ISSN 2706-655X Vol.4, Issue 10: 34-40, DOI: 10.25236/IJFET.2022.041006

DenseNet network-based surface defect detection algorithm for strip steel

Penghui Zhu, Junjie Dai, Haoyuan Chang, Yao Xu, Zhimao Zhang

Nanjing Engineering College, Nanjing, 211167, China

Abstract: *To address the shortcomings of current surface defect detection algorithm with many parameters, slow detection rate and low accuracy, a defect detection algorithm based on DenseNet network is proposed to mitigate the effects of gradient disappearance and gradient explosion with its more aggressive dense connection mechanism, which also reduces the number of parameters to some extent. Meanwhile, the enhancement effect of SENet network on the effective features is utilized to optimize the network model and enhance the accuracy. Using the strip steel surface defect dataset released by Northeastern University as the base defect sample, the enhancement operation is performed on it. The accuracy of the present algorithm tested on this dataset is as high as 99.44%, which is higher than that of the unimproved DenseNet network, and in terms of computational speed, the present algorithm is only 1.9ms/frame, which obtains a significant improvement compared to the DenseNet network.*

Keywords: *Defect Detection; DenseNet Network; SENet Network*

1. Introduction

In the production process, the material will inevitably be affected by a variety of influences, such as production processes, the surrounding environment, machine defects, etc., the surface of the material will often appear scratches, abrasion and other defects, these defects will not only affect the product aesthetics, but also affect the performance of the product, resulting in economic losses and safety hazards. Traditional visual inspection methods suffer from low accuracy, low efficiency, and too much reliance on human labor, which can make employees' health damaged by long hours of work and cannot meet the needs of industry. Traditional visual inspection methods suffer from low accuracy, low efficiency, and too much reliance on human labor, which can make employees' health damaged by long hours of work and cannot meet the needs of industry.

Pulsed eddy current inspection technique [1], which is not widely available due to its large equipment, although it has a fast detection rate, high accuracy, and can be used in high temperature states, narrow areas of the workpiece, and deep hole walls (including pipe walls). Metal magnetic memory detection does not require a special magnetization device, using its own magnetization phenomenon, can quickly and accurately determine the stress concentration area, and high sensitivity, but it is a weak magnetic signal detection method, the signal is vulnerable to material factors interference. Common machine learning methods use manual extraction of low-level features such as HOG [2] and SIFT [3], and then classify the features with classifiers such as AdaBoost [4], which have defects such as time-consuming feature extraction, low generalization ability, low detection accuracy, and slow detection speed.

With the development of computer technology, the emergence of deep learning has provided a new direction for material surface inspection. Jianfu Xing [5] collected 10 categories of common surface defects of strip steel, and the categories with small number of samples were expanded by means of data enhancement so as to construct a sample library of steel plate defect images. Experiments were conducted using an AlexNet convolutional neural network based model, which improved the classification accuracy. However, this algorithm suffers from saturation phenomenon when the input value is very large or very small,i.e., the gradient of these neurons is close to 0 and there is a gradient disappearance problem; Zhou [6] et al. proposed a CNN-based defect detection algorithm to solve the problem of low feature generalization, however, the image input channel of this algorithm is a single channel, the size of each layer feature map is too small, the feature loss is serious, and the detection accuracy is low.

In order to solve the above existing problems, the dense linking algorithm of DenseNet network is used to solve the problem of AlexNet gradient disappearance and gradient explosion, and the special

ISSN 2706-655X Vol.4, Issue 10: 34-40, DOI: 10.25236/IJFET.2022.041006

learning ability of SENet for features is also used so that DenseNet network can enhance the favorable characteristics to improve the detection accuracy and speed, and a convolutional neural network algorithm for defect detection of strip steel of different materials is proposed.

2. DenseNet Network

2.1 Mechanism of DenseNet network

In the field of computer vision, convolutional neural network (CNN) has been called the mainstream, with the development of technology, when training deep convolutional neural network, the problem of gradient disappearance and gradient explosion often occurs, causing the network difficult to train, then a milestone event in the history of CNN appeared: the emergence of ResNet model. ResNet can train a deeper CNN model, so as to achieve After this, many models such as the recent GoogLeNet, VGG-19, and Incepetion models were developed to address this problem. However, each layer of these networks only learns the features of the previous layer or two layers of that layer, which has low utilization of shallow feature maps and low network learning performance [7]. The DenseNet network, which won the best paper award at CVPR 2017, can effectively solve this problem.

