PLANT SPECIES RECOGNITION BASED ON HYPERSPECTRAL PLANT CANOPY IMAGES WITH DEEP LEARNING

Yu-Kai Chen[1], Keng-Hao Liu[1], and Chinsu Lin[2]

[1]Department of Mechanical and Electro-Mechanical Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan
[2]Department of Forestry and Natural Resources, National Chiayi University, Chiayi, Taiwan

ABSTRACT

Plant species recognition is an important topic in forest remote sensing. In the past, people have used laboratory plant images to design different recognition methods. Some applied machine learning algorithms to perform species recognition on the plant RGB images, while others started to use deep learning (DL) approaches, such as convolutional neural network (CNN), to improve recognition performance. However, two issues arise: Firstly, many species with similar appearance are difficult to be correctly classified by CNN under the limited spectral information of RGB imaging. Secondly, the existing CNN-based classification models are designed for particular datasets and thus are not suitable for plant images. To tackle these issues, this paper proposes a revolutionary framework that combines hyperspectral imaging (HSI) and DL technologies to perform plant species classification with a large number of species. We collected canopy images of 100 plant species via a Visible-NIR hyperspectral camera, and built an plant canopy dataset which consists of 3250 training images and 3250 test ones. Furthermore, we designed a lightweight CNN model that utilizes both 3D CNN and 2D CNN modules to implement spectral information fusion and spectral-spatial features extraction, called Hybrid-CNN. Experimental results show that Hybrid-CNN achieved at least 98% in overall accuracy rate and 0.98 in Kappa coefficient, and significantly outperformed the reference classification models.

Index Terms— Plant species recognition, deep learning, convolutional neural network, hyperspectral imaging

1. INTRODUCTION

Identifying species of individual plants or trees over an inventory plot as well as a forest stand is essential for automatic mapping of plant distribution in forest remote sensing. Therefore, in recent years, many people have begun to develop image-based plant species recognition methods. In the early days, they identified species through the local features of plants, such as roots, stems, leaves, flowers, and fruits images, etc., via ML classifiers such as SVM or kNN [1-5]. However, the recognition performance of ML approaches relies on the quality of feature engineering process, which is determined by researchers' subjective experience and knowledge. Recently, with the rise of deep learning (DL), convolutional neural network (CNN) has become the major method in the field of image processing and computer vision. The CNN can learn useful recognition features, find the correlation between different components within the image, and perform accurate predictions under a single framework. The DL approaches can be applied to larger scale or complex images in which more spatial features can be referenced in the classification process and thus resulting in higher accuracy. By virtue of the superior performance of CNN, it has also begun to be adopted in plant image recognition [6-11].

Despite the success obtained by using CNN, there are still two remaining challenges. 1. High interclass similarity issue: limited by the poor spectral information of the RGB imaging system, many species with similar appearance (color, leaf shape, or leaf veins) are difficult to be correctly classified by CNN. It happens particularly when the number of species is large. 2. Model fit problem. The existing CNN classification models are designed for the image datasets with large differences in color appearance or monotone composition between different categories. Therefore, they may not be suitable for classifying plant images. Furthermore, the architecture of those networks is usually very large, and thus requires a large amount of training data to reach convergence in the training process.

To overcome these issues, this paper proposes a revolutionary framework that combines hyperspectral imaging (HSI) and DL technologies to perform plant species classification with a large number of species. With the support of fine spectral resolution and near-infrared (NIR) information provided by HSI and the robust spatial feature extraction/integration capability of CNN, the classification performance can be improved. We first collected canopy images of 100 plant species via a visible-NIR hyperspectral camera [12] with a spectrum range of 470-900nm, and built an HSI plant canopy dataset which consists of 3250 training images and 3250 test ones. Each image has a size of 400x400x150. In order to make more use of spectral information in HSI, we designed a compact CNN model that utilizes both 3D-CNN and 2D-CNN network modules to implement spectral information fusion and spectral-spatial features extraction, called Hybrid-CNN. In the experimental setting, we selected three existing network models for comparison: AlexNet [??], LtCNN [11], and MobileV3 Small [15]. Since the spectral correlation in a HSI cube is high, we adopted several band selection (BS) methods to reduce the dimensionality of HSI cubes to lower the burden of model training and avoid curse of dimensionality issue. The experimental flowchart is shown in Fig. 1.

