Optical Image Damage Detection Technology Based on Convolutional Neural Networks

Y Wang¹, M Yu¹, S Zhong¹, L Wang¹, and M Xu¹

¹University of Electronic Science and Technology of China, Chengdu 611731, China.

wangyanlkj@sina.com

Abstract. In the daily work of technical defense of universities, many optical devices need to be managed, and there are many components that are easily damaged in these devices. Therefore, timely detection of component damage in optical equipment is of great significance for reducing equipment losses and improving work efficiency. The traditional detection method is to use professional equipment for testing after removing the equipment. In this way, for the technical defense work of universities, the real-time online of the equipment cannot be guaranteed and the use cost of the equipment will increase. In the field of computer vision, deep learning and convolutional neural network technology have begun to play an important role with the wave of artificial intelligence. Therefore, this subject attempts to apply this technology to the damage detection of technical defense equipment in universities, thereby improving work efficiency and reducing equipment loss.

1. Introduction

First of all, the images used in this paper are produced in practical applications of optical equipment. There are very few damage images in the images obtained by the imaging of the optical device, and the internal components involved in the optical device are numerous. If the optical device is destroyed artificially in order to obtain the experimental data of the imaging damage imaging of the optical component, obtaining the experimental data in this way is very expensive and cannot be allowed. On the other hand, since the design of a classifier based on a convolutional neural network will be used in this paper, the classification of damaged and non-damaged areas will be realized. During the training process of the convolutional neural network classifier, a large amount of data is required to support the training to ensure that the trained network model can fully learn the information in the damaged image, so as to use these information to classify the image.

Therefore, in this paper, the existing data set is segmented first[1]. According to the analysis of the existing data set, the area in the existing data set can be divided into a damaged area, a non-damaged area, and a suspected damaged area. Through this classification, a better classification can be performed for all image regions in the data set. By analyzing the existing data set, we find that the size of the damage area is concentrated between $10 \times 10$ and $15 \times 15$ pixels.
2. Data set production

As shown in Figure 1 above, pixel matrix mapping is performed on the pre-processed damage images of optical components, and the area of the damage points is observed. According to this matrix, the overall damage image of optical components is divided according to the size of the damage point. The size of the original damage image is $600 \times 600$. We design the damage discrimination area to be $30 \times 30$ pixel areas. During the data set production process, there were 317 area images selected as the damage area. Since the size of the damaged area is smaller than the size of the recognition area block, its position in the recognition area block is not fixed. Therefore, each damage area is placed in 9 different positions, and the surrounding pixels are collected according to their positions. As a result, 2853 damage point samples were obtained. In order to ensure the balance of the sample size, the number of non-damaged area samples selected in this experiment was 2658. The total number of samples at this time is 5511 image blocks.

In order to avoid the irreversible consequences of the subsequent training of the classifier due to the number and distribution of sample selection, this paper expands the sample number. This subject has fully investigated the public standard datasets of other pictures in the field of image processing, such as the MNIST [5] dataset, the CIFAR 10 [8] image dataset, and the CIFAR 100, NIST SD19 [6] dataset, and GTSRB [7] dataset. Finally, in the original damage sample data, 200 damage points data and 200 background data (data without any operation) are extracted for expansion. After expansion, there are 1600 damaged samples and 1600 background samples, and a total of 3,200 make up the test set Test_1. After processing, the above 400 test samples were deleted from 5511 samples in the original sample set, and then expanded into a data set containing 40888 samples. Among them, there are 21,224 damaged sample data and 19,664 non-damaged data. After this processing, the training set has a total of 40,888 sample data. The training data set formed at this time is recorded as Data_1.

Since all the training sets and the samples in the test set do not coincide with each other, this guarantees the validity of the test set used to test the results.

![Figure 1](image1.png)  
**Figure 1** Mapping the damage area to a pixel matrix

![Figure 2](image2.png)  
**Figure 2** Data set expansion example
3. Convolutional neural networks model

In terms of the online detection of damage images of optical components, our goal is to identify whether damage has occurred through a given image or an image area. And we hope that the image or image block can still be identified after simple deformation, rotation, translation, etc.

![Figure 3](image-url)

**Figure 3** Classification example diagram

The classifier used in this paper is a classifier designed based on the convolutional neural network theory. It does not need to use the original method to extract features from the image and then train the machine learning classifier. The convolutional neural network only needs to process the image data set, then design its network model, optimize its parameters, and then improve the accuracy.

In this paper, the main application of the convolutional neural network model is shown in Figure 4 below:

![Figure 4](image-url)

**Figure 4** Network model used in this paper

After the damaged image block is input in this paper, it is processed by a convolution layer. And then processed further by the pooling layer, the convolutional layer and the pooling layer appear twice alternately, then pass through two fully connected layers, and finally connect an output layer. The specific model parameters used in the network model are as follows:

| Layer | Input | C1   | S1   | C2   | S2   | Fc1 | Fc2 | Output |
|-------|-------|------|------|------|------|-----|-----|--------|
| Size  | 30*30 | 28*28| 14*14| 12*12| 6*6  | 172 | 36  | 2      |

The size of the convolution kernel used in the model is 3x3, the pooling layer filter size is 2x2. There are 16 convolution kernels extracted in the convolution layer. The convolution kernel used by the two convolutional layers is the same size as the pooling filter used by the two pooling layers.

