

# AIRPORT TARGET RECOGNITION AND GROUND DETECTION MANAGEMENT BASED ON NEURAL NETWORK

Zhao Bai

*Civil Aviation Flight University of China, Guanghan, Sichuan 618307, China.*

[bzhaomail@163.com](mailto:bzhaomail@163.com)

---

**Abstract** - Modern airport management system is complex. For airport managers, it is increasingly difficult to identify aircraft and airport ground. High-resolution high-altitude reconnaissance equipment provides an effective tool for airport supervision. In order to effectively recognize aircraft and airport ground in reconnaissance image, this paper briefly introduced convolutional neural network (CNN) and airport target recognition model based on CNN. In order to improve the recognition and location accuracy of the recognition model, the loss function of the model was improved. Then, two recognition models were simulated and analyzed by using MATLAB software. It was found that the improved recognition model was better than the former one in the aspect of visual effect of recognition; the improved model was closer to the real value in traffic density recognition of airport operation state; the accuracy and recall rate of dispatch management opinions based on traffic density feedback were higher; the time required for recognition and management was shorter. In summary, the improved identification management model based on CNN can effectively identify airport aircraft, and quickly feedback accurate management dispatching opinions according to the identified airport operation status.

**Keywords:** Convolutional Neural Network, Airport Recognition, Loss Function, Aircraft Recognition.

---

## 1. Introduction

Since the Wright brothers invented the aircraft, with the development of science and technology, the aircraft has become the fastest and relatively safest means of transportation [1]. The popularity of airplanes has greatly increased the scope of transportation logistics and promoted the development of civil aviation industry. For civil aviation aircraft, the airport is a very important transport platform, which plays an important role in taking off, parking and transferring passengers [2].

However, the expansion of the scale of civil aviation makes the management of airports more complex and improper management of airports can easily lead to major accidents. Airport scale is usually relatively large, and manual inspection will inevitably lead to increased labour costs.

The manpower is usually not enough, and moreover 24 h online manpower cannot be guaranteed. Therefore a system which can automatically detect the airport aircraft and ground detection system becomes more and more necessary [3]. The emergence of high-resolution high-altitude reconnaissance equipment also contributes to the generation of airport automatic identification system by providing high-resolution airport and aircraft images. The traditional image target recognition algorithm [4] firstly preprocesses an image, then

extracts the features, and compares them with the pre-set artificial features, so as to achieve the effect of target recognition. The recognition effect depends on the pre-treatment effect and the accuracy of artificial features. The application of the traditional image target recognition algorithm not only has a large amount of calculation, but also has a poor recognition speed and accuracy. The emergence of convolution neural network (CNN) can effectively solve the above problems, and with the increase of network structure level and learning amount, the recognition effect will gradually increase. Zhang et al. [5] detected airports in satellite images using CNN and shared image features using transfer learning method to achieve small sample learning.

The simulation results showed that the detection rate of that method was as high as 88.8%. Chen et al. [6] proposed a fast R-CNN automatic detection method. Firstly, the candidate airport area was generated based on CNN, and then another CNN was used for airport detection. The simulation results of airport data set demonstrated the effectiveness of the method. Xia et al. [7] detected remote sensing image using single-lens multi-box detector (SSD), solved the problem of limited labeling data by transferring the trained parameters of convolution neural network, and found that that the detection rate of this method was 83.5%, which is faster than that of R-CNN. In this study, CNN and the airport

target recognition model based on CNN were briefly introduced. In order to improve target recognition and positioning accuracy, the loss function of the model was improved. Then the two recognition models were simulated and analyzed by using MATLAB software.

