An Easy Method for Identifying 315 Categories of Commonly-Used Chinese Herbal Medicines Based on Automated Image Recognition Using AutoML Platforms

Wenting Chen (✉ wntingchen@163.com)

Guangzhou University of Chinese Medicine <https://orcid.org/0000-0001-7912-154X>

Jiayun Tong

Guangzhou University of Chinese Medicine

Rui He

Guangzhou University of Chinese Medicine

Peiting Chen

Guangzhou University of Chinese Medicine

Zixin Chen

Guangzhou University of Chinese Medicine

Xiaoli Liu

Guangzhou University of Chinese Medicine

Research

Keywords: Chinese herbal medicines (CHMs), Identification, Automated machine learning (AutoML), Image recognition, Huawei ModelArts, Baidu EasyDL

Posted Date: December 23rd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-132203/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Version of Record: A version of this preprint was published at Informatics in Medicine Unlocked on January 1st, 2021. See the published version at <https://doi.org/10.1016/j.imu.2021.100607>.

Abstract

Background: The identification and authentication of Chinese herbal medicines (CHMs) are directly related to their safety and efficacy in clinical treatment. However, the limited number of qualified professionals with expertise fails to meet the demand of the vast CHMs market. To make the CHMs identification more convenient and accurate, this study aimed at assessing the feasibility of the state-of-art automated machine learning (AutoML) technology in CHMs image recognition.

Methods: This study presented an experimental AutoML model built on the one-stop Huawei ModelArts platform instead of a handcrafted neural network. A rich and representative dataset of 31,460 images consisting of 315 categories of commonly-used CHMs was built and used for the model creation. Furthermore, the Huawei ModelArts model was compared with a model built on the Baidu EasyDL platform using the same dataset to investigate their ability to recognize CHMs images. Three professionals were also invited to recognize images of 315 categories of CHMs.

Results: During the model evaluation, high accuracies of 99.2% and 98.4% were achieved by ModelArts and EasyDL, respectively. In the subsequent held-out tests, the accuracies of ModelArts and EasyDL models were 91.2% and 91.85%, respectively. Both models performed very well individually and no statistically significant difference was found in model performance between these two platforms. However, the model-training time was only approximately 41 minutes on ModelArts platform but 118 minutes on EasyDL. The mean accuracy of the manual recognition for 315 CHMs was $97.46 \pm 1.58\%$.

Conclusion: Results revealed that AutoML technology is a fast and simple approach and has great practical potential in the field of CHMs image recognition. Since the Huawei ModelArts platform requires less training time, we recommend it as a priority.

Background

Chinese herbal medicines (CHMs) are important materials in the prevention and treatment of disorders or diseases in traditional Chinese medicine. Most CHMs originate from natural or cultured plants, with some from cultured animals and minerals, and the popularity of CHMs increases day by day globally [1, 2]. For most consumers with no professional identification knowledge, it is almost impossible to identify hundreds of different CHMs even commonly used in the treatment of diseases. This may create opportunities for some retailers to maximize profits by adding fake, non-official herb parts, or inferior plant materials [3]. Moreover, some CHMs are easily-confused even for professionals because of their similarities in terms of shapes, textures, and colors [4]. However, the identification and authentication of Chinese herbs are directly related to their safety, quality, and efficacy in clinical medication [5]. Adulterations and misidentification of CHMs can cause adverse effects or even fatality [6]. Therefore, the accurate identification of CHMs is crucial to regulating the chaotic CHMs market and controlling the quality of CHMs.

The identification methods of CHMs mainly include macroscopic identification, microscopic identification, physicochemical identification, and biological identification. The latter three methods often require sophisticated equipment only available in laboratories [7]. As a traditional method to identify CHMs, empirical macroscopic identification has been proved to effectively reflect the authenticity and quality of CHMs in thousands of years of practice and is widely adopted by the community pharmacies due to its feasibility [8]. Macroscopic identification is based on appearance characters such as shape, size, color, texture, smell, and taste. This method heavily relies on the knowledge of human professionals which is somehow subjective and varies among individuals. Thus, how to make the CHMs identification more feasible, objective, and accessible attracts more and more attention from practitioners.

Previous studies on CHMs identification based on image recognition were reported, and they mainly focused on building the CHMs image recognition model with handcrafted algorithms using low-level image features such as shape [9], color [10, 11], and texture [12]. However, these features are easily affected by image appearance and often require high-resolution images captured by sophisticated camera devices. With the rapid development of computer science, machine learning has become an excellent approach to perform image recognition tasks in a wide variety of fields such as clinical medicine [13], molecular biology [14], or plant recognition [15]. This cutting-edge technology has also been used on automated CHMs identification and has shown promising application prospects: CHMs identification Models built with machine learning algorithms such as convolutional neural networks and transfer learning can automatically learn representative features that may even be ignored by humans from a large amount of image data through operations such as convolution and pooling. These models have been proved to be robust and can achieve high precision [16–18].

