Identification of Mangrove Invasive Plant *Derris trifoliate* Using UAV Images and Deep Learning Algorithms

**Mengxi Liu**, Graduate Student Member, IEEE, Haojun Deng ©, and Wen Dong ©

**Abstract**—*Derris trifoliate*, one of the most notorious invaders of mangroves in South China, seriously threatens the growth of mangroves and the stability of local ecosystem. In order to effectively control the spread of *Derris trifoliate* and stabilize the mangrove ecosystem, it is necessary to identify the distribution and simulate the occurrence of *Derris trifoliate* in mangrove forests. While previous methods for invasive plant mapping based on satellite images were limited by low temporal and spatial resolution, the newly born unmanned aerial vehicle (UAV) images have provided a fast, fine-grained, and low-cost alternative for rapid invasive plant monitoring. However, extracting *Derris trifoliate* in mangrove forests based on UAV images is still a challenge. For one thing, it is difficult to collect enough data for model training, since the *Derris trifoliate* is hard to be distinguished in mangroves. For another thing, few existing methods can meet the requirements of high computing efficiency and low memory consumption of UAVs. Therefore, we proposed two lightweight deep learning networks based on DenseNet and VGG, namely, the lightweight DenseNet (LDN) and the lightweight VGG (LVG), and investigated the capability of LDN and LVG in identifying *Derris trifoliate* from mangrove forests with small amounts of data. This study has verified the effect of the lightweight deep learning algorithms for the accurate detection of mangrove invasive plants *Derris trifoliate* from UAV images.

**Index Terms**—Deep learning (DL), invasive plant, mangrove, remote sensing, unmanned aerial vehicle (UAV) images.

I. INTRODUCTION

**M**ANGROVE, which refers to evergreen trees and shrubs growing in the intertidal regions of the tropics and subtropics [1], is of great ecological and economical significance [2], [3]. However, many mangrove ecosystems are now facing dangerous invaders, namely, various invasive plants in mangroves. The widespread invasive plants have posed great threats not only to the mangrove forests but also to the local biodiversity [4]. Among them, *Derris trifoliate*, a common climbing companion species in mangroves, has become one of the most notorious mangrove invaders and severely endangered mangrove habitats [5]. Therefore, the accurate and rapid identification of *Derris trifoliate* is of great significance for the protection and restoration of mangrove ecosystems.

In the past, manual investigations have been done regularly in many mangrove habitats to detect the occurrence of invasive plants and prevent the spread of them [6], [7]. However, this work is difficult to be carried out in a timely and rapid manner due to the special growth environment of mangroves, not to mention the large amount of manpower required. The remote sensing technology can provide a new way to identify various invasive plants by obtaining a wide range of Earth observation data, such as the satellite imagery [8], [9], [10]. Nevertheless, the identification results and efficiency are greatly limited by the spatial and temporal resolutions of the satellite imagery used in these studies. On the one hand, due to the low spatial resolution, it is difficult to distinguish invasive plants from native plants in satellite images, which is easy to cause missed and false detection; on the other hand, the low temporal resolution of satellite images cannot guarantee the timely detection of invasive vegetation, which may cause huge losses. Therefore, how to identify invasive plants timely and accurately in the complex mangrove ecosystem is an urgent problem to be solved.

Recently, with the advantages of flexibility, safety, low cost of data collection [11], [12], the unmanned aerial vehicle (UAV) has been widely used in many fields, such as topographic surveying, precision agriculture, and ecological monitoring [13], [14]. Working at a height of ten to several hundred meters, the UAV can detect centimeter-level details of ground objects. In addition, due to its high flexibility, the UAV can quickly obtain the information of the required target area once needed. Therefore, the UAV imagery can greatly compensate for the insufficient for the low spatial and temporal resolution of the satellite imagery [15]. In view of the advantages of UAV images, recent studies have attempted to explore the potential of UAV technology in environmental protection, including invasive plant identification [7], [16].

Although UAV images have been used more and more widely in the past few years, specialized methods for mapping mangrove invasive plants, especially the *Derris trifoliate*, based on UAV
images are still rarely studied. The traditional method and the latest deep learning (DL) method have achieved good results based on satellite images, but their complex network structures usually require large computing power and memory consumption, which are unaffordable on UAVs. Therefore, in order to satisfy the real-time monitoring of mangrove forests, how to build more efficient models for UAVs is an urgent problem to be solved.

Meanwhile, the identification of *Derris trifoliate* is also limited by insufficient samples. Most DL models need enough samples to ensure that the models are fully trained to obtain good performance. However, there are few datasets available for *Derris trifoliate* at present. Owing to the unique growth environment of mangroves, it is really a high cost to collect adequate image data. In addition, since *Derris trifoliate* usually attaches to the canopy of mangrove plants and has a similar appearance to mangrove plants, it is difficult to distinguish and annotate *Derris trifoliate* samples. Therefore, how to solve the problem of limited training samples is the key to accurately identify *Derris trifoliate*.

