Available online at http://www.ssstj.sci.ssru.ac.th

Suan Sunandha Science and Technology Journal

Thai Herb Identification with Medicinal Properties Using Convolutional Neural Network

Lawankorn Mookdarsanit[1*], Pakpoom Mookdarsanit[2]

[1*]Business Computer Department, Faculty of Management Science, Chandrakasem Rajabhat University
[2]Computer Science Department, Faculty of Science, Chandrakasem Rajabhat University
[1,2]39/1 Ratchadaphisek Rd, Khwaeng Chandrakasem, Khet Chatuchak, Bangkok 10900, Thailand
Corresponding author e-mail: *lawankorn.s@chandra.ac.th

Received: *4 July, 2019*/ **Revised:** ***17 July, 2019***/ **Accepted:** *25 July, 2019*

Abstract

This paper built an intelligent computer model to identify Thai herb from a single image using convolutional neural network. Thailand is one of the world herbal sources. We used 2,700 herbal images with their medicinal properties to train the computer model that covered 11 well-known Thai herbs: Siamese Rough-bush, Cumin, Holy Basil, Sweet Basil, Cha Muang, Kaffir-lime Leaf, Siamese Morning-glory, Pandanus Leaf, Mint, Chinese Kale and Chaplu, respectively. The feature extraction framework and model architecture were done by Fast Region Convolution Neural Network (Fast R-CNN) and Visual Geometry Group Network (VGGNet) that produced the recall as higher than 0.75 and the precision as higher than 0.80.

Keywords: Herb Identification, Leaf Recognition, Convolutional Neural Network, VGGNet, Fast R-CNN

1. Introduction

Thai medicinal herbs or herbal medicines refer to some natural plants that can be extracted some various pharmaceutical substances. These substances are considered as organic compounds that are absolutely safe to cure by naturopathy, also known as "Thai traditional home remedy". From the historical stone inscription, King Ramkhamhaeng of Sukhothai has inscribed the reference treatise that healers cultivated and used these herbal plants to cure the king's people (Masao, 1908). Furthermore, Simon de La Loubère – a French diplomat wrote a cultural and living description about Suvarnabhumi zone to the Emperor Louis XIV of France mentioned some contextual information about the traditional herbal and pharmaceutical plants (Love, 1994). Later the first medicinal prescription in Thailand was originated by the King Ramathibodi III (or King Narai) of Ayutthaya, also known as "the original King Narai's pharmacopeia" (Hodges, 1999). Until now, Thailand is one of the famous herbal sources that exports various medicinal herbs as the main materials for many international pharmaceutical companies. In a nutshell, Thai herbal plants can be seen as "Thai treasure and worthiness" that should be conserved to the young generation.

Despite Thailand as one of the world's herbal sources, many Thai youths do not perceive/identify Thai herbs and their essential properties clearly. This paper built an intelligent computer model based on object detection to identify Thai herbs and their properties. Even some herbal identification methods (Pornpanomchai & Rimdusit, 2011; Arun, et al., 2013; Janani & Gopal, 2013; Satti, et al., 2013; de Luna, et al., 2017) are available; these works are based on traditional hand-crafted models (in term of "Bag of words") with dimensionality reduction (Stephen, et al., 2019). Since the hand-crafted model was proven to be outperformed by the convolutional neural network (CNN) in term of classification accuracy (Zhao, et al.,2018; Zheng, et al. ,2018), (AlexNet) (Krizhevsky, et al., 2012) in the Large Scale Visual Recognition Challenge (ILSRVC) as well as food identification (Mookdarsanit, et al., 2018). These days, many applications have gradually replaced the hand-crafted models with CNN. A well-known CNN approach was applied for Chinese-herbal identification (Sun & Qian, 2016) based on Region-CNN (R-CNN) (Girshick, et al., 2014). The concrete R-CNN architecture has its multistage complex pipeline and is trained by SVM classifiers that are known as expensive, especially in memory