While machine learning has shown great potential in CHMs identification, building an effective classification model remains a burdensome task. The model creation often requires immense resources: high capacity memory, strong graphics processing units, the help of professionals during the development, and long training time [19]. Even though pre-trained neural architectures such as AlexNet [20], VGGNet [21], and ResNet [22] can reduce the amount of labeled data needed to train a model, data labeling and model fine-tuning can still be very laborious and time-consuming, which greatly impedes the development of machine learning models in both academics and industry. To overcome these challenges, automated machine learning (AutoML) has emerged as a new sub-area in machine learning. With automated model selection [23], neural architecture search [24], and feature engineering [25], AutoML not only simplifies the creation and application of machine learning models but also greatly reduces the turnaround time and improves the accuracy of the customized models by removing human errors [26]. Nowadays, there are some products of AutoML available in the industry, which provide end-to-end AutoML pipelines to reduce user intervention during model development: the user simply provides data, and the AutoML system automatically determines the approach performed best for this particular application [23]. This technology has been shown to produce encouraging results in various studies with only small numbers of images [27, 28]. With the help of AutoML, researchers can focus more on solving problems with more application and business value.

In this study, to make CHMs identification more convenient and accurate, experimental machine learning models were created and evaluated using state-of-art AutoML platforms available on the internet instead of handcrafted machine learning algorithms. Given the subsequent applications potential, the handy device— smartphone was used to capture images for AutoML model creation. Thus, this work will greatly lower the barrier of CHMs identification to allow ordinary people to identify different CHMs in their daily life, and it possesses great potential in commercial applications.

Materials And Methods

Materials collection

Most of the CHMs decoction pieces used in this study were commercial samples purchased from CHMs markets, a small number of samples were collected from the field. All of these materials were authenticated by Haibo Huang (Prof.) and Jiayun Tong (Ph.D.).

Dataset Construction

In previous studies on the CHMs identification, recognition models were generally trained using images with a single slice of Chinese herb on a clear background [12], or multiple slices heaped together on a cluttered background [16]. The latter often contained unrelated information that could seriously downgrade the prediction accuracy, even more, introduce bias. Actually, there are hundreds of commonly-used CHMs in the market, which mainly come from different parts of plants, with some from animals and minerals. And before they were applied in the disease treatment, all the CHMs have to go through different processing procedures which further diversify the characters (shapes, colors, and textures) within the same category, or assimilate characters of different categories. However, the quality and representativeness of the dataset play a decisive role in the AutoML model creation [29]. That it is, a rich and representative dataset is definitely needed to establish a high-performance CHMs identification model. Therefore, all the images used in this study were captured under a clear background with a single slice or non-overlapping multiple slices of the CHMs placed on light and untextured background without clutter. This enabled us to combine our professional knowledge to guide the algorithm to find more relevant features from the CHMs images, thus, to learn more like a human. To fully investigate the feasibility of this method, an image dataset with 315 categories of commonly-used CHMs (listed in Additional file 1: Table S1) was established, which contains many images of easily-confused CHMs. Different types of these easily-confused CHMs were summarized as follows:

1. Adulterants (Fig. 1a1-5): fake and genuine CHMs are often mixed in the market and hard to distinguish because of their highly similar characters, such as *Ziziphi Spinosae Semen* (Fig. 1a1-3) and its counterfeit *Ziziphi Mauritanae Semen* (Fig. 1a4-5).
2. CHMs with similar colors (Fig. 1b1-d5): some herbs are similar in colors, such as *Scrophulariae Radix* (Fig. 1d3) and *Rehmanniae Glutinosae Radix* (Fig. 1d1), which are easily confused because of their black color after going through similar processing procedures. Some CHMs from minerals or animals with indistinctive characters can also be very hard to distinguish (Fig. 1c1-5).

3. CHMs originated from closely related plants (Fig. 1e1-5): These plants are highly similar in morphology, the medicinal parts of each plant become even more difficult to identify after being chopped into slices, such as CHMs from the genus *Ardisia* (Fig. 1e1-5). The transverse sections of the different CHMs slices often show similar color in the bark or wood and have scattered or radial dots due to the secretory cavities.
4. CHMs applied in whole or aerial parts of the plants (Fig. 1f1-5): this type of CHMs often contains different parts of the plants—roots, stems, leaves, flowers, and fruits, which often mix and require careful examination.
5. CHMs small in size (Fig. 1g1-5): compared with other types of CHMs, this type of CHMs consists of seeds and fruits that are small or tiny in size (diameter < 5 mm).
6. CHMs using the same medicinal part — bark (Fig. 1h1-5): they often have similar appearance and texture: flat or curved in shape, the outer surface with or without scars, and the fracture surface granular or fibrous.

