

PAPER • OPEN ACCESS

A Convolutional Neural Network for Airport Security Inspection of Dangerous Goods

To cite this article: Qiang Gao *et al* 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **252** 042042

View the [article online](#) for updates and enhancements.

A Convolutional Neural Network for Airport Security Inspection of Dangerous Goods

Qiang Gao^{1,*}, Zhen Li^{2,a} and Jun Pan^{1,b}

¹School of Airport Management, Guangzhou Civil Aviation College, Guangzhou, China

²School of Cosmetics and Art Design, Guangdong Food and Drug Vocational College, Guangzhou, China

*Corresponding author e-mail: gaoq11@qq.com, ^acn.lizhen@qq.com,

^b381788030@qq.com

Abstract. According to the problems of heavy workload, low efficiency, easy fatigue misjudgement with artificial recognition and imbalance of dangerous goods image dataset in airport security inspection caused the low recognition accuracy, a convolution neural network automatic recognition model based on oversampling for dangerous goods is proposed. Firstly, the oversampling technique is used to equalize the dataset of dangerous goods image, and then the image is inputted into the convolution neural network model composed of four convolution layers and one full-connection layer for training. The stochastic deactivation optimization technique is introduced in the training to get better recognition effect. The experimental results on a dangerous goods image dataset of public security in 2017 show that the recognition accuracy of the model can reach 90.7% after equalization, which is 33.4% higher than that before equalization. In addition, the recognition accuracy of the model is 5.8%, 7.2% and 5.4% higher than that of GoogleLeNet, AlexNet and ResNet respectively. The model has high recognition accuracy and good real-time performance, which is of positive significance to improve the level of airport security intelligence.

1. Introduction

Airport security inspection is an important link to ensure the safe flight of civil aviation, and is related to social stability. At present, the international situation is complex and severe. At the same time, the airport passenger flow is explosive growth, and dangerous goods are more diversified, which have brought great challenges to the high quality and efficiency of security inspection. At present, China's airport safety inspection of dangerous goods recognition also depends on the X optical machine operator to manually identify. Because of the different angle of luggage and the density and volume of objects, the X-ray machine image characteristics of dangerous goods are very different. Even the same dangerous goods, the images presented in the X-ray machine are not the same. This brings great difficulties to the machine operator to accurately identify dangerous goods, and is very easy to judge errors, which causes a hidden danger to the safety of civil aviation. In addition, the machine operator to identify dangerous goods image belongs to a typical repetitive task, a long time of high-intensity,



intense work, the brain and body of machine operator are easy to fatigue, which has a great impact on the accuracy and efficiency of dangerous goods recognition.

Convolutional neural network (CNN) [1] can automatically learn image features, which are widely used in face recognition [2, 3], vehicle detection [4, 5], traffic sign recognition [6] and object detection [7, 8] image recognition fields. Due to the excellent performance of CNN models such as GoogleNet AlexNet [10] and ResNet [11] in the image recognition contest, many researchers have applied these models directly or improved to image recognition in related fields.

Although CNN has achieved good results in image recognition in different fields, research and application in the field of airport security and dangerous goods have not been found. In addition, traditional CNN assumes that the number of images in each category in the dataset is balanced. When designing models, usually only attention to the recognition accuracy and real-time of the model is considered, and the influence of the imbalance of the dataset on the recognition effect is not considered. Based on this, this paper proposes an over-sampling-based convolutional neural network automatic recognition model for dangerous goods, and compares it with the recognition effects of GoogLeNet, AlexNet and ResNet models to verify the validity of the model.

2. Convolution neural network model based on oversampling

This section will elaborate on the over-sampling-based convolutional neural network automatic recognition model for dangerous goods. Firstly, the implementation of oversampling technology is described, and it is used to realize the equalization of the dataset of the security dangerous goods. Then the CNN model proposed in this paper is introduced from the aspects of convolutional layer, pooling layer, fully connected layer, optimization technology and training process.