Suan Sunandha Science and Technology Journal

consumption and runtime (Mookdarsanit, et al., 2019). This paper used the Fast R-CNN (Girshick, 2015) architecture to identify the Thai herb from an image. For model creation, many Thai-herbal images with their medicinal properties were trained to the Visual Geometry Group Network (VGGNet) architecture (Simonyan & Zisserman, 2014). Firstly, the multiple regions of interests (RoIs) of an image as the feature maps were extracted by a two-stage framework (known as Fast R-CNN). After that, the feature maps as RoI vector were trained to Visual Geometry Group Network (VGGNet) architecture. Finally, an unknown herbal image was identified by the built model.

In this paper, we used 2,700 arbitrary herbal images as our primary dataset of the 11 medicinal herb types: Siamese Rough-bush, Cumin, Holy Basil, Sweet Basil, Cha Muang, Kaffir-lime Leaf, Siamese Morning-glory, Pandanus Leaf, Mint, Chinese Kale and Chaplu, as shown in Figure 1." The environmental backgrounds of herbal objects were one-tone scene and taken by iPhone 4s, Sony E PZ 16-50mm F3.5-5.6 OSS and Huawei P30 Pro.

This paper was organized into 5 sections (including introduction). The section 2 describes Fast R-CNN Framework. VGGNet Architecture was explained in section 3 and 4. Finally, the conclusion is in section 5.

2. Fast R-CNN Framework

Fast R-CNN is an end-to-end detection which is either speed or accuracy. Since the concept of R-CNN shares the computation of convolution over many region proposals and the Region of Interest (RoI) Poolings are added in the middle between the last convolutional (CONV) layer and the first Fully-connected (FC) layer.

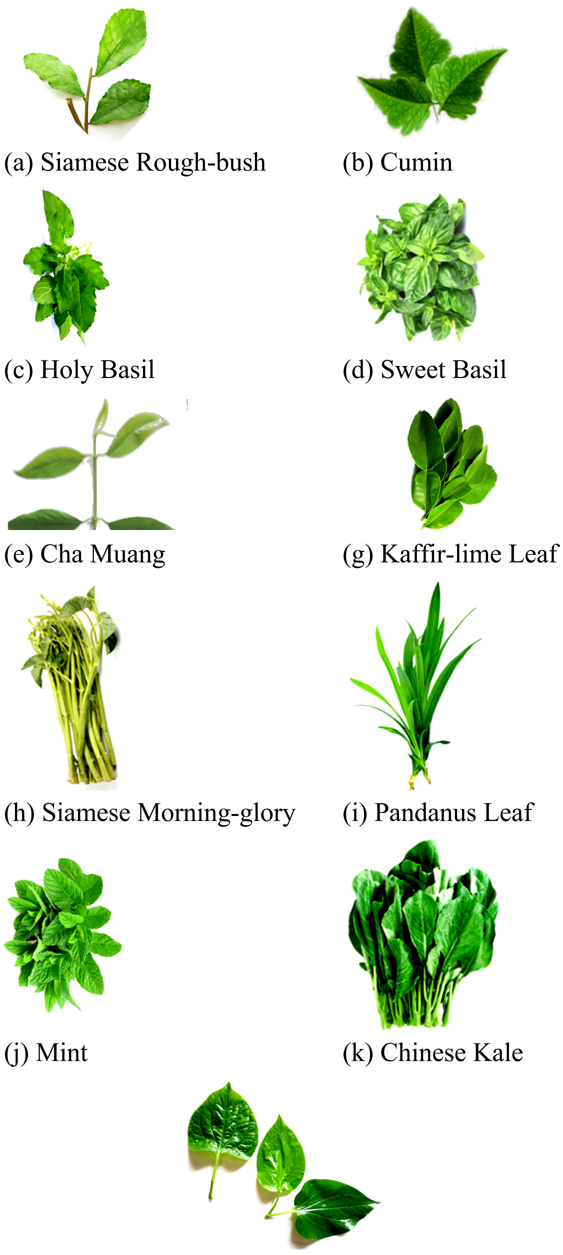
2.1 Convolution with ReLU

Thai herbal image passes the convolution with filters and receptive fields. All CNN filters within the same convolutional layer have the same size but different parameters.

