

Research Article

Toward Recognition of Easily Confused TCM Herbs on the Smartphone Using Hierarchical Clustering Convolutional Neural Network

Kun-Chan Lan ¹, Tzu-Hao Tsai ¹, Min-Chun Hu,² Juei-Chun Weng,¹ Jun-Xiang Zhang,¹ and Yuan-Shiun Chang³

¹The Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan 701, Taiwan

²The Department of Computer Science and Information Engineering, National Tsing Hua University, Hsinchu 300044, Taiwan

³The Department of Chinese Pharmaceutical Sciences and Chinese Medicine Resources, China Medical University, Taichung 40402, Taiwan

Correspondence should be addressed to Kun-Chan Lan; klan@csie.ncku.edu.tw

Received 9 May 2022; Revised 30 December 2022; Accepted 5 April 2023; Published 5 May 2023

Academic Editor: Ammar AL-Farga

Copyright © 2023 Kun-Chan Lan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background and Objective. The use of Chinese herbal medicines (CHMs) for treatment plays an important role in traditional Chinese medicine (TCM). However, some herbs are easily confused with the others because their shapes/textures look similar and they could have totally different utilities. Recently, deep learning has attracted great attention for the application of image recognition and could be useful for TCM herb identification. **Methods.** For recognizing easily-confused TCM herbs on a smartphone, we propose a CHM recognition system using hierarchical clustering convolutional neural networks (HCNNs) based on the affinity propagation clustering method. **Results.** We implement our system on the smartphone and show recognition accuracy close to 98%, based on a dataset of 65 kinds of herbs (including 12 easy-confused herbs pairs). We also investigate the effect of different parameters (e.g., selection of clustering algorithms for HCNNs, types of smartphone, and number of layers in the neural network) on the system performance. **Conclusions.** In this work, we proposed a hierarchical clustering convolutional neural network (HCNN) method to distinguish similar TCM herbs with a high accuracy. We also showed the usefulness of applying the data augmentation techniques when implementing the proposed system for a variety of smartphones.

1. Introduction

Chinese herbal medicines (CHMs) play an important role in TCM. Chinese herbs primarily come from different parts of the plants, including leaves, roots, stems, flowers, and seeds. The core idea of CHM is to restore the balance of the human body to achieve a state of health. While CHM has become an increasingly popular method of treatment globally, for most people, it is difficult to recognize different Chinese herbs and know the properties of each kind of herb. Moreover, some herbs are easily confused with the others because their shapes/textures look similar to the others but they could have totally different utilities. For example, Astragali Mongolici Radix [1] is commonly used in CHMs treatment

because of its efficacy in strengthening the immune system. However, some people will sell Hedysari Radix [2] as a replacement for Astragali Mongolici Radix because its flavor is tastier and the price is cheaper than Astragali Mongolici Radix. Hedysari Radix has similar shapes/textures as Astragali Mongolici Radix, but with lower efficacy for boosting the immune system. Other examples of easily-confused herbs are Diocoreae Rhizoma [3] and Manihot Esculenta [4]. The former is commonly used to maintain the function of the lung and kidney while the latter could be poisonous if not properly used. In Table 1, we listed some commonly-used Chinese herbs [5].

Chinese herbs are commonly used for food preparation and play a very vital role in Chinese medicine. For the

TABLE 1: Different types of Chinese herbs.

Categorization		Herb names (example)	Treating diseases
Four natures	Hot, warm	Zingiberis rhizome	Abdominal cramp, diarrhea
		Angelicae Sinensis Radix	Chronic constipation, menstrual disorders
	Cool, cold	Rehmanniae Radix	Cardiovascular diseases
		Gypsum	Bronchial asthma
	Neutral	Poria	Promoting urination, edema
Five flavors	Acrid	Gypsum	Bronchial asthma
	Sweet	Cyperis rhizoma	Emotional disorders
	Bitter	Armeniacae semen	Dissolve phlegm, relieve cough
		Atractylodis rhizoma	Palpitation, edema
	Sour	Schisandrae Fructus	Spontaneous sweating, night sweating
		Scrophulariae Radix	Chronic pharyngitis
	Salty	Pumex	Sticky sputum
	Ecklonia thallus	Goiter, scrofula	

reasons of safety and efficiency, it is important to have proper recognition of these herbs. However, given that some herbs have similar shapes/textures, most people find it difficult to recognize them without extensive experiences or expert knowledge. Therefore, it might be necessary to develop a system to facilitate people to recognize the herbs and understand the properties of these herbs.

Although there are many illustration handbooks of Chinese herbs around, it is time-consuming and inefficient to use these books to distinguish these easily-confused herbs. On the other hand, given the popularity of the smartphone, it can serve as a convenient vision-based-measurement (VBM) [6] instrument for recognizing the herbs. While a few prior studies attempted to use computer vision techniques for herb recognition [7–10], their results are generally limited in the following aspects: (1) relying on hand-crafted features, (2) based on a small data set (e.g., only 18 herbs in [7, 8]), (3) not targeting on easily-confused herbs, and (4) low recognition accuracy. In this work, we aim to build a system on top of the smartphone based on convolutional neural network (CNN) for recognition of easily confused TCM herbs. More specifically, we proposed a hierarchical CNN method to classify easily confused herbs by first clustering similar herbs into a group (using the affinity propagation algorithm [11]) and building a CNN model for these groups. We then train a CNN model for each group to classify herbs in the same group.

The contributions of this paper are in two folds. First, we set out to develop a system for automatic recognition of easily confused CHMs on the smartphone. Users just need to take pictures of CHMs and the system will show information of the herbs on the phones. The proposed system could potentially be used for the following applications: (1) knowing whether the herb is genuine or not and (2) understanding the properties of the herbs. As far as we know, this is the first TCM herb recognition App implemented on a phone. Second, we proposed a hierarchical CNN (HCNN) method for recognizing 24 easily-confused herbs. Our initial results show classification accuracy close to 98% (a 5% improvement in comparison to the naive CNN). Note that, while our proposed HCNN architecture is not entirely new, we investigate the effect of different parameters (e.g., selection of clustering algorithms for HCNN, types of

smartphone, and different CNN models) on the system performance. We believe that these insights could be of interests to readers of this journal.

2. Related Work

2.1. Herb Recognition Based on Its Smell and Taste. Luo et al. [12] developed an electronic nose that simulates biological olfactory organs to achieve the physiological function of the nose through machine learning. Their work can identify 6 types of Pungent CHMs. In addition, they proposed a method using the electronic tongue to identify taste information of five different CHMs [13]. However, these prior works have the same problem. First of all, such instruments are not easy to be built or obtain. In addition, some time-consuming preprocessing needs to be performed first before one can employ such approaches to identify different herbs (e.g., herbs need to be grinded into powders and heated for 30 minutes).

2.2. Traditional Vision Techniques for Herb Recognition. Herb recognition using computer vision techniques is generally more cost-effective than methods based on taste or smell. Some of them are based on hand-crafted features. Tao et al. [7, 8] utilized texture to classify 18 different CHMs. Cai et al. [9] and Liu et al. [10] used color, texture, and shape feature descriptors to identify 3 and 8 different CHMs, respectively. Finally, there are also some prior work on leaves and flowers recognition [14–17], using techniques such as local binary pattern (LBP) [18], histogram of oriented gradients (HOG) [19], and scale-invariant feature transform (SIFT) [20].

2.3. CNN for Herb Recognition. The problem in using traditional hand-crafted feature descriptors is that one needs to know what kinds of features are appropriate for classification. However, it might be difficult to find the representative features to identify the differences between a set of easily confused herbs. Deep learning has recently become increasingly popular, and many studies have shown that it can outperform many traditional machine learning methods for various image recognition tasks [21]. In particular, CNN has

attracted strong interest from both academia and industry since the ImageNet dataset became available. Sun and Qian [22] used CNN for CHMs recognition by collecting a total of 5,523 images from 95 categories. The average accuracy rate of their results is about 71%. They did not particularly consider the use of CNN for recognizing easily confused herbs though (e.g., as shown in Table 2, there are only two easily confused herbs pairs in their dataset). In our work, we propose a method based on HCNN to distinguish 12 pairs of easily confused herbs.

2.4. Hierarchical CNN. The concept of hierarchical CNN was introduced in some prior work. Yan et al. [23] implemented HD-CNN (hierarchical deep CNN) that breaks down an image recognition task into two levels. To separate simple classes from each other, an HD-CNN first uses a CNN classifier to classify the image data into K coarse categories. More complicated classes are redirected downstream to fine classifiers with divisions that concentrate on confusing classes. This work showed an improvement of 2.28% on the accuracy rate based on CIFAR100 and ImageNet datasets. They used spectral clustering to cluster their data into K coarse categories.

Mao et al. [24] evaluated their HCNN approach on the German traffic sign recognition benchmark (GTSRB). They proposed a CNN-oriented family clustering (CFC) algorithm to partition the traffic signs into K families. In these studies, the number of clusters (i.e., K) needs to be pre-defined which are more suitable for the fixed dataset like ImageNet. In our work, we employ affinity propagation (AP) to cluster easily confused herbs. AP does not require the number of clusters to be determined in advance. Given that currently there is no large herb image database available (the image data used in this study are all created by ourselves), AP is more suitable to us since we can then expand our database over time without worrying about changing our algorithm.