2. CNN

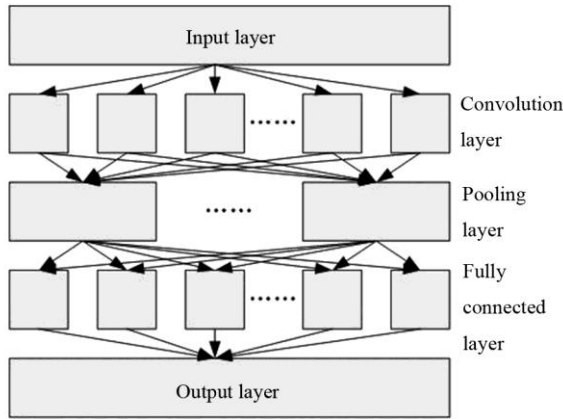


Figure 1: Structure of CNN

As shown in Figure 1, CNN [8] is divided into input layer, output layer, convolution layer, pooling layer and fully connected layer, in which the convolution layer and pooling layer are the key parts of the whole neural network. The process of forward propagation in convolution layer is called convolution operation, its output is called characteristic graph, and its convolution is like linear weighting operation. Its formula [9] is:

$$u = Wx + b \tag{1}$$

where  $u$  is a convolution characteristic graph,  $W$  is a weight parameter,  $x$  is convolution kernel input, and  $b$  is offset. Then  $u$  will obtain the activated convolution characteristic through activation function, then the complete convolution formula of the convolution layer is:

$$x_j^l = f \left( \sum_{j \in M} x_i^{l-1} \cdot W_{ij}^l + b_j^l \right) \tag{2}$$

where  $x_j^l$  is the output characteristic graph in the  $l$ -th convolution layer after the activation of of the  $j$ -th convolution kernel,  $x_i^{l-1}$  is the characteristic output of the  $i$ -th convolution kernel in the last layer,  $W_{ij}^l$  is the weight parameter between the  $i$ -th convolution kernel and the  $j$ -th convolution kernel,  $b_j^l$  is the offset of the  $j$ -th convolution kernel in the  $l$ -th layer,  $M$  is the number of convolution kernels in the  $l$ -th convolution layer, and  $f(\bullet)$  is an activation

function. In this study, non-linear sigmoid function was used.

When the convolution kernel is convoluted on the input image, its parameters remain unchanged, and the relevant calculation formula [10] is:

$$\begin{cases} OS = \frac{(IS - FS + 2P)}{s} + 1 \\ FS = h \cdot w \cdot c \\ Pa = (h \cdot w \cdot c + 1) \times n \\ Nerve = OS^2 \times n \end{cases} \tag{3}$$

where  $OS$  stands for the size of output characteristic graph,  $IS$  is the size of the input image,  $FS$  is the size of convolution kernel,  $P$  is zero-filling around,  $s$  is the step size of convolution operation of convolution kernel,  $h, w, c$  are the height, width and number of channels of the convolution kernel,  $Pa$  is the number of parameters in a single convolution layer,  $n$  is the number of convolution kernels, and  $Nerve$  is the number of neurons in the single convolution layer.

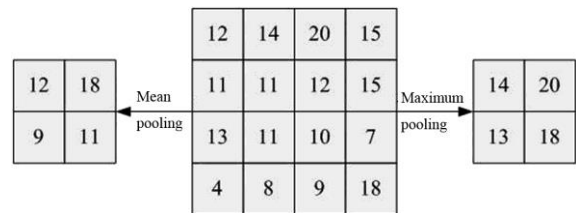


Figure 2: The schematic diagram of mean and maximum pooling of the pooling layer

The convolution characteristic graph of an original image can be obtained after the convolution operation of the convolution layer. Then it is transferred forward to the pooling layer for dimension reduction statistics. Common pooling operations include maximum pooling and mean pooling [11], as shown in Figure 2. The size of the target box is  $2 \times 2$ , and the sliding step size is 2.

The original image in a size of  $4 \times 4$  is converted into a characteristic graph in a size of  $2 \times 2$  after pooling, which greatly reduces the characteristic dimension. Moreover the translation invariance of the pooling operation improves the fault tolerance rate, and the pooling layer can transform different regions in different sizes into regions with the same size in processing multiple objects.

The layer behind the pooling layer is the fully connected layer whose function is to match the feature data extracted from the previous convolution layer and pooling layer with the labels in the samples, that is, set collection and feature classification.