In each convolutional layer, all Region of Interests (RoIs) of a leaf are extracted by complementary similarity measures that consist of color similarity, texture similarity, size integration and grown-region consistency.

For color similarity ($s_{color}(r_i, r_j)$) by *(1)*, each region (r_i) measured by 25 bins of color histogram $C_i = \{c_i^1, c_i^2, \ldots c_i^n\}$ using L_1 norm. The configuration of $n = 75$ is tuned in case of 3 (RGB) color channels.

$$s_{color}(r_i, r_j) = \sum_{k=1}^{n} \min(c_i^k, c_j^k) \qquad (1)$$



(a) Siamese Rough-bush (b) Cumin (c) Holy Basil (d) Sweet Basil (e) Cha Muang (g) Kaffir-lime Leaf (h) Siamese Morning-glory (i) Pandanus Leaf (j) Mint (k) Chinese Kale (l) Chaplu

Figure 1. Some medicinal taxonomy of Thai herbs

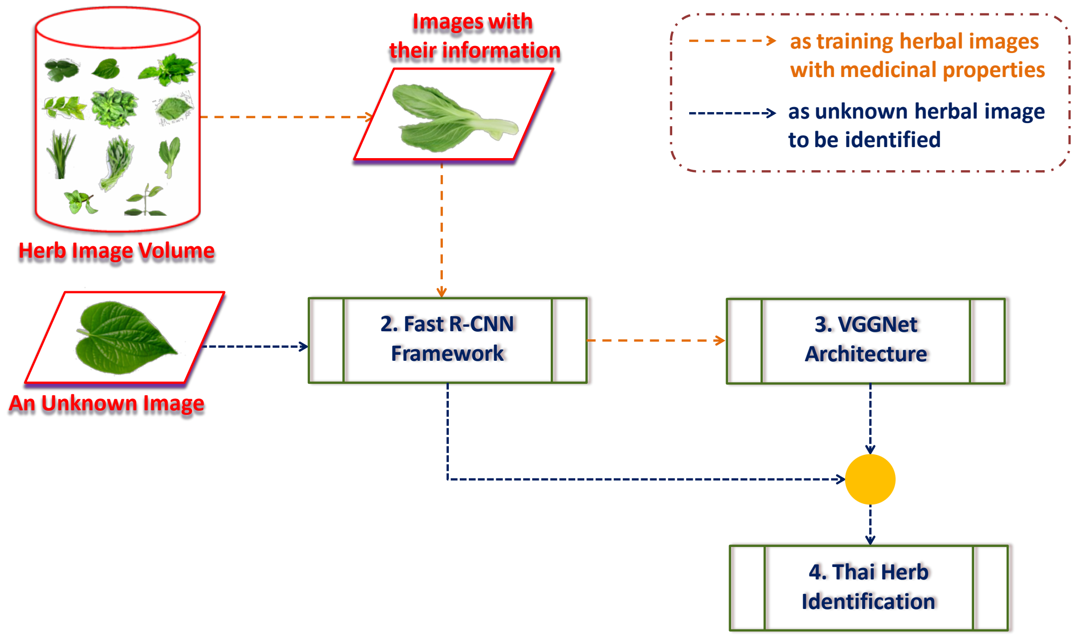


Figure 2. The overall framework of Thai herb identification with medicinal properties

For texture similarity ($s_{texture}(r_i, r_j)$) by *(2)*, the Gaussian derivatives in 8 orientations is done by $\sigma = 1$ in each color channel. The 10 bins are allocated for each orientation of each color channel $T_i = \{t_i^1, t_i^2, ... t_i^n\}$ where $n = 240$ for 3 channels.

