Mathematical Biosciences and Engineering

MBE, 18(2): 1121–1135.
DOI: 10.3934/mbe.2021060
Received: 29 August 2020
Accepted: 14 December 2020
Published: 11 January 2021

http://www.aimspress.com/journal/MBE

Research article

Efficiently deep learning for monitoring *Ipomoea cairica (L.) sweets* in the wild

Fei Tang[1,2], Dabin Zhang[1,*] and Xuehua Zhao[2]

[1] College of Mathematics and Information, South China Agricultural University, Guangzhou 510642, China
[2] Shenzhen Institute of Information Technology, Shenzhen 518172, China

* **Correspondence:** E-mail: zdbff@scau.edu.cn.

Abstract: *Ipomoea cairica (L.) sweets* are an invasive weed which has caused serious harm to the biodiversity and stability of the ecosystem. It is very important to accurately and rapidly identifying and monitoring *Ipomoea cairica (L.) sweets* in the wild for managements taking the necessary strategies to control the *Ipomoea cairica (L.) sweets* to rapidly grow in the wild. However, current approaches mainly depend on manual identification, which result in high cost and low efficiency. Satellite and manned aircraft are feasible assisting approaches, but the quality of the images collected by them is not well since the ground sampling resolution is low and cloud exists. In this study, we present a novel identifying and monitoring framework and method for *Ipomoea cairica (L.) sweets* based on unmanned aerial vehicle (UAV) and artificial intelligence (AI). In the proposed framework, we low-costly collected the images with 8256 × 5504 pixels of the monitoring area by the UAV and the collected images are split into more small sub-images with 224 × 224 pixels for identifying model. For identifying *Ipomoea cairica (L.) sweets*, we also proposed a novel deep convolutional neural network which includes 12 layers. Finally, the *Ipomoea cairica (L.) sweets* can be efficiently monitored by painting the area containing *Ipomoea cairica (L.) sweets*. In our experiments, we collected 100 raw images and generated 288000 samples, and made comparison with LeNet, AlexNet, GoogleNet, VGG and ResNet for validating our framework and model. The experimental results show the proposed method is excellent, the accuracy is 93.00% and the time cost is 7.439 s. The proposed method can achieve to an efficient balance between high accuracy and low complexity. Our method is more suitable for the identification of *Ipomoea cairica (L.) sweets* in the wild than other methods.

Keywords: *Ipomoea cairica (L.) sweets*; invasive weed; image identification; deep learning

1. Introduction

As a perennial vine, *Ipomoea cairica (L.) sweets* are viewed as the noxious invasive weeds in China and are called as plant killer because of heavily covering other plants in wild lands [1]. *Ipomoea cairica (L.) sweets* like climbing, can rapiddly grow, propagate, and cover in its area. Consequently, *Ipomoea cairica (L.) sweets* can result in the serious harm to the biodiversity and stability of the ecosystem [2]. How to identify and monitor the *Ipomoea cairica (L.) sweets* in the wild is becoming very important since it can help the governments to take the necessary strategies and methods in the identified area to prevent and control the growth of the *Ipomoea cairica (L.) sweets* for biodiversity and stability of the ecosystem.

Currently, identifying and monitoring *Ipomoea cairica (L.) sweets* mainly depends on manual approaches of low efficiency and high cost [3]. With the quick development of image processing and artificial intelligence (AI), it is possible to help managements to make the precise efficient faster measures by efficiently identifying and monitoring *Ipomoea cairica (L.) sweets* [4]. Hyperspectral images have been used to identify the invasive species for monitoring and control their rapidly growth. Liu, et al. used the high-resolution imagery from Google Earth to monitor the Spartina alterniflora and showed that this approach can help governments to control Spartina alterniflora invasion [5], West et al. tested five models to identify Tamarik in satellite remote sensing data [6]. Bustamante et al. used hyperspectral sensors to detect Spartina densiflora Brongn [7]. Jomar et al. used imaging spectroscopy to detect Psidium cattleianum in Hawaiian Forests [8]. Khare et al. assess invasive plant species diversity in Pléiades 1A, RapidEye and Landsat-8 data [9]. However, it is difficult for these methods to detect the *Ipomoea cairica (L.) sweets* since the high-resolution hyperspectral images collected by aircraft and satellite are very high cost. For free low-resolution spectral data, some key information in these images is lost. As a result, the hyperspectral images are not suitable to monitor *Ipomoea cairica (L.) sweets* in the early stages.

The number of study has proved that the high-resolution digital images can be used to plants identification in various environments and shows their excellent results [10], especially the high-resolution images collected by low-altitude UAV because of retained good characteristics. Ge et al. used the aerial photographs to determine the Tamarix parviflora [11]. Jones et al. used the image analysis-based approach to detect Japanese Knotweed s.l. taxa in high-resolution, infra-red aerial images [12]. Dorigo et al. proposed an identifying method based on remote sensing to find F. japonica from low-cost images [13]. Sandino et al. used unmanned aerial vehicles (UAVs) and machine learning to detect the buffel grass and spinifex [14]. Hung et al. used feature learning to classify the invasive weed species based on the images from UAVs [15]. However, the detecting accuracy need be further improved, the efficient method with high accuracy need be developed for identifying invasive weed in the wild.

Now deep learning shows the excellent ability in image recognition fields and is widely applied to the object detection in agricultural. Pound et al. demonstrated deep approaches can offer excellent ability to identify root and shoot feature [16]. Alvaro et al. presented deep learning method to identify the tomato diseases [17]. Amanda et al. presented the transfer learning and deep CNN for

pest damage and diseases detection, their results showed that their method can offer a fast, affordable method for digital disease detection [18]. Lu et al. used the deep CNN to identify the counts of maize tassels [19]. Though deep learning has widely applied to the agriculture fields, it still is difficult for identifying *Ipomoea cairica (L.) sweets* since it is very similar between *Ipomoea cairica (L.) sweets* and phorophyte.