The introduction should be succinct, with no sub-headings. Limited figures may be included only if they are truly introductory, and contain no new results.

3. System Framework

We first started our experiments using a naive CNN to recognize some easily confused herbs. But then, we realized that some of these herbs look very similar and we were unable to obtain good results for these herbs. Therefore, in this work, we implement a hierarchical CNN method for these easily confused herbs.

Figure 1 shows our hierarchical clustering CNN architecture. In the training phase, we first cluster similar-looking herbs into the same group. Next, we create a two-layer CNN. The first layer is to create a model for cluster classification while the second layer is to classify herbs in the same cluster. In the testing phase, the system will first decide which cluster the input image belongs to, and then use the trained model in the second layer to recognize the herb within the identified cluster.

In this paper, we apply the AP algorithm [11] to cluster similar herbs into a group. AP is based on the concept of message passing between data points so that each data will

find the most suitable ones as its exemplars (i.e., cluster center or cluster head) and how much they are suitable as exemplars. Unlike traditional clustering algorithms such as k -means, AP does not require the number of clusters to be determined in advance. More specifically,

- (1) For our training data, we randomly sample some images from each kind of herb, and then, we extract their features to calculate the similarity matrix [11] as the input of the AP algorithm.
- (2) After performing the AP algorithm, each data point (i.e., herb) decides its exemplar. It is possible that the same kind of herb might choose different exemplars, so we adopt a majority-vote mechanism to decide the final exemplar for each kind of herb.
- (3) If the final exemplar of two kinds of the herb are the same, we cluster them into the same group.

In the first layer of our CNN-based clustering model, we utilize an open-source deep learning framework named Caffe [25], and we pretrain our CHMs model over 1 million ImageNet images of 1,000 categories. The architecture of CaffeNet is shown in Figure 2. It consists of eight layers, of which the first five layers are convolutional layers. Three max-pooling layers follow the first, second, and fifth convolutional layers, respectively. The last three layers are fully-connected layers. The number of neurons in the last fully-connected layer of our clustering model is set to the number of herb groups. The function of the first layer of our CNN-based clustering model is to decide which group the input herb belongs to.

The function of the second layer of our HCNN model is to recognize the target herb from an herb group. In the second layer, for each herb group, we train a CNN model similar to the first layer. In other words, there are multiple second-layer models, and each of them is corresponding to an herb group. Each herb group will contain at least one herb based on the clustering results.

4. Results

4.1. Experimental Environments. We evaluated our proposed hierarchical CNN model using CaffeNet [25] based on the AlexNet model [21], and all experiments were run on 64 bit Ubuntu 14.04 with an INTEL i7-4790 CPU, a GEFORCE GTX 1060 GPU, and 16 GB RAM. In particular, BLVC CaffeNet [25] is used for training the CNN model. The trained model is later ported to the phone for the testing phase.

4.2. CHM Dataset. In this work, we select 12 pairs of easily-confused CHMs from a book named “Illustrations of Commonly Misused Chinese Crude Drug Species in Taiwan” [26] for our experiments, as shown in Figure 3. The image data used for our experiments were taken by an iPhone6. One hundred images are taken for each herb so that our dataset contains 2,400 images in total. 1,440 images are used for training and the rest are used for testing. The name of the herbs is shown in Table 3. There are a total of twelve easily-confused herbs pairs in this dataset (e.g., A1/A2).

TABLE 2: Comparison of vision-based techniques for herb recognition.

	# of category	# of easily confused herbs pair	Methods	Implemented on smartphones
Tao et al. [8]	18	1	Hand-crafted	No
Liu et al. [10]	8	0	Hand-crafted	No
Sun and Qian [22]	95	2	CNN	No
Ours	65	12	Hierarchical clustering CNN	Yes

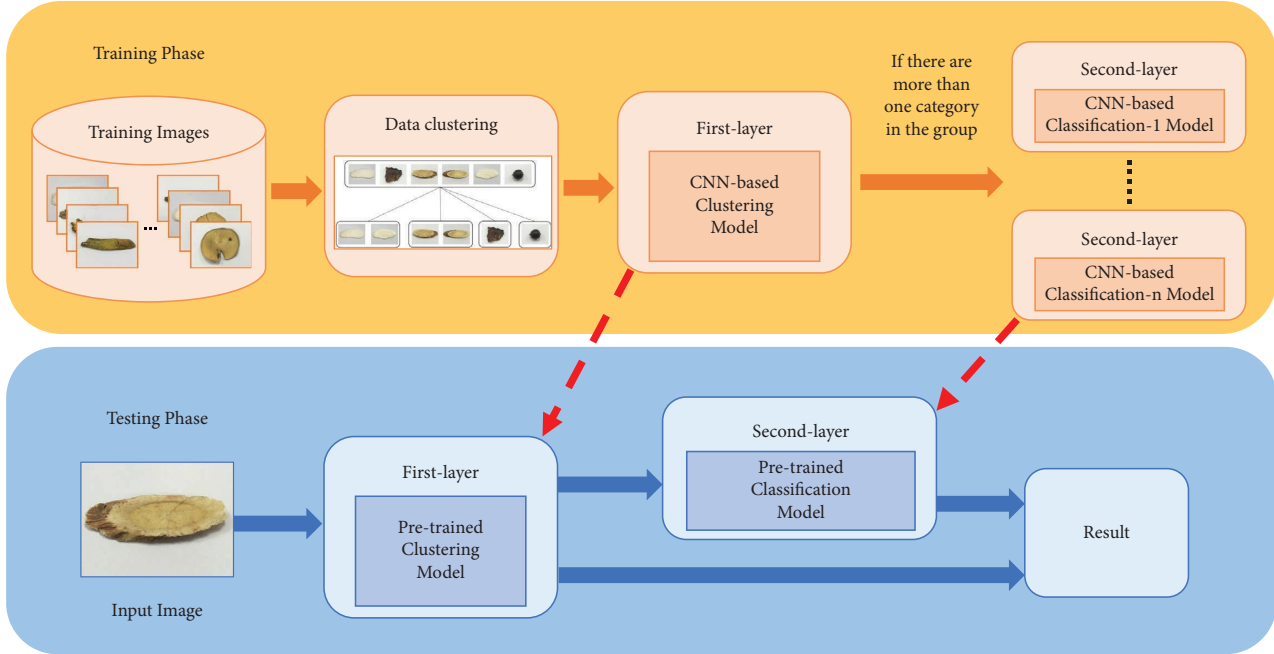


FIGURE 1: Hierarchical clustering CNN framework for CHM recognition.

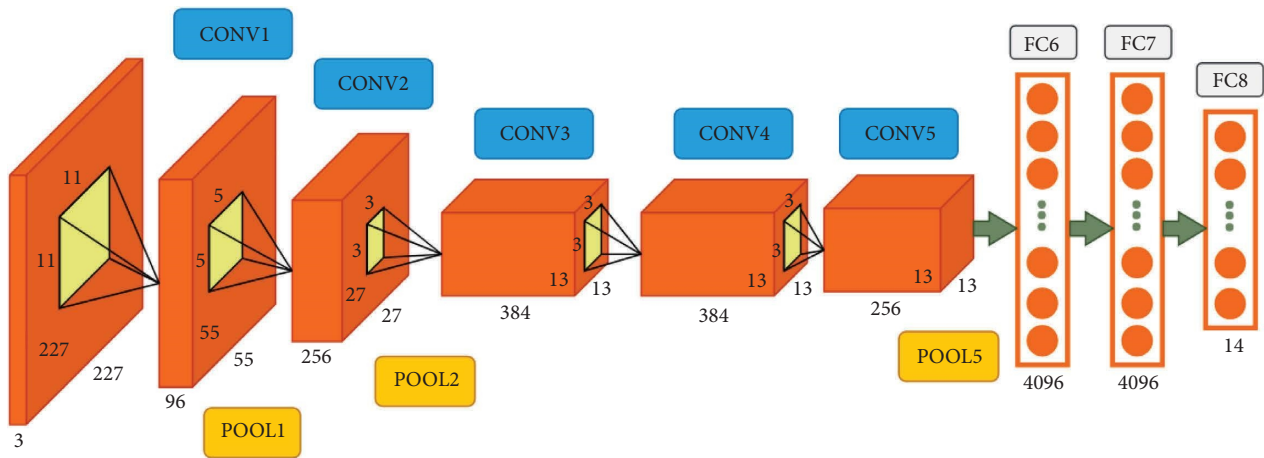


FIGURE 2: The architecture of our CaffeNet model in the first layer.

4.3. *The Benefit of Using Transfer Learning on the Accuracy of Our CNN Model.* BVLC CaffeNet provides an option called fine-tune which allows one to copy the model parameters from a pretrained CNN model (otherwise, all the parameters in the CNN model are initialized with random values). Given our dataset is small, it is expected to be beneficial from using some pretrained parameters to initialize our model.

This is known as the transfer learning method. We utilize the pretrained parameters from the model.