$$s_{texture}(r_i, r_j) = \sum_{k=1}^{n} \min(t_i^k, t_j^k) \qquad (2)$$

For size integration ($s_{size}(r_i, r_j)$) by *(3)*, the related small regions are decided to grow together to ensure the object location for all pyramid scales.

$$s_{size}(r_i, r_j) = 1 - \frac{size(r_i) + size(r_j)}{size(im)} \qquad (3)$$

For grown-region consistency ($fill(r_i, r_j)$) by *(4)* that measures how consistent the mergence between r_i and r_j is. The bounding box ($size(BB_{ij})$) is a representative of integration between r_i and r_j. In case of the hardly touching between r_i and r_j, these regions are not merged.

$$fill(r_i, r_j) = 1 - \frac{size(BB_{ij}) - size(r_i) - size(r_j)}{size(im)} \qquad (4)$$

In a nutshell, the RoI ($x(r_i, r_j)$) are extracted by *(5)*, where $0 \le a_i \le 1$

$$x(r_i, r_j) = a_1 s_{color}(r_i, r_j) + a_2 s_{texture}(r_i, r_j) + a_3 s_{size}(r_i, r_j) + a_4 fill(r_i, r_j) \qquad (5)$$

Finally, the activation functions of these RoIs along gradient-based learning is approximated by Rectified Linear Unit (ReLU), where $\mathrm{Re}LU(x) = \max(0, x)$.

2.2 RoI Max Pooling

Before the max pooling, a linkage between convolutional layers is called the feature maps that contain RoIs. The feature maps ($Map_{Feature}$) from previous layers are embedded with kernels/receptive fields (k_{ij}^l) and biases (b_{ij}^l) that go to ReLu function. Each input feature maps of current layer inherit from previous output feature maps of previous layer as *(6)*, where M_j is a selection of input maps.

$$x_j^l = Map_{Feature}\left(\sum_{i \in M_j} x_i^{l-1} \bullet \left(k_{ij}^l + b_{ij}^l\right)\right) \qquad (6)$$

Suan Sunandha Science and Technology Journal

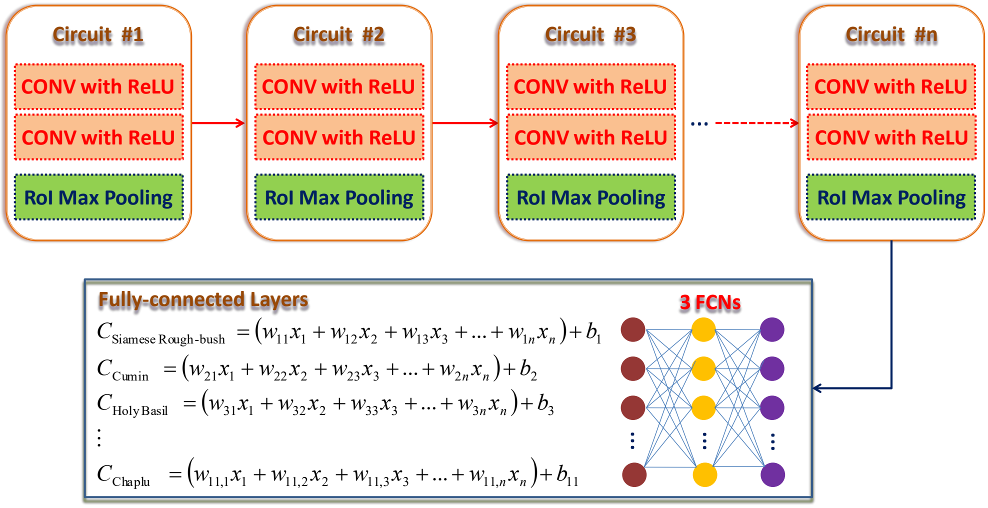


Figure 4. The Visual Geometry Group Network (VGGNet) Architecture

Later, the max pooling is used to reduce the dimension of output maps from previous layer that also can be implemented under the architecture of multi-scale pyramid pooling as shown in Figure 3.