Moreover, unlike the parameter sharing of the convolution layer, the parameters of the fully connected layer account for most of the parameters of the whole network model.

### 3. Detection Model Based on CNN

With the development of civil aviation industry, in addition to the expansion of airports, the number of aircraft has increased greatly, which makes the management of airports more complex. Manual scheduling alone has been difficult to meet the management needs. Airport operation management includes on-site command and dispatch, runway maintenance, airport security maintenance, etc.

This study mainly focused on the command and dispatch in airport operation management. In the peak period of work, the traffic flow in the airport increases sharply. If the airport is not commanded in time, it will cause regional congestion and affect the normal work of the airport. Therefore, quickly identifying and judging the operation status of the airport and giving timely opinions can effectively improve management efficiency.

The recognition indicators for judging the operation status of the airport include the speed of traffic, traffic flow, traffic density and so on. Combining with the main application objects of CNN model, this study selected the traffic density as the identification indicator of airport operation status.

Traffic density refers to the number of aircraft per unit distance in an area at some point, and its formula is:

$$\rho = \frac{n}{l} \tag{4}$$

where  $l$  is the length of the recognition area, i.e., the size of the image taken by the airport identification equipment in this study, and  $n$  is the number of aircraft in the identification area.

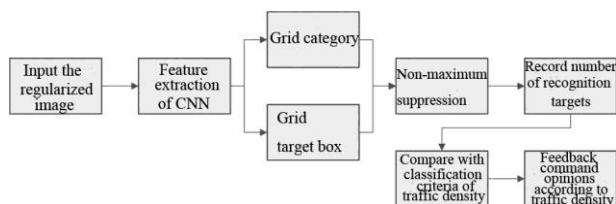


Figure 3: The flow of basic algorithm of CNN-based recognition management model

As shown in Figure 3, when the model is trained and used for airport aircraft recognition and management, their basic algorithm flow [12] is similar.

Firstly, the inspected pictures are regularized to reach the input size set by the model, and then the image is segmented into  $M \times M$  grids.

Secondly, CNN is input. After feature extraction of the convolution layer and pooling layer, each grid generates grid category probability and grid target box, in which grid target box is used for target detection. Several prediction target boxes are generated in each grid, and each prediction box includes the corresponding credibility  $A$ .

The formula [13] is:

$$\begin{cases} A = P_{object} \times IOU \\ IOU = \frac{S}{S_1 + S_2 - S} \\ P_{object} = \begin{cases} 1 & \text{yes} \\ 0 & \text{no} \end{cases} \end{cases} \tag{5}$$

where  $P_{object}$  stands for the probability of whether there is detection object in the grid, 1 for yes and 0 for no,  $IOU$  is the intersection of the prediction target box and actual target box generated by the grid, and  $S_1, S_2, S$  stand for the area of overlap of the prediction target box, actual target box and the area that the two boxes overlap.

The grid category probability divides the grid into  $N$  categories, and then the distribution of the category of identification object in grid,  $B$ , can be obtained by multiplying the probability when a grid includes identification object belongs to a category by credibility. Moreover, it can also detect the fitting degree of the prediction box and actual box.

The formula of  $B$  is:

$$B = P_{class/object} \times A \tag{6}$$

where  $P_{class/object}$  stands for the grid category probability.

Finally, when the grid category probability and target box are generated, the prediction box with the greatest reliability in the grid target box is selected by non-maximum suppression algorithm, and the category probability of the prediction box is calculated using equation (6). Comparing the calculated results with the set threshold, an effective prediction box is selected, and the identification target is judged according to the category of the prediction box.

The number of prediction boxes which are identified as aircraft is recorded, which is used as the number of aircraft identified in the input image, that is, the number of aircraft in the recognition area.

As the identification cameras used in airport have the same specification and the images are regularized before entering the recognition model, the size of the recognition images is consistent in theory. The number of aircraft counted in the previous step can be directly treated as the traffic density. By comparing traffic density with classification criteria, the current operation status of airports can be judged.

