Approach on Recognizing Uncovered Earth Region from Aerial Images Based on Multi-feature Fusion

Wenxiang Zhang1,*, Yingchun Zhong1, Linfeng Xie1, Bo Wang2 and Zhiyong Luo3

1 School of Automation, Guangdong University of Technology, Guangzhou, Guangdong Province, China
2 Department of Electro-mechanical Engineering, Guangdong Machinery Technician College, Guangzhou, Guangdong Province, China
3 Guangzhou Ufly Technology Co., Ltd., Guangzhou, Guangdong Province, China

*Corresponding author email: 2111904374@mail2.gdut.edu.cn

Abstract. It is one of the main hidden dangers of power transmission lines accidents if there is the uncovered earth under or near power transmission lines. It can give the important early warning message to prevent the accidents through recognizing the uncovered earth region from aerial images of Unmanned Aerial Vehicle (UAV) power transmission lines inspection. Due to the low recognizing precision of Mask RCNN CNN (Mask Convolution Neural Network), this paper proposed an approach to recognize the uncovered earth region from aerial images of UAV by image feature fusion. The HOG and LBP features of aerial images were extracted and their dimension were reduced. Then these two features were fused by different weights. The experiments show that, (1) the average precision of recognizing the uncovered earth region can be above 80%, which is the lowest requirement to use; (2) the weights of two features should make the orders of magnitude of the two features as equal as possible. The approach is application for the first image filtering by the UAV airborne platform because it not only has enough good recognizing precision but also is very rapid, which provides a novel way for objective recognition from UAV aerial images.

1. Introduction

Unmanned aerial vehicle (UAV)-based inspection is a major development trend for power transmission line inspection[1]. During the inspection, UAV takes the aerial images of power lines and recognizes the hardware defects and hidden dangers from the aerial images for the first filtering. Then those images which probably include the hardware defects and hidden dangers will be sent to the remote server to detect again. This mode can greatly reduce the image data transmission on UAV and improve the efficiency.

It is one of the main hidden dangers of power transmission lines accidents if there is the uncovered earth under or near power transmission lines. The reasons include that: (1) The uncovered earth indicates that there would be probably building operation under or near the power transmission lines. During the construction, the bucket or boom of construction vehicle would probably enter the safety distance of power transmission lines while they lift, which is very easy to result in the breakdown accident. (2) If there is the uncovered earth under or near power transmission lines, there are hidden dangers of soil loss or collapse, which is very easy to result in the collapse of electric tower. Therefore, it is one of the main tasks to find out the uncovered earth during power transmission lines inspection by UAV.
When the UAV is used for power transmission lines inspection, the inspection efficiency and the safety of inspector will be improved greatly. Sampedro Carlos, et al (2014) used HOG (Histograms of Oriented Gradients) features to train two MLP (multi-layer perceptron) neural networks, the first network was used for background-foreground segmentation, and the second one was used for classification of four different types of electricity towers, which achieved good results[2]. Zhenbing Zhao, et al (2016) have proposed a new method in which the state of insulators were represented using the CNN (Convolution Neural Network) model with a polyhedral feature extraction method[3]. Xiao W, et al (2016) have proposed using histograms of the stripe direction and stripe length to describe the characteristics of bird nest, and applied them to the detection and identification of bird nest in the overhead net system of high-speed rails[4]. The multi-class of defects of the electrical devices were detected by a modified region-based CNN in literature [5](2018). Y. Bazi (2019) have proposed a two-branch neural network architecture for multi-label classification in UAV imagery[6]. S. Vemula and M. Frye (2020) have proposed a Mask RCNN Powerline detector based on powerline-detection segmentation algorithm, that was deployed on an UAV[7]. Yingchun Zhong, et al (2020) have used the YOLO V3 (You Only Look Once) to detect the bird nest on the electric tower from aerial images of transmission line[8]. D. Sadykova, et al (2020) have used the YOLO(You Only Look Once) to detect insulators from aerial images, it is a cost-effective solution under the conditions of varied object resolution, an uncluttered background and illumination conditions[9]. All these researches have demonstrated that it is the trend to recognize the objects or defects from images. However, there is little literature about how to detect the uncovered earth from aerial images of UAV (Unmanned aerial vehicle) power lines inspection.