Experimental results show that the proposed Hybrid-CNN can achieve at least 98% in overall accuracy rate and 0.98 in Kappa coefficient under most BS settings, and significantly outperformed the reference classification models. We also visualize the classifying results with CAM algorithm [??] to localize the key features that the CNN focuses on.

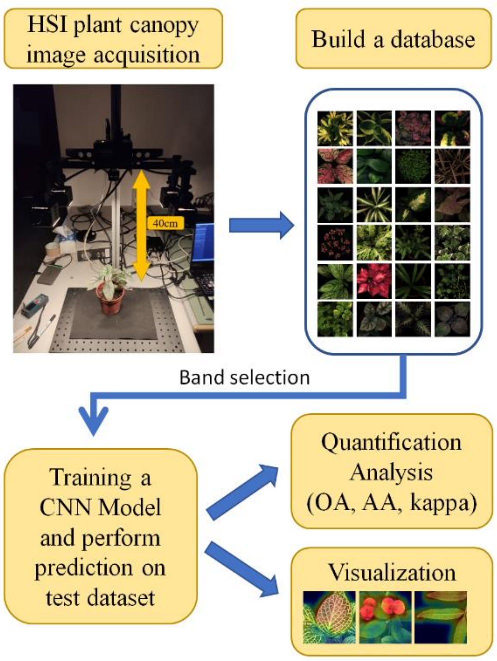

Figure 1. The experimental flowchart.

2. IMAGE ACQUISITION AND DATASET PREPARATION

2.1. HSI image acquisition

This study aims to recognize plant species based on plant morphology via features of leaf color, size, contour, surface, venation, and phyllotaxy. One hundred species of foliage plants from 100 species and 46 families were collected to produce plant images for analysis. To increase the leaf features and plant geometry diversity in the images, at least 2 or 3 plant individuals were gathered for replications.

We used the IMEC Snapscan VNIR imaging system [12] capture the HSI images of the species. This system composes of the image sensor, spectral filters, optical lens, and some other components that can acquire hypercube datasets up to a full-image size of 3,600 × 2,048 pixels covering a spectrum range from 458 nm to 913 nm. The system's spectral resolution is 2.8 nm (equivalent to 161 bands).

Figure 2. Example images of the 100 plant species.

In the image acquisition, the camera is mounted on a tripod facing downward to the plant at a distance of 40 to 60 cm. Four 50 w/12 V halogen lamps were deployed, one on each side of the plant at a 45-degree elevation angle from the horizontal plane. A black material was used to minimize the background material reflectance effects on the target reflectance. The aperture of the camera was set to f5.6. Due to the vertical and horizontal variations of the leaves locations, changing the orientation of the plant led to changes in light intensity over the crown area and therefore helped to increase the diversity of the sample images. The plant was set to rotate 90 degrees to generate diverse hyperspectral images of the same plant. With the combination of two camera-target distances and four plant orientations, eight HSI raw images of every individual plant of the 100 species were acquired. Each hypercube raw image has a dimension of 1,200 ×1,200 ×161 bands. Due to the significant noise in the wavelengths at both ends of the sensor, several noisy bands were removed and totally 147 bands with a spectrum range of 468–898 nm were used for the analysis.

2.2. Dataset preparation

To increase the total number of images for training a CNN model, each 1,200×1,200 HSI image is evenly segmented into nine non-overlapping 400×400 sub-images. Those sub-images with a noticeable shadow or insufficient leaves, e.g., the leaf/background ratio does not exceed 60%, were removed. As a result, 65 sub-images were inspected and retained for each species, and a total of 6,500 HSI reflectance images were generated. In our experiment, we adopted 1:1 ratio for training and test image. That is, 3,250 images were used to train CNNs and the remaining 3,250 ones were used for test.