To better address aforementioned challenges, we present a novel identifying and monitoring framework and method for *Ipomoea cairica (L.) sweets* based on unmanned aerial vehicle (UAV) and artificial intelligence (AI). In the proposed framework, we can low-costly collected the images with 8256 × 5504 pixels of the monitoring area by the UAV, and the collected images are split into more small sub-images with 224 × 224 pixels for identifying model. For identifying *Ipomoea cairica (L.) sweets*, we propose a novel deep CNN (ICSNet) which consists of 12 layers. Finally, the *Ipomoea cairica (L.) sweets* can be monitored by painting the area containing *Ipomoea cairica (L.) sweets*. Our experiment results show our method can automatically monitoring and eradicating *Ipomoea cairica (L.) sweets* in wilds.

In summary, the main contributions of this work are as follows:

1) A novel identifying and monitoring framework and method based on unmanned aerial vehicle (UAV) and artificial intelligence (AI) are proposed for Ipomoea cairica (L.) sweets

2) A deep CNN (ICSNet) is proposed to identifying Ipomoea cairica (L.) sweets, which can achieve the tradeoff between accuracy and time.

3) The framework and method are validated in our experiments and the experimental results show our method has the excellent performance.

The remainder is as follows. Section II presents our methods. Section III describes the experimental results and discussions. The conclusions are in Section IV.

2. Method

2.1. Proposed detecting framework for Ipomoea cairica (L.) sweets

For monitoring the *Ipomoea cairica (L.) sweets*, we propose a novel detecting framework for *Ipomoea cairica (L.) sweets*, which is shown in Figure 1. In our framework, monitoring the *Ipomoea cairica (L.) sweets* includes the following steps:

First step: The UAV is used to collect the images with 8256 × 5504 pixels of monitoring area, the whole images of monitoring area consists of sub-images with 8256 × 5504 pixels. The image sample is shown in Figure 1(b).

Second step: The collected images with 8256 × 5504 pixels are split into many more small images with 224 × 224 pixels. The samples are shown in Figure 1(c).

Third step: For each image with 224 × 224 pixels, we use the proposed deep CNN model to identify if the *Ipomoea cairica (L.) sweets* are contained in the images. This is corresponding to Figure 1(d).

Four step: Based on above identified results, we paint the images containing *Ipomoea cairica (L.) sweets* in red color. This is corresponding to Figure 1(e).

Five step: The painted images with 224 × 224 pixels is reassembled, as a result, the areas containing *Ipomoea cairica (L.) sweets* are covered by red color in the images with 8256 × 5504 pixels.

This is corresponding to Figure 1(f). Finally, all the *Ipomoea cairica (L.) sweets* are identified in monitoring area.

For our framework and algorithm, in fact we can review it as a special case of object detection algorithm. For R-CNN and Fast R-CNN, they usually include three main steps: determining candidate region, feature extraction and classification. But slowly detecting speed is the lack of these algorithms. For our algorithm, we improve detection speed by in advance splitting the image into the little sub-image with fixed same size. Consequently, we can achieve the required monitoring accuracy by adjusting the size of sub-image, and detection time also can meet the actual requirements. Although, the object detection algorithms such as Yolo and SSD, which achieves the classification and positioning in one step, can improve the detection speed, their detection accuracy is lower than R-CNN and Fast R-CNN.

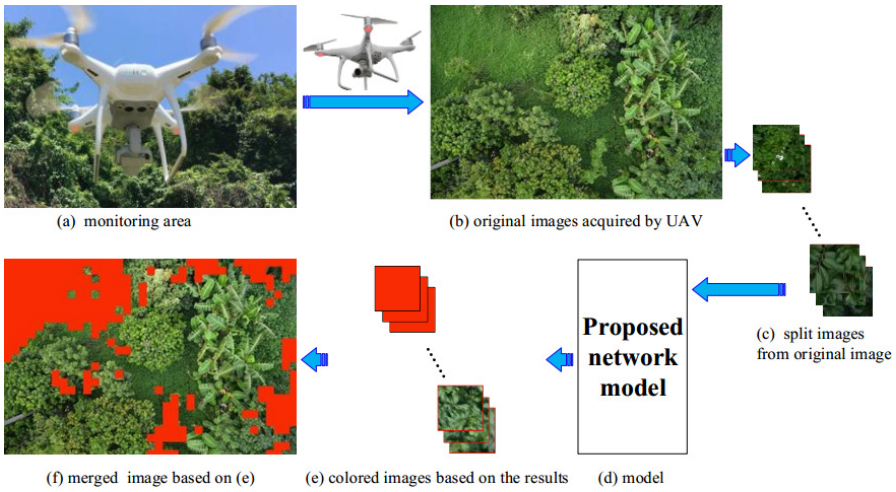


Figure 1. Proposed detecting framework for *Ipomoea cairica (L.) sweets*.

2.2. Proposed deep CNN model for Ipomoea cairica (L.) sweets

The multilayer neural network only consists of input, output and hidden layers. The actual needs decide the number of the hidden layers. Based on the original multilayer neural network, a more effective feature learning part is added in network. The part is the convolution layer and pooling layer which is set in front of fully connected layer. This is convolution neural network (CNN) which is shown in Figure 2. CNN deepen the number of layers in neural network and deep learning can be realized.

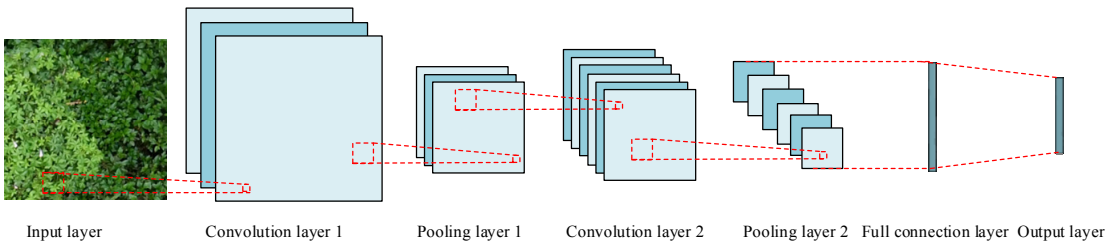


Figure 2. Convolution neural network.

