Applying the Artificial Neural Networks and Unmanned Aerial Vehicle Images to Classify Various Crop Types

Wan-Li Sun , Shu-Ting Kuo, Chien-Erh Wang, Wan-Chen Lin
Ming-Yun Chang , , Yu-Cheng Hsiao, Re-Yang Lee
No. 100, Wenhwa Rd., Seatwen, Taichung, Taiwan 40724, (R.O.C.)
e-mail:wanli9925@gmail.com

Keywords: Unmanned Aerial Vehicle(UAV),
Artificial Neural Networks(ANN),Convolutional Neural Network(CNN).

Abstract

At present, the government is mostly manual to investigate agricultural information, but this method is lengthy and time-consuming. In order to improve the efficiency of agricultural survey, artificial intelligence and deep learning technology are increasingly widely used in recent years, compared with the traditional method of image classification. The artificial neural networks (ANN) can improve the accuracy of interpretation through learning of image characteristics. Therefore, this study adopted the images taken using Unmanned Aerial Vehicles (UAVs) on the southwest side of Yan's Nest District in Kaohsiung City, Taiwan. The ANN and convolutional neural network (CNN) were used to interpret various crop types respectively. The efficiencies of using these two methods for classifying crop types were then evaluated. The results can provide more efficient ways for agricultural surveys.

1. INTRODUCTION

Agricultural policies and production and marketing strategies have long been unfavorable in Taiwan, resulting in a widespread imbalance in agricultural production and marketing. To mitigate this situation, scholars have employed remote-sensing technologies in agricultural research. For example, Huang et al. (2008) used the color images captured by Formosat-2 to interpret the texture and spectral properties of pineapples, a notably sensitive crop (Chang and Shen, 2011), and to predict the area of rice production. The results indicated that the predictions of possible yield at the time of harvest were satisfactorily accurate. Conventional aerial photography produces photographs that cover large areas; however, aerial photograph is easily affected by factors including the weather and clouds and thus has limited applicability under certain conditions. By contrast, an unmanned aerial vehicle (UAV) offers advantages including high mobility, nonsusceptibility to clouds, and relatively low cost and has thus become a focus in research. Baker (2012) used UAV-obtained images with a resolution of 10 cm to classify soil in vineyards located in southern France and employed multiresolution segmentation to then effectively improve the management of vineyards. Lee et al. (2018) used an UAV to photograph fields of lychee and longan and obtained satisfactory interpretation results, concluding that UAVs are beneficial tools for determining the cultivation area of agricultural land. Given the aforementioned strengths of remote sensing technologies and UAVs, this study employed an UAV to capture photos that served as a basis for image interpretation.

The high-resolution images obtained by UAVs contain a large volume of data; therefore, using a conventional automated interpretation method to classify such images would possibly cause an agricultural survey to take an excessively long period of time. Recently, artificial intelligence (AI) and deep learning have been widely employed in various domains such as travel time prediction (Chen et al., 2019), automated real-time monitoring of water levels (Li et al., 2018), and facial recognition (Chen 1et al., 2013). The neural network is the core of AI; such networks comprise various artificial neural, are used to solve complex nonlinear problems, and are highly adaptable, fault tolerant, reliable, and able to organize, learn, and associate. Scholars have predominantly used back-propagation neural networks (BPNNs) or convolutional neural networks (CNNs) for image interpretation. For example, Hsieh et al. (1997) used texture analysis and a BPNN to interpret images of cabbage seedlings, and Zhou et al. (2017) employed a CNN to classify and interpret the main organs of tomatoes. The present study employed a BPNN and a CNN to interpret images captured by an UAV and compared the two networks in terms of their interpretation accuracy, determining which is more suitable for crop interpretation.

2. MATERIALS AND METHODS

2.1 Study Area

The study area was located on the southeast side of Yanchao District in Kaohsiung City, Taiwan. Because of the unique geology and imbalanced fertility of soil in this district, agricultural productivity is easily affected by the weather. This district was thus selected for an agricultural survey with the aim of formulating appropriate agricultural strategies and policies.

The cropland in Yanchao District has a size of 3,026 ha and mostly comprises plains. Crop production accounts for 16% of all fruit production, 0.93% of all grain production, and 4.35% of all vegetable production in Kaohsiung. The land in Yanchao District comprises mud volcano and limestone soil and exhibits a terrain higher to the east and lower to the west. The sloped area in the east constitutes one third of the district’s total area, and the sloped terrain limits agricultural activities to certain crops, such as lychee, longan, watermelon, and bamboo shoot. Crops suitable for growing in limestone soil are usually long-term crops, such as guava, jujube, and pomelo.