Inspired by Weng's study [17], all the images used in this study were collected by our team with a smartphone camera. For some CHMs like small seeds or fruits (Fig. 1g1-5), a micro-lens was equipped on the top of the mobile phone camera lens to obtain high-resolution images. After eliminating low-quality and repetitive images, a total number of 31,460 CHMs images were collected for dataset construction, with about 100 images (varies from 94 to 105) in an exclusive folder for each category.

Dataset Pre-processing And Split

After the images were collected, the length-width ratio of each image was resized to 1:1. To reduce the training time of models, the resolution of each image was downsampled from 3024 × 3024 pixels or 3456 × 3456 pixels to 850 × 850 pixels.

The dataset was split into different subsets according to the hierarchical structure as shown in Fig. 2. Firstly, the original dataset was split into two subsets: one for the model building and another for the subsequent held-out test. Then, the model-building dataset was further split into a training set and an internal validation/testing set by AutoML platforms as described in the next subsection. The held-out dataset included 6260 images with a mixture of all categories, which were extracted from the original dataset and avoided exposure to the AutoML models earlier than the held-out test.

Building Image Recognition Models With Automl Platforms

Huawei ModelArts [30] provided by Huawei Cloud (Huawei, Shenzhen, China) was chosen to build our CHMs image recognition model. Its main features include:

- ModelArts ExeML provides a customized, code-free development platform for beginners with no coding knowledge, and it can help users to build a customized, high-precision model quickly and flexibly.
- Apply multiple pre-trained models and self-developed deep learning framework to build models that can achieve excellent performance using a small amount of data.

To build our CHMs image recognition models, the following steps were taken according to the website tutorials [31]:

- Entered the ModelArts platform, created an OBS bucket and uploaded the images for the model-building to the bucket.
- Created an image classification model and imported all the training data from the OBS bucket.
- Set the parameters as default on the Model Configuration user interface (UI): Training Set Ratio (0.8) and Validation Set Ratio (0.2), Max Inference Time (300 milliseconds), Max Training Time (1 hour).

After the training job was submitted, ModelArts automatically searched for the best algorithm, neural architecture, and hyperparameters based on the training dataset.

For comparison, Baidu EasyDL [32] provided by Baidu Brain (Baidu, Beijing, China), a platform similar to Huawei ModelArts, which also provides a user-friendly interface and code-free development, was also chosen to build an image recognition model with the same training dataset.

To build our image recognition model with EasyDL, the following steps were taken for the data preparation and model configuration:

- Entered the classic EasyDL platform and created an image classification model.
- Created a new dataset for model building on the platform's data center and uploaded the images in the form of a zip file.
- On the Model Training UI:
 - Selected Public Cloud application programming interface (API) as deployment so that we can call the API to use the batch Services for the subsequent held-out test.
 - Selected the AutoDL Transfer as the training algorithm which is more suitable for fine classification scenarios like CHMs classification in this study.
- Started the training and the dataset created for the model-building is automatically divided into a training set (70%) and a test set (30%) during the process.

Held-out Test

A subsequent held-out test was conducted with the held-out dataset after each model was built. The numbers of images used to create these two models and the held-out test were summarized in Table 1. The held-out dataset was created with 6260 images that were avoided exposure to the AutoML models

during model-training. Due to the free calls limitation of the API provided by the EasyDL platform (1000 times), 945 pictures randomly extracted from the same held-out dataset were used to evaluate the EasyDL model. The schematic representation of the AutoML model creation is depicted in Fig. 3. On the ModelArts platform, the test was carried out in a batch mode by deploying the model and using the online batch prediction services. On the EasyDL platform, the test was carried out by deploying the model to the Public Cloud, publishing the model as API, and calling the API to use batch prediction services.

Table 1
Detailed information on model building and held-out test

Model	Number of Categories	Number of Images (Model Building)	Training Set Ratio	Validation Set Ratio	Number of Images (Held-out test)
ModelArts	315	25200	0.8	0.2	6260
EasyDL	315	25200	0.7	0.3	945

Classification Performance Measures

The goal of a machine learning algorithm is to learn from training data and predict class labels for testing data. Therefore, the assessment method is a key factor in evaluating the classification performance and guiding classifier modeling. In this study, four measures, namely the accuracy, the precision, the recall, and the F1-score were chosen to evaluate the models' performance. For overall model evaluation, the accuracy is one of the most commonly used measures for evaluating classification performance, and it is defined as a ratio between the correctly classified samples number to the total number of samples. For each category, the precision represents the proportion of positive samples that were correctly classified to the total number of positive predicted samples; the recall represents the positive correctly classified samples to the total number of positive samples; the F1-score represents the harmonic mean of the precision and the recall. The F1-score value ranges from zero to one, and high values of the F1-score indicate high classification performance [33, 34].