The basic idea is the same as that of ResNet, except that the ResNet model is based on establishing a "short-circuit connection" between the front and back layers, as shown in Figure 1, while the DenseNet model is based on establishing a dense connection between all the front layers and the back layers, as shown in Figure 2.

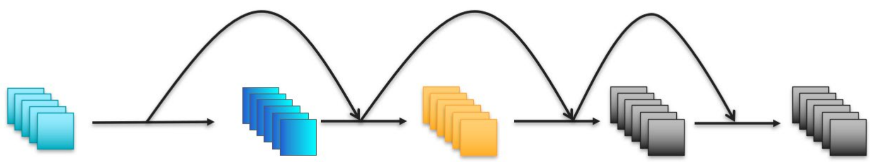

Figure 1: Short-circuit connection mechanism of ResNet network

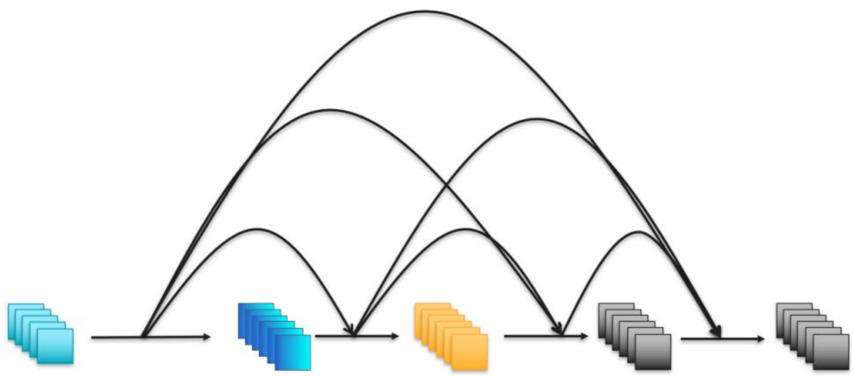

Figure 2: Dense connectivity mechanism of DenseNet network

As you can see from the above figure, ResNet is that each layer is short-circuited to some previous layer (usually 2 to 3 layers) and connected by element-level summation, while in DenseNet, each layer is connected to all previous layers in the channel dimension.

In terms of the equation, the output of the general network at the l layer is

$$X_l = H_l(X_l - 1) \tag{1}$$

And for the ResNet model, the identity function from the upper layer of inputs is added to this.

$$X_l = H_l(X_l - 1) + X_{l-1} \tag{2}$$

ISSN 2706-655X Vol.4, Issue 10: 34-40, DOI: 10.25236/IJFET.2022.041006

In DenseNet, all previous layers are connected as input:

$$X_l = H_l([X_0, X_1, \ldots, X_{l-1}]) \tag{3}$$

H_l is the nonlinear transformation function, which is a combined operation that may contain multiple BN, ReLU, Pooling and Conv operations.

2.2 Structure of DenseNet network

Densenet network is mainly composed of two parts: dense block and transition layer.

The dense block is an intuitive representation of the linking mechanism of the Densenet network, and its main role in the network structure is to establish the relationship between input and output, reduce the dissipation of gradients, and enhance the transmission of features, and its internal structure is shown in Figure 3.

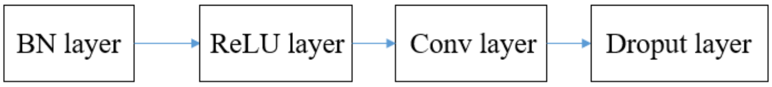

Figure 3: Internal structure of dense block module

The transition layer is usually in between the dense block, which plays the role of reducing the number of channels of the feature map, and contains several BNs, ReLU, convolution, pooling, etc. The structure is shown in Figure 4.