For the above problems, this study aims to explore the potential of UAV images and lightweight DL models in identifying *Derris trifoliate*. Specifically, we proposed two lightweight models based on VGG and DenseNet, namely, lightweight DenseNet (LDN) and lightweight VGG (LVG) for the identification of *Derris trifoliate*. Besides, in order to support model training, a dataset of *Derris trifoliate* in mangrove forests has been collected and annotated, which contains a total number of 800 samples. Comparative experiments have been conducted on our dataset among the proposed lightweight models and the bag of visual words (BoVW) model, a traditional machine learning algorithm, which have proved the effectiveness of the LDN and LVG. With the advantage of fewer parameters, the LDN and LVG models can not only be trained with a few annotated samples, but also be implemented to UAVs, due to the great improvements on computing and memory efficiency. Moreover, different feature combination strategies were compared to explore the assistance of features extracted from UAV images (such as Gabor and HSV), to identify *Derris trifoliate*.

The contribution of this article can be summarized in the following three parts.

1) A dataset based on UAV imagery for invasive plant *Derris trifoliate* in mangrove forests has been provided to support DL model training.
2) Two lightweight DL networks, LDN and LVG, have been constructed for the real-time monitoring of *Derris trifoliate* on UAV platforms.
3) Comparative experiments have been conducted on the proposed dataset and networks, which proved the effectiveness of the LDN and LVG on *Derris trifoliate* identification from UAV images.

The rest of this article is organized as follows. Section II introduces related works. Section III illustrates the study area and materials. Section IV describes the detail information of the proposed methods. The experiments and results are demonstrated in Section V. Discussions are presented in Section VI. Finally, Section VII concludes this article.

II. RELATED WORKS

A. Conventional Methods

Conventional invasive plant identification methods usually solved the problem by constructing linear regression model [19] or spectral index [9] from spectral data and using machine learning to classify invasive plants, in which nearest neighbor [16], support vector machine (SVM) [16], random forest [20], and BoVW [21] were mostly used. Developed from machine learning, DL algorithms can generally reach much better performance than conventional methods, by constructing more complex architectures and automatically extracting deeper feature information [22], which have received more attention these years.

B. DL Methods

A convolutional neural network (CNN) is the most commonly used model in DL, such as VGG-Net [24], AlexNet [25], and ResNet [26]. In recent years, some new CNN models such as DenseNet [27] were proposed and also widely used in many fields, including land use classification [28], image scene classification [29], [30], and road extraction [31]. DL models have the advantages of strong robustness, high adaptability, high detection usability, and so on [32]. Previous studies have attempted to identify invasive species based on DL models [33], [34]. However, the common DL network occupies too much memory and has low computational efficiency, making them hard to be implemented to mobile platforms for real-time monitoring [35], [36]. Moreover, DL algorithms also rely on sufficient labeled samples to optimize a large number of parameters and avoid overfitting. However, the collection of enough data is often extremely hard and expensive. Hence, how to reduce the dependence of the model on the number of samples and improve the efficiency is a problem worth discussing.

III. STUDY AREA AND MATERIALS

A. Study Area

The Qi’ao Island Mangrove Natural Reserve (22°23’N-22°27’N, 113°36’E-113°39’E), located in the northeast of Zhuhai City, Guangdong Province (see Fig. 1), is one of the largest artificial mangrove forests in China. Owing to the coastal conditions and the southern subtropical marine climate, the reserve has become a good ecological habitat for mangrove plants and an important mangrove gene bank in the Pearl River Delta.

However, the *Derris trifoliate* also grows vigorously on a large scale in the Reserve, which attaches to the trunk and covers the canopy of mangrove plants to hinder their photosynthesis [37], leading to the lack of light and death of the mangrove plants (see Fig. 2). The *Derris trifoliate* has seriously affects the growth of mangrove plants and damages the value of mangrove forests [38]. Therefore, the Qi’ao Island is selected as the study area in this article to learn and identify *Derris trifoliate* from the mangrove forests.
The UAV images were obtained by DJI UAV equipped with the RGB camera [see Fig. 3(a)] on June 28, 2018 and October 28, 2018. The equipment parameters of the camera are as follows: the focal length is 8.8 mm, the pixel number is $2.0 \times 10^7$, and the pixel size is 2.412 $\mu$m. In this study, we collected the UAV images in seven subsets of our study site, named A1–A3 and B1–B4 subsets. Each subset had an approximate size of 5400×3000. The A1–A3 sample subsets are used to select training and testing samples, and the B1–B4 sample subsets are applied to test the generalization ability of our proposed DL algorithms.

