some of them are about foods with the description. The numbers of food images have been exponentially increased on the social media. With the concept of data mining, the more data produces the more correctness. Hence, all food images on the social media are large-scale (also called “Big data”) that cover enough for all foods and dishes around the world. For future work, it is possible to further analyze nutrients; to compute the calories of Thai-foods from a single image. Imagine that how it is more convenient if the recipe or cookbook will be automatically shown to user from a food image. In coming soon, many street foods will be estimated their ingredients and calories using a mobile phone.
6. References