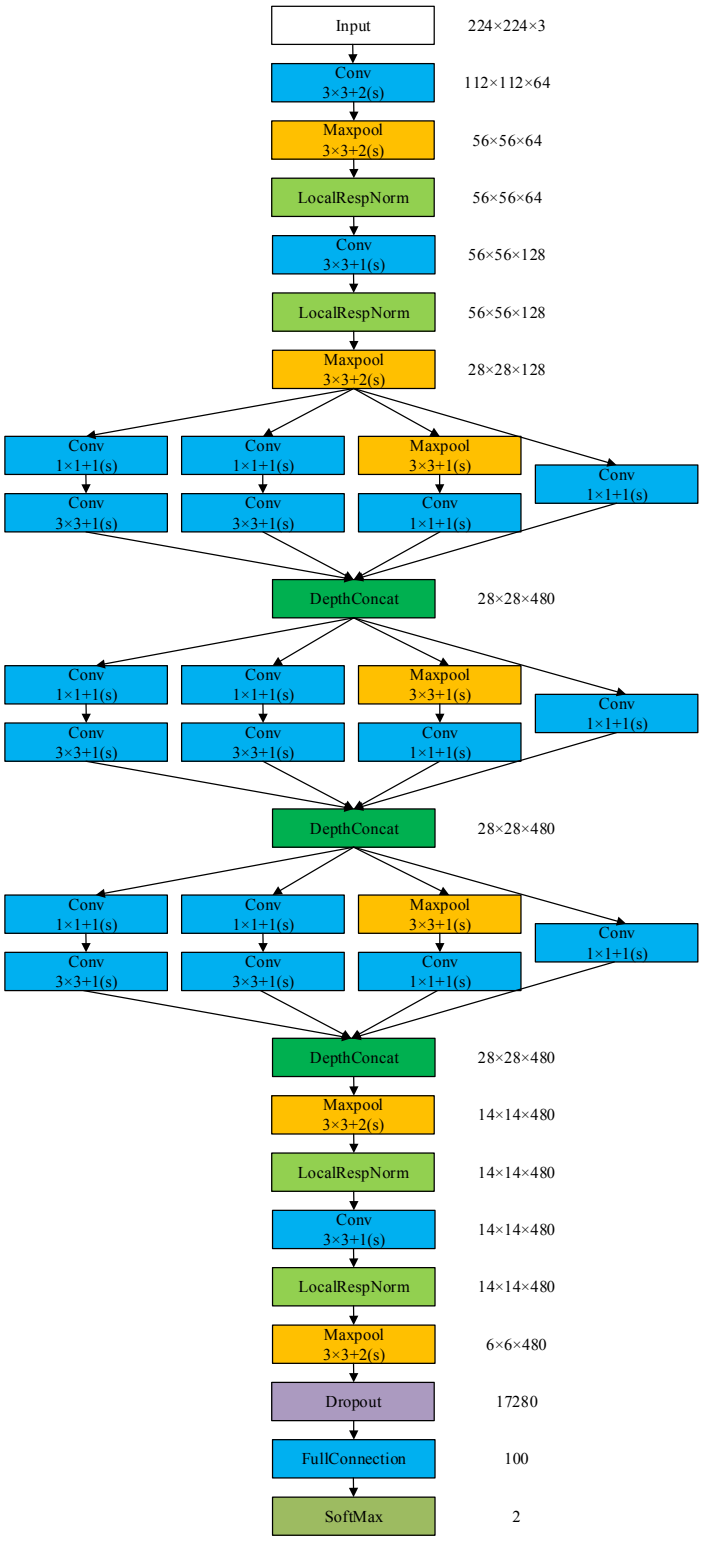


Figure 3. The architecture of ICSNet.

Here, we build a 12-layer deep CNN network model for identifying *Ipomoea cairica (L.) sweets*, which mainly consists of convolutional layer, pool layer, inception module, dropout layer, full connected layer. The input is the RGB image with the size of 224 × 224 × 3. The output consists of two

neurons. The proposed network model is named ICSNet for short. The architecture of ICSNet is illustrated in Figure 3.

First layer: It is a convolution layer, which is with 64 3 × 3 filters and the stride is 2. The ReLU (rectified linear units) are used as the activation function for improving the convergence speed and simple gradient calculation. The ReLU can nonlinearly map output results of convolution layer. The output results of second layer is the data with the size of 112 × 112 × 64.

The pooling operator is connected after this layer for extracting the main features of a certain area, reducing the number of parameters, and preventing the model from over fitting. The maximum pooling with a 2 × 2 window and 2 strides is used in this layer. The pooling result is the data with the size of 56 × 56 × 64.

After pooling, the LRN (Local Response Normalization) is used to normalize the local response. The LRN enhances the larger response value and reduce the smaller response value. Thus, the LRN also can improve the generalization of the model. The output result of the LRN is the data with the size of 56 × 56 × 64.

Second layer: It is a convolution layer, which is with 128 3 × 3 filters and the stride is 1. The ReLU is selected as the activation function. The LRN is connected after this layer, the output results is the data with the size of 56×56×128. Then above data is processed by pooling operator with a 3 × 3 window and 2 strides. The output result of this layer is the data with the size of 28 × 28 × 128.

Third to eighth layer: These layers consist of three inception modules which consist of seven convolution modules and a pooling module. They are divided into the following four branches.

The first branch try convolution with 1 × 1 window on the output result of previous layer. The 1 × 1 convolution can efficiently enhance the expression ability, and at the same time, it can increase and reduce the dimension of the output channel. All four branches of the inception module use 1 × 1 convolution to perform low-cost (much less computation than 3 × 3 cross-channel feature transformation). Then 3 × 3 convolution is connected.

In the second branch, 1 × 1 convolution is used first, then 5 × 5 convolution is connected. The third branch is to use 3 × 3 maximum pooling with 2 strides and then connect 3 × 3 convolution. The fourth branch uses 1 × 1 convolution. Finally, the four branches are merged through an aggregation operation at the end, and the output is the data with the size of 28 × 28 × 128.

After first inception modules, second and third inception modules are set in turn. The output is the data with the size of 28 × 28 × 480. Then, the pooling and LRN are performed, the data with the size of 14 × 14 × 480 is output.

Ninth layer: It is a convolution layer with 480 3 × 3 filters and their stride is 1. The ReLU is selected as the activation function. The result of this layer is the data with the size of 14 × 14 × 480. After convolution, the data is performed by LRN. Then, the maximum pooling with a 3 × 3 window and 2 strides is used. The output result is the data with the size of 6 × 6 × 480.

Tenth layer: This layer is dropout layer which is used to improve the generalization capability. The size of output result is 17,280.

Eleventh layer: This layer is the full connected layer with 100 neurons.