Manual Prediction

To investigate the similarities and differences between the AutoML model recognition and manual identification, three professionals were invited to identify each CHM category using the same images from the held-out dataset.

Results

Models evaluation and held-out test

Two CHMs image recognition models were built with the same model-building dataset on ModelArts and EasyDL, respectively. After the model was built, each platform presented an evaluation report based on predicted results of the Validation Set (ModelArts) or Testing set (EasyDL). Performances of two models

in terms of accuracy, F1-score, precision, and recall were summarized in Table 2. During the model evaluation, high accuracies of 99.2% and 98.4% were achieved by ModelArts and EasyDL, respectively. In the held-out tests, the accuracies of the ModelArts and the EasyDL Models were 91.21% and 91.85%, respectively. Figure 4 shows some examples of CHMs categories correctly recognized by the ModelArts model.

Table 2
Comparison of model performance between ModelArts and EasyDL

Model	Accuracy (%)	F1-score (%)	Precision (%)	Recall (%)
ModelArts (evaluation)	99.20	99.19	99.23	99.21
EasyDL (evaluation)	98.40	98.40	98.50	98.50
ModelArts (held-out)	91.21	89.88	90.91	91.19
EasyDL (held-out)	91.85	90.35	90.44	91.85

To understand how well the AutoML algorithm performs in recognizing different categories of easily-confused CHMs mentioned in Sect. 2.1, the average F1-score, precision, and recall for each category of CHMs in the ModelArts held-out test were summarized in Table 3. Results showed that the ModelArts model performed generally well in recognizing most of the easily-confused CHMs with the average precision surpassing 0.81. To visualize the performance of the ModelArts model in recognizing easily-confused CHMs, the F1-score values of both model evaluation and held-out test of each category were plotted into a heatmap (Fig. 5a). These results showed that the model performed well in recognizing most of the easily-confused CHMs (Fig. 5a-g). Surprisingly, some CHMs categories got very high F1-score values, for example, Astragali Complanati Semen, Vaccariae Semen, Trichosanthis Radix, Rehmanniae Radix Praeparata, Psoraleae Fructus, Cannabis Fructus, Schisandrae Chinensis Fructus, Mume Fructus, and Moutan Cortex got F1-scores of 1.0 in both evaluation and held-out test. However, some CHMs originated from the bark of plants such as Periplocae Cortex and Lycii Cortex only obtained F1-scores around 0.5 (Fig. 5h), which is relatively low. And there were a few CHMs categories misidentified by the ModelArts model (some examples shown in Fig. 6).

Table 3
The average F1-score, precision, and recall of different types of CHMs

Types of CHMs	Number of Categories	F1-score	Precision	Recall
Genuine and Adulterant	2	0.9750	0.9762	0.9750
White color	16	0.9236	0.9042	0.9563
Black color	5	0.9846	0.9905	0.9800
Whole plant or aerial parts	30	0.9276	0.9194	0.9533
From relative plants	5	0.9527	0.9476	0.9661
Bark	14	0.7777	0.8109	0.7821
Small seeds and fruits	10	0.9749	0.9545	1.0000

To further investigate the performance differences between the ModelArts and the EasyDL models, the computation time of the models was calculated by subtracting the time of the start from the end of model training, and a two-tailed distribution, paired t-test was performed based on the held-out test results. No significant difference in models' performance was found. However, it cost the ModelArts approximately only 41 minutes to train the classification model with default computational power, while almost two times longer (118 minutes) as it for the EasyDL.

AutoML recognition vs. manual identification

To better understand the differences between the AutoML recognition model and identification by humans, three professionals were invited to identify all the 315 CHMs categories using the same images from the held-out dataset. In the held-out test of ModelArts, 1.90% (six categories) of CHMs: the images of Cornu Cervi Pantotrichum, Lasiosphaera Calvatia, Arisaematis Rhizoma, Cirsii Herba, Polygonati Odorati Rhizoma, and Arisaematis Rhizoma Preparatum were completely misrecognized with the precision of zero. In contrast, the false recognition rate of the manual prediction for 315 CHMs was $2.54 \pm 1.58\%$ (the mean accuracy of $97.46 \pm 1.58\%$).

