Research on Ceramic Sanitary Ware Defect Detection Method Based on Improved VGG Network

Bowen Teng, Hongjian Zhao, Pei Jia, Junfeng Yuan, Chenghua Tian

Beijing Machinery Industry Automation Research Institute Co., Ltd, Beijing 100120, China

Abstract. The defect detection of ceramic sanitary ware is often done manually, which is inefficient and unhealthy. The development of deep learning technology makes a non-contact and high-efficient method of inspection possible. In this paper, based on VGG-16 network, we proposed an optimized method for ceramic sanitary ware defect detection. We carried out pre-processing to denoise and enhance the original image data, improved activation function of MReLU, and used transfer learning method to train the model. The results of the test on the ceramic sanitary ware defection data set showed that the accuracy can reach 97.48%, which is 6.46% higher than that of the model using the ReLU activation function. The detection speed can reach 7 fps, which could meet the requirements of industrial online real-time production.

1. Introduction

China is the largest ceramic sanitary ware producer in the world. There are many production enterprises in ceramic industry. However, the automation degree of the entire ceramic industry is low, which means that the ceramic industry belongs to the traditional high pollution, high strength, toxic and harmful industry. The defect detection of ceramic sanitary ware products mainly relies on manual work and requires technical personnel to hold it with their hands and look at it with their eyes. This inefficient detection method will not only affect the surface integrity of the finished products, but also do harm to the human body, especially increasing the working time. The technical personnel will be tired, which seriously affects the accuracy of defect detection.

With the development of deep learning technology [1], it has a prominent performance in the defect detection field. As a classical deep learning model, convolutional neural network has many advantages, such as excellent work efficiency, real time, fault tolerance, learning ability and parallel processing ability, which are not available in the traditional methods. On the one hand, it can deal with some situations where the environmental information is complex, the background is not clear, and the logic rules are not clear. On the other hand, it can also identify the missing or distorted samples. The convolutional neural network has a very good adaptive ability and image recognition effect [2]. Because of these advantages, convolutional neural network is widely used in target detection and recognition. The application of convolutional neural network in sanitary ware defect detection has a higher recognition rate and wider applicability than the traditional detection and identification methods.

In daily life, every family needs to use ceramic sanitary ware products. Facing with such a huge demand, manufacturers must strictly monitor product defects in the production process, improve product quality, and adopt new production technologies to reduce costs. In this
manuscript, we used the convolutional neural network method to detect the defects of ceramic sanitary ware to improve the efficiency and accuracy of defect detection, and increase the automation degree and intelligence of the industry.

1.1. Convolutional Neural Network

Convolutional Neural Network, first proposed by Yann LeCun [3-4] et al., is a typical deep learning model. Convolutional Neural Network has three characteristics: local receptive field, weight sharing and down sampling. These three features enable the Convolutional Neural Network to effectively extract image features and greatly improve the results of the convolutional neural network, which make it favored by more and more researchers.

The basic structure of convolutional neural network consists of three parts: The Convolutional Layer, the Pooling Layer, and the Fully Connected Layer. The convolutional layer extracts features from multiple learnable convolution kernels, and maps the underlying information to high-level features by means of local connection. The size and number of convolution kernels are important for the recognition ability of convolutional neural networks. The pooling layer is usually followed by the convolution layer, which is mainly used to reduce the size of the feature map and the computation of the network, and improve the operation speed of the network. Maximum pooling and average pooling are the two most common pooling methods. The feature map is divided into equal and non-overlapping regions. Take a maximum value of each region or average all elements. The fully connected layer is usually arranged in the last layers of the convolutional neural network. The input is the feature map obtained by convolutional layer and the pooling layer. The fully connected layer fuses the extracted high-level features into a probability distribution, so that the network can obtain the classification result according to the probability.

The VGG-16 convolutional neural network is a type of convolutional neural network that won the ILSVR (ImageNet) championship in 2014 and has excellent model migration capabilities [6]. VGG-16 consists of 13 convolutional layers, 5 pooling layers, and 3 fully connected layers. The model structure is shown in figure 1. The main features of VGG-16 are simple: (1) the convolution kernel parameters adopted by convolutional layers are the same, that is, the size of the convolution kernel is 3*3; (2) the pool kernel parameters adopted by pooling layers are the same; (3) the model is composed of several convolutional layers and pooling layers, and it is easy to form a deep neural network structure.