Twelfth layer: This layer is the Softmax activation function layer with 2 neurons.

In our ICSNet, we adopt the inception module from GoogLeNet, but only three inception modules are built in ICSNet. And we also add more convolution layers than AlextNet, but the number of convolution layers is less than VGGNet and ResNet. Consequently, the tradeoff between accuracy and time consuming can be achieved by our ICSNet. This is basic idea of our work.

3. Experiments

3.1. Data description

We use the DJI Phantom 4 Advanced UAV to collect the color images with 8256 × 5504 pixels of *Ipomoea cairica (L.) sweets* in the monitoring area. The flight altitude is approximately 15 m above the ground. In our experiments, we set the distance according to the following three rules: 1) the clear images can be acquired; 2) the identifying accuracy can meet the actual requirements; 3) the UAV can be easily controlled. So we set the value of distance 15m above the ground according to our experience. The camera angle is adjusted in horizontal 45° for good front view since *Ipomoea cairica (L.) sweets* are a climbing weed. Finally, 100 raw images with 8256 × 5504 pixels are acquired. Figure 4 shows four samples of the scaled-down raw images.

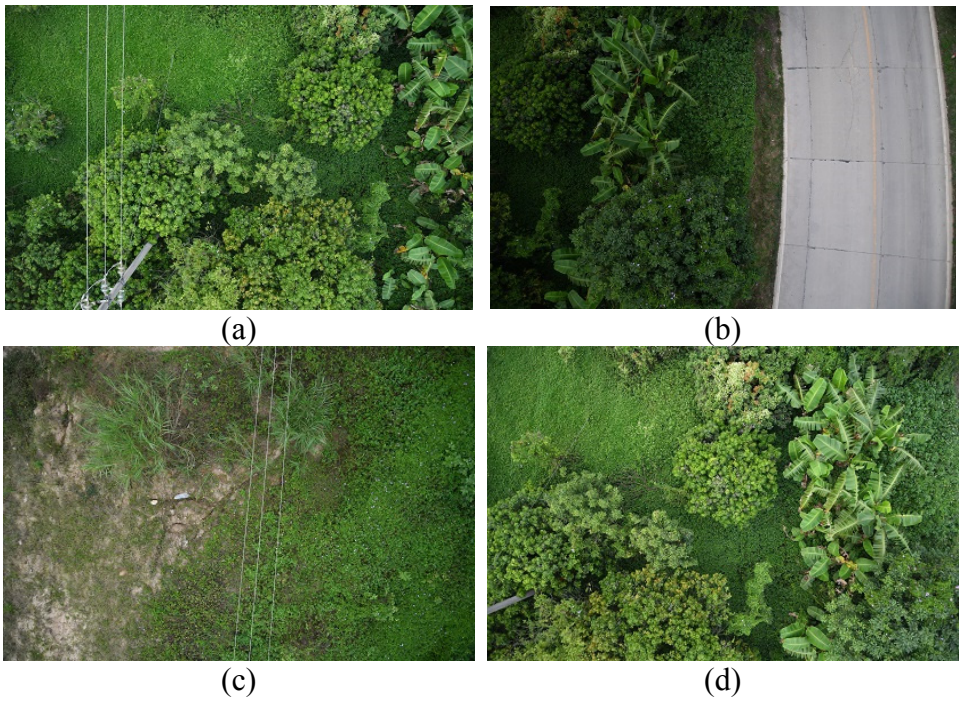


Figure 4. Four samples of the scaled-down raw images.

We produce the dataset according to the following steps:

(1) The raw images are automatically divided into small images with 224 × 224 pixels by computer program. Figure 5 shows the samples of the images. Finally, 32000 samples with 224 × 224 pixels are selected, in which 24,000 samples contain *Ipomoea cairica (L.) sweets* and 8000 samples do not contain *Ipomoea cairica (L.) sweets*.

(2) The new samples are generated by horizontally and vertically rotating sample. Accordingly, 32000 samples are expanded to 96,000 samples.

(3) The new samples are generated by subtracting the mean of R, G and B from R, G, B value of sample, and Gray conversion. Finally, 288,000 samples are obtained to form the dataset in our experiments in which 216,000 samples contain Ipomoea cairica (L.) sweets and 72,000 samples do not contain *Ipomoea cairica (L.) sweets*.

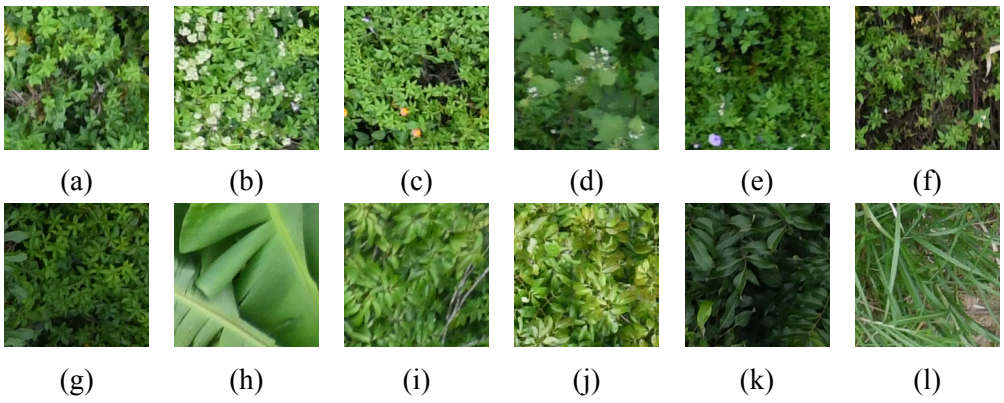


Figure 5. The sample of the images with 224 × 224 pixels. (a–g) contain *Ipomoea cairica (L.) sweets*; (h–l) do not contain *Ipomoea cairica (L.) sweets*.

3.2. Evaluation Measure

In our experiments, the Accuracy, Precision, Recall, F1, AUC, min-batch size, computing cost and the number of layers are used to validate the performance of methods.