3. THE DESIGN OF THE HYBRID-CNN

We built a compact CNN model, referred to as Hybrid-CNN, which can extract both spectral and spatial features in an efficient way. It is composed of several components as shown in Fig. 3. The major there components are: 3DCNN module, 2DCNN module, and classification module.

The 3DCNN module includes two 3D kernels of different sizes, which can fuse the spectral information as well as extract low-level spatial features. The 2DCNN module consists of two Inception block, which is responsible for extracting mid-level and high-level spatial features from plant image. The classification module is composed of one GAP layer, two FC layers, and followed by a Softmax classifier. It makes predictions about plant species based on the features obtained in the previous two modules. The detail of the architecture of Hybrid-CNN is listed in Table I.

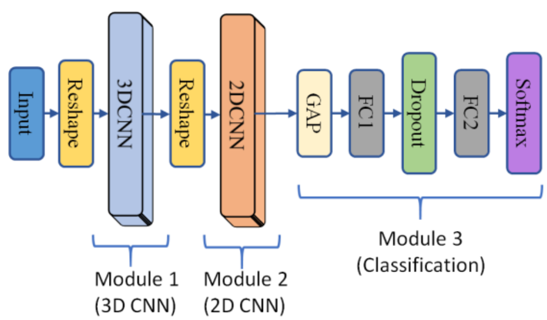

Figure 3. The architecture of the Hybrid-CNN.

For training our Hybird CNN, we adopted Cross Entropy (CE) as the loss function. The formula is as follows:

$$Loss_{CE} = -\frac{1}{Q}\sum_i^Q \sum_{c=1}^M y_{i,c} \times \log(p_{i,c}), \qquad (1)$$

where Q denotes the number of samples, M denotes the number of classes (species), $y_{i,c}$ is the binary indicator, and $p_{i,c}$ is the predicted probability.

4. EXPERIMENTS

4.1. Experimental setting

All the experiments were conducted under the following hardware environment: INTEL® CORE™ I9-9900 CPU, GEFORCE RTX 2080TI 11GB GPU, 64G DDR4 RAM, and WINDOW 10 operating system. For software environment, PYTHON 3.6 was used to write the code, and TENSORFLOW GPU 1.8.0 was used as the framework for establishing the neural network.

There classification CNN networks were selected for performance comparison with our Hybrid-CNN: AlexNet [15], MobileNetV3 Small [16], and LtCNN[11]. We followed [11] to reduce the parameters of AlexNet for better performance on classifying on our plant images. Table.2 lists the number of parameter of all four CNNs. It can be found that Hybrid-CNN has the fewest parameters and only one-tenth of AlexNet.

Table 1. The definition of each block in Hybrid-CNN.

Parameter	Kernel size	Depth	Strides	Output
Input				224x224xL
Reshape				224x224xLx1
Convolution	5x5x5	16	(2,2,1)	112x112xLx16
Max pooling	2x2x2	16	(2,2,1)	56x56xLx16
Convolution	3x3x3	32	(1,1,1)	56x56xLx32
Max pooling	2x2x2	32	(2,2,1)	28x28xLx32
Reshape		Lx32		28x28x(Lx32)
Inception module1		288		28x28x288
Inception module2		480		14x14x480
Max pooling	3x3	480	(2,2)	7x7x480
Global average pooling		480		1x1x480
FC layer 1		256		1x1x256
Dropout (40%)		256		
FC layer 2		100		1x1x100
Softmax		100		1x1x100

4.2. Band selection

To avoid curse of dimensionality issue occurred by the high spectral information redundancy in HSI, we applied four band selection (BS) methods to reduce data dimensionality: RGB, NIR, RGB+NIR, and UBS-9. RGB selects the bands with spectrum 472.9nm, 536.4nm, 602.8nm. NIR selects one red edge bands and two near-infrared ones whose spectrum are 749.9nm, 799.0nm, and 850.9nm. RGB+NIR is the pure combination of RGB and NIR, including a total of 6 bands. UBS-9 uniformly selected 9 bands in the entire spectrum.