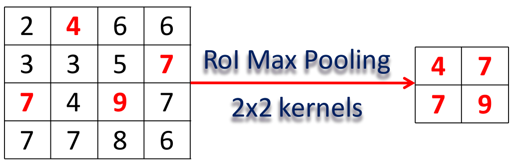


Figure 3. The 2x2 kernels of RoI max pooling

3. VGGNet Architecture

Visual Geometry Group Network (VGGNet) is a learning-based hierarchical architecture, including the convolution layers with ReLU activation functions and RoI Max-pooling layers in terms of vectors with scalar values. Finally, the Fully-connected layers build the class of each herb ($C_{herbal\ name}$) that use the concept of biases (b_i) and weights (w_{ij}) coupled with the extracted RoIs as parameters by *(7)*.

$$
\begin{aligned}
C_{\text{Siamese Rough-bush}} &= (w_{11}x_1 + w_{12}x_2 + w_{13}x_3 + \ldots + w_{1n}x_n) + b_1 \\
C_{\text{Cumin}} &= (w_{21}x_1 + w_{22}x_2 + w_{23}x_3 + \ldots + w_{2n}x_n) + b_2 \\
C_{\text{Holy Basil}} &= (w_{31}x_1 + w_{32}x_2 + w_{33}x_3 + \ldots + w_{3n}x_n) + b_3 \\
&\vdots \\
C_{\text{Chaplu}} &= (w_{11,1}x_1 + w_{11,2}x_2 + w_{11,3}x_3 + \ldots + w_{11,n}x_n) + b_{11}
\end{aligned} \tag{7}
$$

As shown in Figure 4, the initialization of VGGNet training is done by 2,700 herbal images with their textual medicinal properties. All features of each image (like color, texture, size and grown-region similarities) are mapped in term of 134M parameters by Fast-CNN and classified by 3 Fully-connected layers.

4. Thai Herb Identification

An unknown herbal image inputs identify itself with its medicinal properties by CNN-based model according to Figure 5.

The unknown herbal image is extracted the RoIs using Fast-RCNN Detection that consists of recursive convolutional layers and RoI Max Pooling layers. Later, the VGGNet architecture considers these extracted RoI features with other parameters like biases and weights to identify the herb and its medicinal properties.

The unknown herbal image identification is statistically computed by the conditional probability of FCN-based output classes by *(8)*.

$$
P_{FCN}\left(C_{herbal\ name} \mid x_i, (w_{i1}, w_{i2}, \ldots, w_{in}, b_1, b_2, \ldots, b_{11})\right) = \frac{e^{w_{ij}+b_j}}{\sum_{j=1}^{11} e^{w_{ij}+b_j}} \tag{8}
$$

where $C_{herbal\ name}$ is any target classes of the 11 herbal types, the max probability *max*($P_{FCN}(C_{herbal\ name})$) is selected to identify the herb with its properties.

For the model evaluation, we used the recall and precision (as *(9)* and *(10)*) to measure our CNN-based model to identify the herb and its medicinal properties.