ImageNet work [21] is used to initialize our CHMs model. Figure 4 shows the average accuracy with and without the use of transfer learning. It clearly shows that the model can quickly converge with much higher accuracy when the fine-tune option is enabled.

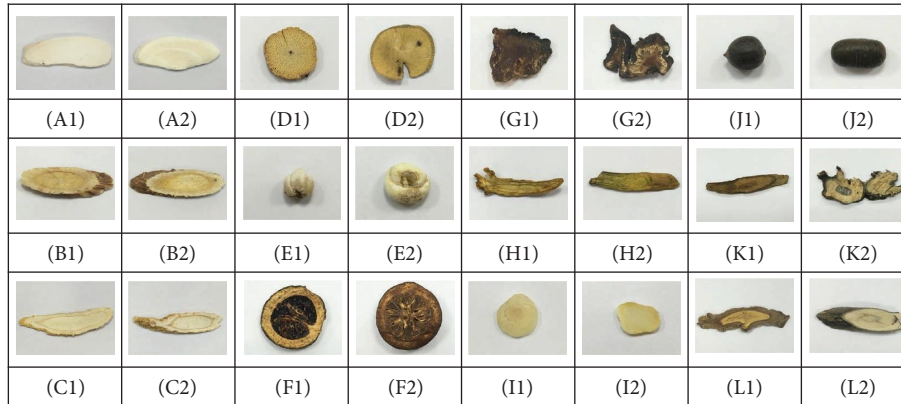


FIGURE 3: Twenty four kinds of CHMs collected by ourselves.

TABLE 3: The name of 24 kinds of herbs.

Numbers	Names
A1	Diocoreae Rhizoma
A2	Manihot esculenta
B1	Astragali Mongolici Radix
B2	Hedysari Radix
C1	Ginseng Radix
C2	Panacis Quinquefolii Radix
D1	Clematidia Armandii Caulis
D2	Aristolochia Manshuriensis Caulis
E1	Fritillariae Cirrhosae Bulbus
E2	Fritillariae Ussuriensis Bulbus
F1	Poncirus Trifoliata Fructus
F2	Aurantii Immaturus Fructus
G1	Aconiti Radix
G2	Aconiti Kusnezoffii Radix
H1	Scutellariae Baicalensis Radix
H2	Scutellariae Amoena Radix
I1	Pinelliae Rhizoma
I2	Typhonium Rhizoma
J1	Nelumbinis Fructus
J2	Caesalpinia Fructus
K1	Cyathulae Radix
K2	Strobilanthes Radix
L1	Isatis Radix
L2	Baphicacanthus Radix

4.4. Comparison of the Hand-Crafted Method with CNN for Herb Recognition. Some prior studies employed hand-crafted features for the herb recognition. In this study, we compare three different feature extraction methods, including HOG [19], LBP [18], and BOW SIFT [20], with CNN. For HOG and LBP implementation, the cell size is set to 32. Because the number of SIFT feature points in each image is not fixed, we first extract the SIFT feature points of all the training data and run them through K-means clustering (with the center set to 200) so that all images can have the same dimensional vector. Finally, a pretrained SVM [27] model to be used for herb classification.

For the CNN experiment, we first rescale the image size to $256 * 256$ and then randomly crop $224 * 224$ patches from these images to increase the number of training data and reduce overfitting. We enable the fine-tune and set the

number of neurons to be 24 in the last layer to match the number of easily confused herbs in our data. Our model is trained using stochastic gradient descent with a batch size of 60 samples (we set the momentum to 0.9, weight decay to 0.0005, and gamma to 0.1). An equal learning rate is used for all layers and the start learning rate is initialized at 0.0001.

Table 4 shows the results using hand-crafted methods and the CNN method. We employ five-fold cross-validation to calculate the accuracy. Among the traditional hand-crafted methods, LBP achieves the highest accuracy at 86.85%. This is not surprising since texture is an important feature for CHMs and LBP is powerful for texture classification. The accuracy of using CNN is 95.69%, which is much better than that of all the hand-crafted feature descriptors methods.

In addition, we compare two different CNN models, CaffeNet [25] and VGG16 [28]. The latter is used by a prior study for CHMs recognition [28]. VGG16 uses a deeper structure than CaffeNet so that it takes more time for training and testing, as shown in Table 5 (the same parameters are used for both the models). The execution time of VGG16 is about three times longer but the accuracy of both models is similar. Therefore, we decide to use a simple model like CaffeNet in this work because it runs faster with acceptable accuracy, as shown in Table 4.

Figure 5 shows the accuracy of detection for all 24 different herbs using a confusion matrix. We find that some herbs are more easily mistaken for another herb, such as B1 and B2 as well as C1 and C2. These recognition errors are reasonable though since they are easily confused even for human eyes. In the next section, we show that these classification errors can be improved with the proposed hierarchical clustering CNN method.

4.5. Performance of Hierarchical Clustering CNN Method.

In this work, we employ the use of HCNN to reduce recognition errors. We propose the use of the AP algorithm [11] to automatically cluster similar herbs into a group. In this section, we compare the performance of AP clustering and manual clustering which is based on an illustrated handbook of easily confused herbs [26]. As shown in Figure 6, twenty-

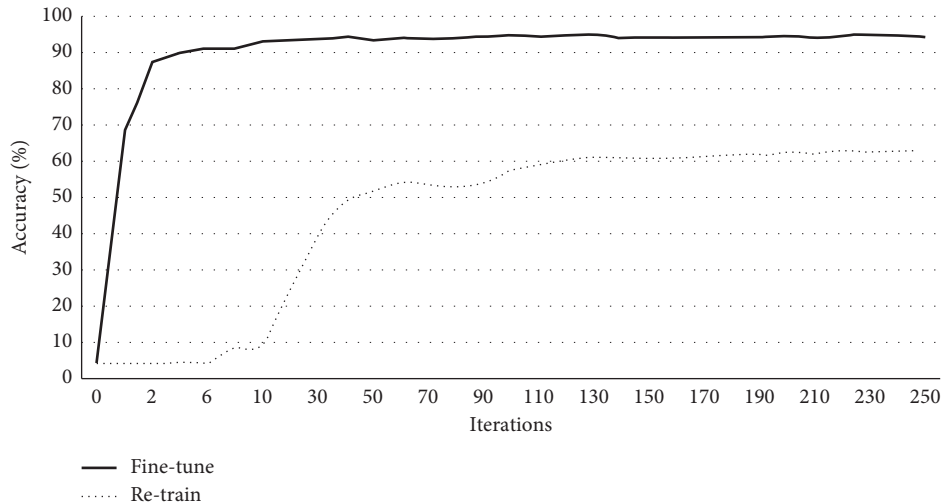


FIGURE 4: The effect of fine-tune on the performance of the CNN model.

TABLE 4: Accuracy of the hand-crafted method vs. CNN method.

Methods	Accuracy (%)
CNN (CaffeNet)	95.65
CNN (VGG16)	95.63
LBP + SVM	86.85
HOG + SVM	75.31
BOW SIFT + SVM	70.83

TABLE 5: Training and testing time of the CaffeNet and VGG16 models.

Methods	Training time (s)	Testing time (s)
CNN (CaffeNet)	408	1.9306
CNN (VGG16)	1.704	5.9577

four herbs are divided into 14 groups according to the illustrated handbook.

For the implementation of the AP algorithm, we calculate the similarity matrix based on LBP features because it has the best classification accuracy among all hand-crafted features we tried. We randomly sample 20 images for each herb to run AP. The results of AP clustering are shown in Figure 7 for comparison with the result of manual clustering in Figure 6. Figure 8 shows the confusion matrix of the cluster classification using CNN based on the AP algorithm (i.e., if a herb is classified into the correct cluster), and Figure 9 shows the confusion matrix for classifying the herb within a cluster based on the classification results shown in Figure 8.

Furthermore, we compare the performance of the AP algorithm with other clustering algorithms. Table 6 shows the herb classification accuracy from ten experiments using the other automatic clustering algorithm including k-means [29] and spectral clustering [30]. Both of them require the number of clusters to be predetermined before running the algorithm. Here, we let the number of clusters (i.e., K value) be 14 which matches with the manual clustering using the illustrated handbook. Table 6 shows that the AP algorithm

has a more stable and higher accuracy. For spectral clustering, some of the results are even worse than naive CNN (i.e., CNN without hierarchical clustering, as shown in Table 4). This is due to that spectral clustering first uses a Laplacian matrix to reduce the dimension, and then employs the k-means algorithm to do the clustering. Therefore, it might lose some information during the dimension reduction. In addition, its results are sensitive to the decision of the initial K value. A bad choice of K might lead it to a local optimal solution.

Table 7 shows a detailed comparison of the recognition accuracy for each herb between the CNN method and the proposed HCNN method. We find a significant improvement for some herbs such as B1 and H2 in addition to a general improvement of average accuracy (about 2%) when the HCNN method is employed. The results from AP clustering are very similar to that of manual clustering based on the illustrated handbook (which is considered as the ground truth for herb clustering in this study).