For test samples, *TP* denotes the number of correctly identified samples including *Ipomoea cairica (L.) sweets*, *TN* denotes the number of correctly identified samples without *Ipomoea cairica (L.) sweets*, *FP* denotes the number of incorrectly identified samples including *Ipomoea cairica (L.) sweets*, *FN* denotes the number of incorrectly identified samples without *Ipomoea cairica (L.) sweets*. The Accuracy, Precision, Recall, F1 are respectively written as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \tag{1}$$

$$\text{Precision} = \frac{TP}{TP + TN} \times 100\% \tag{2}$$

$$\text{Recall} = \frac{TP + TN}{TP + FN} \times 100\% \tag{3}$$

$$\text{F}_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100\% \tag{4}$$

AUC is Area under the Curve of ROC (Receiver Operating Characteristics). The min-batch size is the number of samples in one batch. The time is the identifying time of samples. The number of layers is the number of convolutional and full connected layers, which can denote the complexity of the model. For obtaining valid results, the *k*-fold cross-validation (CV) is used to evaluate the algorithms and *k* is 10 in our experiments.

3.3. Comparative algorithms

To efficiently validate the proposed method, we make comparisons with LeNet, AlexNet, GoogLeNet, VGG-19 and ResNet, all of which are deep CNNs [20].

(1) LeNet is firstly proposed by LeCun et al. in the 1990s, which includes three convolution layers, two sampling layers, two fully connected layers. At that time, it is difficult to run LeNet due to the limit of computation and memory capacity [21].

(2) AlexNet is proposed by Alex et al. and won the ILSVRC 2012 [22]. AlexNet achieves higher identifying accuracy than all the traditional machine learning algorithms. It is significant breakthrough for machine learning for classification.

(3) GoogLeNet is proposed by Christian of Google and is the winner of ILSVRC 2014 [23], in which inception layers, including different receptive areas with different kernel sizes capturing sparse correlation patterns, are integrated into CNN.

(4) VGG is proposed by Simonyan et al. in 2014, in which the convolution filter is a 3 × 3 filter and the stride is 2. VGG-11, VGG-16 and VGG-19 respectively include 11, 16 and 19 layers. The Softmax layer is the final layer for classification [24].

(5) ResNet is proposed by Kaiming et al. in 2015 to address to the vanishing gradient problem and shows the excellent ability of classification [25]. The popular ResNet consists of 49 convolution layers.

All the algorithms in our experiments are coded in Matlab R2019a. The values of initial parameters are randomly generated for networks. The parameters are set as follows: the number of epochs, learning rate drop factor, learning rate drop period, batch size are 200, 0.0004, 0.6, 30 and 128, respectively. The operating system of PC is Window 7 and the GPU is a NVIDIA with 8 GB memory.

3.4. Results and discussions

The results of six algorithms are listed in Table 1 and Figure 6. Compared with ICSNet, ICSNet_n do not include inception modules. As we can see, the accuracy of ICSNet is 93.00%, the highest among six algorithms. The identifying time of ICSNet is 7.439 s, ranking the third shortest among six algorithms. The number of layers of ICSNet is 12, ranking the third lest among six algorithms. The accuracy of ResNet is equal to our algorithm, but their identifying time is more than that of our algorithm. The time of LeNet and AlexNet is less than that of our algorithm, but their accuracy is obviously lower than our algorithm. The Precision of our algorithm is 93.20%, only is lower than that of ResNet. The Recall of our algorithm is 93.30%, equal to that of ResNet ranking the first highest among six algorithms. The F_1 of our algorithm is 93.00%, equal to that of ResNet, ranking the highest among six algorithms. The AUC of our algorithm is 94%, the second highest among six algorithms. Consequently, there is the best tradeoff between accuracy and time. Compared with LeNet, AlexNet, the number of layers of ICSNet is more, accordingly, the accuracy of ICSNet becomes higher but the time cost is only increased a little. The reason is that the size of filters is 5 × 5 in LeNet, the size of filters is 5 × 5 or 11 × 11 in AlexNet, the size of filters is 3 × 3 or 1 × 1 in ICSNet. For our algorithm, the size of filters is smaller than LeNet and AlexNet.

We also see that, the more the number of the layers of networks, the larger the min-batch size is. But the VGG-19 is just the opposite due to its too much layers. The inception module induces the number of parameters and computing cost of the GoogLeNet and ICSNet, and remains more high identifying ability.

Table 1. The results of six algorithms.

Algorithms	Accuracy (%)	Precision (%)	Recall (%)	F_1 (%)	AUC	Min-batch size	Time(s)	The number of Layers
ICSNet	**93.00**	93.20	**93.30**	**93.00**	94.00	128	7.439	12
ICSNet_n	80.50	81.60	80.50	80.50	81.44	128	6.451	7
LeNet	80.50	81.42	80.64	80.50	81.00	128	6.511	7
AlexNet	85.00	84.35	84.41	84.38	85.00	128	6.953	8
VGG-19	91.00	92.30	92.00	92.15	91.50	28	16.342	19
GoogLeNet	92.50	92.64	92.50	**93.00**	93.00	80	13.192	22
ResNet-18	91.10	92.42	92.15	92.24	92.00	20	13.012	18
ResNet-50	92.15	92.50	92.15	92.50	92.50	20	20.160	50
ResNet152	**93.00**	**93.45**	**93.30**	**93.00**	**94.20**	20	31.879	152