Given the UAV images, the actual distribution area of Derris trifoliate in the subset areas was manually annotated by visual interpretation by experts with rich field experiences. Then, the training samples were selected with a window of 150×150 pixels from the A1–A3 sample subsets. When the pixels labeled as Derris trifoliate account for more than 10% in a patch, the patch will be classified as “Derris trifoliate”; otherwise, it will be classified into “other plants.” Finally, a total number of 400 samples were collected, including 150 samples of Derris trifoliate and 250 samples of other plants in the dataset. Some example samples have been demonstrated in Fig. 4. In addition, considering that DL model training requires sufficient samples, data enhancements including flip and rotation were applied to increase and augment training samples.

IV. METHODOLOGY

To meet the computing and storage efficiency of real-time monitoring on Derris trifoliate, this study proposed two lightweight DL algorithms: LDN and LVG. The detail structure of the LDN and LVG will be introduced in the following.
Fig. 4. Examples for samples selected from the UAV images. (a) Image samples of *Derris trifoliate*. (b) Image samples of other plants (including *Kandelia candel*, *Acanthus ilicifolius*, *Aegiceras corniculatum*, etc.).

Fig. 5. Structure of a DenseBlock with three layers of dense connection. In the DenseBlock of DenseNet, BN-reLu-Conv represents BN-reLu-1×1 Conv–BN-reLu-3×3 Conv, denoted as 1×1 conv–3×3 conv.

A. Lightweight DenseNet

The convolutional layer (Conv) in DenseNet was connected behind the batch normalization (BN) and rectified linear unit (ReLU). The Conv was used to extract local information of the image by traversing the whole image with a convolution kernel. The basic structure of DenseNet generally includes the DenseBlock and several transition layers. Each layer in the DenseBlock accepted all the feature maps of the \(x_0\) to \(x_{l-1}\) layers, so that any two layers were used to build a densely connected structure (see Fig. 5). The 1×1 conv and 3×3 conv connect to form the bottleneck layer in the DenseBlock of DenseNet, so the layer mentioned below refers to the bottleneck layer. The output of the \(l\)th layer can be defined as

\[
    x_l = H_l([x_0, x_1, \ldots, x_{l-1}])
\]

where \(x_0, x_1, \ldots, x_{l-1}\) denotes the feature map of the zeroth layer to the \(l\)th layer, [...] denotes the concatenation operation of this layers, and the specific structure of \(H_l\) is 1×1 conv–3×3 conv of the \(l\)th layer.

The feature maps were repeatedly used in the densely connected structure to extract deep features, so that this structure can help to learn more complex patterns. In addition, the direct connections between all layers are conducive to the flow of gradients and information, which can effectively avoid the gradient vanishing when the network layer is deep, and has strong generalization and antioverfitting ability when the training data were scarce [39]. The densely connected structure can also help to fit the data better when there are fewer feature maps learned at each layer. Hence, it only needs fewer parameters and can reduce the redundancy of the structure [31]. The structure of 1×1 conv–3×3 conv also greatly reduced the number of parameters [26] and improved the computational efficiency [27]. The number of output features of each layer in the DenseBlock denoted the growth rate. The transition layer consisted of a 1×1 conv and a 2×2 average pool. This structure can be used to reduce the number of feature layers and the size of the feature map.

As shown in Fig. 6, the LDN first utilized a 7×7 Conv to capture the low-level features of the input image, followed by a 3×3 max pooling layer. Second, three DenseBlocks and two transition layers were applied to extract high-dimensional features. Then, after a global average pool layer, the output was flattened into a vector. Finally, a fully connected layer was employed to obtain the probability. In this study, the growth rate of LDN was set to 4. Each layer in the DenseBlock of LDN still used the structure of 1×1 conv–3×3 conv. The number of convolution kernels in 1×1 conv was equal to twice the growth rate, and the number of convolution kernels in 3×3 conv was equal to the growth rate. The number of feature layers was expanded by 1×1 conv to be equal to twice the growth rate, and the number of feature layers was compressed by 3×3 conv to be equal to the growth rate. The numbers of the output feature layers of each DenseBlock were equal to the growth rates. The specific network parameters of LDN were shown in Table I.