Aim to the problem of recognizing the uncovered earth region from UAV aerial images, this paper proposed an approach. Two different image features were extracted. Then they were fused through respective weights and employed to recognize the uncovered earth region.

2. Image Data Set

2.1. Image Acquisition
The sample images of this investigation were taken by LID-20c Hasu aerial camera on Mavic 2 of DJI. The resolution of each original image was 5472*3684 pixels. A total of 1493 original images were collected. And all these images were normalized to 1280*1040 pixels.

2.2. Establishment of Image Data Set
The number of original images is not enough. We have augmented the original images by rotating and zooming[10] to increase the diversity of image data, reduce the over fitting in the process of model training, and improve the generalization ability of the model. After image augmentation, the data set included 4479 images. Then 3359 images were chosen arbitrarily as the training data set and rest 1120 images were testing data set.

2.3. Image Annotation
All images in data set were labeled by software LabelImg[11]. The type and location of uncovered earth regions in aerial images were labeled.

3. Establishment of Recognition Model

3.1. Framework
The Mask Convolution Neural Network (Mask RCNN) has excellent performance in objects recognizing and segmenting from images. So, we first tried to use the Mask RCNN to recognize the uncovered earth region. However, the recognizing precision of Mask RCNN was too low to use. Then we proposed to use the features fusion method to recognize the uncovered earth regions from aerial images. The framework structure is shown in figure 1.
According to figure 1, after dividing the \(i^{th}\) image \((i = 1 \ldots n)\) into several same scale subblocks, we tried to use three different models to extract the features of uncovered earth region. The Model No.1 uses the Hist Of Grey (HOG) feature to illustrate the uncovered earth regions. The Model No.2 uses the Local Binary Pattern (LBP) feature to illustrate the uncovered earth regions.

### 3.2. Model No.1

The HOG feature is an operator describing the edge and shape features of an image. The gradient of pixels in an image are calculated by the discrete differential template \([-1 \ 0 \ 1]\) and its transpose template \([-1 \ 0 \ 1]^T\):

\[
G_x(x, y) = H(x + 1, y) - H(x - 1, y) 
\]

\[
G_y(x, y) = H(x, y + 1) - H(x, y - 1) 
\]

\[
G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} 
\]

\[
\theta(x, y) = \tan^{-1} \left( \frac{G_y(x, y)}{G_x(x, y)} \right) 
\]

where, \(G_x(x, y)\) is horizontal gradient, \(G_y(x, y)\) is vertical gradient, and \(G(x, y)\), \(\theta(x, y)\) are the magnitude and angle of gradients respectively. The histogram of the values of gradient of all pixels neighborhood are calculated, which is the HOG feature.

The scale image from data set is divided into 10*10 subblocks. The scale of each subblock is 1280*1040 pixels. Then the dimension of HOG feature of subblock is 6480. The Principal Component Analysis (PCA) is employed to reduce the dimension of HOG feature.

So, the steps of Model No.1 include:
1. HOG feature of every subblock is extracted.
2. The dimension of HOG feature is reduced by PCA method.
3. The HOG features of all subblocks are combined as the feature of the \(i^{th}\) image.
4. All images in training data set are processed as the same steps and the features of all images formed the training feature matrix.
5. A SVM(Support Vector Machine) is trained by the training feature matrix.

### 3.3. Model No.2

The LBP feature can describe the texture character of image, which often uses a window with 3*3 pixels. When the window masks the image, the gray scale value of central pixel in the window is the threshold and the gray scale values of its 8 neighborhoods are compared with the threshold. Then the LBP feature of a binary code with 8 bits for the central pixel of image is generated.
Obviously, the original LBP feature is unable to meet the requirement of different scale because it uses just a window with 3*3 pixels. The LBP feature in Uniform Pattern was proposed by Ojala[14], which enlarges the LBP feature to the scale with any size and reduces its dimension:

$$U(G_p) = s(g_{p-1} - g_c) - s(g_0 - g_c) + \sum_{p=1}^{p-1} s(g_p - g_c) - s(g_{p-1} - g_c)$$

(5)

where, $g_c$ is the gray scale value of central pixel of its neighborhood, $g_p$ corresponds to the gray scale value of equidistant pixels $P$ on the circumference with radius $R$.