4.6. The Effect of the Number of CNN Layers. The above results are based on CNN architecture of 8 layers. A recent trend is to perform model compression (e.g., by reducing the number of layers of a deep neural network) for resource-limited devices like smartphones [31]. We next want to explore the use of a smaller number of layers for CNN training. Specifically,

- (1) Eight layers: 5 convolutional layers and 3 fully-connected layers (the original)
- (2) Six layers: 5 convolutional layers and 1 fully-connected layer
- (3) Four layers: 3 convolutional layers and 1 fully-connected layer.

We find that the recognition accuracy drops as we reduce the number of CNN layers. The average accuracy is about 94% for 6-layer CNN and 90% for 4-layer CNN.

	A1	A2	B1	B2	C1	C2	D1	D2	E1	E2	F1	F2	G1	G2	H1	H2	I1	I2	J1	J2	K1	K2	L1	L2	
A1	200	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100.00%
A2	4	196	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98.00%
B1	0	0	162	27	8	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	81.00%
B2	0	0	14	186	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	93.00%
C1	0	0	4	0	180	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	90.00%
C2	0	0	0	0	17	183	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	91.50%
D1	0	0	0	0	0	0	196	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98.00%
D2	0	0	0	0	0	0	4	196	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98.00%
E1	0	0	0	0	0	0	0	0	193	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	96.50%
E2	0	0	0	0	0	0	0	0	11	189	0	0	0	0	0	0	0	0	0	0	0	0	0	0	94.50%
F1	0	0	0	0	0	0	0	0	0	0	200	0	0	0	0	0	0	0	0	0	0	0	0	0	100.00%
F2	0	0	0	0	0	0	0	0	0	0	4	196	0	0	0	0	0	0	0	0	0	0	0	0	98.00%
G1	0	0	0	0	0	0	0	0	0	0	0	0	196	4	0	0	0	0	0	0	0	0	0	0	98.00%
G2	0	0	0	0	0	0	0	0	0	0	0	0	14	186	0	0	0	0	0	0	0	0	0	0	93.00%
H1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	187	13	0	0	0	0	0	0	0	0	93.50%
H2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	177	0	0	0	0	8	1	0	0	88.50%
I1	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	188	9	0	0	0	0	0	0	94.00%
I2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	195	0	0	0	0	0	0	97.50%
J1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	199	1	0	0	0	0	99.50%
J2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	200	0	0	0	0	100.00%
K1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	198	0	2	0	99.00%
K2	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	198	0	0	99.00%
L1	0	0	4	0	2	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	189	4	0	94.50%
L2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	200	0	100.00%

FIGURE 5: Classification results for 24 easily-confused herbs using the naive CNN method. The ones marked in gray are classification errors. For example, 4 images of A2 are classified into A1.

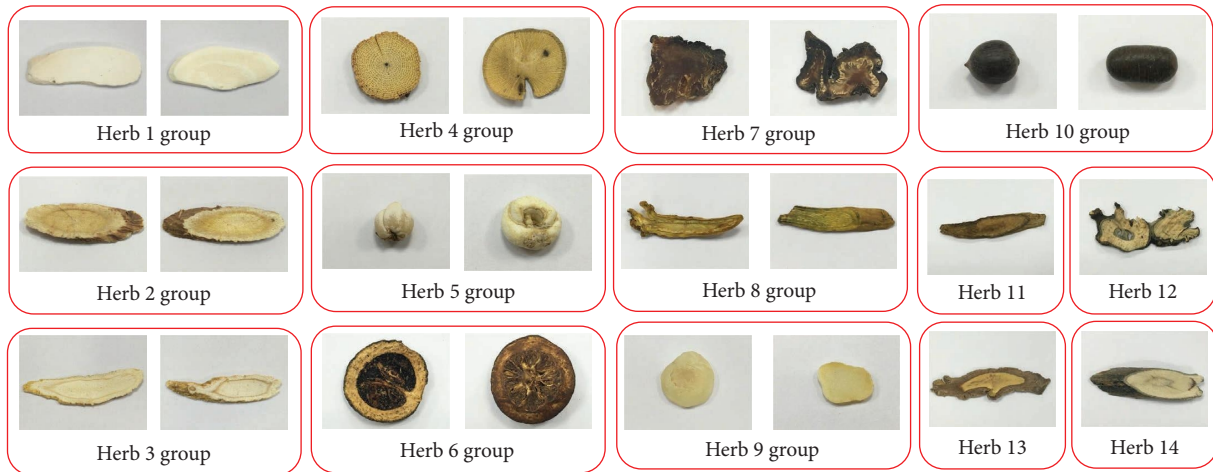


FIGURE 6: Manual clustering of 24 CHMs into 14 groups based on the illustrated handbook (group 1 to 10 are the set of easily-confused herbs, while groups 11, 12, 13, and 14 only contain a single herb).

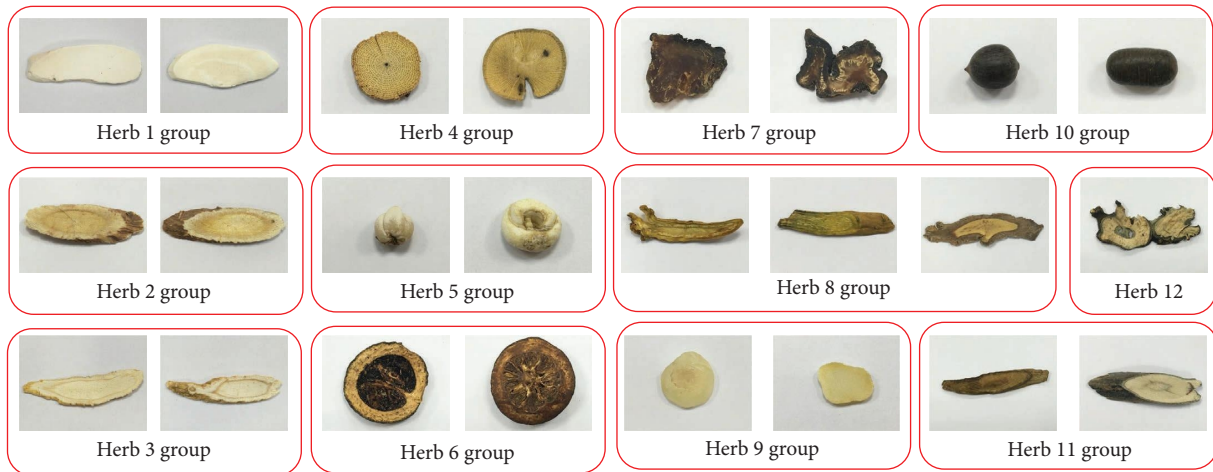


FIGURE 7: AP clusters 24 kind of CHMs into 12 groups (group 12 contains only a single herb).

	1	2	3	4	5	6	7	8	9	10	11	12
1	400	0	0	0	0	0	0	0	0	0	0	0
2	0	398	2	0	0	0	0	0	0	0	0	0
3	0	1	399	0	0	0	0	0	0	0	0	0
4	0	0	0	400	0	0	0	0	0	0	0	0
5	0	0	0	0	400	0	0	0	0	0	0	0
6	0	0	0	0	0	400	0	0	0	0	0	0
7	0	0	0	0	0	0	400	0	0	0	0	0
8	0	0	2	0	0	0	0	594	0	0	4	0
9	0	0	0	0	1	0	0	0	399	0	0	0
10	0	0	0	0	0	0	0	0	0	400	0	0
11	0	0	0	0	0	0	0	0	0	0	400	0
12	0	0	0	0	0	0	0	0	0	0	0	200

FIGURE 8: The confusion matrix for the clustering model using AP algorithm. The erroneous classification is marked in gray.

	A1	A2	B1	B2	C1	C2	D1	D2	E1	E2	F1	F2	G1	G2	H1	H2	I1	I2	J1	J2	K1	K2	L1	L2	
A1	200	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100.00%
A2	0	200	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100.00%
B1	0	0	184	14	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	92.00%
B2	0	0	10	190	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	95.00%
C1	0	0	1	0	189	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	94.50%
C2	0	0	0	0	14	186	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	93.00%
D1	0	0	0	0	0	0	198	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	99.00%
D2	0	0	0	0	0	0	1	199	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	99.50%
E1	0	0	0	0	0	0	0	0	197	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98.50%
E2	0	0	0	0	0	0	0	0	4	196	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98.00%
F1	0	0	0	0	0	0	0	0	0	0	200	0	0	0	0	0	0	0	0	0	0	0	0	0	100.00%
F2	0	0	0	0	0	0	0	0	0	2	198	0	0	0	0	0	0	0	0	0	0	0	0	0	99.00%
G1	0	0	0	0	0	0	0	0	0	0	0	0	198	2	0	0	0	0	0	0	0	0	0	0	99.00%
G2	0	0	0	0	0	0	0	0	0	0	0	0	5	195	0	0	0	0	0	0	0	0	0	0	97.50%
H1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	187	13	0	0	0	0	0	0	0	0	93.50%
H2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	188	0	0	0	0	0	0	0	0	94.00%
I1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	199	0	0	0	0	0	0	0	99.50%
I2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	196	0	0	0	0	0	0	0	98.00%
J1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	199	1	0	0	0	0	0	99.50%
J2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	200	0	0	0	0	100.00%
K1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	200	0	0	0	100.00%
K2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	199	0	0	99.50%
L1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	189	4	94.50%
L2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	200	0	100.00%

FIGURE 9: The confusion matrix for classifying the herb within a cluster using AP algorithm to cluster herbs. The one in gray indicates a classification error.