$$
recall = \frac{TP}{TP + FN} \tag{9}
$$

Suan Sunandha Science and Technology Journal

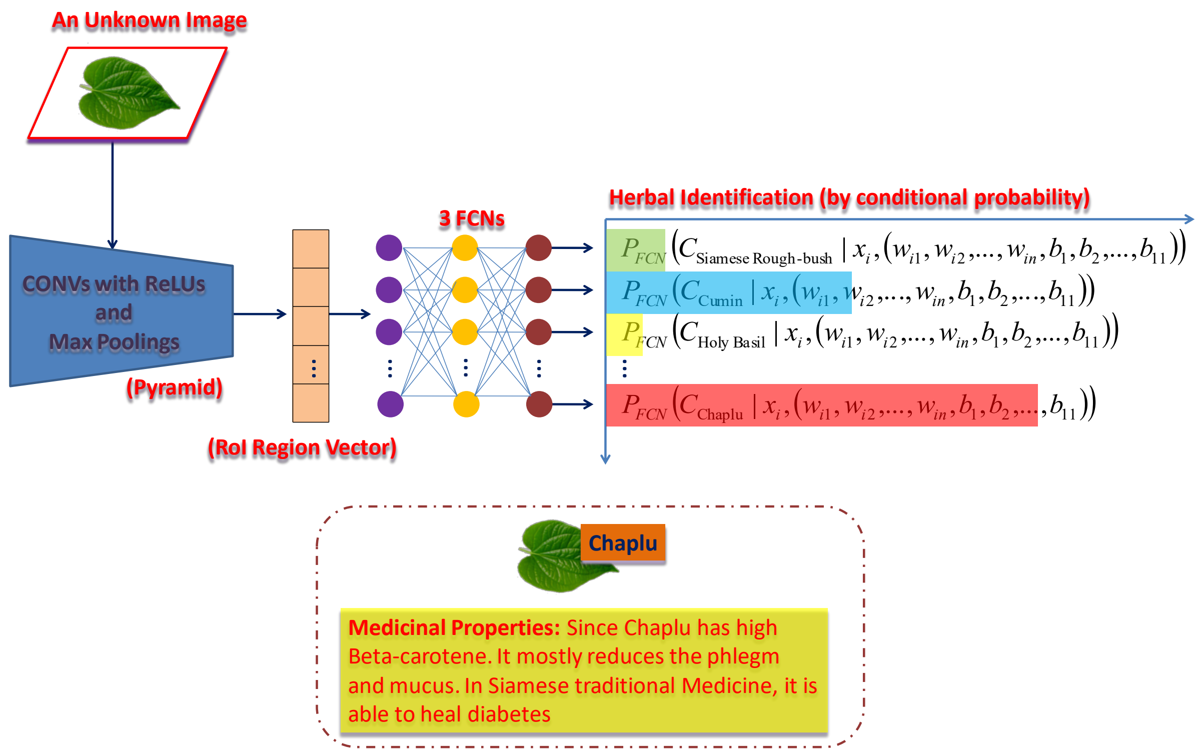


Figure 5. Herb identification with medicinal properties

$$precision = \frac{TP}{TP + FP} \tag{10}$$

where *TP* (True Positive) means "a number of herb A(s) and the model identifies correctly as the herb A", *TN* (True Negative) means "a number of not other herbs, the model identifies correctly as not the herb A", *FP* (False Positive) means "a number of not other herbs, the model identifies wrongly as the herb A" and *FN* (False Negative) means "a number of herb A(s), the model identifies wrongly as not the herb A"

5. Results and Discussion

The results with amount of herbal images are shown in Table 1. Pandanus Leaf has their own physical distinctiveness in shape and texture that totally differs from other herbs. Therefore it is the highest correctness in term of recall precision. However, Cumin and Chaplu sometimes look very similar that easily causes the misidentification. As well as Holy Basil, Sweet Basil and Mint that are not too much different then physically affects the differentiation between them.

Arbitrary herbal images in this research are taken from different recorders like mobile phones and cameras therefore the same herbal images may have different operations like resolution, distribution, texture or background. Thus, the herbal identification may produce some errors. Overall errors (computed by *FP*+*FN*) are shown in Figure 6.

Table 1. The herb identification's evaluation

Herbal Name	Amount of Images	Recall	Precision
Siamese Rough-bush	245	0.85	0.82
Cumin	245	0.78	0.86
Holy Basil	248	0.81	0.83
Sweet Basil	247	0.89	0.92
Cha Muang	245	0.95	0.92
Kaffir-lime Leaf	240	0.95	0.99
Siamese Morning-glory	245	0.91	0.92
Pandanus Leaf	244	0.99	1.00
Mint	245	0.82	0.82
Chinese Kale	248	0.95	0.94
Chaplu	248	0.88	0.81