Discussion

In this study, an automated identification method for commonly-used CHMs was established. Firstly, a representative CHMs image dataset with more than 300 categories of CHMs has been constructed. Then, classification models on one-stop AutoML development platforms—Huawei ModelArts and Baidu EasyDL were built. Overall, both models performed well in recognizing different CHMs images with the accuracy

surpassing 98% in the evaluation and 91% in the held-out test. Results showed that although the CHMs slices within each category varied in shapes or colors, the ModelArts model can successfully recognize them after trained using images containing key features and enough details of CHMs for identification (examples in Fig. 4). The difference between the accuracy of the model evaluation and the held-out test also indicates that the overfitting problem exists in both models. Thus, carrying out a following held-out test is necessary.

Compared with other machine learning techniques, AutoML technology provided by ModelArts or EasyDL platform requires zero programming knowledge and offers a user-friendly interface on web-based applications. For the EasyDL platform, it possesses a unique significance in that, by using the data augmentation strategy on this platform, the dataset can be augmented by altering the appearance of images, and the model training effect can be enhanced. For example, images can be cropped, rotated, blurred, and flipped to optimize the model's training abilities to recognize the test images. Since the quality of the images captured by the smartphone camera is easily affected by different environmental factors such as lighting conditions, camera position, and different camera parameters, this strategy can make the model more robust under different scenarios [17]. But the detailed information including the total number of images after augmentation is not provided by the platform. As this feature is not available on the ModelArts platform, and to simplify the experimental operation, the augmentation strategy on EasyDL was set as default, that is to say, such strategy did not use to artificially enlarge the dataset.

It is also important to note that both platforms charge for different services such as cloud storage, model training, or model deployment but offer free services for a limited time during the developing period [35, 36]. For model deployment, both platforms allow users to deploy the model for serving in several ways: real-time, batch, or edge services. The real-time services allow users to deploy a model as a web service to provide real-time test UI and monitoring capabilities. The batch services can perform inference on a batch mode that allows high-throughput prediction. The Edge services provide users a complete edge computing solution, in which cloud applications are extended to the edge. By leveraging edge-cloud synergy, users can manage applications remotely and process data nearby. In this study, the batch services were applied to conduct the held-out test, in which the EasyDL gives limited 1000 times free API calls to users for batch prediction, while the ModelArts provides online batch service free in a limited time (1 hour).

Although based on our analysis we determined there was no significant difference between the two AutoML platforms in terms of model performance, the ModelArts model was set as the baseline model during our research for the following reasons: firstly, the time to train the ModelArts model is far shorter than the EasyDL model with the default computational setting. Although it is possible to decrease the model-training time of EasyDL by setting a different algorithm provided by the platform, this might sacrifice the precision of the model simultaneously. Secondly, as the free API calls provided by the EasyDL platform are only 1000 times, only 945 images in total (three images per category) can be uploaded for the held-out test of the EasyDL. Therefore, with far more images (6260) in the held-out

dataset, the ModelArts model was set as the baseline model and the total cost for the experimental AutoML model was \$0.

The performances of the experimental AutoML models in this study and the models in the previous studies were compared in Table 4 [16–18]. Results showed that both Weng’s and Wu’s models achieved high accuracies, which were 96% and 97%, respectively. However, both models were built with only 11 categories of CHMs. In another study, Sun’s model built with 95 categories of CHMs achieved an average recognition precision of only 71% and all of the images used for model-building were downloaded from Google. The reason may be that each image contains multiple slices of CHMs heaped up together under complex background, which means those images often contain unrelated information that could seriously downgrade the prediction accuracy, even more, introduce bias [37]. By contrast, the models in this study are more promising in automated CHMs identification with a high accuracy performance using the AutoML algorithm instead of handcrafted machine learning algorithms. The method proposed in this study can greatly simplify the operation as well as reduce the time to develop a CHMs recognition model and improve its performance.

Table 4
Comparison of accuracies among various models

Author	Model	Number of Categories	Accuracy (%)
This paper	ModelArts	315	99
(evaluation)			
This paper	EasyDL	315	98
(evaluation)			
This paper (held-out)	ModelArts	315	91
This paper (held-out)	EasyDL	315	92
Sun Xin[16]	VGG16-Net	95	71
Weng JC[17]	CafeNet	11	96
Wu Chong[18]	ResNet50	11	97

