New Method of Airport Pavement Health Inspection Based on MobileNet-SSD and Mask R-CNN

Wentong Guo¹, Hongyuan Fang¹, Niannian Wang¹*

¹School of Water Science and Engineering, Zhengzhou University, Zhengzhou, 450000, China
*Corresponding author: Niannian Wang (e-mail: wnnian@zzu.edu.cn).

Abstract. This paper presents a rapid detection and pixel size measurement method for airport pavement apparent disease and Foreign Object Debris. Firstly, MobileNet-SSD algorithm is used for object identification. Then Mask R-CNN algorithm is used to segment the target image and measure the pixel size of the target object. Finally, the detailed information of the target object is obtained. The experimental verification shows that the recognition speed of this method reaches 65 frames per second, and the pixel segmentation accuracy reaches 96 %, which can meet the requirements of airport pavement health inspection.

1. Introduction
Airport pavement apparent disease (Cracks, holes, etc.) and Foreign Object Debris (FOD) (Stone, leaves, etc.) are important factors threatening Aircraft take-off and landing. With the development of Deep Learning and Image Processing technology, significant progress has been made in the autonomous identification of airport pavement apparent disease and Foreign Object Debris by computers. Shi [1] proposed Crack-Forest algorithm which used random forest to learn the characteristics of cracks. It generated airport pavement apparent disease detector which can identify complex topological characteristics. Li [2] et al. constructed a convolutional neural network model based on semantic segmentation, and used sliding window method to realize airport pavement apparent disease block detection. Yang [3] et al. proposed a crack detection model based on feature pyramid and fused multi-convolution layer information to achieve fast crack segmentation. Zou [4] et al. constructed a neural network airport disease detection model of encoder and decoder, and fused the target information of each scale to realize the effective combination of different size feature information. The team of China Institute of Aeronautical Radio and Electronics [5] proposed the identification and localization technology of airport runway foreign body based on Faster-RCNN, and used the Faster-RCNN algorithm framework to identify foreign body types.

In order to improve the detection speed and pixel measurement accuracy of airport pavement apparent disease and FOD, this paper adopts the image processing method based on MobileNet-SSD and Mask R-CNN. Firstly, MobileNet-SSD algorithm is used to preliminarily identify the image, and then Mask R-CNN algorithm is used to extract topology information and measure pixel size of airport pavement apparent disease and FOD. The experimental results show that this method can effectively extract airport pavement apparent disease and FOD pixel information.
2. MobileNet - SSD Target Recognition

2.1 MobileNet-SSD algorithm analysis
In this study, MobileNet-SSD [6] was used as the surface apparent disease and FOD identification tool. MobileNet-SSD is an improved version based on SSD algorithm [7] and MobileNet algorithm [8], and its performance has advantages in speed and accuracy [9]. Figure 1 is the network structure of MobileNet-SSD. The reason why MobileNet-SSD can achieve real-time accurate recognition is that:

1. MobileNet-SSD replaces the mapping convolution layer applied to image feature extraction in traditional SSD algorithm with deep separable convolution in MobileNet, and divides standard convolution into deep convolution and point convolution. Deep convolution applies a single convolution to each input channel, and convolution is performed on each input channel to obtain the convolution value of a single channel. Point convolution combines the output value of deep convolution to obtain the final convolution value. The computational speed of neural network is improved by reducing the complexity of convolution operation without reducing the accuracy.

2. In MobileNet-SSD network, the full connection layer of MobileNet is removed, and eight standard convolution layers are added after the 13th convolution kernel conv13 to broaden the acceptance range of feature images. In order to prevent the gradient from disappearing, BatchNorm layer and activation function (ReLU6) are introduced into each layer of the classification task network. In the process of model training, two super-parametric width multipliers and resolution multipliers are introduced to reduce the input and output channels and feature map sizes.