B. Lightweight VGG

The VGG is a classic convolutional neural network with simple network structure. VGG can decrease the amount of parameters and deepen the network depth, which is conducive to the extraction of deep features [24]. The block of VGG is made up of several 3×3 convolution layers, which replaces a convolution layer with larger convolution window. The proposed LVG maintains the ratio of layers and the characteristics of VGG (see Fig. 7). The LVG was composed of five blocks, and the numbers of convolutional layers in each block were set to 2, 2, 3, 3, and 3, respectively. The LVG employed the same number

| Layers       | LDN                  |
|--------------|----------------------|
| Convolution  | 7×7 conv, stride 2   |
| Pooling      | 3×3 max pool, stride 2 |
| DenseBlock(1)| 1×1 conv, 3×3 conv × 2 |
| Transition Layer(1) | 1×1 conv |
| DenseBlock(2)| 1×1 conv, 3×3 conv × 4 |
| Transition Layer(2) | 1×1 conv, 3×3 conv × 2 |
| DenseBlock(3)| 1×1 conv, 3×3 conv |
| Classification Layer | global average pool |

2-D fully connected, softmax
of convolution kernels in the convolutional layers in the same block. For each block, the number of convolution kernels was set to 3, 4, 6, 6, and 8, respectively, and the number of output feature layers of each block was set to 3, 4, 6, 6, and 8, respectively. The window size of all the convolutional layers in the LVG was set to 3×3. A max pooling layer with the size of 2×2 was connected to each block to reduce the size of the feature image following each block. Finally, the classification result was obtained by two fully connected layers. Each convolutional layer of the LVG consisted of Conv, BN, and reLu. Two neurons were utilized in the softmax layer to determine whether an input image belongs to *Derris trifoliate* or other plants.

### V. Experiments and Results

#### A. Experimental Settings

1) **Configurations:** The size of input subimages for the model was set to 150 × 150. After data augmentation, we used 800 samples to train the model and test the identification performance in the subsets. In addition to the proposed LDN and LVG, the BoVW model, a traditional machine learning algorithm, was also adopted for comparison, which has been widely used in weed species classification [21]. During the training process, an initial learning rate of 0.01 was employed, with an Adam optimizer to help faster model convergence. The batch size of 50 was adopted for each iteration, and a total number of 180 iterations were applied. After that, fivefold cross-validation was used to compare model performance.

2) **Bag of Visual Words:** BoVW consists of the scale-invariant feature transform (SIFT) and the SVM classifier. The extraction of the SIFT feature is the key step of BoVW. The SIFT feature vector generally has 128 dimensions [40], and it has good stability on the transform, rotation, scale scaling, brightness change, visual change, and noise [41], [42]. In this study, SIFT features were first extracted from each subimage in the training sample subsets. Then, all the SIFT features were clustered using the K-means algorithm. Each cluster center represents a visual word, N cluster centers represent N visual words and form a visual dictionary of N dimensions, and N was set to 200, 500, and 1000 in turn. The SIFT feature was marked as the visual word with the highest similarity, using Euclidean distance to measure the similarity. The ith (i ∈ [1, N]) visual words correspond to x, SIFT features, and each subimage corresponding to the visual dictionary will generate an N-dimensional feature vector (x1, . . . , xN). The feature vector was normalized and constituted the visual features of the subimage. Finally, the SVM classifier was used for image classification based on the visual dictionary.

3) **Image Feature Extraction:** Considering that the input image features will have an impact on the effectiveness of the DL network, this study compared different features extracted from UAV images to identify *Derris trifoliate*. For the high-resolution UAV images, in addition to spectral band features, the hue, saturation, value (HSV) color space and Gabor features are commonly used for information extraction and classification [43], [44]. The Gabor features generally represent texture information, which include the local characteristics of different spatial frequencies, positions, and directions by setting different parameters [45], [46]. Besides, the Gabor features are not sensitive to the brightness and posture changes [47]. In this study, we employed the HSV features to enhance the color information of UAV image for *Derris trifoliate* identification. The addition of the Gabor features can enhance the geometric information of

![Fig. 6. Network structure of the proposed LDN. The probability of the input image being categorized into *Derris trifoliate* and other plants.](image1)

![Fig. 7. Network structure of the proposed LVG. Different colors represent different numbers of kernels used in each convolution. Each convolution is made up of 3×3 Conv-BN-reLu in the LVG.](image2)
UAV image, so that the efficiency of spatial feature extraction by the CNN can be improved, and its dependence on training samples can be reduced [46]. In this study, we combined the HSV and Gabor features with three spectral band features and formed three feature combinations, including RGB+Gabor, RGB+HSV, and RGB+HSV+Gabor. Three feature combinations were, respectively, selected as input features to train the proposed lightweight DL networks and used to compare the accuracy of the *Derris trifoliate* identification. The strategy of using feature extraction is applied in an end-to-end way.