So, the steps of Model No.2 include:
(1) The LBP feature in Uniform Pattern of every subblock is extracted.
(2) The HOG features of all subblocks are combined as the feature of the ith image.
(3) All images in training data set are processed as the same steps and the features of all images form the training feature matrix.
(4) A SVM is trained by the training feature matrix.

### 3.4. Model No.3

The HOG feature after dimension reduction has good invariance to the optical and geometric shape of the image. But it is weak to represent the edge. The LBP feature in Uniform Pattern has the gray scale invariance and rotation invariance of the image. But it often fails to extract the complete information of image. So, this paper fuses these two features in order to describe not only the local texture information but also the edge information of the image. The fusing formula is:

$$F = \alpha HOG + \beta LBP$$

### 4. Experiments

#### 4.1. Preparation

The Mask RCNN was employed to recognize the uncovered earth region from image, whose hardware configuration was: CPU: i7-6700K, RAM 32G, GPU: GTX1070, Video Memory: 8G. The software configuration was: operating system: Ubuntu16.04, CUDA, cuDNN, Keras, Tensorflow and Python 3.6.

The hardware configuration to extract HOG and LBP features was: CPU: Core i5-8250U, RAM: 8G, GPU: MX150, Video Memory: 2G. The software configuration was: operating system: Windows 10, Skimage, Sklearn and Python 3.6.

The mean Average Precision and F1 are employed to evaluate the experimental results[15,16].

#### 4.2. Experiments and Results Analysis

##### 4.2.1. Recognition Experiments of Models.

The Mask RCNN, Model No.1, 2 and 3 were trained and testified by the same training and test data sets. The recognition precision and other evaluation criterion are clearly described in table 1.

| Model No.                                      | AP      | F1      | MTD /ms  | Training time/min | Volume of parameters /M |
|------------------------------------------------|---------|---------|----------|-------------------|-------------------------|
| Mask RCNN after parameters adjustment         | 57.66%  | 55.61%  | 160      | 1560              | 255                     |
| Model No.1                                    | 74.05%  | 71.63%  | 57       | 52                | 0.8                     |
| Model No.2                                    | 65.82%  | 62.67%  | 55       | 27                | 0.4                     |
| Model No.3                                    | 81.13%  | 78.37%  | 62       | 60                | 1                       |

Comparing four models in table 1, we know that: (1) The AP and F1 of Mask RCNN are less than other 3 models obviously even though the parameters of Mask RCNN are adjusted. The parameters volume of Mask RCNN is too large to use on airborne platform of UAV. Additionally, the Mean Time
to Detect (MTD) of Mask RCNN is nearly three times of other 3 models. (2) The AP and F1 of Model No.1 are better than the Model No.2’s, which implies that the HOG feature with dimension reduction has obvious better discriminative capability than the LBP. But these two models are unusable because their APs are all less than 80%. (3) The AP of Model No.3 is over 80% and the volume of parameters is just 1M, that is, the AP of Model No.3 pass blue line and the volume of parameters is small for using on the airborne platform. So we believe that the Model No.3 is suit for using on the airborne platform of UAV for the images first filtering.

Two images were selected as the cases to show the detect results of Model No.3, as shown in figure 2.

![Figure 2. Recognizing cases of Model No.3.](image)

4.2.2. Variation Rule of Weights in Model No.3. The weights are the most important factor to influence the precision of Model No.3. In order to explore the variation rule of weights, many experiments have been done. The results are illustrated, as in figure 3.

![Figure 3. Variation rule of weights.](image)

In figure 3, the horizontal ordinate is the $\alpha$ in equation (6) and the vertical ordinate is AP. There are four curves in figure 3. Different curve means different value of $\beta$. According to Figure 3, when the $\beta=0.0015$ and $\alpha=1.1$, the max value of AP appears. The trends of all four curves show that the best value of AP always appears while $\alpha=1.0-1.1$. Additionally, the value of $\beta$ has little affection on AP. Furthermore, the experiments reveal that the optimal weight combination should make the orders of magnitude of the two features as equal as possible.