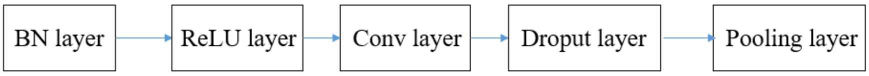

Figure 4: Internal structure of transition layer

2.3 SENet Network

The SENet network is a network model proposed by Momenta and Jie Hu et al. of Oxford University in 2017, which aims to improve the representation of the network by convolving the interdependencies between feature channels, which can reduce errors to a great extent. Their network structure is shown in Figure 5.

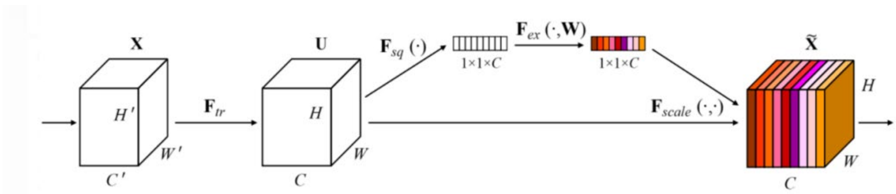

Figure 5: Dense connectivity mechanism of the SENet network

The SENet network is a model based on the idea of ResNet network with the input set of features as:

$$X = [x_1, x_2, \ldots, x_C], X \in R^{H \times W \times C} \tag{4}$$

The network structure consists of two parts: squeeze and stimulus

Squeeze: the dimension of the original feature map is H × W × C, among them H is highly, W is the width, C is the number of channels. The extrusion operation is to put H × W × C compressed to 1 × 1 × C real number sequences.

$$Z = [z_1, z_2, \ldots, z_C], Z \in R^{1 \times 1 \times C} \tag{5}$$

$$Z_C = F_{sq}(X_C) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} X_C(i,j) \tag{6}$$

ISSN 2706-655X Vol.4, Issue 10: 34-40, DOI: 10.25236/IJFET.2022.041006

Among them, F for feature compression processing, which is generally implemented using global average pooling, and after compression, this one-dimensional parameter gains the big picture view of $H \times W$, It allows for a wider range of sensory horizons.

Excitation: after obtaining the squeezed 1D covariates, the excitation operation is performed on them to obtain the interdependence between the channels.

$$s = F_{ex}(Z,W) = \sigma(g(Z,W)) = \sigma(W_2\delta(W_1 Z)) \tag{7}$$

In the above equation, F_{ex} represents motivating operations to learn to use global information selectively, emphasizing useful features and suppressing useless ones. δ is the RELU activation function, σ is the Sigmoid activation function, here the RELU activation function embedded in the Sigmoid activation function is used to train the model, which helps to limit the complexity of the model; two FC fully connected layers are introduced to constitute the bottleneck parameterized gating mechanism. W1 for dimensionality reduction, W2 for ascending.

Finally, the original features are recalibrated to obtain the final output feature map $\widetilde{X}$

$$\widetilde{X} = F_{Scale}(x_C, s_C) = x_C \times s_C \tag{8}$$

Among them, $F_{Scale}(x_c, s_c)$ indicates the mapping characteristic.

2.4 Densenet network model based on SENet network improvement

The core of the SENet network is to add structure between the input and output of Layers, and due to this flexibility, this network structure can be directly applied to other convolutional networks. In the dense block of the Densenet network, each layer has the problem of feature redundancy in the connection with subsequent layers, while the SENet network can effectively enhance the favorable features and suppress the useless features, and together with its flexibility, the SENet network can be embedded into the Densenet network to complete the optimization of the network structure, whose structure is shown in Figure 6.

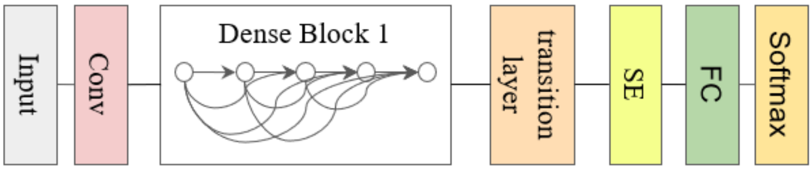

Figure 6: Improved Densenet network model

3. Experiment

3.1 Construction of experimental environment

In order to get accurate experimental results, the training, testing and other characters are done on a single computer. The processor is Intel(R) Core(TM) i7-10750H CPU @ 2.60GHz 2.59 GHz, the GPU is NVIDIA GeFore GTX 1650, the Windows version is Windows 10 Home Edition, the internal version of the operating system is 19044.2130, and Python 3.9 is used For data processing, the deep learning framework is Pytorch,

3.2 Network Parameters

To improve the computational speed and reduce the influence of useless information, global maximum pooling is used here instead of global average pooling, with Conv2d denoting the two-dimensional convolutional layer and BatchNorm2d denoting the BN layer. The specific parameters of each layer of the network are shown in Table 1.