2.1. Dataset equalization

The traditional machine learning recognition algorithm assumes that the number of samples in the dataset is similar, but in reality, the number of various types of samples is not balanced, usually a large number of classes are called majority classes, otherwise known as the minority class. As the traditional recognition algorithm ignores the imbalance of the number of samples in the dataset, recognition algorithms often biased towards the majority class, and the minority classes have a high false recognition rate. Generally speaking, the imbalance ratio of different kinds of samples exceeds 4:1, and the recognition algorithm cannot meet the recognition requirements because of the imbalance of data. The number of samples of the dangerous goods dataset used in this paper is shown in row 1 of Table 1.

As shown in row 1 of Table 1, the imbalance of the number of different categories in the dangerous goods data set is obvious. Therefore, it is necessary to deal with the imbalance of the dataset before designing the convolutional neural network model.

This paper uses Synthetic Minority Over-sampling Technique (SMOT) to implement dataset equalization. SMOT is a common oversampling technique, which randomly selects several nearest neighbor samples for each minority sample, and randomly selects points on the line between the samples and these nearest neighbors to generate new minority samples without repetition. SMOT makes the classification plane of minority classes extend to the space of majority classes, which can effectively avoid the over-fitting problem caused by random replication of samples. After SMOT equalization, the number of samples is shown row 2 of Table 1.

Table 1. The number of dangerous good with disequilibrium and equilibrium.

| | Explosives | Ammunition | Blunt | Control apparatus | Kindling material | Guns | sharps | Dangerous articles | pyrotechnics |
|----------------|------------|------------|-------|-------------------|-------------------|------|--------|--------------------|--------------|
| Disequilibrium | 264 | 299 | 43 | 986 | 744 | 325 | 99 | 188 | 241 |
| Equilibrium | 845 | 943 | 631 | 986 | 744 | 975 | 692 | 752 | 823 |

2.2. Convolutional Neural Network Model

After the Dangerous Goods image dataset is equalized by SMOT technology, which is will be inputted into convolutional neural network model to identify. The model consists of five hidden layers, the first four layers are Conv1, Conv 2, Conv 3 and Conv 4 convolutional layers, each of which is divided into convolutional layer (C), active layer (R) and pooled layer (P), After flattening the output result of P4 in Conv 4, it is connected to the fifth-layer fully-connected layer FC 5, and finally outputs the classification probability of the nine types of dangerous goods, and its structure is as shown in Figure 1.

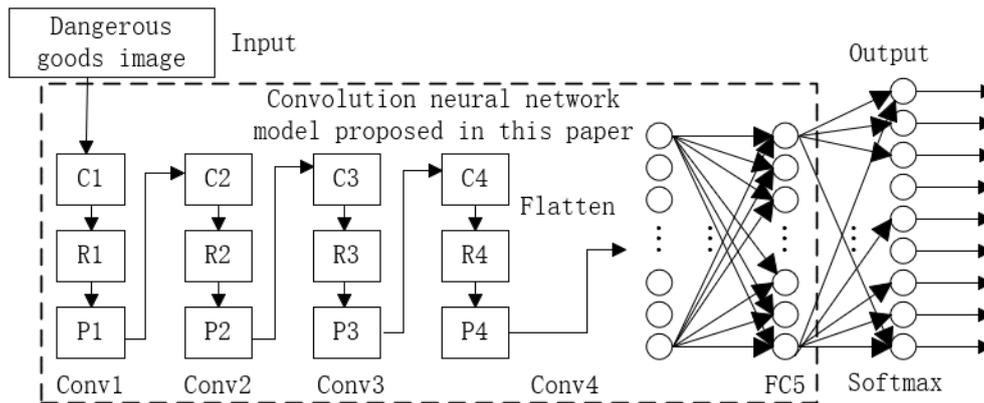


Figure 1. Convolution neural network model.

Convolutional layer: The size of the dangerous goods input to C1 is $100 * 100 * 3$. C1 and C2 use a $5*5$ size convolution kernel, and the number of convolution kernels is 32 and 64 respectively; C3 and C4 use $3*3$ convolution kernels, and the number of convolution kernels is 128; C1, C2, C3, C4 have a convolution movement step of 1.