5. Conclusion

At present, there is rare research on the recognizing of uncovered earth region from UAV inspection aerial images. After trying to use the methods of Mask R-CNN, HOG+SVM and LBP+SVM, this paper proposed to fuse the HOG and LBP features of UAV inspection aerial images in order to recognize the region of uncovered earth. The experiments show that, (1) the average precision of recognizing the uncovered earth region can be above 80%, which is the lowest requirement to use; (2) the weights of two features should make the orders of magnitude of the two features as equal as possible. The approach proposed in this paper, which is application for the first image filtering by the UAV airborne platform because it not only has enough good recognizing precision but also is very rapid, which provides a novel way for objective recognition from UAV aerial images.
Acknowledgments
This paper is supported by the key project of science and technology of Guangzhou City, Guangdong Province, China (No. 202007040004), the Natural Science Foundation of China (No.61975248) and the Natural Science Foundation Project of Guangdong Province, China (No. 2018A0303130137).

References
[1] Dong G, Chen X, Wang B, et al. Inspecting transmission lines with an unmanned fixed-wings aircraft. In 2012 2nd International Conference on Applied Robotics for the Power Industry (CARPI) (pp. 173-174). IEEE
[2] Sampedro C, Martinez C, Chauhan A, et al. A supervised approach to electric tower detection and classification for power line inspection. 2014 International Joint Conference on Neural Networks (IJCNN). 2014: 1970-1977
[3] Zhenbing Zhao, Xu G, Qi Y, et al. Multi-patch deep features for power line insulator status classification from aerial images. International Joint Conference on Neural Networks (IJCNN). 2016: 3187-3194
[4] Xiao W, Yuan P, Peng Q, et al. Detection of bird nests in overhead catenary system images for high-speed rail. Pattern Recognition. 2016, 51: 242-254
[5] Lan M, Zhang Y, Zhang L, et al. Defect detection from UAV images based on region-based CNNs. IEEE International Conference on Data Mining Workshops (ICDMW). 2018: 385-390
[6] Y. Bazi, "Two-Branch Neural Network for Learning Multi-label Classification in UAV Imagery," IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 2019, pp. 2443-2446, doi: 10.1109/IGARSS.2019.8898895
[7] S. Vemula and M. Frye, "Mask R-CNN Powerline Detector: A Deep Learning approach with applications to a UAV," 2020 AIAA/IEEE 39th Digital Avionics Systems Conference (DASC), San Antonio, TX, USA, 2020, pp. 1-6, doi: 10.1109/DASC50938.2020.9256456
[8] Yingchun Zhong, Siyu Sun, Shuai Lyu, et al. Recognition of Bird’s Nest on Transmission Tower in Aerial Image of High-voltage Power Line by YOLOv3 Algorithm. Journal of Guangdong University of Technology. 2020, 037(003):42-48
[9] D. Sadykova, D. Pernebayeva, M. Bagheri and A. James, "IN-YOLO: Real-Time Detection of Outdoor High Voltage Insulators Using UAV Imaging," in IEEE Transactions on Power Delivery, vol. 35, no. 3, pp. 1599-1601, June 2020, doi: 10.1109/TPWRD.2019.2944741
[10] Weihang Wei, Qianli Ma, Zhiliang Gao, et al. Application of improved Convolutional Neural Network in image smoke detection. Intelligent Computer and Application. 2019, 9(03):47-51
[11] Gevaert C, Persello C, Elberink S O, et al. An automated technique for basemap updating using UAV data. Joint Urban Remote Sensing Event (JURSE). 2017: 1-4
[12] He K, Gkioxari G, Dollár P, et al. Mask r-cnn. Proceedings of the IEEE international conference on computer vision. 2017: 2961-2969
[13] Pan S, Sun S, Yang L, et al. Content retrieval algorithm based on improved HOG. 3rd International Conference on Applied Computing and Information Technology/2nd International Conference on Computational Science and Intelligence. 2015: 438-441
[14] Ojala T, PietikäInen M, Harwood D. A Comparative Study of Texture Measure with Classification Based on Feature Distribution. Pattern Recognition. 1996, 29(1):51-59.
[15] Feichtenhofer C, Pinz A, Zisserman A. Detect to track and track to detect. Proceedings of the IEEE International Conference on Computer Vision. 2017: 3038-3046
[16] Ren S, He K, Girshick R, et al. Object detection networks on convolutional feature maps. IEEE transactions on pattern analysis and machine intelligence. 2016, 39(7): 1476-1481
[17] Baifeng Ning. Pin Fault Detection in UAV Inspection Image Based on Deep Learning. Computer Measurement & Control. 2019, 27(11): 25-29