Available online at http://www.ssstj.sci.ssru.ac.th

Suan Sunandha Science and Technology Journal

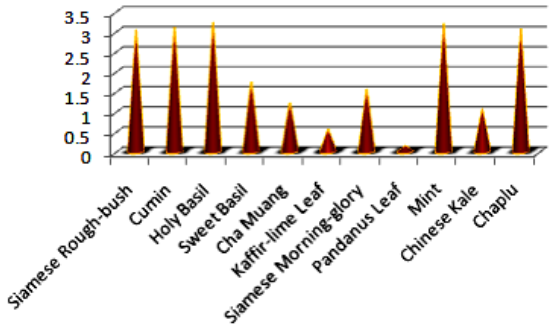


Figure 6. Overall errors (*FP+FN*) in percentage

As shown in Figure 7, the comparison between those traditional hand-crafted herbal identification (Pornpanomchai & Rimdusit, 2011; Arun, et al., 2013; Janani & Gopal, 2013; Satti, et al., 2013; de Luna, et al., 2017) and CNN-based model (Sun & Qian, 2016) and ours are compared based on the amount of data (Alom, et al., 2018). If there are a small-sized number of images to train the computer model, the hand-crafted based recognition has more performance than CNN. On the other hand, CNN is totally better if a large-scale amount of images are trained to the model.

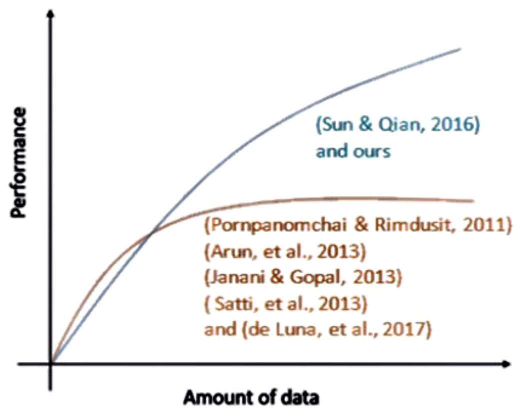


Figure 7. Comparison between hand-crafted and CNN approaches

According to the big data, the performance of CNN approaches have become more compatible in large-scale problems, known as the more volume, veracity, velocity and variety of image data. In case of image data insufficiency, a large number of images with the same distribution can be autonomously generated by transfer learning like Generative Adversarial Network (GAN) with Auto Encoder (AE) (Giuffrida, M.V., et al., 2017). Moreover, there are another some secondary data about CNN-based comparison (Liu, et al., 2018) between R-CNN (herbal identification by Sun & Qian) and Fast R-CNN in Table 2.

Table 2. Comparison between R-CNN and Fast R-CNN framework

Framework	Dataset Pascal-VOC07	Dataset Pascal-VOC12
R-CNN based (by Sun & Qian, 2016)	0.58	0.53
Fast R-CNN based (ours)	0.70	0.68

6. Conclusion

Thai medicinal herbs are one of the treasure since Sukothai era (known as "Thai traditional home remedy"). It is not surprise that why Thailand is one of the biggest herbal agricultural sources for many well-known pharmaceutical companies. Many Thai youths do not perceive clearly Thai traditional herbs and their essential properties. In this paper, we built a CNN-based model to identify the herb and its medicinal properties from an unknown image. The Fast-RCNN framework was applied to extract CNN-features from 2,700 herbal-tagged images with their medicinal properties. IT covers 11 herbal types: Siamese Rough-bush, Cumin, Holy Basil, Sweet Basil, Cha Muang, Kaffir-lime Leaf, Siamese Morning-glory, Pandanus Leaf, Mint, Chinese Kale and Chaplu. the learning model was built by Visual Geometry Group Network (VGGNet). For the model evaluation, the overall recall was higher than 0.75 and the precision was higher than 0.80. However, the same herb taken from different devices with various resolutions may produce the misidentification.

For future work, the improvement of image resolution can be done using the power-law-adaptive sample generation by Generative Adversarial Network (GAN) coupled with Auto-encoder (AE). Moreover, these self-supervised approaches are able to generate more than 10,000-100,000 large-scale images from the collection of original primary images by randomized adversarial perturbation.