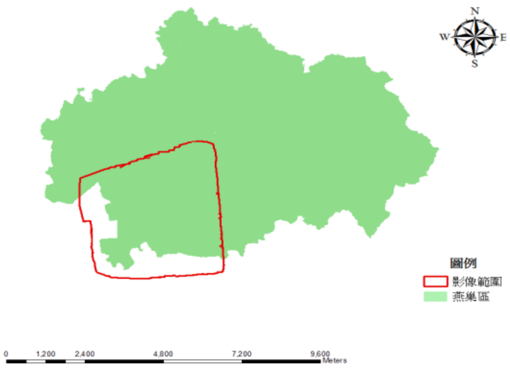


Figure 1. Study area

2.2 Image Collection

This study captured images using a SonyA7R II mounted on the CW-10, a vertical take-off and landing (VTOL), fixed-wing UAV (Fig. 2). The CW-10 has a composite wing structure integrating fixed-wing and quadrotor systems and facilitating VTOL, of which fixed-wing UAVs are incapable. Moreover, the UAV incorporates the strengths of a fixed-wing system, namely endurance and long flight distance, with that of a rotor system, namely VTOL. Its VTOL ability enables the CW-10 to take off and land without a runway or air space, greatly improving its flexibility. In particular, the UAV can perform flying tasks in mountainous areas in which expansive runways are absent, rendering it suitable for use in Yanchao considering this district’s terrain.

Figure 2.UAV (CW-10)

2.3 Image segmentation

This study employed region-based multiresolution segmentation for image segmentation; in consideration of spectral and shape variation, the segmentation parameters were determined according to the different image characteristics.

eCognition, a method of multiresolution segmentation, involves grouping a pixel with similar neighboring pixels to form various areas until the result approaches the defined scale parameter. The scale parameter is the upper limit of pixel heterogeneity; a small parameter indicates low heterogeneity and hence small segments. The overall level of heterogeneity is calculated according to the parameter settings, namely the weights of the different types of information considered; for example, Shape denotes the weights of a spectrum and shape, whereas Compactness denotes the weights of compactness and smoothness. Different parameter settings can be used to yield different segmentation results for various purposes.

In eCognition, the same parameters result in different levels of heterogeneity for different images. Therefore, the parameters must be optimized for every image, and the segmentation result must be validated with reference to the ground truth to ensure intrasegment homogeneity and prevent oversegmentation. This study optimized the parameters manually, employed UAV images and ground truths for validation, and maximized the area of segments under the condition of intrasegment homogeneity. The following optimal parameter settings were identified:

(1) Image Layer Weights: the default was used, with the weight of each band being 1.

(2) Scale Parameter: set to 500.

(3) Shape: set to 0.1.

(4) Compactness: the default of 0.5 was used.

Except for the Scale Parameter, these parameter settings are consistent with those employed by Park and Park (2015), who also used eCognition for crop classification. The difference in the Scale Parameter is possibly because the present study employed images with different resolution to that in Park and Park (2015), which was 8 cm. Future studies on agricultural interpretation may refer to these parameter settings to facilitate efficient identification of optimal parameter settings.

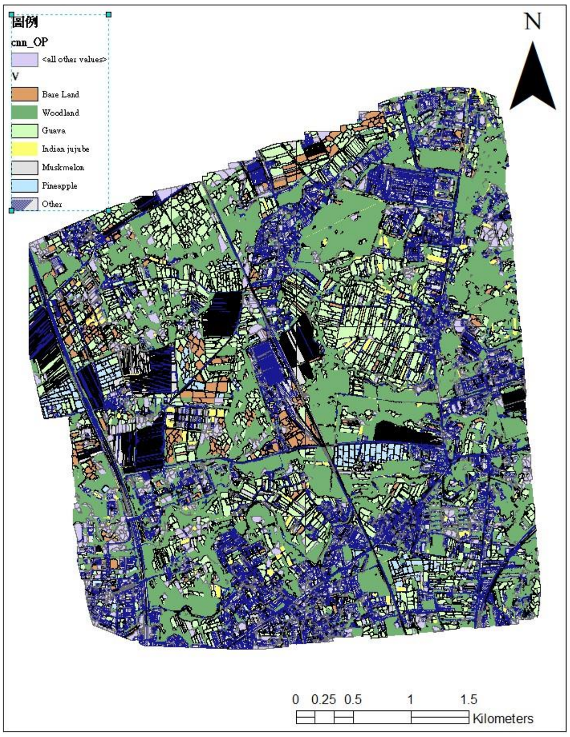
2.4 Image Interpretation

This study employed a BPNN and CNN to perform image interpretation, and the two sets of results were compared to determine the network with higher accuracy. The basic structure of a BPNN is an input layer, hidden layer, and output layer. The input layer receives signals from the external environment. The hidden layer reveals the relationship between the inputs and outputs; the number of units processed cannot be determined using any standard method and is usually determined through constant testing. The output layer presents the output variables of the network. The following numbers of inputs were used during the BPNN training process in this study: 460 neurons for the input layer; five hidden layers with 1,000, 2,000, 2,000, 1,000, or 500 neurons in the layer, respectively; and eight neurons for the output layer, with eight corresponding categories. A total of 460 inputs and 20,000 times of training were used to obtain the final BPNN.