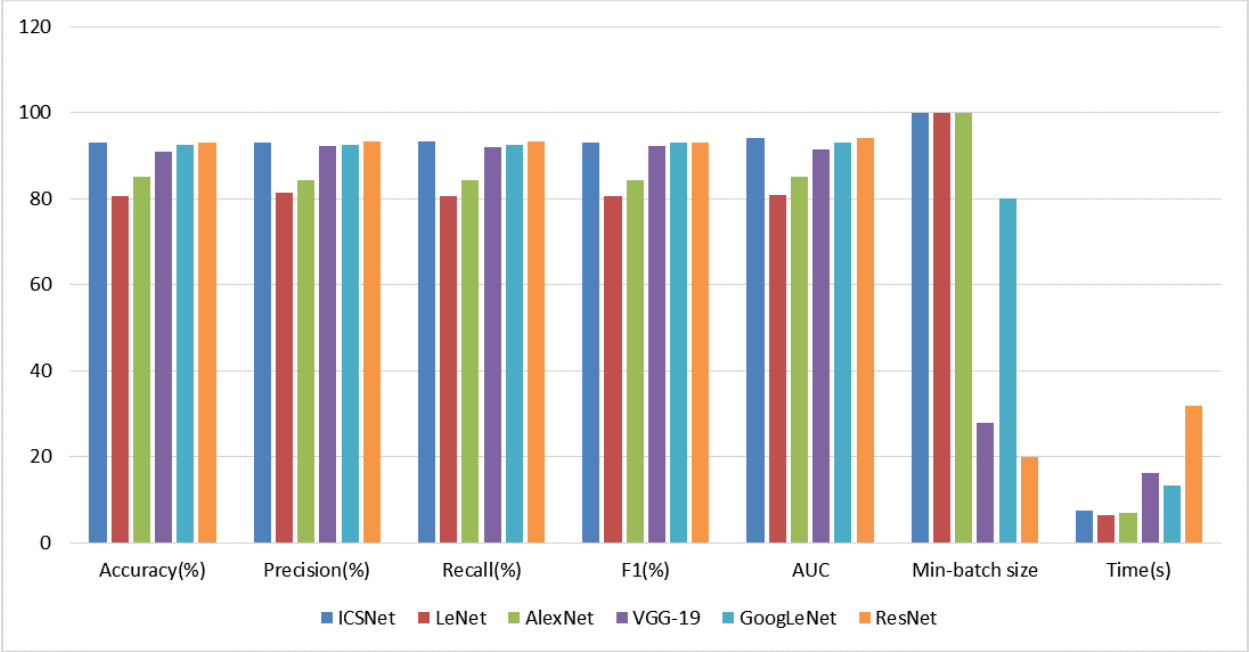
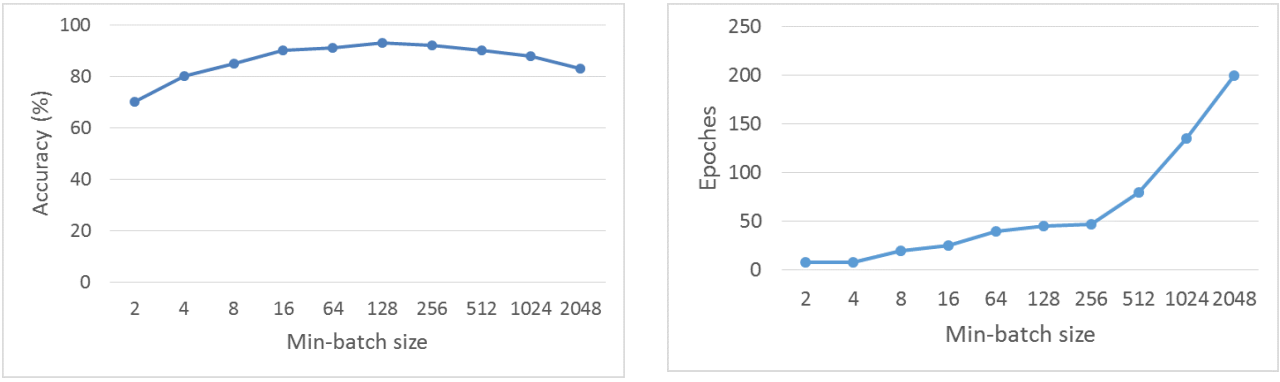


Figure 6. The results of algorithms.



(a) (b)

Figure 7 The relation between Min-batch size and Accuracy, Epoches.

Table 2. The results of our algorithms.

Fold	Class	Number	A	B	Accuracy (%)
1	A	21600	20079	1521	93.00
	B	7200	495	6705	
2	A	21600	20036	1564	93.00
	B	7200	452	6748	
3	A	21600	20415	1185	94.48
	B	7200	405	6885	
4	A	21600	20440	1160	94.41
	B	7200	539	6661	
5	A	21600	19814	1786	90.12
	B	7200	748	6452	
6	A	21600	20596	1004	95.42
	B	7200	315	6885	
7	A	21600	19491	2109	89.60
	B	7200	886	6314	
8	A	21600	19795	1805	91.22
	B	7200	729	6471	
9	A	21600	21122	478	97.57
	B	7200	221	6979	
10	A	21600	19966	1634	91.18
	B	7200	906	6294	
Average					**93.00**

Figure 7 shows the relation between Min-batch size and Accuracy, Epoches. As we see in Figure 7(a), when the Min-batch size is 128, the accuracy is the highest. This indicates the Min-batch size is relation to the accuracy of algorithm. In Figure 7(b), the larger is the Min-batch size, the larger is the number of epochs. When the Min-batch size is more than 256, the number of epochs is sharp large. The larger is the number of the epochs, the more is the consumed time when the algorithm converges. We also can see, when the Min-batch size is 128, the accuracy of algorithm is high, and the number of epochs is relatedly small.

The Table 2 lists the accuracy of each fold in test sets in the 10-folds CV and the confusion matrix of each fold. In Table 2, "A" denotes the image containing Ipomoea cairica (L.) sweets, "B" denotes the image without Ipomoea cairica (L.) sweets.

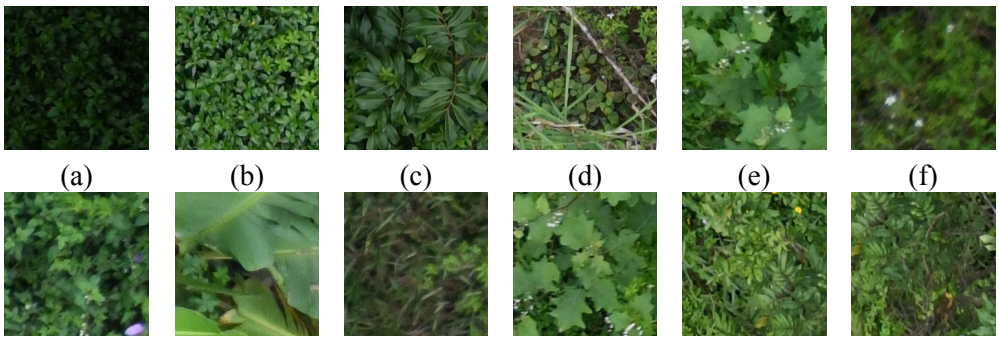


Figure 8. The incorrectly identified samples of *Ipomoea cairica (L.) sweets*.