Activation layer: R1, R2, R3, and R4 all use the ReLU activation function, which turns all negative values to zero while the positive values remain unchanged. This unilateral suppression operation makes the ReLU activation function have the advantages of sparse activation and fast convergence, and its function expression is shown in equation (1).

$$y = f(x) = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{if } x \leq 0 \end{cases} \quad (1)$$

Pool layer: P1, P2, P3 and P4 all use the maximum pool size with $2*2$ size and 1 step size.

Full connection layer: FC5 takes the results of P4 flattening as input and contains 1024 neurons; FC5 is fully connected to the final output layer, which contains 9 neurons.

In the training process, when the convolution and pooling operations are performed, the image feature map size changes as shown in Table 2.

Table 2. Three Scheme comparing.

| Convolution layer | Input feature map | Output feature map |
|-------------------|-------------------|--------------------|
| Conv1 | $100*100*3$ | $50*50*32$ |
| Conv2 | $50*50*32$ | $25*25*64$ |
| Conv3 | $25*25*64$ | $13*13*128$ |
| Conv4 | $13*13*128$ | $7*7*128$ |

2.3. Model optimization technique

2.3.1. Loss function. The loss function of the model using cross entropy cost function, cross entropy cost function can not only overcome the problem that the update weight of variance cost function is too slow, but also avoid gradients to dissipate, which is shown in equation (2).

$$C = -\frac{1}{n} \sum_x [y \ln a + (1 - y) \ln(1 - a)] \quad (2)$$

Among them, C is the loss value, n is the total number of training samples, a is the actual output of the neurons, and y is the expected output.

2.3.2. Optimization function. The optimization function of the model is Adaptive Moment Estimation (Adam) function. Adam dynamically adjusts the learning rate of each parameter by using the first-order moment estimator and the second-order moment estimator of the gradient. It has faster convergence speed and more effective learning effect. Moreover, it can correct the problems that exist in other optimization techniques, such as the disappearance of learning rate, the slow convergence or the large fluctuation of loss function caused by the parameter updating of high variance.

2.3.3. Random inactivation. In order to improve the generalization ability and reduce the risk of over-fitting, dropout random inactivation strategy is adopted at the fully connected layer of the model. Dropout is a regularization method to prevent over-fitting of neural network. It randomly discards the output value of the hidden layer part of the neural node during the process of neural network training, and does not need to update the weights associated with the nodes when the weights are updated by the back propagation. The working principle of dropout is shown in Figure 2.

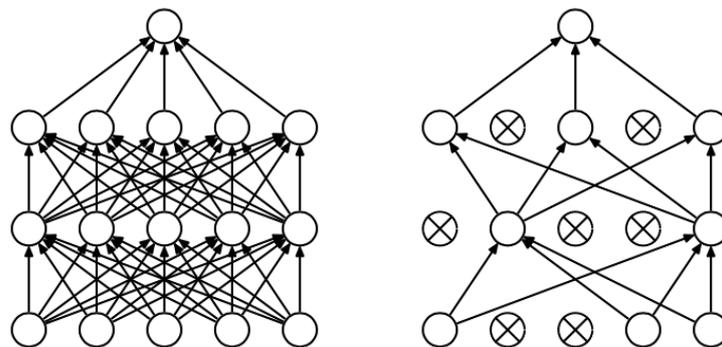


Figure 2. Dropout work diagram.

3. Experimental verification

3.1. Experimental platform and dataset

This experiment was carried out on a computer with Intel Corei5-3230M CPU, 2.60 GHz main frequency, 8G memory and W Windows 10 (64 bit) operating system, used Python 3.5 + TensorFlow 1.6 to implement the convolution neural network.

To verify the validity of the proposed model, the dataset used in this experiment is from the 2017 Public Security No.1 Bureau (National Final Edition) Dangerous Goods Image Library, which contains 3225 pairs of 9 types of dangerous goods images. Including explosives, ammunition, blunt, control apparatus, kindling material, guns and other weapons, sharps, dangerous articles and pyrotechnics.

The dataset from the dangerous goods image library of 2017 Public Security No.1 Bureau (National Final Edition) is called dataset I (9 categories, 3,225 images), and the dataset II (9 categories, 7,391 images) will be obtained by balancing the dataset I with SMOT technology. Datasets I and II are divided into training sets and verification sets. The images of dangerous goods in dataset I and II is shown in Figure 3.