ISSN 2706-655X Vol.4, Issue 10: 34-40, DOI: 10.25236/IJFET.2022.041006

Table 1: Parameters of the improved Densenet network

Layer（type)	**Output Shape**	**Param**
Conv2d-1	[-1, 64, 112, 112]	9,408
BatchNorm2d-2	[-1, 64, 112, 112]	128
ReLU-3	[-1, 64, 112, 112]	0
MaxPool2d-4	[-1, 64, 56, 56]	0
BatchNorm2d-5	[-1, 64, 56, 56]	128
…	…	…
Conv2d-364	[-1, 32, 7, 7]	36,864
BatchNorm2d-365	[-1, 1024, 7, 7]	2,048
Linear-366	[-1, 6]	6,150

3.3 Experimental procedure

The experimental steps are shown in Figure 7.

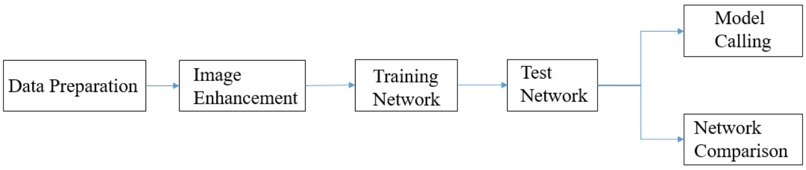

Figure 7: Experimental flow

1) Data preparation

A dataset of strip surface defects collected from the surface defect database [8] published by Northeastern University (NEU) was used as samples, which contains six typical defects, namely rolled oxide (RS), plaque (Pa), cracking (Cr), pitting surface (PS), inclusions (In) and scratches (Sc), each containing 300 samples. Figure 8 shows a partial plot from the base sample.

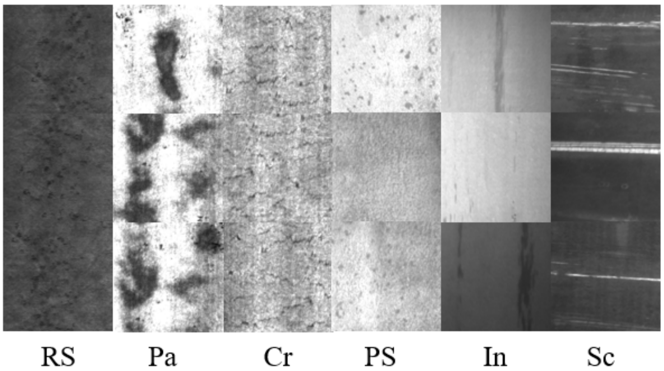

Figure 8: Defective sample data graph

The resolution of each graph is 200×200. 240 samples from each defective sample set are randomly selected and incorporated into the training set train, and the remaining 60 samples are incorporated into the test set test.

2) Data Enhancement

The 1440 sample images in the training set are rotated by 90°, 180°, and 270° to achieve image enhancement, and the pixel values of each channel are normalized before input to the network.

3) Training Network

Using Adam optimizer iterative optimization, the loss function expression is as follows:

$$Loss = -\left[y\log\hat{y} + (1-y)\log\left(1-\hat{y}\right)\right] \tag{9}$$

Among them, y is the expected probability, $\hat{y}$ is the predicted probability, Loss is the loss value.

ISSN 2706-655X Vol.4, Issue 10: 34-40, DOI: 10.25236/IJFET.2022.041006

The network is trained 90 times to obtain images of accuracy and loss values, as shown in Figure 9.