Results showed that the ModelArts model can correctly recognize most of the CHMs images that are even confusing for human professionals. However, it still has certain limitations in differentiating some highly similar CHMs images such as Periplocae Cortex and Lycii Cortex. This is not surprising because they (Fig. 1g1-2) look very similar in the images. Interestingly, we have also found that the AutoML models misrecognized some CHMs (applied using different parts of plants) images that are distinct and not difficult to identify by human professionals with material objects (Fig. 6b, h). As AutoML solutions are

generally black boxes [38], we speculate that this might be due to the difference between the ways AutoML and humans learn to recognize different CHMs. As humans, we often combine multiple senses during the CHMs identification process and utilize different information such as the size, weight, smell, and even taste of different CHMs important features to identify them. However, AutoML learned mainly through the images we presented which only contain visual information and are limited in scope. Different CHMs may look highly similar when captured by the camera at certain angles. Sometimes there were only slice fragments in the images, which makes them harder to distinguish (Fig. 6). Eventually, results showed that the identification accuracy of the AutoML algorithm was close to that of three professionals. And both the AutoML algorithm and the professionals failed to identify some visually-confusing CHMs images. This means human professionals also found it hard to differentiate some CHMs images with little difference (Fig. 6). This is understandable since even some CHMs materials are easily confused in real life.

Conclusions

In this study, an experimental AutoML model for automated CHMs identification was built and evaluated using Huawei ModelArts. With high accuracies of 99.2% and 91.2% from the model evaluation and the held-out test, respectively, a conclusion can be drawn that this CHMs image recognition model has successfully learned to recognize more than 300 different CHMs. By providing user-friendly UI and flexible deployments, the current state-of-art AutoML technology makes the development of the CHMs recognition model simpler and faster. Thus, this work has the potential to greatly lower the cost of commercial applications in the CHMs recognition field and the barrier of CHMs recognition with flexible deployments offered by the AutoML platforms. However, significant challenges still exist in the CHMs image recognition with AutoML, particularly in the recognition of easily-confused CHMs that are highly similar in images or even their material objects. To further improve the model performance, our team will enrich the training dataset by capturing images of CHMs from different angles to obtain more detailed features of each herb and applying the data augmentation strategy to make the model more robust.

List Of Abbreviations

CHMs: Chinese herbal medicines; AutoML: Automated machine learning; UI: User interface; API: Application programming interface.

Declarations

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Fund for Improving Medical Service and Security Capacity of the Department of Social Security from Ministry of Finance (Grant No. (2019)39), and the Xinhua Project from Guangzhou University of Chinese Medicine (Grant No. XH20170107).

Authors' contributions

JYT and WTC conceived and designed this study and collected herbal materials. WTC, PTC, ZXC, and XLL collected the image data. WTC created the experimental models and performed data analysis. WTC wrote the manuscript. JYT monitored the conduct of this study. JYT and RH revised the manuscript. All authors have read and approved the final manuscript.

Acknowledgements

We would like to thank Mr. Shuoyu Xu (Ph.D.) from the BIO-TOTEM technology Co., Ltd and Mr. Guokai Huang (Licensed Pharmacist) from Guangdong Institute for Drug Control for their professional advice, also Mrs. Zhidong Xu from School of Pharmaceutical Sciences, Guangzhou University of Chinese Medicine for the help in the collection of herbal materials.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Availability of data and materials

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

References

1. Eisenberg DM, Davis RB, Ettner SL, Appel S, Wilkey S, Van Rompay M, et al. Trends in alternative medicine use in the United States, 1990-1997: results of a follow-up national survey. *Jama*. 1998;280(18):1569-75.
2. Ernst E, White A. The BBC survey of complementary medicine use in the UK. *Complementary therapies in medicine*. 2000;8(1):32-6.
3. Zhang J, Wider B, Shang H, Li X, Ernst E. Quality of herbal medicines: Challenges and solutions. *Complementary Therapies in Medicine*. 2012;20(1-2):100-6.
4. Ming J, Chen L, Cao Y, Yu C, Huang B, Chen K. Rapid Identification of Nine Easily Confused Mineral Traditional Chinese Medicines Using Raman Spectroscopy Based on Support Vector Machine. *Journal of Spectroscopy*. 2019. <https://doi.org/10.1155/2019/6967984>