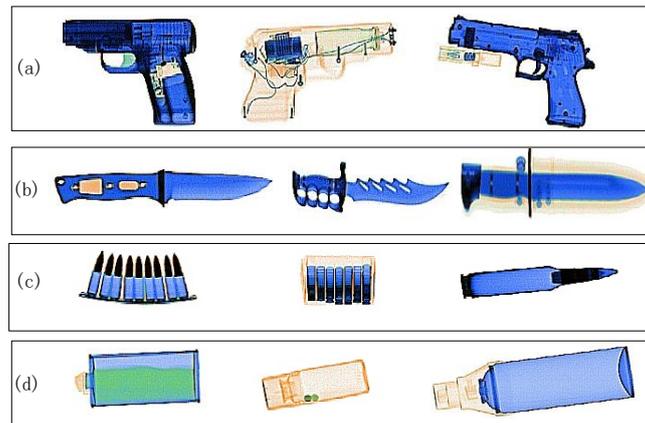


Figure 3. The image of dangerous goods in dataset I and II (part) (portion), (a) guns (b) control apparatus, (c) ammunition, (d) dangerous articles.

3.2. Validity verification

Dataset I and II are respectively input into the convolutional neural network model proposed in this paper for training and verification to test the validity of the model. The experimental parameters are: 80% training set and 20% verification set, 25 batches, 50 iterations and 0.3 dropout.

After 50 rounds of training, the loss of the training set and the recognition accuracy of the verification set are respectively shown in Figure 4 and Figure 5.

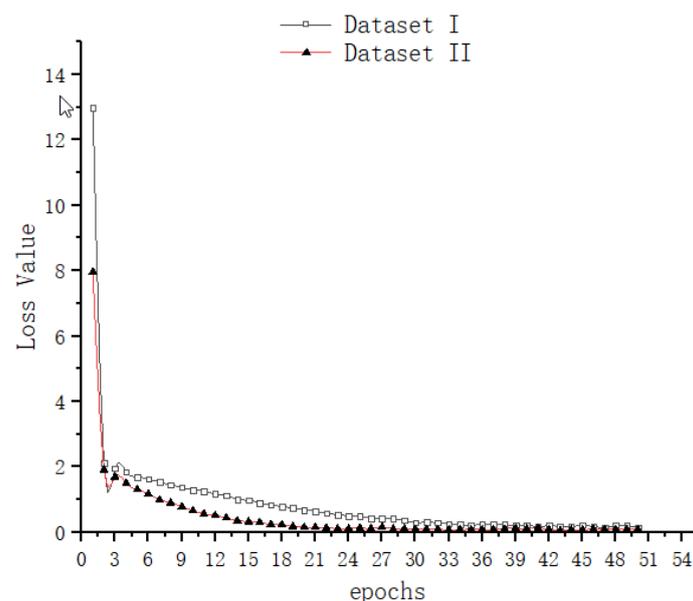


Figure 4. Loss value reduction chart.

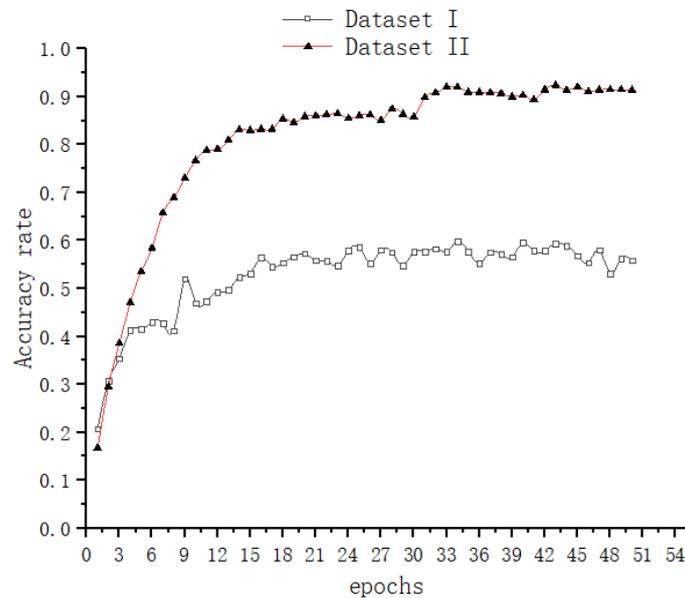


Figure 5. Recognition accuracy rate.