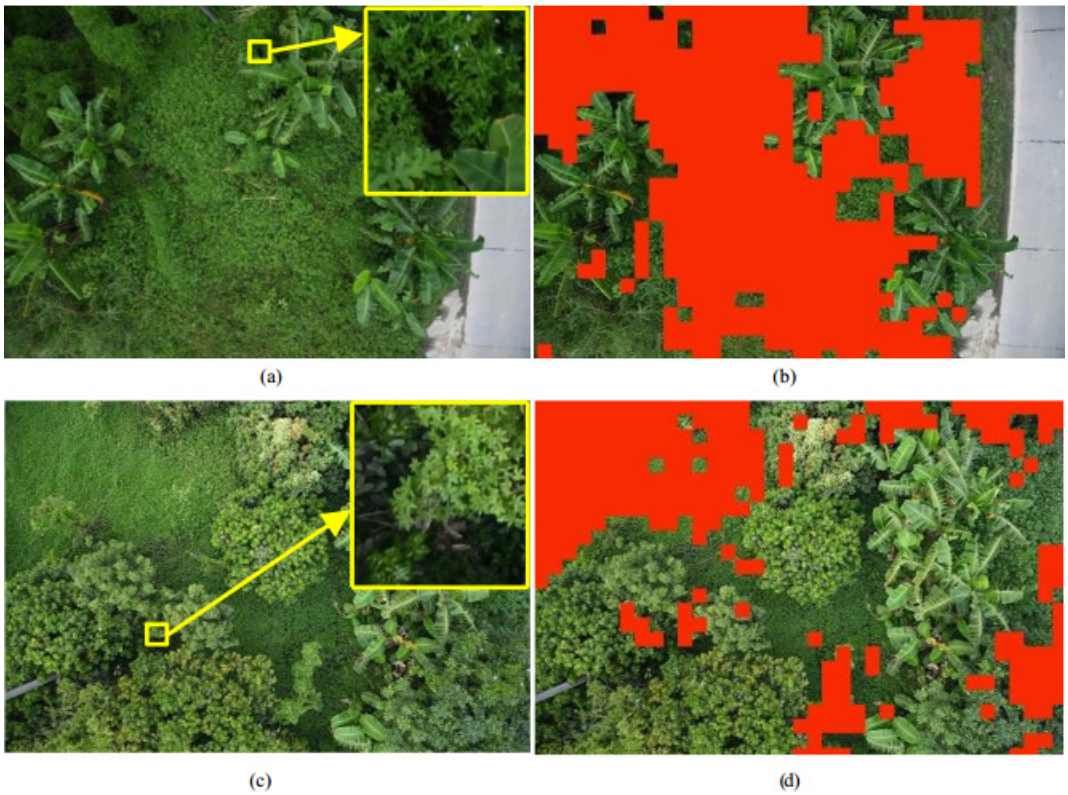


Figure 9. The raw images and images in which the area of *Ipomoea cairica (L.) sweets* is painted red.

Compared with LeNet, AlexNet, the number of layers of ICSNet is more, accordingly, the accuracy of ICSNet becomes higher but the time cost is only increased a little. Compared with GoogLeNet and ResNet, the time cost is obviously reduced. Since the layers of network is closely related to the time cost, our algorithm is designed considering their tradeoff.

The ICSNet aims to accurately identify *Ipomoea cairica (L.) sweets* and *non-Ipomoea cairica (L.) sweets* in the wild. Some incorrectly identified samples are shown in Figure 8. We can see that, in Figure 8 (b–e), (g), (k) and (l), the parts of *Ipomoea cairica (L.) sweets* are clear, but their proportions are very small. It is reason that they are incorrectly identified by our algorithms. For Figure 8 (f), (d) and (i), the bad quality of images affects the performance of our algorithm. The samples in Figure 8 (a) is too dark, so the misidentification is excusable. To solve the problem, in future work, we need consider two aspects: (1) Set the parameters of camera of UAV for good quality of acquired images. (2) Determine the right distance between UAV and Ipomoea cairica (L.) sweets for good quality of acquired images.

Based on the results of our algorithm, we can draw the area of *Ipomoea cairica (L.) sweets* in the raw images or map to help management to efficiently monitor the *Ipomoea cairica (L.) sweets* and take measures to protect the environment. The raw images are shown in Figure 9 (a) and (c), and the corresponding images, in which the area of Ipomoea cairica (L.) sweets is painted red, are shown in Figure 9 (b) and (d). In Figure 9 (a) and (c), the images in big yellow box are zoomed, which are corresponding the images in small yellow box. As we can see in Figure 9, the area of *Ipomoea cairica (L.) sweets* can be intuitively found.

4. Conclusions

Ipomoea cairica (L.) sweets is an invasive weed which has caused serious harm to the biodiversity and stability of the ecosystem. It is very important to accurately and rapidly identifying *Ipomoea cairica (L.) sweets* in the wild for management taking the necessary strategies to control the *Ipomoea cairica (L.) sweets* to rapidly grow in the wild. In this study, we propose an identification framework and method for *Ipomoea cairica (L.) sweets* based on UAV and AI.

The UAV can help us to low-costly acquire lots of high quality of images of monitoring areas. The deep CNN is designed to efficiently identify *Ipomoea cairica (L.) sweets* in the acquired images. To validate the efficiency of proposed method, we generated 288000 samples in our experiments, and made comparison with LeNet, AlexNet, GoogleNet, VGG and ResNet. The experimental results showed that our method was excellent, the identifying accuracy is 93.00% and the identifying time is 7.439 s. Our work can help to promote the performance of online monitoring systems for *Ipomoea cairica (L.) sweets*.

In our experiments, we set the distance according to the following three rules: 1) the clear images can be acquired; 2) the identifying accuracy can meet the actual requirements; 3) the UAV can be easily controlled. In addition, we also consider the sensor resolution. So we set the value of distance 15m above the ground according to our experience. When the value of distance is higher than 15 m, the collected images became more and more vague. This will affect the accuracy of model. When the value of distance is lower than 15 m, although we can collect clear images, this will consume the more time to collect the data and process data.

For our detection frame and model, the size of images can be set as the possible predefined size. The size of image affects not only the accuracy of model but also the monitoring accuracy. It is important to design an excellent model and algorithm in the special application. So, in future, we will further study the question.