4) Accuracy Evaluation: Three metrics including the overall accuracy (OA), producer’s accuracy (PA), and Kappa coefficient were used for quantitative evaluation in this study. OA refers to the proportion of correctly classified samples to all the samples. PA reflected the omission of *Derris trifoliate*. The higher PA is, the fewer omission of *Derris trifoliate* is. The Kappa coefficient is used to measure the consistency between the prediction results and the ground truths, whose value ranges from −1 to 1. The formula of OA, Pa, and Kappa can be denoted as

\[
OA = \frac{TP + TN}{TP + TN + FP + FN} (2)
\]

\[
PA = \frac{TP}{TP + FN} (3)
\]

\[
Kappa = \frac{(TP + FN) \times (TP + FP) + (FP + TN) \times (FN + TN)}{(TP + TN + FP + FN)^2} (4)
\]

where TP, TN, FP, and FN represent true positive, true negative, false positive, and false negative, respectively.

**B. Identification Result of Derris Trifoliate**

Table II summarized the average OA (Ave OA) and the standard deviation (SD) of the fivefold cross-validation results of each model. It can be seen that the LDN has the highest average OA of 93% but with the lowest SD of 1.6%, indicating that the LDN has the best recognition ability and stability among all models. The second-ranked model is the LVG with an average OA of 90% and an SD of 2.3%. The BoVW has the worst performance, with only 84% OA on average, and a relatively large SD of 6.7%. The parameters of the LVG and LDN were also provided in Table II for more comprehensive analysis, which refers to the total number of parameters need to be trained in the network. Both the LVG and LDN have shown very low number of parameters, which demonstrates their advantages in memory occupation and proves their potential in fast and real-time detection for *Derris trifoliate*.

In the study site, the canopy closures of *Derris trifoliate* and mangrove stands were high. In addition, the morphological characteristics of *Derris trifoliate* and mangrove stands were similar, which made it very difficult to extract *Derris trifoliate* from mangrove forests. As shown in Table III, the highest identification accuracy of BoVW was 85% in the A2 subset, while for two DL algorithms, the identification accuracy of LDN and LVG reached 93% and 90%. For the A1–A3, B1, and B4 sample subsets, the identification accuracy of the LDN was higher than that of LVG, while for the B2 and B3 sample subsets, the accuracies of LVG and LDN were close. The LDN has higher identification accuracy and more stable performance than that of LVG. As shown in Fig. 8, for the identification result of BoVW in most of the sample subsets, there was a large area of misclassification of the *Derris trifoliate*. Compared to BoVW,
the identification results of both LDN and LVG could effectively identify *Derris trifoliate* and accurately map the growth areas of *Derris trifoliate* in the sample subsets. Compared to LVG, the area of LDN-based identification result was closer to the actual distribution of *Derris trifoliate*, and the LDN was less likely to misclassify other plants into *Derris trifoliate*.

Nevertheless, for the areas with poor light conditions and severe confounding, the LDN and LVG were also prone to misclassification. For the A1, A3, and B1 sample subsets with shadow areas, there was less misclassification of LDN than LVG. As shown in Fig. 8, there is a large area of the other plants wrongly identified as *Derris trifoliate* by using LVG, while this phenomenon was alleviated by the LDN and the identification accuracies were significantly improved.

### C. Exploration of Feature Combination

Considering that the identification results of the LDN were better and more stable than those of LVG, this study used the LDN to further analyze the availability of different image features, including the RGB features, the derivative Gabor, and HSV features. For LDN, the OA, Kappa coefficient, and PA of *Derris trifoliate* identification with the three feature combinations were shown in Table IV. As shown in Table IV, the identification accuracies of the RGB+HSV features were slightly higher than those of the RGB+Gabor features and RGB+HSV+Gabor features for all the sample subsets except the A2 sample subset. The PAs of RGB+HSV features were lower than those of RGB+Gabor features and RGB+HSV+Gabor features. The results showed that the RGB+Gabor features and RGB+HSV+Gabor features have stronger recognition ability than RGB+HSV features, and the addition of Gabor features could decrease omissions in the *Derris trifoliate* identification results. The PAs of *Derris trifoliate* identification with the RGB+Gabor and RGB+HSV+Gabor features were improved in all the sample subsets, compared with those of RGB features. Among them, the RGB+Gabor features gave the highest PA for each sample subset. While for the RGB+HSV features, the PAs of *Derris trifoliate* identification were only increased in the B1–B4 sample subsets. This result indicated that the addition of Gabor features is beneficial to decrease the omissions of *Derris trifoliate*.