Figure 1. Network structure of VGG-16
1.2. ReLU Activation Function
The traditional activation functions of convolutional neural network are Sigmoid function and Tanh function [7], but both of them have the fatal disadvantage of gradient dispersion, so at present, the activation function of neural network model generally adopts unsaturated modified linear function[8]. Rectified linear unit (ReLU) is one such function and its mathematical expression is:

\[ y(x) = \max(0, x) \]  (1)

The ReLU function curve is shown in figure 2. If the value of the input is less than or equal to zero, then the output is zero; if the value of the input is greater than zero, then the output remains unchanged. This practice of directly forcing some data to zero creates a degree of sparsity [9]. It not only reduces the direct interdependence of parameters, but also alleviates the occurrence of over-fitting problem. Compared with the traditional activation function, the calculation of ReLU is faster because it does not include division and exponential operation. However, the output value after convolution operation is often negative, if all negative values are replaced by zero after ReLU function, then a large amount of input information will be lost after convolution, which is very disadvantageous to feature extraction, and is easy to affect the accuracy of the final identification.

![ReLU function](image)

**Figure 2.** ReLU function

2. Image Defect Detection Method

2.1. Image Preprocessing
Because of the complex production environment and various electromagnetic interference, it is easy to produce a lot of noise in the collected images of ceramic sanitary ware, which will seriously affect the identification result of convolutional neural network defect detection. And VGG-16 neural network requires a uniform input size of 224*224*3, so we need to adjust the size of the collected image and carry out image pre-processing to denoise and enhance, which mainly includes the following steps:

1. Filtering: Use adaptive median filtering method to remove noise;
2. Enhancement: Use Linear enhancement to enhance image defect features.

The preprocessing effects flow is shown in the following figures.
2.2. Improved Algorithm

2.2.1. Transfer Learning. The convolutional neural network has a large number of parameters that need to be trained. The training process needs a large number of labeled data and training time. And the convolutional neural network learning effect is not good due to the shortage of defect samples, which is easy to lead to the sub-fitting and over-fitting problems [10]. Transfer learning can just solve the problem that the samples are few and the network cannot fit. The principle of transfer learning is that, at the beginning, the neural network is trained under a large amount of data in the similar field, and relatively good network model parameters are obtained, then the trained parameters are directly applied to the target network for feature extraction, in the end, fine-tune the network parameters to obtain excellent experimental results. Transfer learning can often achieve better results, even better than retraining in many cases, and the model has a relatively strong generalization ability.

In this paper, we pre-trained VGG-16 network on ImageNet [11]. After training, we discarded the last layer parameters, added a custom classification network in the top layer of VGG-16 network, and replaced the original classification layer of VGG-16.

2.2.2. Modified Rectified Linear Unit Activation Function. Inspired by the ReLU function, an improved Modified ReLU (MReLU) is proposed in this paper. The mathematical expression of the Modified ReLU function is as follows:

\[ f(x) = \max(z_{\text{min}}, x) \]

Where \( x \) represents the input value of the function and \( f(x) \) represents the output of the function and \( z_{\text{min}} \) is a constant in the MReLU activation function. The MReLU activation function is shown in figure 4. As you can see from the figure, when the input \( x > z_{\text{min}} \), the output remains the same; when the input \( x < z_{\text{min}} \), the output is \( z_{\text{min}} \). Comparing with ReLU function, we can see that MReLU keeps some negative values, and solves the problem of information loss of ReLU function to some extent.
As shown in figure 5, VGG-16 and the parameters of the training model were transferred to the convolutional layer and the potting layer of the ceramic sanitary ware defect detection model. The main steps are as follows:

1. Collect the original data set of ceramic sanitary ware defects, and select the images needed to be detected;
2. Due to the large amount of noise in the original image data set, the pre-processing operations such as filtering and denoising, feature enhancement and data enhancement are needed before the input;
3. Establish the improved model and adjust the hyper-parameters;
4. Train the improved model;
5. Use the test set to test the effects of the model.

Figure 5. Overall flow diagram of the improved model
3. Experimental results and analysis

3.1. Experimental Environment
The convolutional neural network is based on the current mainstream deep learning framework, Tensor Flow, and is written in Python. The system hardware environment is as follows: ubuntu 16.04 operating system, Intel i7-8750H processor, Nvidia 1060 GPU, 16 GB RAM.

3.2. Data Set
The collected images were mainly divided into four types of defects: waviness ripples, cracks, stains and pinholes. We collected 200 images of every defect, and there are 800 images in total. Convolutional neural network training requires a large amount of training data, and the current sample size is not sufficient to support the network to complete the training and learning steps. Therefore, we used data augmentation method to expand the data, which enabled more effective feature extraction of convolutional neural network, and prevented model overfitting, and improved classification accuracy [12]. The original data was processed by random rotation, horizontal reversal and random clipping. The sample size was expanded to 4000 images. In addition, in order to complete the training of neural network effectively, the data set must be randomly divided into training set, verification set and test set according to the ratio of 8:1:1, the resulting model training data set distribution is shown in the following table.