TABLE 6: The recognition accuracy of 10 results based on three automatic clustering algorithms.

	K-means (%)	Spectral clustering (%)	Affinity propagation (%)
1	96.83	96.29	97.65
2	96.21	95.79	97.48
3	97.38	96.21	97.85
4	96.85	95.25	97.08
5	96.96	95.48	97.85
6	96.10	95.58	97.88
7	96.15	95.25	97.48
8	96.33	96.13	97.23
9	95.77	96.08	97.65
10	96.85	95.63	97.23
Avg.	96.54	95.77	97.54

TABLE 7: The classification accuracy of CNN and the hierarchical CNN method for each herb.

	CNN (%)	HCNN by AP algorithm (average) (%)	HCNN based on manual clustering (%)
A1	100.0	100.0	100.0
A2	98.0	100.0	100.0
B1	81.0	90.8	92.0
B2	93.0	94.9	95.0
C1	90.0	93.2	94.5
C2	91.5	93.0	93.0
D1	98.0	99.0	99.0
D2	98.0	99.5	99.5
E1	96.5	98.5	98.5
E2	94.5	98.0	98.0
F1	100.0	100.0	100.0
F2	98.0	99.0	99.0
G1	98.0	98.9	99.0
G2	93.0	97.3	97.5
H1	93.5	93.5	94.5
H2	88.5	94.8	97.0
I1	94.5	98.9	99.5
I2	97.5	98.0	98.0
J1	99.5	99.6	99.5
J2	100.0	99.7	100.0
K1	99.0	99.7	99.5
K2	99.0	99.6	99.5
L1	94.5	95.3	96.0
L2	100.0	100.0	100.0
Avg.	95.65%	97.54	97.85

A significant improvement for some herbs such as B1 and H2 in addition to a general improvement of average accuracy (about 2%) when the HCNN method is employed.

Nevertheless, these results are still better than traditional methods using hand-crafted features.

4.7. The Effect of Different Brands of Smartphones. The camera parameters (e.g., resolution, image size, and color) of different smartphones can be quite different. Figure 10 shows the herb images taken by 4 different smartphones, including iPhone6, Samsung S7, Xiaomi, and Asus Pad-Phone. Figure 11 shows the color distributions from these phones. We can see that the images taken by iPhone are more similar to those from Samsung but quite different from images taken by Xiaomi and Asus phones. Therefore, if the training data are taken by iPhone and tested on other brands of phones, the recognition results could be poor, as shown in Table 8 (in this experiment, the training data and testing data were collected from different phones).

Data augmentation (DA) is a common way to improve the results of CNN by artificially creating more training data from the original dataset through various transformations of the original images. In this study, we implement four simple different data transformations on the original iPhone dataset, including rotation, resizing, and changes in brightness and histogram equalization, as shown in Figure 12.

Tables 9 and 10 show the performance of using the one single data augmentation (DA) method as well as combining multiple DA methods, respectively. By comparing Table 9 with Table 10, we can see that the data augmentation method is generally helpful to improve the recognition accuracy (up to about 9% for Asus phone) if we only have the training data

from one single type of phone (i.e., iPhone in our case). However, different types of transformations might have different effects on different phones. As shown in Table 9, providing additional data can improve the results for Xiaomi and Asus phones that does not help much for iPhone and Samsung S7. In particular, data augmentation through histogram equalization might reduce the recognition accuracy for iPhone and Samsung S7.

Finally, we consider adding additional training data from other smartphones. Specifically, 1,440 images were taken from each smartphone (60 for each herb) for training the CNN model. As shown in Table 11, introducing additional training data from all the smartphones obviously can produce a better improvement than the only use of data augmentation. However, given that it might not be feasible to collect the training data from all the smartphones in the world, data augmentation is still a good way to improve the performance of CNN. In addition, we currently are exploring the utilization of generative adversarial networks (GANs) [32] to generate synthetic herb images for data augmentation as an on-going work.

4.8. Visualization. In order to understand what features in the herb images our proposed hierarchical CNN model considers important for recognizing a herb, we employ the layer-wise relevance propagation (LRP) algorithm [33] to visualize which pixels in the input images contribute most strongly to the classification. LRP decomposes the output of the network into the sum of the relevance of the pixels in the

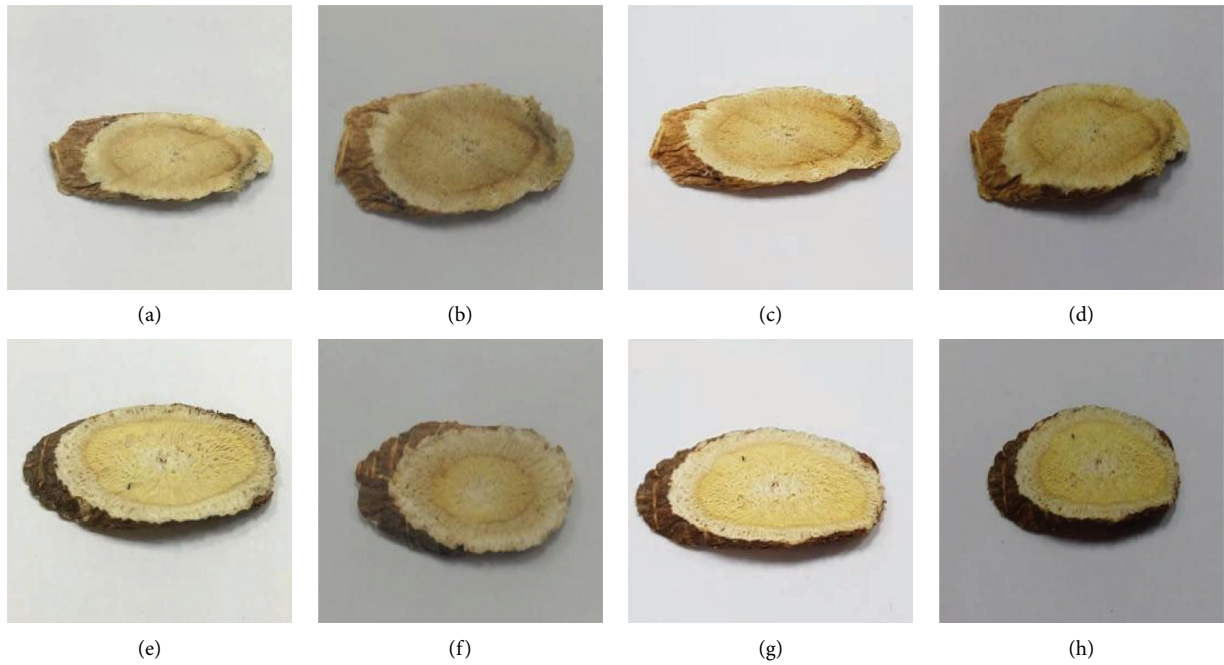


FIGURE 10: (a) and (e) are taken by iPhone6. (b) and (f) are taken by Xiaomi. (c) and (g) are taken by Samsung S7. (d) and (h) are taken by Asus PadPhone.

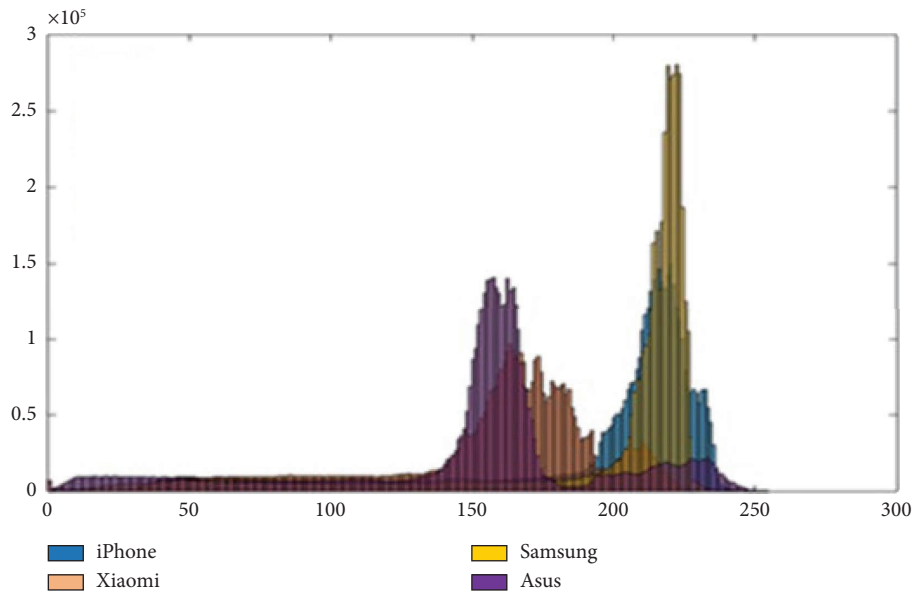


FIGURE 11: The color distribution of 4 smartphones. X axis represents the brightness of the pixel and y axis represents the number of pixels.