Experimental analysis: The experimental results show that the proposed convolutional neural network model in dataset II drops faster and tends to zero than in dataset I, which shows that the model has better convergence in data set II, and proves the effectiveness of SMOT in dataset equalization. In addition, in each iteration, the recognition accuracy of dataset II is higher than that of dataset I. After 33 iterations, the average recognition accuracy of dataset II is stable at 90.7, which is 57.3% higher than that of dataset I, which proves the validity of the convolutional neural network model proposed in this paper.

3.3. Method effect comparison

To further verify the validity of a convolutional neural network model (myCNN) proposed in this paper, the recognition effect of myCNN is compared with that of GoogLeNet, AlexNet and ResNet on dataset II. The experimental results are shown in Table 3.

Table 3. Dataset II recognition accuracy rate.

| Model name | Recognition accuracy |
|------------|----------------------|
| GoogLeNet | 84.9% |
| AlexNet | 83.5% |
| ResNet | 85.3% |
| myCNN | 90.7% |

Experimental analysis: From table 2, we can see that the recognition accuracy of myCNN is 5.8%, 7.2% and 5.4% higher than that of GoogLeNet, AlexNet and ResNet respectively, which further proves the validity of myCNN.

4. Conclusion

In this paper, the problem of automatic recognition of dangerous goods in airport security inspection based on imbalanced dataset is studied, and a convolutional neural network model based on oversampling is proposed. The model can accurately recognize the images of dangerous goods with imbalanced distribution, and provides a method for automatic recognition of dangerous goods. The

next step is to migrate the model to a complex environment (such as multi-hazard occlusion, overlap, etc.), and to study the validity of the model and the new problems.

Acknowledgments

This work was financially supported by Science and Technology Program of Guangzhou Civil Aviation College, China (17X0205, 18X0202); Science and Technology Program of Guangdong Food and Drug Vocational College, China (2016YZ028) fund.

References

- [1] Ziler M D, Fergus R. Visualizing and understanding Convolutional networks [C] // European Conference on Computer Vision. Springer, Cham, 2014, pp. 818-833.
- [2] Chao stone, with ethylene weeks, the pools. Adjustable supervision function convolution integration of human neural network multilayer facial feature recognition gender [J/OL]. 2019, 36(4). [2018-02-07]. <http://www.aocmag.com/article/02-2019-04-012.html>.
- [3] Wu Wei, Qiu Weigen. Face recognition based on improved deep convolutional neural network. Computer Engineering and Design, 2017, 38 (8), pp.2246-2250.
- [4] Cao Shiyu, Liu Yuehu, Li Xinzhao. Vehicle Target Detection Based on FastR- CNN. Journal of Graphics and Graphics, 2017, 22 (5), pp.671-677.
- [5] Songhuan Sheng, Zhang Xiangqing, Zheng Baofeng, strict application of the Detection Complex Teng scene depth vehicle target learning process-based computer, 2018, 35(4), pp.1270-1273.
- [6] Wang Xiaobin, Jin Jie, Liu Wenju. Traffic Sign Recognition Based on Optimized Convolutional Neural Network. Computer Application, 2017, 37(2), pp.530-534.
- [7] Girshick R, Donahue J, Darrell T, et Al. Rich feature Hierarchies for accurate object detection and semantic Segmentation [C] // IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2013, pp.580-587.
- [8] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6), pp.1137-1149.
- [9] Szegedy C, Liu W, Jia Y, et al. Going deeper with Convolutions [C] // IEEE Computer Society. IEEE Conference on omputer Vision and Pattern Recognition, 2014, pp.1-9.
- [10] Krizhevsky A, Sutskever I, Hinton G E. Imagenet Classification with deep convolutional neural networks [C] // Advances in Neural Information Processing Systems, 2012, pp1097-1105.
- [11] He K, Zhang X, Ren S, et al. Deep residual learning for Image recognition [C] // Proceedings of the IEEE Conference on omputer Vision and Pattern Recognition, 2016, pp.770-778.