5. Liu S, Chuang W, Lam W, Jiang Z, Cheng Y. Safety surveillance of traditional Chinese medicine: current and future. *Drug safety*. 2015;38(2):117-28.
6. Ernst E. Adulteration of Chinese herbal medicines with synthetic drugs: a systematic review. *J Intern Med*. 2002;252(2):107-13.
7. Chen K, Huang L, Liu Y. Development history of methodology of Chinese Medicines' Authentication. *China Journal of Chinese Materia Medica*. 2014;39(7):1203.
8. Leong F, Hua X, Wang M, Chen T, Song Y, Tu P, et al. Quality standard of traditional Chinese medicines: comparison between European Pharmacopoeia and Chinese Pharmacopoeia and recent advances. *Chinese Medicine*. 2020;15(1):1-20.
9. Zhu L, Li X, Zhang Y, Pu H, Wu X. Image retrieval method for Chinese herbal medicine based on shape features and texture features. *Computer Engineering and Design*. 2014;35(11):3903-7.
10. Chen S, Lu W, Wang F. Image Recognition of Chinese Herbal Pieces Based on Color Matching Template. *Chinese Journal of Experimental Traditional Medical Formulae*. 2020;26(6):158-62.
11. Xie S, Zhou M, Zhou J. Research of OpenCV-based Recognition System of Herbal Decoction Pieces. *Lishizhen Medicine and Materia Medica Research*. 2018;29(2):510-2.
12. Tao Q, Lin Z, Zhang X, Wang Y, Qiao Y. Research on Identification Model of Chinese Herbal Medicine by Texture Feature Parameter of Transverse Section Image. *Modernization of Traditional Chinese Medicine and Materia Medica-World Science and Technology*. 2014;16(12):2558-62.
13. Wang H, Roa AC, Basavanahally AN, Gilmore HL, Shih N, Feldman M, et al. Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features. *Journal of Medical Imaging*. 2014;1(3):034003.
14. Shao Y, Chou K. pLoc_Deep-mVirus: A CNN Model for Predicting Subcellular Localization of Virus Proteins by Deep Learning. *Natural Science*. 2020;12(6):388-99.
15. Lee SH, Chan CS, Wilkin P, Remagnino P, editors. Deep-plant: Plant identification with convolutional neural networks. 2015 IEEE international conference on image processing (ICIP); 2015: IEEE. doi:10.1109/ICIP.2015.7350839.
16. Sun X, Qian H. Chinese Herbal Medicine Image Recognition and Retrieval by Convolutional Neural Network. *PLoS One*. 2016;11(6):e0156327.
17. Weng JC, Hu MC, Lan KC, editors. Recognition of Easily-confused TCM Herbs Using Deep Learning. the 8th ACM; 2017. <https://doi.org/10.1145/3083187.3083226>
18. Wu C, WU C, Huang Y, Wu C, Chen H. Intelligent Identification of Fritillariae Cirrhosae Bulbus, Crataegi Fructus and Pinelliae Rhizoma Based on Deep Learning Algorithms. *Chinese Journal of Experimental Traditional Medical Formulae*. 2020;26(21):195-201.
19. Chauhan K, Jani S, Thakkar D, Dave R, Bhatia J, Tanwar S, et al., editors. Automated Machine Learning: The New Wave of Machine Learning. 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA); 2020: IEEE. doi:10.1109/ICIMIA48430.2020.9074859.

20. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*. 2017;60(6):84-90.
21. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 3rd International Conference on Learning Representations (ICLR)2015. <http://arxiv.org/abs/1409.1556> Accessed 10 Apr 2015.
22. He K, Zhang X, Ren S, Sun J, editors. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2016. doi: 10.1109/CVPR.2016.90.
23. Hutter F, Kotthoff L, Vanschoren J. *Automated machine learning: methods, systems, challenges*. 1rd ed. Springer Nature; 2019.
24. Zoph B, Le QV. Neural architecture search with reinforcement learning. *International Conference on Learning Representations (ICLR)2017*. <https://arxiv.org/abs/1611.01578>. Accessed 15 Feb 2017.
25. Katz G, Shin ECR, Song D. Exploreskit: Automatic feature generation and selection. 2016 IEEE 16th International Conference on Data Mining (ICDM); 2016: IEEE. doi: 10.1109/ICDM.2016.0123.
26. Truong A, Walters A, Goodsitt J, Hines K, Bruss CB, Farivar R. Towards automated machine learning: Evaluation and comparison of automl approaches and tools. 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI); 2019: IEEE. doi: 10.1109/ICTAI.2019.00209.
27. Zeng Y, Zhang J. A machine learning model for detecting invasive ductal carcinoma with Google Cloud AutoML Vision. *Comput Biol Med*. 2020;122:103861.
28. Sawaki R, Sato D, Nakayama H, Nakagawa Y, Shimada Y. ZF-AutoML: An Easy Machine-Learning-Based Method to Detect Anomalies in Fluorescent-Labelled Zebrafish. *Inventions*. 2019;4(4):72.
29. Luo C, Li X, Wang L, He J, Li D, Zhou J. How does the data set affect cnn-based image classification performance? 2018 5th International Conference on Systems and Informatics (ICSAI); 2018: IEEE. doi: 10.1109/ICSAI.2018.8599448
30. HUAWEI CLOUD. ModelArts. <https://www.huaweicloud.com/en-us/product/modelarts.html>. Accessed 01 July 2020.
31. HUAWEI CLOUD. Introduction to ExeML. https://support.huaweicloud.com/intl/en-us/exemlug-modelarts/modelarts_21_0001.html. Accessed 01 Jul 2020.
32. Baidu Brain. EasyDL. <https://ai.baidu.com/easydl/>. Accessed 16 Sep 2020.
33. Sokolova M, Japkowicz N, Szpakowicz S, editors. Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. *Australasian joint conference on artificial intelligence*; 2006: Springer.
34. Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. *Information processing & management*. 2009;45(4):427-37.
35. Baidu Brain. EasyDL Pricing. <https://ai.baidu.com/ai-doc/EASYDL/8k4w58167>. Accessed 16 Sep 2020.
36. HUAWEI CLOUD. HUAWEI CLOUD Pricing. <https://www.huaweicloud.com/intl/en-us/product/price.html>. Accessed 01 Jul 2020.