4.3. Classification results

Table 3 lists the classification performance of all the CNNs in overall accuracy (OA), precision, Macro F1-score, and Kappa coefficient. There are several major observations. Firstly, it shows that the proposed Hybrid-CNN can achieve at least 99% in overall accuracy rate and 0.99 in Kappa coefficient under most BS settings. The results also illustrate the high potential of using HSI materials and the well-designed CNN model for plant species identification. Secondly, our Hybrid-CNN significantly outperformed the three reference CNN models in all BS settings. It is worth mentioning that LtCNN and MobileNet V3 Small are both lightweight CNN models, however, their performance is

Table 3. Classification performance of using different CNN models in three BS settings.

CNN model	Feature Selection	OA (%)	Precision	Macro F1-score	Kappa coefficient
AlexNet [15]	RGB(3)	75.13	0.8112	0.7531	0.7488
	NIR(3)	55.10	0.6013	0.5427	0.5465
	RGB+NIR(6)	76.49	0.8244	0.7743	0.7625
	UBS-9 (9)	80.09	0.8473	0.8038	0.7989
Mobile-NetV3 Small [16]	RGB(3)	68.80	0.7293	0.6932	0.6848
	NIR(3)	56.95	0.6272	0.5771	0.5651
	RGB+NIR(6)	76.89	0.8187	0.7822	0.7665
	UBS-9 (9)	76.21	0.8159	0.7762	0.7597
LtCNN [12]	RGB(3)	93.78	0.9418	0.9374	0.9372
	NIR(3)	80.58	0.8170	0.8013	0.8038
	RGB+NIR(6)	95.96	0.9614	0.9594	0.9592
	UBS-9 (9)	97.04	0.9727	0.9703	0.9701
Hybrid-CNN	RGB(3)	98.92	0.9895	0.9891	0.9891
	NIR(3)	87.81	0.8945	0.8763	0.8769
	RGB+NIR(6)	**99.16**	**0.9921**	**0.9917**	**0.9916**
	UBS-9 (9)	**99.16**	0.9920	**0.9917**	**0.9916**

obviously not as good as Hybrid-CNN. This shows that the 3D CNN module can better capture the key features for identifying plants, and justifies the use of 3D CNN for processing hyperspectral images. Thirdly, using different BS settings did have an impact on classification performance. In most CNN models, the prediction performance ranking was: UBS-9≥RGB+NIR＞RGB＞NIR. It is verified that using visible bands (RGB) are more discriminative than NIR bands. However, using NIR bands as auxiliary information is beneficial to classification results.

4.4. Visualization results

We adopted class activation mapping (CAM) [14] to visualize the spatial locations of the discriminative features captured by Hybrid-CNN for the plant images. Fig. 4 shows three plant test images successfully classified by Hybrid-CNN with the corresponding heat maps produced by CAM. It is confirmed that Hybrid-CNN did learn the characteristics of phyllotaxy, leaf/flower colors, leaf veins, leaf shape, leaf margin, and leaf overlapping pattern, and serve them as the basis for classification.

5. CONCLUSION

This paper proposes a framework using a well-designed deep neural network hyperspectral plant canopy images for accurate plant species recognition. We first built a HSI plant dataset composed of the images of 100 plant species. Later, we designed a lightweight network named Hybrid-CNN for plant species classification. Experimental results demonstrated the excellent classification performance brought by this framework.

6. ACKNOWLEDGEMENT

This work was supported by the National Science and Technology Council (NSTC), Taiwan, in Grant No.: 111-2221-E-110-030-MY2 and 110-2221-E-415-007-MY2.