![Identification results of *Derris trifoliate* using LDN for the A1–A3 and B1–B4 sample subsets with different features.](image)

![The area identified as *Derris trifoliate* by model](image)

The identification results of *Derris trifoliate* areas in all the sample subsets are shown in Fig. 9. For the B1–B4 sample subsets, the identification accuracies of the RGB+HSV features were higher than the original RGB features, while for the A1–A3 sample subsets, in which the train samples were collected, the OA and the Kappa coefficient of the RGB+Gabor, RGB+HSV and RGB+HSV+Gabor features have not been significantly improved. The addition of HSV features improved the

### TABLE IV

Table showing the accuracies of the LDN model using different feature combinations on sample subsets.

| Feature       | Metric | A1  | A2  | A3  | B1  | B2  | B3  | B4  |
|---------------|--------|-----|-----|-----|-----|-----|-----|-----|
| RGB           | OA     | 84% | 93% | 89% | 82% | 91% | 83% | 84% |
|               | Kappa  | 0.64| 0.74| 0.74| 0.53| 0.66| 0.61| 0.67|
|               | PA     | 71% | 68% | 77% | 75% | 57% | 78% | 77% |
| RGB+Gabor     | OA     | 86% | 92% | 89% | 86% | 93% | 86% | 85% |
|               | Kappa  | 0.57| 0.68| 0.73| 0.58| 0.69| 0.64| 0.69|
|               | PA     | 53% | 60% | 75% | 80% | 61% | 81% | 86% |
| RGB+HSV       | OA     | 86% | 93% | 86% | 78% | 93% | 83% | 84% |
|               | Kappa  | 0.63| 0.74| 0.68| 0.45| 0.75| 0.60| 0.67|
|               | PA     | 76% | 70% | 87% | 89% | 82% | 85% | 80% |

The area identified as *Derris trifoliate* by model. The area labelled manually as *Derris trifoliate*.
Derris trifoliate is necessary for the investigation were clearer, and the mixing degrees between Derris trifoliate and its surrounding plants were not obvious, the addition of HSV and Gabor features led to reduced OA and Kappa coefficient. For the A1 sample subset in a complex scene where the canopy structure characteristics of Derris trifoliate and its surrounding plants were lower. This result further proved that the Gabor features can be used to extract canopy structure characteristics, which were helpful to accurately identify Derris trifoliate from mangrove stands.

VI. DISCUSSION

A. Advantages of High-Spatial-Resolution UAV Images

Timely and effective monitoring of mangrove invasive plant is important for biodiversity and ecosystem balance. The accurate and rapid identification of invasive plant from UAV platforms is useful for mangrove monitoring and management [13]. Experiments performed on synthetic features showed that the proposed lightweight DL method is more accurate and robust, with a higher PA than that on UAV RGB features. This performance was superior to that of the traditional machine learning method to which it was compared. BoVW was clearly less robust, as its identification accuracy was dependent on illumination. The DL method presented in this article provides an approach to invasive plant identification that effectively and efficiently addresses the shadow problem. The method presented here was applied over a relatively limited area measured by an RGB camera in a very short time. These pieces of evidence make the UAV feasible in limited area. However, we noticed that many factors might negatively influence the performance, such as shadow discussed in the following sections. UAV-based invasive plant identification has also been used to monitor vegetation conditions. This is the first study aimed at identifying the Derris trifoliate patches in mangrove forests. Compared with Spartina alterniflora [59], which belongs to the climbing vine, the identification of invasive mangrove species from mangrove forests has always been difficult and is the focus of research.

Mangroves are mainly distributed in the tidal flats of the land–sea junction. Compared to traditional field surveys, low-altitude UAV remote sensing can provide quickly and accurately capture the distribution and intrusion status of mangroves. However, it should be noted that because the Derris trifoliate is a climbing vine, the leaves are similar to mangroves, which will cause confusion in the recognition results. The use of drone remote sensing to identify Derris trifoliate is necessary for the investigation and regular maintenance of mangrove resources in the study area. The flexibility of the UAV was exploited in a novel way to map invasive plant from close-range sensing images. This proposed approach successfully reduced the shadow problem in close-range images and overcame the limitations of commercial RGB cameras, by combining the RGB images with the HSV color space.

The distribution of Derris trifoliate extracted from proximal sensing image samples may not accurately describe the whole scene. This limitation is due to surface spatial heterogeneity, limited spectral quantification information from RGB images, and variations of environment-dependent and camera-related factors. These distributions can be derived from fine-resolution RGB images collected at several sampling points according to the procedure described above.

B. Improvement of Lightweight DL for Invasive Plant Identification

DL methods have significant advantages for the information extraction of high-resolution remote sensing data, and they have been widely employed to the UAV images. These two developed DL methods of LDN and LVG in this study provided higher identification accuracy of Derris trifoliate. Compared with traditional methods, DL methods can extract deep features to identify ground objects in complex scenes. For local-scale intrusion recognition research, the amount of sample data usually available is small, and complex deep models are prone to overfitting. For intrusion research, rapid identification is the key, so it is necessary to select a highly efficient DL model. The LDN and LVG show significant advances on traditional machine learning BoVW to produce the accurate identification of Derris trifoliate. With the limited training samples, the lightweight DL methods can achieve satisfactory results in the recognition of Derris trifoliate even for the shadow areas. The lightweight optimization strategy proposed in this study can ensure higher classification accuracy, while requiring less calculation time and reducing dependence on samples and experimental equipment; this is consistent with the existing literature research. Compared with recent DL methods, the LDN and LVG can quickly and accurately identify Derris trifoliate with shorter calculation time and less calculation amount, which is of great significance for mangrove invasive plant investigations.