An activation function similar to ReLU is used [2] [3] [4]. This activation function is simple but works well. This is because when training the classification model, there are usually few features related to the target features. Therefore, the sparse model can better perform deep mining on relevant features and fit the corresponding training data. Its expression is as follows:
\[ \phi(x) = \begin{cases} x & \text{if } x > 0 \\ 0.1x & \text{if } x \leq 0 \end{cases} \]

Compared with other activation functions, the activation function of this experiment has the following advantages: For linear functions, activation functions are more expressive. For non-linear functions, there is no problem of gradient disappearance because the gradient of the non-negative interval is constant.

4. Result

In order to verify the effect of the network structure model used in this paper on the online detection of optical component damage. The experiments in this paper are compared with traditional HOG + SVM-based machine learning methods and deep belief networks (DBN) in the field of image processing. The size of each layer of the network structure used in the deep belief network in this paper is 1225-1000-1000-800-200-2. For these three structures, this paper uses the Data_1 training set to train them, and then test with Test_1 test data set.

After training, this paper use the Test_1 test dataset to test its trained classifier model. The experimental results are as follows:

**Table 2** Test Results of CNN Model in This Paper

|     | Damage | Normal |
|-----|--------|--------|
| Right | 1493   | 1424   |
| Wrong | 107    | 176    |

**Table 3** Results of DBN Model Test

|     | Damage | Normal |
|-----|--------|--------|
| Right | 1407   | 1317   |
| Wrong | 193    | 283    |

**Table 4** Results of HOG + SVM Model Test

|     | Damage | Normal |
|-----|--------|--------|
| Right | 1189   | 1104   |
| Wrong | 411    | 496    |

According to the test results, the performance indicators of each classifier are shown in Table 5 below:

**Table 5** The Performance Indicators of The Classifier

| Classifier Name | CNN Model in This Paper | DBN | HOG+ SVM |
|-----------------|-------------------------|-----|---------|
| Accuracy        | 91.17%                  | 85.13% | 71.66%  |
| Precision       | 89.45%                  | 83.25% | 70.56%  |
| Recall          | 93.31%                  | 87.94% | 74.31%  |
| F-Measure       | 0.9223                  | 0.8651 | 0.7296 |

As can be seen from the above classifier performance indicators. The performance indicators of the classifier used in this paper are better than the performance indicators of the other two comparative experimental classifiers. In addition, in order to explore the influence of parameters in the convolutional neural network model on the performance of the trained classifier, this paper establishes another network...
model. The size of the first layer of the convolution kernel is 3×3. The size of the second layer of the convolution kernel is the same as in this paper. The size of the merged filter template is 2×2. The number of convolution kernels used for feature extraction in each layer is 9. Its structure is as follows:

**Table 6** The Parameters of Network After Changes

| Layer | Input | C1   | S1   | C2   | S2   | F1   | F2   | Output |
|-------|-------|------|------|------|------|------|------|--------|
| Size  | 30*30 | 24*24| 12*12| 10*10| 5*5  | 25   | 10   | 2      |

After training the above model using the data set in Data_1, and test it with Test_1. The experimental results are as follows:

**Table 7** Results of Convolutional Neural Network Test After Parameter Changes

|       | Damage | Normal |
|-------|--------|--------|
| Right | 1412   | 1358   |
| Wrong | 188    | 242    |

After calculation, the accuracy is 86.56%, the precision is 85.37%, the recall is 88.25%. It can be found that for the deep belief network and HOG + SVM structure, the network model of the above comparative experiment still has a slight advantage. However, compared with the structure in this paper, the effect obtained under the same training set and test data set is not as good as the structure in this paper.

**5. Conclusion**

Obtaining damage images of optical components is costly and difficult. Therefore, in order to meet the needs of the sample data size for the training of the classifier, this paper has sampled at different locations and generated a large amount of pseudo sample data, thereby expanding the damage data set. In this paper, the structure of the convolutional neural network is used to classify and detect the damage images of optical components. Compared with the traditional HOG + SVM machine learning algorithm and deep belief network, the experiments show that the actual detection results are better than these two algorithms. Subsequently, it was compared with the network structure of changing parameters to verify the influence of different structural parameters on the experimental results under the same convolutional neural network.

Through the experiments in this paper, it is confirmed that the method used in this paper has significantly improved the effect of image damage detection. This has an important role in the maintenance of optical equipment involved in technical defense work in colleges and universities. It is also important for daily equipment management. It also plays an important role in the management of technical defense equipment in universities.

**References**

[1] Gonzalez R C , Woods R E . Digital Image Processing (3rd Edition)[M]. Prentice-Hall, Inc. 2007.
[2] Krizhevsky A , Sutskever I , Hinton G . ImageNet Classification with Deep Convolutional Neural Networks[J]. Advances in neural information processing systems, 2012, 25(2).
[3] Redmon J , Divvala S , Girshick R , et al. You Only Look Once: Unified, Real-Time Object Detection[J]. 2015.
[4] Shafiee M J , Chywl B , Li F , et al. Fast YOLO: A Fast You Only Look Once System for Real-time Embedded Object Detection in Video[J]. 2017.
[5] Rosenblatt, F. "The perceptron: a probabilistic model for information storage and organization in the brain." Psychological Review65.6(1958):386-408.
[6] P Smolensky. “Information Processing in Dynamical Systems: Foundations of Harmony Theory "CU-CS-321-86.
[7] Hinton, G. E., S. Osindero, and T. Yw. "A fast learning algorithm for deep belief nets." Neural
Computation 18.7(2006):1527-54.

[8] Cun, Y. Le, et al. "Handwritten Digit Recognition with a Back-Propagation Network." Advances in Neural Information Processing Systems1990:396--404.