Our model also can be regarded as the special GoogLeNet. The reason simplifying GoogLeNet is to achieve a tradeoff between accuracy and time. In addition, we also adopt the idea of VGGNet, that is 3×3 filters is used in our model to further reduce the time complexity. For our algorithm, the idea of ResNet is not adopted. The reason is that we find the performance cannot meet our requirements.

In the future, we will acquire more *Ipomoea cairica (L.) sweets* images to train ICSNet model for further improving the identifying accuracy of model. In addition, we also will build the dynamic monitoring system of *Ipomoea cairica (L.) sweets*.

Acknowledgments

The work is funded by the Scientific Research Team Project of Shenzhen Institute of Information Technology (SZIIT2019KJ022), Special Innovation Project of Guangdong Education Department (2019GKTSCX093) and National Natural Science Foundation of China (71971089).

Conflict of interest

The authors declare that there is no conflict of interests regarding the publication of article.

References

1. Z. Y. Sun, T. J. Zhang, J. Q. Su, W. S. Chow, J. Q. Liu, L. L. Chen, et al., A novel role of ethephon in controlling the noxious weed Ipomoea cairica (Linn.) Sweets, *Sci. Rep.*, **5** (2015), 11372.
2. G. Liu, Y. Gao, F. Huang, M. Yuan, S. Peng, The Invasion of coastal areas in south China by Ipomoea cairica may be accelerated by the ecotype being more locally adapted to salt stress, *Plos One*, **11** (2016), e0149262.
3. S. Shen, Z. Shen, M. Zhao, Big Data Monitoring System Design and Implementation of Invasive Alien Plants Based on WSNs and WebGIS, *Wireless Pers. Commun.*, **97** (2017), 4251–4263.
4. M. Mafanya, P. Tsele, J. O. Botai, P. Manyama, G. J. Chirima, T. Monate, Radiometric calibration framework for ultra-high-resolution UAV-derived orthomosaics for large-scale mapping of invasive alien plants in semi-arid woodlands: Harrisia pomanensis as a case study, *Int. J. Remote Sens.*, **39** (2018), 5119–5140.
5. M. Liu, H. Li, L. Li, W. Man, M. Jia, Z. Wang, et al., Monitoring the Invasion of Spartina alterniflora Using Multi-source High-resolution Imagery in the Zhangjiang Estuary, China, *Remote Sens.*, **9** (2017), 539.
6. A. M. West, P. H. Evangelista, C. S. Jarnevich, N. E. Young, T. J. Stohlgren, C. Talbert, et al., Integrating remote sensing with species distribution models; mapping Tamarisk invasions using the software for assisted habitat modeling (SAHM), *J. Visualized Exp.*, **116** (2016), e54578.
7. J. Bustamante, D. Aragonés, I. Afán, C. Luque, A. Pérez-Vázquez, E. Castellanos, et al., Hyperspectral sensors as a management tool to prevent the invasion of the exotic cordgrass spartina densiflora in the doñana wetlands, *Remote Sens.*, **8** (2016), 1001.
8. J. M. Barbosa, G. P. Asner, R. E. Martin, C. A. Baldeck, F. Hughes, T. Johnson, Determining subcanopy psidium cattleianum invasion in hawaiian forests using imaging spectroscopy, *Remote Sens.*, **8** (2016), 33.
9. S. Khare, H. Latifi, S. K. Ghosh, Multi-scale assessment of invasive plant species diversity using Pléiades 1A, RapidEye and Landsat-8 data, *Geocarto Int.*, **33** (2018), 681–698.
10. L. Tao, C. Cheng, *Plant Identification Based on Image Set Analysis*, In: Artificial Intelligence and Mobile Services–AIMS 2018, Lecture Notes in Computer Science, 2018.
11. S. Ge, R. Carruthers, P. Gong, A. Herrera, Texture analysis for mapping Tamarix parviflora using aerial photographs along the Cache Creek, California, *Environ. Monit. Assess.*, **114** (2006), 65–83.
12. D. Jones, S. Pike, M. Thomas, D. Murphy, Object-based image analysis for detection of Japanese knotweed sl taxa (Polygonaceae) in Wales (UK), *Remote Sens.*, **3** (2011), 319–342.
13. W. Dorigo, A. Lucieer, T. Podobnikar, A. Čarni, Mapping invasive Fallopia japonica by combined spectral, spatial, and temporal analysis of digital orthophotos, *Int. J. Appl. Earth Obs. Geoinf.*, **19** (2012), 185–195.
14. J. Sandino, F. Gonzalez, K. Mengersen, K. J. Gaston, UAVs and machine learning revolutionising invasive grass and vegetation surveys in remote arid lands, *Sensors*, **18** (2018), 605.

15. C. Hung, Z. Xu, S. Sukkarieh, Feature learning based approach for weed classification using high resolution aerial images from a digital camera mounted on a UAV, *Remote Sens.*, **6** (2014), 12037–12054.
16. M. P. Pound, J. A. Atkinson, A. J. Townsend, M. H. Wilson, M. Griffiths, A. S. Jackson, et al., Deep machine learning provides state-of-the-art performance in image-based plant phenotyping, *GigaScience*, **6** (2017), gix083.
17. A. Fuentes, S. Yoon, S. C. Kim, D. S. Park, A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition, *Sensors*, **17** (2017), 2022.
18. A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg, D. P. Hughes, Deep learning for image-based cassava disease detection, *Front. Plant Sci.*, **8** (2017), 1852.
19. H. Lu, Z. Cao, Y. Xiao, B. Zhuang, C. Shen, TasselNet: counting maize tassels in the wild via local counts regression network, *Plant Methods*, **13** (2017), 79.
20. M. Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M. S. Nasrin, et al., The history began from alexnet: A comprehensive survey on deep learning approaches, *arXiv preprint arXiv:1803.01164*, 2018.
21. Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, *Proc. IEEE*, **86** (1998), 2278–2324.
22. A. Krizhevsky, I. Sutskever, G. Hinton, Imagenet classification with deep convolutional neural networks, *Commun. ACM*, **60** (2017), 84–90.
23. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, et al., *Going Deeper with Convolutions*, 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015.
24. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, *Comput. Sci.*, **2014** (2014), 21–30.
25. K. He, X. Zhang, S. Ren, J. Sun, *Deep residual learning for image recognition*, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016.