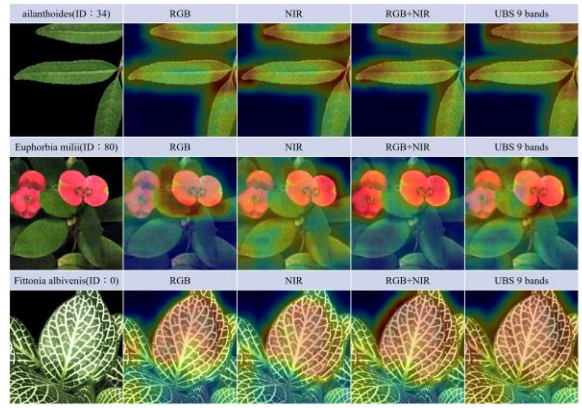

Figure 4. The heat maps of three plant sample images.

7. REFERENCES

[1] H. Kan, L. Jin, and F. Zhou, "Classification of medicinal plant leaf image based on multi-feature extraction," Pattern Recognition and Image Analysis, vol. 27, no. 3, pp. 581–587, 2017.

[2] S. Mahajan, A. Raina, X.-Z. Gao, and A. K. Pandit, "Plant recognition using morphological feature extraction and transfer learning over SVM and AdaBoost," Symmetry, vol. 13, no. 2, 2021, Art. no. 356.

[3] A. Ambarwari, Q. J. Adrian, Y. Herdiyeni, and I. Hermadi, "Plant Species Identification Based on Leaf Venation Features Using SVM," Telkomnika, vol. 18, no. 2, pp. 726–732, 2020.

[4] D.S. Guru, Y.H. Sharath and S. Manjunath, "Texture features and KNN in classification of flower images,", IJCA, Special Issue on RTIPPR (1), pp. 21-29, Aug. 2010.

[5] S. Zhang and K.-W. Chau, ''Dimension reduction using semi-supervised locally linear embedding for plant leaf classification,'' in Proc. Int. Conf. Intell. Comput., pp. 948–955, 2009.

[6] W.-S. Jeon and S.-Y. Rhee, ''Plant leaf recognition using a convolution neural network,'' *Int. J. Fuzzy Log. Intell. Syst*., vol. 17, no. 1, pp. 26–34, Mar. 2017.

[7] M. Berihu, J. Fang, and S. Lu. "Automatic Classification of Medicinal Plants of Leaf Images Based on Convolutional Neural Network," Commun. Comput. Inf. Sci. on Big Data , vol 1496. Springer, Singapore, 2022.

[8] A. Hassan, S. Islam, M. Hasan, S. B. Shorif, T. Habib, and M. S. Uddin, ''Medicinal plant recognition from leaf images using deep learning,'' in Computer Vision and Machine Learning in Agriculture, vol. 2. Singapore: Springer, pp. 137–154, 2022.

[9] M. Dyrmann, H. Karstoft, and H. S. Midtiby, "Plant species classification using deep convolutional neural network," Biosystems Engineering, vol. 151, pp. 72–80, 2016.

[10] S. K. Mahmudul Hassan, and A. Maji, "Identification of Plant Species Using Deep Learning," Front. Comput. Sci, vol 1255. Springer, Singapore, 2021.

[11] K.-H. Liu, M.-H. Yang, S.-T. Huang, and C. Lin, "Plant Species Classification Based on Hyperspectral Imaging via a Lightweight Convolutional Neural Network Model," Front. Plant Sci., vol. 13, 2022.

[12] Imec SNAPSCAN VNIR Hyperspectral Camera System. https://www.imechyperspectral.com/en/cameras/snapscan-vnir

[13] E. J. Heravi, D. Puig, and H. H. Aghdam, ''Classification of foods using spatial pyramid convolutional neural network,'' in Frontiers in Artificial Intelligence and Applications. Amsterdam, The Netherlands: IOS Press, 2016.

[14] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva and A. Torralba, "Learning Deep Features for Discriminative Localization," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 2921-2929, 2016.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Commun. ACM, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.

[16] A. Howard et al., "Searching for MobileNetV3." arXiv, Nov. 20, 2019. doi: 10.48550/arXiv.1905.02244.