7. Acknowledgement

This research was implemented by M-script of Caffe MATLAB. All herbal images (samples) in this research were taken by Sony E PZ 16-50mm F3.5-5.6 OSS and the embedded cameras in iPhone 4s and Huawei P30 Pro. All images and diagrams in this paper were hidden watermarked as our primary dataset. All supports with computational hardware and software in the computer laboratory of Chandrakasem Rajabhat University are acknowledged.

Suan Sunandha Science and Technology Journal

8. References

Alom, M. Z., et al. (2018). The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches. arXiv, 1803.01164, 1-39.

Arun, C. H., et al. (2013). Texture Feature Extraction for Identification of Medicinal Plants and Comparison of Different Classifiers. International Journal of Computer Applications, 62(12), 1-9.

de Luna, R. G., et al. (2017). Identification of Philippine herbal medicine plant leaf using artificial neural network. The 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management, pp. 1-8.

Girshick, R., et al. (2014). Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. International Conference on Computer Vision and Pattern Recognition, pp.580-587.

Girshick, R. (2015). Fast R-CNN. The 2015 International Conference on Computer Vision, pp.1440-1448.

Giuffrida, M.V., et al. (2017). ARIGAN: Synthetic Arabidopsis Plants Using Generative Adversarial Network. The 2017 IEEE International Conference on Computer Vision, pp. 2064-2071.

Hodges, I. (1999). Time in Transition: King Narai and the Luang Prasoet Chronicle of Ayutthaya. Journal of the Siam Society, 87(1), 33-44.

Janani, R. & Gopal, A. (2013). Identification of Selected Medicinal Plant Leaves Using Image Features and ANN. IEEE Conference on Advanced Electronic Systems, pp.238-242.

Krizhevsky, A., et al. (2012). ImageNet Classification with Deep Convolutional Neural Networks. The 25th International Conference on Neural Information Processing Systems, pp. 1097-1105.

Liu, L., et al. (2018). Deep Learning for Generic Object Detection: A Survey. arXiv, 1809.02165, 1-37.

Love, R. S. (1994). Simon De La Loubere: French Views of Siam in the 1680s. Journal of the Siam Society, 82(2), 155-164.

Masao, T. (1908). The New Penal Code of Siam. The Yale Law Journal, 18(2), 085-100.

Mookdarsanit, P., et al. (2018). Name and Recipe Estimation of Thai-desserts beyond Image Tagging. Kasembundit Engineering Journal, 8(Special Issue), 293-303.

Mookdarsanit, S., et al. (2019). SiamFishNet: The Deep Investigation of Siamese Fighting Fishes. International Journal of Applied Computer Technology and Information Systems, 8(2), 40-46.

Pornpanomchai, C. & Rimdusit, S. (2011). Thai Herb Leaf Image Recognition System (THLIRS). Kasetsart Journal, 45(3), 551-562.

Satti, V., et al. (2013). An Automatic Leaf Recognition System for Plant Identification Using Machine Vision Technology. International Journal of Engineering Science and Technology, 5(4), 874-879.

Simonyan, K. & Zisserman A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. The 2014 European Conference on Computer Vision, pp. 84-90.

Stephen, F. T., et al. (2019). Application of Principal Component Analysis as a Data Reducing Technique. Suan Sunandha Science and Technology Journal, 6(1), 27-30.

Sun, X. & Qian, H. (2016). Chinese Herbal Medicine Image Recognition and Retrieval by Convolutional Neural Network. Public Library of Science One, 11(6), 1-19.

Zhao, Z. Q., et al. (2018). Object Detection with Deep Learning: A Review. IEEE Transactions on Neural Networks and Learning Systems, PP(99), 01-21.

Zheng, L., et al. (2018). SIFT Meets CNN: A Decade Survey of Instance Retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(5), 1224-1244.