The basic structure of a CNN is an input layer, convolution layers, a pooling layer, fully connected layers, and an output layer. The CNN was established through the following three steps. (1) Convolution Layer 1 was defined as having 36 weight filters, with each moving filter having dimensions of 3 × 3; the size of the convoluted images was set to be constant at dimensions of 6 × 6 × 1. Convolution Layer 2 converted the original 36 images into 72 images, with the size of the images held constant. Convolution Layer 3 converted the 72 images into 36 images, with their size set constant. (2) The fully connected layers were established with 1296 neurons. (3) The output layer comprised eight neurons and eight corresponding categories. Training data comprising 460 samples were used to train the CNN for 20,000 times.

3. RESULTS AND DISCUSSIONS

3.1 Results



Classification Result（BPNN）

Classification Result (CNN)

3.2 Accuracy Assessment

The image classification process was divided into two parts: interpretation of the aerial images using the BPNN, and such interpretation using the CNN.

According to Tables 1 and 2, the BPNN has overall accuracy of 76.18% and a kappa coefficient of 0.7 for image classification. By contrast, the CNN as an image classifier has overall accuracy of 77.74% and a kappa coefficient of 0.72. The two networks thus both achieve overall accuracy greater than 70%, satisfying the general standard for image classification.

Table 1. Classification Accuracy (BPNN)

	Producer's Acc.	User's Acc.	Overall Acc.
Bare Land	58.5	10.5	76.18
Woodland	92.8	87.9	KAPPA
Guava	72.0	74.9	0.7
Indian jujube	36.6	36.1	
Muskmelon	75.2	97.7	
Pineapple	0	0	
Other	61.2	3.6	

Table 2. Classification Accuracy (CNN)

	Producer's Acc.	User's Acc.	Overall Acc.
Bare Land	58.5	7.83	77.74
Woodland	92.6	87.3	KAPPA
Guava	70.3	78.3	0.72
Indian jujube	36.9	38.2	
Muskmelon	73.6	95.1	
Pineapple	89.4	78.0	
Other	47.4	3.5	

The kappa coefficient reflects the agreement between the ground truth and classification result and ranges between 0 and 1. A high kappa coefficient indicates high precision. The kappa coefficients of the BPNN and CNN were 0.7 and 0.72, respectively (Tables 1 and 2). With reference to the levels of agreement in Table 3 (Landis and Koch, 1977), the BPNN and CNN both exhibit satisfactory precision.

The overall precision of the BPNN was similar to that of the CNN; however, when the precision was examined for each classification category, substantial differences were discovered. For example, when classifying bare land, the accuracy of the BPNN and CNN is 10.5% and 7.83%, respectively; accordingly, the BPNN has higher accuracy than the CNN for bare land classification. When classifying land on which pineapples are cultivated, the accuracy of the BPNN is 0%, whereas that of the CNN is 78%, showing the relatively precise classification ability of the CNN for pineapple land.

Table 3. Classification quality associated to a Kappa statistics value

Kappa coefficient	Interpretation
<0.00	Poor agreement
0.00~0.20	Slight agreement
0.20~0.40	Fair agreement
0.40~0.60	Moderate agreement
0.60~0.80	Substantial agreement
0.80~1.00	Almost perfect agreement

4. **CONCLUSIONS AND SUGGESTIONS**

(一) Conclusions

This study employed an UAV to capture images of the southwest side of Yanchao District in Kaohsiung City. eCognition was employed to segment the images, and a BPNN and CNN were used to classify the images. Comparing the accuracy and efficiency of the two networks, this study obtained the following conclusions and makes these recommendations.

1. Comparison of satellite and UAV images: Because of the dense population in Taiwan, its agricultural activities predominantly comprise intensive smallholder farming. Traditional remote-sensing aerial photography, despite its ability to cover a large area, has limited capacity to gather information because of its relatively low image resolution and susceptibility to factors such as the weather and clouds. By contrast, UAVs have high mobility, are not easily affected by clouds, have high resolution, and can gather information in places that are normally difficult to reach; UAV-obtained data are thus conducive to image classification.
2. Importance of image segmentation: Because images captured using an UAV have high precision, the amount of data they contain is large. Accordingly, eCognition was used in this study to segment the images before image classification to establish hierarchical relationships among objects in an image with regards to spectral analysis, shape texture, and similarity between neighboring objects. Subsequently, object-oriented techniques were used to integrate the images, which was conducive to increasing the interpretation accuracy.
3. Comparison between the BPNN and CNN interpretation results: From the perspective of the rate of correct classifications, the BPNN and CNN have overall accuracy of 76.1% and 77.7%, respectively, only differing slightly. However, in terms of the pineapple producer's accuracy of the two networks, the BPNN has accuracy of 0%, whereas the CNN has accuracy of 89.4%, indicating that the BPNN was barely able to interpret land containing pineapples, whereas the CNN was fairly satisfactory at identifying this crop. Therefore, the CNN might be more suitable than the BPNN for agricultural surveys.