In this study, UAV remote sensing is used to establish a sample set of Derris trifoliate. Based on UAV high-resolution images and lightweight DL, the rapid monitoring of the intrusion status of Derris trifoliate is realized. The lightweight method proposed in this article optimizes the overfitting problem in DL algorithms. This is confirmed by our results, in that combining spectral information with textural and structural information derived from UAV high-resolution imagery clearly affected the invasive plant mapping result and reduced the omission of Derris trifoliate. Moreover, this study showed that shadows have a severe impact on the accuracy at the UAV scale, especially for the structurally complex canopies of mangroves.

VII. CONCLUSION

Mangrove forests suffer severely from the influence of invasive species, which greatly limits its ecosystem balance and
biodiversity. This study proposed two lightweight DL models (LDN and LVG) to detect and delineate invasive small *Derris trifoliate* patches from high-resolution UAV images. By comparing with the conventional machine learning algorithm for *Derris trifoliate* identification, we found that the DL-based LDN and LVG algorithms were superior in constructing spectral and spatial details from UAV images. Compared with LVG, the LDN has obtained the highest accuracy of 93% on *Derris trifoliate* identification. And the *Derris trifoliate* patches delineated by the LDN agreed well with the ground truth. The results have demonstrated the effectiveness and stability of the LDN for *Derris trifoliate* recognition especially in complex scenes. The experiments have also verified the effectiveness of the use of high-resolution UAV images for monitoring the *Derris trifoliate* in the mangrove forest with lightweight DL models. With high-resolution UAV images becoming increasingly available, the proposed models can be promising means to facilitate the timely monitoring of *Derris trifoliate* invasion, which will further benefit effective intervention and protection of invasion plants.

ACKNOWLEDGMENT

The authors would like to especially thank Kai Liu, Jingjing Cao, and Qian Shi for their technical guidance and Yuanhui Zhu, Hui Gong, Deqiang Cai, and Zhiquan Liu for their help in the data acquisition of UAV images and field investigations.

REFERENCES

[1] W. Jia, D. Mao, and Y. Zhang, “A new vegetation index to detect periodically submerged mangrove forest using single-tile Sentinel-2 imagery,” *Remote Sens.*, vol. 11, no. 17, 2019, Art. no. 2043.

[2] S. Bayen, O. Wurli, S. Karupiah, N. Sivasothi, H. K. Lee, and J. P. Obbard, “Persistent organic pollutants in mangrove food webs in Singapore,” *Chemosphere*, vol. 61, no. 3, pp. 303–313, 2005.

[3] M. Jia, Z. Wang, D. Mao, C. Ren, C. Wang, and Y. Wang, “Rapid, robust, and automated mapping of tidal flats in China using time series sentinel-2 images and Google Earth Engine,” *Remote Sens. Environ.*, vol. 253, 2021, Art. no. 112285.

[4] H. Carter, S. Schmidt, and A. Hirons, “An international assessment of mangrove management: Incorporation in integrated coastal zone management,” *Diversity*, vol. 7, no. 2, pp. 74–104, 2015.

[5] X. Huang, “Mangrove forest threatened by derris trifoliate,” *Wetland Sci. Manage.*, vol. 11, no. 2, pp. 26–29, 2015.

[6] D. Huang, X. Sun, X. Guo, and Z. Gou, “Risk analysis of mangrove associate plant derris trifoliate in Hainan,” *Trop. Forestry*, vol. 47, no. 3, pp. 62–65, 2019.

[7] P. Dro et al., “Unmanned aerial vehicles for alien plant species detection and monitoring,” *Int. Arch. Photogrammetry Remote Sens. Spatial Inf. Sci.*, vol. 40, pp. 83–90, 2015.

[8] C.-Y. Huang and G. Asner, “Applications of remote sensing to alien invasive plant studies,” *Sensors*, vol. 9, no. 6, pp. 4860–4889, 2009.

[9] M. R. Mahmud, S. Numata, and T. Hosaka, “Mapping an invasive golden saltcedar (Tamarisk spp.) distribution mapping using multiresolution remote sensing imagery,” *Sensors*, vol. 13, no. 4, pp. 1738–1764, 2014.

[10] A. Krizhevsky, I. Sutskever, and G. Hinton, “ImageNet classification with deep convolutional neural networks,” in *Adv. Neural Inform. Process. Syst.*, vol. 25, pp. 1097–1105, 2012.