Command and dispatch opinions are put forward according to the current airport operation status, including commanding the aircraft with late take-off time to transfer to the non-busy runway or commanding the aircraft to land on the runway with low traffic density.

For the airport aircraft recognition management model proposed in this study, the accuracy of the scheduling management opinions depends on the accuracy of the target recognition part of the model, and the accuracy of target recognition mainly depends on the setting of loss function in the learning process. Loss function, also known as learning objective function in the recognition model in this study, is used to reversely adjust the weight parameters of the neural network according to the error between the predicted value and actual value in the process of training the model. The loss function [14] used in this study can be expressed as:

$$\begin{cases} L(p, u, t^u, v) = L_{cls}(p, u) + g \cdot L_{loc}(t^u, v) \\ g = \begin{cases} 1 & u \geq 1 \\ 0 & u = 0 \end{cases} \\ L_{loc}(t^u, v) = \sum_{i \in (x, y, w, h)} R(t_i^u, v_i) \end{cases} \quad (7)$$

where  $L(p, u, t^u, v)$  is a joint loss function,  $L_{cls}(p, u)$  stands for classification loss,  $L_{loc}(t^u, v)$  stands for regression loss of prediction box,  $v$  stands for the regression value of the prediction box, i.e., the parameter value of the actual target box,  $p$  is the probability of grid category,  $t^u$  is the predicted value of the target box,  $u$  refers to the probability of whether there is detection object, and  $R(\bullet)$  stands for a robust loss function.

However, the loss function mentioned above just classifies prediction object box whose threshold is smaller than  $IOU$  as one category in practical application. Although it is not a big problem in recognition, the positioning is not accurate, i.e., the specified object in the image can be accurately classified, but its performance of indicating the position of object in the image is not good. Therefore, the loss function is improved in this study.  $L_{cls}(p, u)$  is transformed into definite integral under  $IOU$  threshold, then the improved loss function:

$$L(p, u, t^u, v) = \frac{\sum_h \{L_{cls}(p_h, u_h) + g \cdot L_{loc}(t^u, v)\}}{n} \quad (8)$$

where  $p_h, u_h$  stand for the grid category probability under  $IOU$  threshold  $h$  and the probability of whether there is detection target.

## 4. Experimental Analysis

### 4.1 Experimental environment

The experiment was carried out using MATLAB software [15] on a laboratory server with configuration of Windows 7 system, I7 processor and 16G memory.

### 4.2 Experimental data

Aircraft traffic images of different runways in an airport within 7 days were collected. The picture was taken every one hour by the airport cameras installed for the runway management. Moreover the actual number of aircraft and management staffs in the corresponding runway at the same moment was collected and taken as the management and dispatch information. The image, data and management scheduling information of the first six days were used as the training sets, and the image, data and management scheduling information of the seventh day were used as the testing set.

### 4.3 Experimental parameters

As shown in Table 1, three neurons were used in the input layer, and the input size of the node was  $500 \times 500$ ; there were 64 convolution kernels in the convolution layer, and the size of the convolution kernel was  $3 \times 3$ ; there were six convolution layers in the model; every convolution layer was connected with a pooling layer; there were six pooling layers; the pooling operation in the pooling layer was maximum pooling; the size of the pooling box was  $2 \times 2$ , and the sliding step size was 2; there were 4096 neurons in the fully connected layer; the judgement threshold value of target identification was set as 0.5.

Table 1. Structural parameters of CNN-based recognition model

Structure name	Parameter	Structure name	Parameter
Input layer	$500 \times 500 \times 3$	Pooling layer	$2 \times 2$
Convolution layer	$3 \times 3 \times 64$	Fully connected layer	4096

### 4.4 Evaluating indicator

The traffic density classification criteria of airport operation state were as follows. No more than 10 planes in the picture was determined as low traffic density, and the dispatching opinion was that the aircraft take-off and landing are arranged according to the original plan. Ten to twenty aircraft in the picture was determined as medium traffic density, and the dispatching opinion was that the aircraft with the latest take-off time was transferred to other runways and the take-off and landing of the other aircraft were arranged according to the original plan.