(二)Suggestions

Most agricultural surveys are conducted with traditional ground survey approaches and classification methods. The employment of UAVs for photography and a BPNN and CNN for image classification within such surveys remain relatively rare. The traditional methods can lead to excessively long periods of time being spent on surveys and poor crop predictions.

1. The Council of Agriculture, Executive Yuan, should employ UAV technology in agricultural surveys to reduce the labor and resources associated with traditional ground surveys. The aerial images can be converted into parcel maps of the targeted area and analyzed using a geographic information system to increase the analysis efficiency.
2. To prevent objects being missed or misclassified, the following measures can be taken:

1. As many training samples should be used as possible, and the ratio of ground truth data over the total number of pixels should be utilized to select training samples.
2. In BPNN classification, the weights are constantly adjusted to modify the difference between the inputs and expected outputs, thus producing the desired outputs. An excessively large difference between the inputs and outputs may affect the learning efficiency of a BPNN and hence the classification results.
3. In CNN classification, the size of the convolution kernel greatly influences the accuracy of the CNN. When the convolution kernel is excessively large, the complexity of image features will exceed the network's capacity, but when the convolution kernel is excessively small, local features cannot be interpreted effectively. therefore, selection of appropriate convolution kernels and sizes thus improves the classification accuracy of the CNN.

5. REFRENCE

Lee, R. Y., Ou, D. Y., and Hsu, C. H., 2018.Utilizing Unmanned Aerial Vehicle Images to Interpret Crop Types in the Hillside Area, Journal of Photogrammetry and Remote Sensing,23(4):245-256. (in Chinese)

Chou, C.Y., Wu, S.Y., and Chen, C. L., 2018. Application of unmanned aerial vehicle (UAV) photography on agricultural post-disaster investigation – A case study on image interpretation of banana plantation damages by disasters, Journal of Photogrammetry and Remote Sensing, 23(2): 83-101. (in Chinese)

Chang, K. W., and Yuan, S., 2011. Rice Yield Prediction Using Remote Sensing Images, Journal of Aeronautics, Astronautics and Aviation,43(1):55-63. (in Chinese)

Lo, C. L., Chen, C. H., Hu, J. L., Lo, K. R., and Cho, H.J.,2019. A Fuel-Efficient Route Plan Method Based on Game Theory, Journal of Internet Technology, 20(3):925-932. (in Chinese)

Liu, G. R., Lin, T. H., Huang, S. J., and Kuo, T. H.,2003. Applying Airborne Multi-spectral Remotely Sensed, Journal of Photogrammetry and Remote Sensing,8(4):71-82. (in Chinese)

Zou, L., and Zhang, X. F.,2012. Application and Development of the Artificial Intelligence, Netinfo Security, 2012(2):11-13. (in Chinese)

Li, W. L.,2016. Applying airborne hyperspectral images to classify crop types(Master's thesis), Feng-Chia University Department of Land Management， https://hdl.handle.net/11296/78bbvp.(in Chinese)

Huang, C. H., Shiu, Y. S., and Chu, T. H., 2008, Study on sensitive crops production investigating by remote sensing techniques—Case studies of pineapple and banana crops, Global Change and Sustainable Development,2008(60):19-24. (in Chinese)

Hsieh, C. L., Cheng, S.F., and Lin, T. T.,1997, Application Of Image Texture Analysis And Neural Network On The Growth Stage Recognition For Head Cabbage Seedlings, Journal of Agricultural Machinery,6(2):1-13. (in Chinese)

Chen, Y. N., Ho, G. F., Han, C. C., and Fan, K. C.,2013. Face Detection and Recognition Using Multi-level Classifier, Journal of Advanced Technology and Management,3(2):47-61. (in Chinese)

Zhou, Y. C., Xu, T. Y., Zhen, W., and Deng, H., B.,2017, Classification and recognition approaches of tomato main organs based on DCNN, Transactions of the Chinese Society of Agricultural Engineering,33(15):219-226. (in Chinese)

Li, Y. D., Li, C. Y., Wu, W. Z., Chen, Boris P.-T., 2018.A study on real-time water level recognition of closed-circuit television system for disaster management – based on the convolutional neural network of artificial intelligence, Conference for Disaster Management in Taiwan,2018, p563-570. (in Chinese)