TABLE 8: The recognition results of 4 smartphones.

	iPhone (%)	Xiaomi (%)	Samsung (%)	Asus (%)
A1	100.0	100.0	100.0	100.0
A2	95.0	75.0	95.0	70.0
B1	85.0	95.0	85.0	90.0
B2	80.0	70.0	72.5	62.5
C1	92.5	62.5	100.0	70.0
C2	77.5	60.0	60.0	35.0
D1	95.0	82.5	100.0	87.5

TABLE 8: Continued.

	iPhone (%)	Xiaomi (%)	Samsung (%)	Asus (%)
D2	97.5	100.0	100.0	100.0
E1	97.5	87.5	67.5	100.0
E2	97.5	62.5	97.5	85.0
F1	90.0	85.0	100.0	97.5
F2	95.0	77.5	100.0	55.0
G1	92.5	100.0	95.0	100.0
G2	97.5	70.0	97.5	87.5
H1	92.5	65.0	100.0	50.0
H2	85.0	87.5	47.5	95.0
I1	80.0	100.0	97.5	82.5
I2	92.5	60.0	80.0	67.5
J1	100.0	90.0	100.0	77.5
J2	100.0	95.0	100.0	100.0
K1	100.0	87.5	100.0	100.0
K2	92.5	92.5	100.0	90.0
L1	87.5	92.5	87.5	87.5
L2	100.0	90.0	80.0	82.5
Avg.	92.60	82.81	90.10	81.77

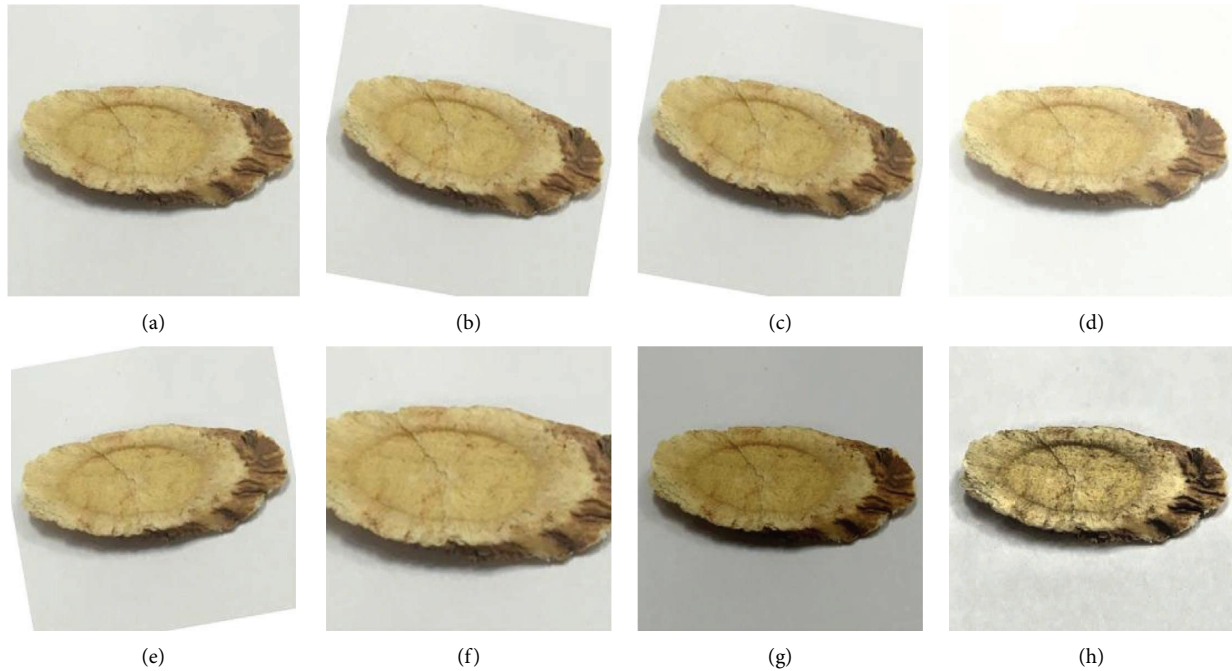


FIGURE 12: Our data augmentation method. (a) Original image, (b) rotate 10 degrees clockwise, (c) zoom-out, (d) increases the brightness, (e) histogram equalization, (f) rotate 10 degrees counterclockwise, (g) zoom-in, and (h) reduce the brightness.

TABLE 9: The result of using different data augmentation (DA) methods.

	iPhone (%)	Xiaomi (%)	Samsung (%)	Asus (%)	Avg. (%)
Without DA	92.60	82.81	90.10	81.77	86.82
(1) Rotation	92.81	84.48	88.13	84.27	87.42
(2) Size	94.58	86.25	89.17	84.79	88.70
(3) Brightness	91.88	87.92	90.31	89.17	89.82
(4) Histogram equalization	89.69	83.23	89.06	83.75	86.43

Data augmentation through histogram equalization might reduce the recognition accuracy for iPhone and Samsung S7.

TABLE 10: The result of combining multiple data augmentation methods.

	iPhone (%)	Xiaomi (%)	Samsung (%)	Asus (%)	Avg. (%)
DA(1)(2)(3)	94.48	88.96	91.02	90.52	91.24
DA(1)(2)(3)(4)	93.33	89.06	90.31	89.06	90.44

TABLE 11: The result of combining multiple data augmentation methods.

	iPhone (%)	Xiaomi (%)	Samsung (%)	Asus (%)	Avg. (%)
Without DA	94.06	93.02	95.31	93.85	94.06
DA(1)(2)(3)	96.04	95.83	96.25	94.90	95.76
DA(1)(2)(3)(4)	95.63	95.52	96.88	94.58	95.65

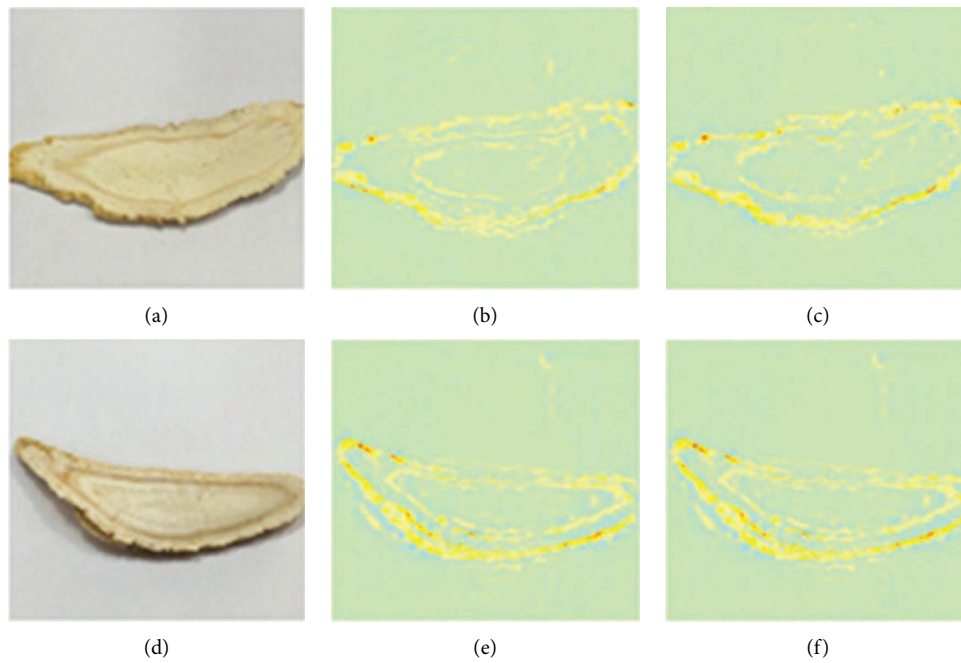


FIGURE 13: Visualization results using LRP for two easily confused herbs (C1 and C2) when CNN and hierarchical CNN are used. (a) Herb C1. (b) Herb C1 using CNN. (c) Herb C1 using hierarchical CNN. (d) Herb C2. (e) Herb C2 using CNN. (f) Herb C2 using hierarchical.