37. Zhu Y, Sun W, Cao X, Wang C, Wu D, Yang Y, et al. TA-CNN: Two-way attention models in deep convolutional neural network for plant recognition. *Neurocomputing*. 2019;365:191-200.
38. Ono JP, Castelo S, Lopez R, Bertini E, Freire J, Silva C. PipelineProfiler: A Visual Analytics Tool for the Exploration of AutoML Pipelines. *IEEE Transactions on Visualization and Computer Graphics (Early Access)*. 2020. doi: 10.1109/TVCG.2020.3030361.

Figures

Figure 1

Different types of easily-confused Chinese herbal medicines: a Genuine and Adulterant CHMs a1-3 Ziziphi Spinosae Semen a4-5 Ziziphi Mauritanae Semen b-c CHMs with white color b1 Fritillariae Thunbergii Bulbus b2 Angelicae Dahuricae Radix b3 Trichosanthis Radix b4 Changii Radix b5 Paeoniae Radix Alba c1 Calamina c2 Draconis Dens c3 Draconis Os c4 Ostreae Concha c5 Gypsum Fibrosum d CHMs with black color d1 Rehmanniae Glutinosae Radix d2 Rehmanniae Radix Praeparata d3

Scrophulariae Radix d4 Schisandrae Chinensis Fructus d5 Mume Fructus e CHMs from relative plants e1
 Ardisiae Crenatae Radix e2 Ardisiae Hanceanae Radix e3 Ardisiae Lindleyanae Radix e4 Ardisiae
 Linangensis Radix e5 Ardisiae Brevicaulis Radix f CHMs applied in whole plant or aerial parts f1
 Hedyotidis Diffusae Herba f2 Scutellariae Barbatae Herba f3 Cirsii Japonici Herba f4 Taraxaci Herba f5
 Eupatorii Herba g CHMs small in size (diameter < 5 mm) g1 Astragali Complanati Semen g2 Vaccariae
 Semen g3 Psoraleae Fructus g4 Cannabis Fructus g5 Cassiae Semen h CHMs using bark h1 Periplocae
 Cortex h2 Lycii Cortex h3 Albiziae Cortex h4 Dictamni Cortex h5 Moutan Cortex

Figure 2

Dataset split for AutoML model building and held-out test

Figure 3

Schematic representation of the AutoML model creation

Figure 4

Examples of CHMs categories correctly recognized by ModelArts: a *Atractylodis Rhizoma* b *Ilicis Cornutae Folium* c *Ginseng Radix et Rhizoma Rubra* d *Citri Reticulatae Pericarpium* e *Cistanches Herba* f *Coptidis Rhizoma* g *Toosendan Fructus* h *Fritillariae Cirrhosae Bulbus*

Figure 5

Prediction results of Auto ML models: a The F1-score of different types of easily-confused CHMs from Fig.1 in the ModelArts evaluation and held-out test b-g Results from the ModelArts held-out test b The F1-score of genuine and adulterant CHMs c The F1-score of CHMs with white color d The F1-score of CHMs with black color e The F1-score of CHMs from relative plants f The F1-score of CHMs applied in whole or

aerial parts of the plants g The F1-score of CHMs small in size (diameter < 5 mm) h The F1-score of CHMs using bark

Figure 6

Examples of CHMs categories misidentified by ModelArts: a Adenophorae Radix b Caryophylli Fructus c Ganodermae Lucidi d Paeoniae Radix Alba e Angelicae Dahuricae Radix f Ardisiae Lindleyanae Radix g Stephaniae Tetrandrae Radix h Aloe i Ganodermae Sinensis j-k Fritillariae Thunbergii Bulbus l Ardisiae Crenatae Radix

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1.xls](#)