![Figure 1 The network structure of MobileNet-SSD](image)

2.2 Analysis of recognition results
In order to measure the performance of the algorithm, this paper uses average precision and Mean Average Precision as the evaluation indexes of recognition accuracy. Average precision (AP) is used to calculate the average precision of a specific category of objects at different recall levels, while Mean Average Precision (mAP) is the average precision of all categories. Each type of target and AP, and mAP results are listed in table 1. The crack AP is slightly lower than other targets, which may be due to the small sample size of the crack relative to other samples. In addition, the model recognition speed reaches 65 FPS / s, which can meet the requirements of real-time airport pavement detection.

| Class     | Crack | Leaf | Stone | Pit hole | Map |
|-----------|-------|------|-------|----------|-----|
| AP        | 77.25%| 87.79%| 89.21%| 92.57%   | 86.71%|
3. Semantic Segmentation of Mask R-CNN

3.1 Mask R-CNN algorithm analysis
In this study, Mask R-CNN [9] is used as a semantic segmentation tool for pavement appearance diseases and FOD. Mask R-CNN is an improved version based on Faster R-CNN [10] and is the most widely used deep learning instance segmentation algorithm. Figure 2 is the network architecture of Mask R-CNN. As shown in the figure, Mask R-CNN consists of four parts: Feature Extraction, Region Proposal Network, RoI pooling, Head Network. The image is input into Feature Extraction to obtain multiple feature maps at different levels. The candidate regions that may contain the detection target are generated by Region Proposal Network, and then the candidate region feature map is obtained by RoI pooling layer. Finally, the candidate regions are corrected, classified and the segmentation results of the detection targets are given through the branches of Head Network.

![Figure 2 Mask R-CNN network architecture](image)

3.2 Model training and result analysis
The training cycle of the Mask R-CNN model was set to 800 times. Iterate 100 steps per cycle. 600,150 and 50 rounds were trained at learning rates of 0.001, 0.0001 and 0.00001, respectively. Due to the large image size, the Batchsize was selected as 1, and the total training duration was 61 hours and 51 minutes. In order to measure the performance of the Mask R-CNN algorithm, this paper uses Pixel Accuracy (PA) to measure the segmentation Accuracy of objects of different categories, and uses Mean Pixel Accuracy (mPA) to measure the average Accuracy of the algorithm for all categories. The calculation formulas of PA and mPA are shown in equation (1) and equation (2). Where \( p_{ij} \) is the correctly classified pixel, \( \text{AllPixel} \) is all pixels of the image, \( k \) is the number of classes, \( p_{ij} \) is the pixel of each class.

\[
PA = \frac{\sum_{i=0}^{k} p_{ii}}{\sum_{j=0}^{k} \sum_{i=0}^{k} p_{ij}} = \frac{\sum_{i=0}^{k} p_{ii}}{\text{AllPixel}} \tag{1}
\]

\[
mPA = \frac{1}{k+1} \sum_{i=0}^{k} \frac{p_{ii}}{\sum_{j=0}^{k} p_{ij}} \tag{2}
\]

Table 2 shows the PA values and mPA values of four types of cracks, pits, gravels and leaves. It can be seen that the PA values of the four targets are greater than 90 %, which can meet the requirements of airport pavement diseases and FOD pixel segmentation. The PA of cracks is slightly smaller than that of other targets, which may be due to the small number of crack samples used in training. Overall, mPA reached 94.55 %, which is sufficient for airport pavement inspection.
Table 2 PA and mPA results for target categories.

| Class     | Crack | Leaf | Stone | Pit hole | mPA |
|-----------|-------|------|-------|----------|-----|
| PA        | 91.09%| 95.72%| 96.65%| 94.75%   | 94.55% |

3.3 Topological Information Extraction and Pixel Size Measurement

In this paper, the topology information extraction and pixel measurement of the target after segmentation are carried out. Firstly, the target mask layer is extracted and the mask image is binarized. Then the gray value of pixels in the non-mask region is set to 255 to filter the background information. Finally, the binarized mask image is completely mapped to the original image to extract the topological information of the target prediction results.