More than twenty aircraft in the picture was determined as high traffic density, and the dispatching opinion was that the aircraft which took off late was transferred to runways with low traffic density and moreover the landing aircraft was guided to land on the runway with low traffic density.

In this study, the accuracy and recall rate were used to evaluate the management scheduling performance of the identification management model.

The formulas for calculating the two indicators are:

$$\begin{cases} Pr = \frac{tp}{tp + fp} \\ Re = \frac{tp}{tp + fn} \end{cases} \quad (8)$$

where Pr refers to the accuracy rate, Re refers to the recall rate, tp is the number of objects which are determined as some category and moreover belongs to the category actually, fp is the number of objects which are determined as some category but belongs to another category, and fn is the number of objects which are determined as another category but belongs to some category actually.

The accuracy rate represents the proportion of the situation when the dispatch opinion made is the same with the actual opinion, and the recall rate represents the proportion of the correct dispatch opinions in actual dispatch opinions of that category.

### 4.5 Experimental results

The results of target recognition of the two models are shown in Figure 4. Both the recognition model based on CNN and the improved recognition model could recognize the aircraft in the airport. But the comparison of the recognition effects shown in Figure 4(a) and 4(b) demonstrated that the improved recognition model was more accurate in positioning, and the incomplete aircraft image was not recognized in Figure 4 (a). It can be seen from the visual comparison results that the improved recognition model had better recognition and detection effect for airport aircraft.



Figure 4: (a) Aircraft recognition results of the recognition model before improvement; (b) aircraft recognition results of the improved recognition model

As shown in Figure 5, in the peak time of the airport, i.e., from 9:00 to 15:00, the traffic density of the airport runway was large. On the whole, the improved recognition model had better recognition effect on the number and density of aircraft in the airport than the recognition model without improvement.

Relatively speaking, in the period of low traffic density, the number and density recognition of the two recognition models was close to the real values; in the period of high traffic density, the number recognition of the improved recognition model was closer to the real value.

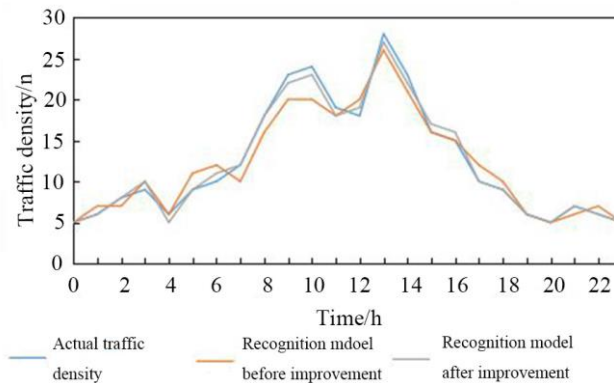


Figure 5: The recognition results of airport traffic density of the two models in the 7<sup>th</sup> day

*Table 2. Management performance of the two models*

Recognition performance	Recognition model before improvement	Recognition model after improvement
Detection time/s	0.9077	0.4712
Accuracy/%	89.12	96.51
Recall rate/%	88.15	94.65

As shown in Table 2, the feedback time of recognition management opinion of the model before improvement was 0.9077 s, the accuracy of management dispatch opinion was 89.12%, and the recall rate was 88.15%; the feedback time of opinion management opinion of the model after improvement was 0.4712 s, the accuracy of management dispatch opinion was 96.51%, and the recall rate is 94.65%.

It can be seen that the improved recognition model not only improved the accuracy and recall rate in the management of dispatch opinion feedback, but also significantly reduced the time required for recognition and detection.

## 5. Conclusion

In this study, the CNN and the airport target recognition model based on CNN were introduced briefly. In order to improve the accuracy of target recognition and location, the loss function of the model was improved. After that, the two recognition models were simulated and analyzed by using MATLAB software. The results were as follows.