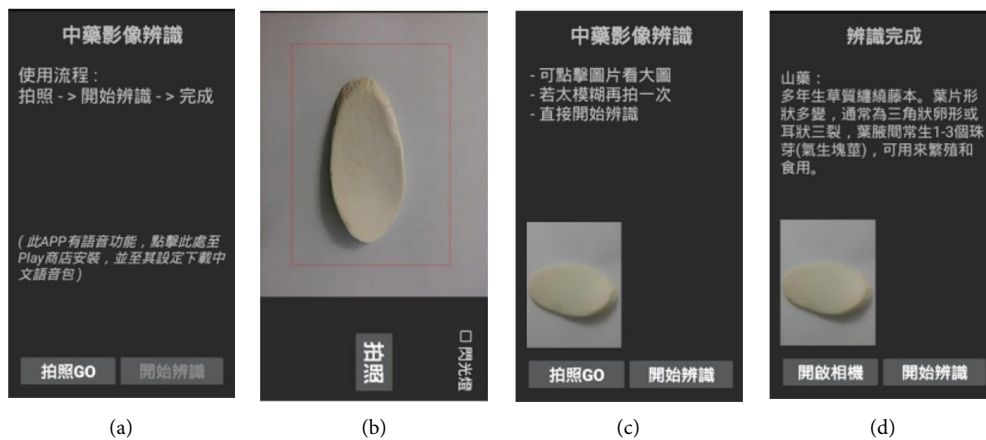


FIGURE 14: The user interface of the proposed system on smartphone. (a) The user instructions. (b) The camera screen. (c) and (d) Interface showing the recognition result and the corresponding information of the recognized herbs.

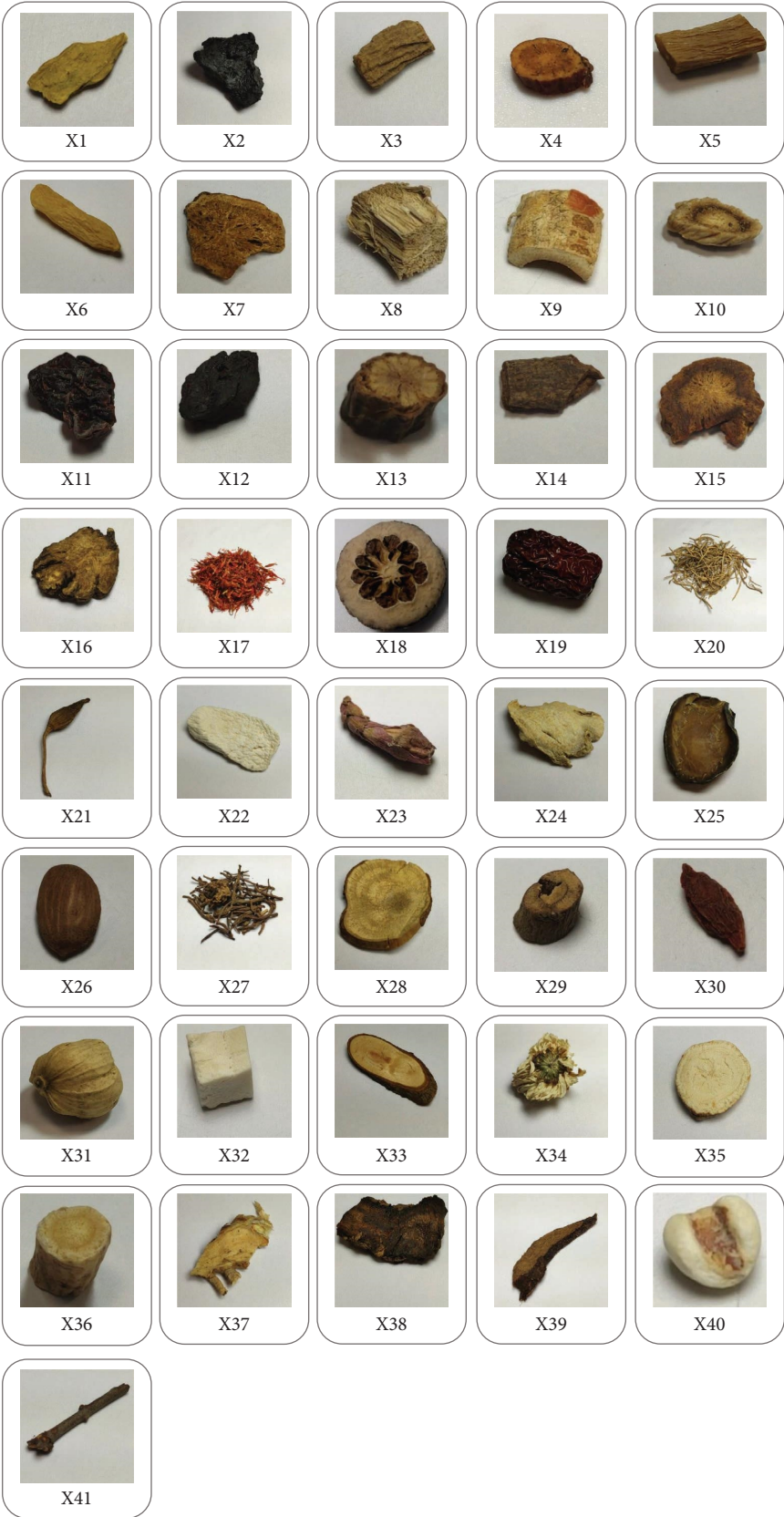


FIGURE 15: Images of additional 41 herbs.

TABLE 12: The name of additional 41 herbs.

Numbers	Names
X1	Scutellariae Radix
X2	Rehmanniae Radix Praeparata
X3	Codonopsis Radix
X4	Glycyrrhizae Radix et Rhizoma
X5	Achyranthis Bidentatae Radix
X6	Ophiopogonis Radix
X7	Atractylodis Macrocephalae Rhizoma
X8	Puerariae Radix
X9	Mori Cortex (桑白皮)
X10	Platycodonis Radix
X11	Corni Sarcocarpium
X12	Reynoutriae Multiflorae Radix
X13	Paeoniae Radix Rubra
X14	Cinnamomi Cortex
X15	Angelicae Sinensis Radix
X16	Chuanxiong Rhizoma
X17	Carthami Flos
X18	Aurantii Fructus Immaturus
X19	Jujubae Fructus
X20	Asari Radix
X21	Forsythiae Fructus
X22	Dioscoreae Rhizoma
X23	Farfae Flos
X24	Zingiberis Rhizoma
X25	Aconiti Lateralis Radix Praeparata
X26	Persicae semen
X27	Asteris Radix et Rhizoma
X28	Linderae Radix
X29	Moutan Radicis Cortex
X30	Lycii Fructus
X31	Myristicae semen
X32	Poria
X33	Cinnamomi Ramulus
X34	Chrysanthemi Flos
X35	Ginseng Radix et Rhizoma
X36	Astragali Radix
X37	Anemarrhenae Rhizoma
X38	Rehmanniae Radix
X39	Drynariae Rhizoma
X40	Coicis semen
X41	Taxilli Herba

input image. R_f is calculated by forward propagation; a pixel-wise relevance scores R_{il} is computed as follows:

$$R_i^l = \sum_j \frac{Z_{ij}}{\sum_{i'} Z_{i'j}} R_j^{l+1}, Z_{ij} = x_i^l w_{ij}^{l,l+1}, \quad (1)$$

where i is a neuron at layer l and P_j runs overall upper-layer neurons.

The results are shown as heatmaps in Figure 13. The pixels in yellow or red mean that they have higher LRP values (and the red pixel has a higher value than the yellow pixel) which are considered to have a strong influence on the classification results. Generally speaking, there are more circular layers (in yellow) in C2 than in C1 when either CNN or hierarchical CNN is used. These circular layers are slightly more observable when the proposed hierarchical CNN is used though.

5. Application

Based on the proposed hierarchical CNN, we implemented a smartphone App (currently only support Chinese) that can automatically recognize CHMs herbs, as shown in Figure 14.

To further validate the performance of our proposed system, in addition to the 24 easily confused herbs, we also collect another dataset which includes 41 herbs, as shown in Figure 15 and Table 12, and integrate these data into this herb recognition App system. We obtain similar results (around 98%) on the recognition accuracy for these additional herb data (the details are not discussed here due to the space limitation).

6. Discussion

In this work, we propose a system that can recognize easily confused TCM (traditional Chinese medicine) herbs on a smartphone with a high accuracy. As far as we know, this is the first smartphone-based system that considers recognition of easily- confused TCM herbs using deep learning techniques. Generally speaking, we observed that a deeper neural network performs better for herb recognition. In addition, we provide an explainable model to show what features in the herb image contribute most strongly to the final results of classification. We found that the recognition accuracy could be affected by the camera parameters (e.g., color histogram) of different brands of smartphones. Different data augmentation techniques were implemented to improve the system accuracy. Finally, we showed that the use of transfer learning is very beneficial where collecting large amount of herb data for training is difficult.

7. Conclusions and Future Work

In this work, we focus on the recognition of easily confused herbs by proposing a hierarchical clustering CNN method that uses affinity propagation to cluster similar herbs into groups. In each group, CNN is then used to extract representative features to distinguish similar herbs. As compared to CNN, our proposed method can improve the detection accuracy by almost 5%. In addition, we study the impact of different brands of smartphones on CHMs recognition accuracy. When the data augmentation is used with more data from different smartphones, we can improve the recognition accuracy from 86.82% to 95.76%. We are currently enriching our herb database so that our system can recognize more CHMs. In addition, we are exploring the use of generative adversarial networks (GANs) [32] to generate synthetic herb images for the data augmentation. Finally, in the future we plan to study the quality of CHMs by extending the App system we developed.