The apparent pavement diseases and FOD size information are measured at the pixel level, including the area, length and average width of the target area. Firstly, the pixel value of the target region is obtained by using the traversal algorithm to count the pixel number of the target region. Then the edge detection algorithm is used to extract the skeleton of the target area, and the pixel value of the target length is obtained by calculating the pixel number of the skeleton. Then the pixel value of the average width is calculated in equation (3), where \( W_{\text{avg}} \) represents the pixel value of the average width, \( A_p \) represents the pixel value of the target area, and \( L_p \) represents the pixel value of the target length. Since the staff mainly focus on the pixel values of the target area of the holes, gravels and leaves, they are not sensitive to the pixel values of the target length and the average width of the target. In order to reduce the computational complexity, the calculation of the pixel values of the target length and the average width of the target is omitted.

\[
W_{\text{avg}} = \frac{A_p}{L_p}
\]  

(3)

4. Verification and analysis

In order to verify the feasibility of this method, airport pavement apparent disease and FOD are identified and pixel information is measured in this paper. Table 3 shows some processing results. It can be seen that the Mobilenet-SSD test can successfully predict the crack, hole, gravel and leaf targets, but there is a problem of slow crack prediction speed, which may be due to the larger crack size and the loading model needs some time. Mask R-CNN can meet the target segmentation work, and can complete the pixel size measurement work, which can provide services for airport pavement apparent disease and FOD early warning evaluation.

Table 3 Part of the inspection system returns results.

| category  | Master map | Mobilenet-SSD recognition | Recognition speed (FPS/s) | Mask R-CNN segmentation | Pixel size       |
|-----------|------------|---------------------------|---------------------------|------------------------|-----------------|
| Crack     | ![Crack](image) | ![Inference time: 18.02ms](image) | 55                        | ![Area:3584 Wide:7 Long:533](image) | Crack:0.94      |
5. conclusion

This paper proposes a pixel-level health detection method for airport pavement based on Mobilenet-SSD and Mask R-CNN. The experimental verification shows that the target recognition speed using this method reaches 65 frames per second, and the pixel segmentation accuracy reaches 96%, which can meet the requirements of airport pavement detection. The next stage needs to improve these two algorithms to make the model have better generalization. In addition, this study will develop corresponding risk assessment algorithms to improve the detailed detection.

References
[1] YONG S, LI M C, ZHI Q Q, FAN M and ZHEN S C 2016 IEEE Transactions on Intelligent Transportation Systems. 17 3434-45
[2] LI Y L and SUN R Y 2019 Laser und Optoelektronik. 56 1-11
[3] YANG F, ZHANG L, YU S J, PROKHOROV D, LING H B and MEI X 2020 Transactions on Intelligent Transportation Systems. 21 1525-35
[4] ZOU Q, ZHANG Z, LI Q Q, QI X B, WANG Q and WANG S 2019 IEEE Transactions on Image Processing. 28 1498-1512
[5] JIANG J, TANG B W, SONG S L, GE C and HE Z Y 2020 Avionics Technology. 53 8-12
[6] REN Y J, YANG J, LIU F T and ZHANG Q X 2019 Journal of Frontiers of Computer Science and Technology. 13 1881-93
[7] TANG C, LING Y S, ZHENG K D, YANG X, ZHANG C, YANG H and JIN W 2018 Infrared and Laser Engineering. 47 302-10
[8] YAN Y F, DUAN H X and HE G W 2019 Computer Technology and Development. 29 192-6
[9] LIU Y, ZHU Z Y and ZHANG B 2019 Ship Electronic Engineering. 39 52-6
[10] REN S, HE K, GIRSHICK R and SUN J 2017 IEEE Transactions on Pattern Analysis and Machine Intelligence. 39 1137-49