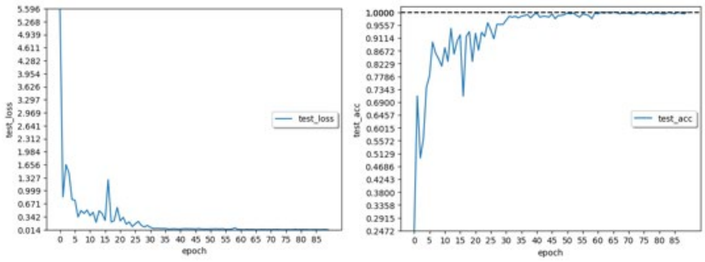

Figure 9: Loss value image and accuracy image

4) Model invocation

A sample is randomly selected from the dataset and the optimal model is invoked to detect it. Figure 10 shows the selected sample images.

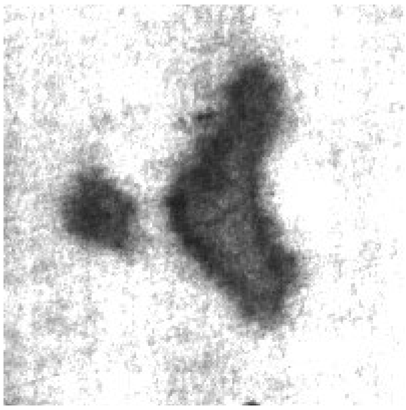

Figure 10: Randomly selected samples

The accuracy after network detection is at 99.44%.

5) Network comparison

Comparing the improved network with the Densenet 121 network. The results are shown in Table 2.

Table 2: Comparison of different network detection results

	Improved network	**Densenet121 Network**
RS	98%	95%
Pa	100%	97%
Cr	100%	96%
PS	99%	95%
In	100%	96%
Sc	100%	97%
Accuracy	99%	96%
Speed (ms/frame)	1.9	3.1

4. Conclusion

For the current surface defect detection algorithm with many parameters, slow detection rate and low accuracy, a defect detection algorithm based on DenseNet network is proposed to mitigate the effects of gradient disappearance and gradient explosion with its more aggressive dense connection mechanism, which also reduces the number of parameters to a certain extent. Meanwhile, the enhancement effect of SENet network on the effective features is utilized to optimize the network model and enhance the accuracy. Using the strip steel surface defect dataset released by Northeastern University as the base defect sample, the enhancement operation is performed on it. The accuracy of the present algorithm tested on this dataset is as high as 99.44%, which is higher than that of the unimproved DenseNet network, and in terms of computational speed, the present algorithm is only 1.9ms/frame, which obtains a significant improvement compared to the DenseNet network. Fully meets the needs of surface defect detection on

ISSN 2706-655X Vol.4, Issue 10: 34-40, DOI: 10.25236/IJFET.2022.041006

production lines.

Acknowledgement

This paper was supported by funds from the Science and Technology Innovation Fund for Students of Nanjing Engineering College (Grant No. TB202201006) and the Practical and Innovative Training Program for Students of Jiangsu University (Grant No. 202211276074Y).

References

[1] Yuan L, Zhang Z, Tao X. The Development and Prospect of Surface Defect Detection Based on Vision Measurement Method[C]// 2016 12th World Congress on Intelligent Control and Automation (WCICA). IEEE, 2016

[2] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C] / /Proc of IEEE Conference on Computer Vision and Pattern Recognition.2005: 886-893.

[3] LOWE D G. Distinctive image features from scale-invariant key points [J]. International Journal of Computer Vision,2004,60 (2):91-110.

[4] Ferreira A J, Figueiredo M A T. Boosting algorithms: a review of methods, theory, and applications [M]/ /Ensemble Machine Learning. New York: Springer, 2012: 35-85.

[5] Xing Jianfu. Convolutional neural network-based surface defect identification and system development for hot-rolled strip steel [D]. Shenyang: Northeastern University, 2017.

[6] Zhou S, Chen Y, Zhang D. Classifification of surface defects on steel sheet using convolutional neural networks. Materiali in Tehnologije 2017; 51(1):123–31.

[7] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions [J]. arXiv:1511.07122, 2015.

[8] K. Song, Y. Yan, "A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects,"Applied Surface Science, vol. 285, 2013: 858-864