Abbreviations

- CHMs: Chinese herbal medicines
- TCM: Traditional Chinese medicine
- HCNNs: Hierarchical clustering convolutional neural networks

HCNN: Hierarchical clustering convolutional neural network
 VBM: Vision-based-measurement
 CNN: Convolutional neural network
 LBP: Local binary pattern
 HOG: Histogram of oriented gradients
 SIFT: Scale-invariant feature transform
 HD-CNN: Hierarchical deep CNN
 GTSRB: German traffic sign recognition benchmark
 CFC: CNN-oriented family clustering algorithm
 AP: Affinity propagation
 DA: Data augmentation
 GANs: Generative adversarial networks
 LRP: Layer-wise relevance propagation algorithm.

Data Availability

The data used to support in this study are available upon request from the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Authors' Contributions

Kun-chan Lan conceived and designed the experiments; Juei-Chun Weng, Jun-Xiang Zhang, and Tzu-Hao Tsai performed the experiments and analyzed the data; Min-Chun Hu helped on the interpretation of data; the acquisition of the herb data is provided by Yuan-Shiun Chang; both Kun-chan Lan and Tzu-Hao Tsai wrote the final version of this paper. All authors read and approved the final manuscript.

Acknowledgments

This research has received funding from Ministry of Science and Technology (MOST), Taiwan under the grant number 111-2221-E-006-120-.

References

- [1] X. Chengyong, W. Yuguo, F. Jian, Q. Li, X. Ran, and D. Yongqi, "Effect of optimal combination of Huangqi (Radix astragali Mongolici) and ezhu (rhizoma curcumae phaeocaulis) on proliferation and apoptosis of A549 lung cancer cells," *Journal of Traditional Chinese Medicine*, vol. 38, no. 3, pp. 351–358, 2018.
- [2] W. L. Zhang, R. C. Y. Choi, J. Y. X. Zhan et al., "Can hedysari Radix replace astragali Radix in danggui buxue tang, a Chinese herbal decoction for woman ailment?" *Phytomedicine*, vol. 20, no. 12, pp. 1076–1081, 2013.
- [3] M. J. Kim, H. Sung, and K. E. Hong, "Effects of dioscoreae rhizoma (SanYak) on peripheral neuropathy and its safety," *Journal of Pharmacopuncture*, vol. 16, no. 3, pp. 7–10, 2013.
- [4] E. Rivadeneyra-Domínguez and J. F. Rodríguez-Landa, "Preclinical and clinical research on the toxic and neurological effects of cassava (*Manihot esculenta* Crantz) consumption," *Metabolic Brain Disease*, vol. 35, no. 1, pp. 65–74, 2020.
- [5] Y. Yang, *Chinese Herbal Medicines: Comparisons And Characteristics*, Elsevier Health Sciences, London, UK, 2009.
- [6] S. Shirmohammadi and A. Ferrero, "Camera as the instrument: the rising trend of vision based measurement," *IEEE Instrumentation and Measurement Magazine*, vol. 17, no. 3, pp. 41–47, 2014.
- [7] O. Tao, Y. Zhang, Q. Chen, Y. Wang, and Y. Qiao, "Extraction of texture feature parameter of transverse section in Chinese herbal medicine by gray-level co-occurrence matrix," *World Science and Technology-Modernization of Traditional Chinese Medicine*, vol. 16, pp. 2531–2537, 2014.
- [8] O. Tao, Z. Lin, X. Zhang, Y. Wang, and Y. Qiao, "Research on identification model of Chinese herbal medicine by texture feature parameter of transverse section image," *World Science and Technology-Modernization of Traditional Chinese Medicine*, vol. 16, pp. 2558–2562, 2014.
- [9] G. Cai, Y. Zhao, H. Yuan et al., "Separation of Potentilla anserina active site (total saponin) and anti-DHBV DNA action in ducks," *Central South Pharmacy (China)*, vol. 1.
- [10] C. Liu, X. Wu, and W. Xiong, "Chinese herbal medicine classification based on BP neural network," *Journal of Software*, vol. 9, no. 4, pp. 938–944, 2014.
- [11] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *Science*, vol. 315, no. 5814, pp. 972–976, 2007.
- [12] D. Luo, D. Fan, H. Yu, and Z. Li, "A new processing technique for the identification of Chinese herbal medicine," in *Proceedings of the 2013 International Conference on Computational and Information Sciences*, pp. 474–477, IEEE, Shiyang, China, June 2013.
- [13] L. Dehan, W. Jia, C. Yimin, and G. Hamid, "Classification of Chinese Herbal medicines based on SVM," in *Proceedings of the 2014 International conference on information science, electronics and electrical engineering IEEE*, vol. 1, pp. 453–456, Sapporo, Japan, April 2014.
- [14] H. M. Zawbaa, M. Abbass, S. H. Basha, M. Hazman, and A. E. Hassenian, "An automatic flower classification approach using machine learning algorithms," in *Proceedings of the 2014 International conference on advances in computing, communications and informatics (ICACCI)*, pp. 895–901, IEEE, Delhi, India, September 2014.
- [15] M. A. Islam, M. S. I. Yousuf, and M. M. Billah, "Automatic plant detection using HOG and LBP features with SVM," *International Journal of Computer*, vol. 33, no. 1, pp. 26–38, 2019.
- [16] C. Narvekar and M. Rao, "Flower classification using CNN and transfer learning in CNN-Agriculture Perspective," in *Proceedings of the 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)*, pp. 660–664, IEEE, Thoothukudi, India, December 2020.
- [17] T. V. Janahiraman, L. K. Yee, C. S. Der, and H. Aris, "Leaf classification using local binary pattern and histogram of oriented gradients," in *Proceedings of the 2019 7th International Conference on Smart Computing & Communications (ICSCC)*, pp. 1–5, IEEE, Sarawak, Malaysia, June 2019.
- [18] P. Dhar, "A new flower classification system using LBP and SURF features," *International Journal of Image, Graphics and Signal Processing*, vol. 11, no. 5, pp. 13–20, 2019.
- [19] M. E. Pothen and M. L. Pai, "Detection of rice leaf diseases using image processing," in *Proceedings of the 2020 Fourth International Conference on Computing Methodologies and*

- Communication (ICCMC)*, pp. 424–430, IEEE, Erode, India, March 2020.
- [20] V. K. Jyothi, D. S. Guru, and Y. H. Kumar, “Classification of natural flower videos through sequential keyframe selection using SIFT and DCNN,” in *Proceedings of the International Conference on Recent Trends in Image Processing and Pattern Recognition*, pp. 305–318, Springer, Singapore, July 2018.
- [21] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [22] X. Sun and H. Qian, “Chinese herbal medicine image recognition and retrieval by convolutional neural network,” *PLoS One*, vol. 11, no. 6, Article ID 156327, 2016.
- [23] Z. Yan, H. Zhang, R. Piramuthu et al., “Hierarchical deep convolutional neural networks for large scale visual recognition,” in *Proceedings of the IEEE international conference on computer vision*, pp. 2740–2748, Cambridge, MA, USA, June 2015.
- [24] X. Mao, S. Hijazi, R. Casas, P. Kaul, R. Kumar, and C. Rowen, “Hierarchical CNN for traffic sign recognition,” in *Proceedings of the 2016 IEEE intelligent vehicles symposium (IV)*, pp. 30–135, IEEE, Gothenburg, Sweden, June 2016.
- [25] Y. Jia, E. Shelhamer, J. Donahue et al., “Caffe: convolutional architecture for fast feature embedding,” in *Proceedings of the 22nd ACM international conference on Multimedia*, pp. 675–678, New York, NY, USA, November 2014.
- [26] Y. Chang and Y. Ho, *Illustrations of Commonly Misused Chinese Crude Drug Species in Taiwan*, Ministry of Health and Welfare, Taipei, Taiwan, 1 edition, 2015.
- [27] C. C. Chang and C. J. Lin, “LIBSVM: a library for support vector machines,” *ACM transactions on intelligent systems and technology (TIST)*, vol. 2, no. 3, pp. 1–27, 2011.
- [28] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2014, <https://arxiv.org/abs/1409.1556>.
- [29] H. H. Bock, “Clustering methods: a history of k-means algorithms,” in *Selected contributions in data analysis and classification. Studies in Classification, Data Analysis, and Knowledge Organization*, P. Brito, G. Cucumel, P. Bertrand, and F. de Carvalho, Eds., Springer, Berlin, Heidelberg, pp. 161–172, 2007.
- [30] U. Von Luxburg, “A tutorial on spectral clustering,” *Statistics and Computing*, vol. 17, no. 4, pp. 395–416, 2007.
- [31] Y. Cheng, D. Wang, P. Zhou, and T. Zhang, “A survey of model compression and acceleration for deep neural networks,” 2017, <https://arxiv.org/abs/1710.09282>.
- [32] A. Aggarwal, M. Mittal, and G. Battineni, “Generative adversarial network: an overview of theory and applications,” *International Journal of Information Management Data Insights*, vol. 1, no. 1, pp. 100004–102021, 2021.
- [33] S. Bach, A. Binder, G. Montavon, F. Klauschen, K. R. Müller, and W. Samek, “On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation,” *PLoS One*, vol. 10, no. 7, Article ID 130140, 2015.