In terms of the visual effect of recognition, both recognition models could recognize aircraft effectively, but the improved recognition model could locate the target more accurately and moreover could recognize the target with slightly incomplete image. The improved recognition model was closer to the real value in traffic density recognition than the recognition model without improvement, especially in the rush hour. The improved recognition model not only had higher accuracy and recall rate, but also needed shorter recognition time.

## Acknowledgements

This study was supported by Scientific Research Fund of Civil Aviation Flight University of China: Study on Performance Evaluation of Airport in Economic Attributes (grant number J2012-31).

## References

[1] Zhao D, Shi J, Wang J, Jiang Z. "Saliency-constrained semantic learning for airport target recognition of aerial images," *Journal of Applied Remote Sensing*, 2015, 9(1):096058.  
 [2] Zhu D, Wang B, Zhang L. "Airport target detection in remote sensing images: a new method based

on two-way saliency," *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(5):1096-1100.  
 [3] Xiao Z, Gong Y, Long Y, Li D, Wang X, Liu H. "Airport detection based on a multiscale fusion feature for optical remote sensing images," *IEEE Geoscience & Remote Sensing Letters*, 2017, PP(99):1-5.  
 [4] Camps O, Gou M, Hebble T, Karanam S, Lehmann O, Li Y, Radke R, Wu Z, Xiong F. "From the lab to the real world: re-identification in an airport camera network," *IEEE Transactions on Circuits & Systems for Video Technology*, 2016, 27(3):1-1.  
 [5] Zhang P, Niu X, Dou Y, Xia F. "Airport detection on optical satellite images using deep convolutional neural networks," *IEEE Geoscience & Remote Sensing Letters*, 2017, 14(8):1183-1187.  
 [6] Chen F, Ren R, Van de Voorde T, Xu W, Zhou G, Zhou Y. "Fast automatic airport detection in remote sensing images using convolutional neural networks," *Remote Sensing*, 2018, 10(3):443.  
 [7] Xia F, Li H Z. "Fast detection of airports on remote sensing images with single shot multibox detector," *Journal of Physics: Conference Series*, 2018, 960:012024.  
 [8] Peng W, Cao Y, Shen C, Liu L, . "Temporal pyramid pooling-based convolutional neural network for action recognition," *IEEE Transactions on Multimedia*, 2017, 27(12):2613-2622.  
 [9] Ding J, Chen B, Liu H, Shen H T. "Convolutional neural network with data augmentation for SAR target recognition," *IEEE Geoscience and Remote Sensing Letters*, 2016:1-5.  
 [10] Sanchez-Perez L A, Sanchez-Fernandez L P, Shaout A, Suarez-Guerra S. "Airport take-off noise assessment aimed at identifying responsible aircraft classes," *Science of the Total Environment*, 2016, 542(Pt A):562-577.  
 [11] Zhu M M, Xu Y L, Ma S P, Tang H, Xin P, Ma H Q. "Airport detection method with improved region-based convolutional neural network," *Acta Optica Sinica*, 2018, 38(7):0728001.  
 [12] Frana G B, De Almeida M V, Rosette A C. "Self-nowcast model of extreme precipitation events for operational meteorology," *Atmospheric Measurement Techniques Discussions*, 2015, 8(10):10635-10661.  
 [13] Fei G, Teng H, Sun J, Wang J, Hussain A, Yang E "A new algorithm of SAR image target recognition based on improved deep convolutional neural network," *Cognitive Computation*, 2018(5):1-16.  
 [14] Kim S. "Infrared variation reduction by simultaneous background suppression and target contrast enhancement for deep convolutional neural network-based automatic target recognition" *Optical Engineering*, 2017, 56(6):063108.  
 [15] Pei J, Huang Y, Sun Z, Zhang Y, Yang JM, Yeo TS. "Multiview synthetic aperture radar automatic target recognition optimization: modeling and implementation," *IEEE Transactions on Geoscience & Remote Sensing*, 2018, PP (99):1-15.