[11] J. Tian et al., “Development of spectral-phenological features for deep learning to understand Spartina alterniflora invasion,” *Remote Sens. Environ.*, vol. 7, no. 6, pp. 2094–2107, Jun. 2014.

[12] J. Zhang, C. Lu, X. Li, H.-J. Kim, and J. Wang, “A full convolutional network based on DenseNet for remote sensing scene classification,” *Math. Biosci. Eng.*, vol. 16, no. 5, pp. 3345–3367, 2019.

[13] Y. Xu, Z. Xie, Y. Feng, and Z. Chen, “Road extraction from high-resolution remote sensing imagery using deep learning,” *Remote Sens.*, vol. 10, no. 9, 2018, Art. no. 1461.

[14] X. Liu, “Research progress of optical remote sensing image object detection based on deep learning,” *Acta Autom. Sinica*, vol. 47, pp. 2078–2089, 2021.

[15] Z. Sun, W. Jing, and X. Qiao, “Identification and monitoring of Mikania micrantha outbreak points based on UAV remote sensing,” *Trop. Geograph.*, vol. 39, no. 1, pp. 482–491, 2019.

[16] J. Tian et al., “Development of spectral-phenological features for deep learning to understand Spartina alterniflora invasion,” *Remote Sens. Environ.*, vol. 242, 2020, Art. no. 111745.

[17] M. Siam, M. Gamal, M. Abdel-Razek, S. Yogamani, M. Jagersand, and H. Zhang, “A comparative study of real-time semantic segmentation for autonomous driving,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2018, pp. 587–597.

[18] M. Tremel et al., “Speeding up semantic segmentation for autonomous driving,” 2016.

[19] Y. Huang, “Present situation and management countermeasures of mangrove wetland protection in Dongzhaiang,” *Trop. Forestry*, vol. 45, no. 2, pp. 49–52, 2017.

[20] S. Wang, Y. Zheng, Y. Peng, and G. Chen, “Health assessment of Qia’o Island mangrove wetland ecosystem in Pearl River Estuary,” *Chin. J. Appl. Ecol.*, vol. 21, no. 2, pp. 391–398, 2010.
J. Dolz, K. Gopinath, J. Yuan, H. Lombaert, C. Desrosiers, and I. B. Ayed, “HyperDense-Net: A hyper-densely connected CNN for multi-modal image segmentation,” *IEEE Trans. Med. Imag.*, vol. 38, no. 5, pp. 1116–1126, May 2019.

D. G. Lowe, “Object recognition from local scale-invariant features,” in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, 1999, pp. 1150–1157.

F. Dellinger, J. Delon, Y. Gousseau, J. Michel, and F. Tupin, “SAR-SIFT: A sift-like algorithm for SAR images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 453–466, Jan. 2015.

D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

M. Wu, R. Chen, and Y. Tong, “Shadow elimination algorithm using color and texture features,” *Comput. Intell. Neurosci.*, vol. 2020, 2020, Art. no. 2075781.

R. Cucchiara, C. Grana, M. Piccardi, A. Prati, and S. Sirootti, “Improving shadow suppression in moving object detection with HSV color information,” in *Proc. IEEE Intel1. Transp. Syst.*, 2001, pp. 334–339.

T. S. Lee, “Image representation using 2D Gabor wavelets,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 10, pp. 959–971, Oct. 1996.

Y. Chen, L. Zhu, P. Ghamisi, X. Jia, G. Li, and L. Tang, “Hyperspectral images classification with Gabor filtering and convolutional neural network,” *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2355–2359, Dec. 2017.

S.-S. Liu, Y.-T. Tian, and C. Wan, “Facial expression recognition method based on Gabor multi-orientation features fusion and block histogram,” *Acta Autom. Sinica*, vol. 37, no. 12, pp. 1455–1463, 2011.

**Mengxi Liu** (Graduate Student Member, IEEE) received the B.S. degree in geographic information science from Sun Yat-sen University, Guangzhou, China, in 2019, and the M.S. degree in cartography and geographic information systems from the School of Geography and Planning, Sun Yat-sen University, in 2022. Her research interests include intelligent understanding of remote sensing images, change detection, and domain adaptation.

**Haojun Deng** received the B.S. degree in geographic information science from Sun Yat-sen University, Guangzhou, China, in 2020, and the M.S. degree in cartography and geographic information systems from the School of Geography and Planning, Sun Yat-sen University, in 2022. His research interests include machine learning and deep learning in phenology and urbanization.

**Wen Dong** received the B.S. degree in environment engineering from China Agricultural University, Beijing, China, in 2005, and the Ph.D. degree in cartography and geographical information system from the Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences (CAS), Beijing, in 2010. She is currently a Research Associate with the State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, CAS. Her research interests include geographic spatiotemporal data mining and application of geographic information systems.