Table 1. Data set

| Data set     | Defect sample number |
|--------------|----------------------|
| Training set | 3200                 |
| Verification set | 400               |
| Test set     | 400                  |

3.3. Results and Analysis
In order to verify the influence of MReLU on the precision of ceramic sanitary ware defect detection, a set of contrast experiments were designed and carried out. One group of experiments used ReLU activation function and the other group used MReLU activation function. All the hyper-parameters are the same except the activation function. Use the cross-entropy loss function and adopt the Adam optimizer to gradually reduce the learning rate to train the neural network. As the number of iterations increases, the learning rate is gradually reduced, which not only avoids the situation that the network cannot converge because of the high learning rate, but also avoids the problem that the training time of the network increases sharply because of the low learning rate. The weight of the network is set to a truncated normal distribution with a mean of 0 and a variance of 0.2, the initialization bias is 0.1, the Dropout value is set to 0.5, the learning rate is set to 0.001, the batch size is 32, and the number of network iterations is 5000. The results are shown in the table below:

Table 2. Corresponding classification accuracy of activation function

| Network model   | Classification accuracy /% | Detection Time / s |
|-----------------|-----------------------------|-------------------|
| VGG16+ReLU      | 91.02                       | 0.139             |
| VGG16+MReLU     | 97.48                       | 0.148             |

It can be seen from table 2 that the accuracy of the original activation function ReLU is 91.02%, and the accuracy of the network is 97.48% after using the improved MReLU, which is 6.46% higher than the former. The detection time is 0.139s and 0.148s, respectively, and the detection time is almost the same. It can be seen that the improved activation function significantly improves the accuracy of defect detection of ceramic sanitary ware with little effect on the detection time.
4. Conclusion
In this paper, a new defect detection method for ceramic sanitary ware based on VGG-16 model is proposed to solve the problem that the defects of ceramic sanitary ware are various and cannot be described accurately. And an improved MReLU activation function is proposed to solve the problem of the loss of the information of the ReLU activation function. Data augmentation techniques are used to increase the sample size and transfer learning is used to improve the efficiency of model learning. Through experiments, the accuracy of defect detection can reach 97.48%, which is 6.46% higher than that of using ReLU model, which proves that the method has strong ability of defect detection. The detection time is shorter and can reach 7 fps. It can meet the requirements of on-line real-time detection in industrial production and has a wide application prospect.

On the other hand, the model has many parameters and requires high performance of the device, which cannot be satisfied in many embedded platforms. The model compression should be considered in the following work.

Acknowledgments
Fund Project: The National Key Research and Development Program of China, Project No.: SQ2018YFB130146.

References
[1] Lecun Y, Bengio Y, Hinton G. Deep learning [J]. Nature, 2015, 521 (7553): 436.
[2] Zhang H, Wang KF, Wang FY. Progress and Prospects of the Application of Deep Learning in Target Visual Detection [J]. Acta Automatica Sinica, 2017, 43 (08): 1289 - 1305.
[3] Lecun Y, Bottou L. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86 (11): P. 2278 - 2324.
[4] Cun Y L, Boser B, Denker J S, et al. Handwritten Digit Recognition with a Back-Propagation Network [J]. Advances in neural information processing systems, 1997, 2 (2): 396 -- 404.
[5] Li H, Lin Z, Shen X, et al. A convolutional neural network cascade for face detection[C]/ 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2015.
[6] Ghamdi M A, Li M, Abdel-Mottaleb M, et al. Semi-supervised Transfer Learning for Convolutional Neural Networks for Glaucoma Detection[C]/ ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019.
[7] Gulcehre C, Moczulski M, Denil M, et al. Noisy Activation Functions [J]. 2016.
[8] Clevert, Djork-Arné, Unterthiner T, Hochreiter S. Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs) [J]. Computer ence, 2015.
[9] Peng Y. Design of Defect Detection System for LCD Screen of Electric Energy Meter Based on Convolutional Neural Network [D]. Zhengzhou University, 2019.
[10] Yang H, Harrison R, Yi Z F, et al. Changing Perceptions of Forest Value and Attitudes toward Management of a Recently Established Nature Reserve: A Case Study in Southwest China [J]. Forests, 2015, 6 (9): 3136 - 3164.
[11] Krizhevsky A, Sutskever I, Hinton G. ImageNet Classification with Deep Convolutional Neural Networks [J]. Advances in neural information processing systems, 2012, 25 (2).
[12] Cui X, Goel V, Kingsbury B. Data Augmentation for Deep Neural Network Acoustic Modeling [C]/ IEEE, 2015:1469 - 1477.