Defect Detection of Using Variant CNN in the Processing of Cover Glass, Touch Screen and Display under Parallel Light

Xiangying Wei[1]*, Wei Feng[23]*, Qujiang Lei[1]†, Guangchao Gui[1], Shangzhi Le[1], Jiahao Zhong[1], and Weijun Wang[1]

[1]Guangzhou Institute of Advanced Technology, Chinese Academy of Sciences, Guangzhou, China
[2]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China
[3]University of Chinese Academy of Sciences
*Co-first authors
†Corresponding author: qj.lei@aaia-ai.org

Abstract—**Due to the development of mobile phone technology, the precision requirements for various parts of the mobile phone processing industry are getting higher and higher. Under the screen fingerprint, curved screen and other functions, so that the phone cover glass, touch screen, display processing needs micron-level precision. Although the production process to minimize dust, powder and other interference, but there will still be some product defects. This paper discusses how to do defect detection better in the processing of mobile phone screen, and proposes a defect detection method based on variant CNN to improve the detection accuracy.**

Keywords-convolutional neural network; defect detection; high precision; parallel light

I. INTRODUCTION

From the invention of the first wireless mobile communication device in 1902 to the birth of the first mobile phone in 1983. From the iPhone, the first touch-screen smartphone with apps in 2007, to today's curved screen phones. Modern mobile phones are constantly changing people's lives. Mobile phones, as a key to intelligent life, are constantly gaining more functions and need smaller volume. And in order to accommodate more parts and more functions, the accuracy of each part of the phone is also increasing. In the challenges of screen fingerprint and curved screen, the precision requirement of smart phone for cover glass, touch screen and display become higher.

With the arrival of Industry4.0 (which is the era of promoting industrial revolution by using information technology, namely the era of intelligence), industrial production tries to avoid the direct involvement of human in production. A big step has been taken towards automating production and liberate the workforce. Nowadays, the precision of automated assembly lines can reach the standard of industrial production. But the testing process in production still needs to be done manually. [1]

In the actual production of the factory, in order to ensure the accuracy of the final product, each production step has its specific specifications. To ensure that each part of the product meets the production specifications, defect detection is necessary after each production step. If unqualified products are not detected, they will not only lead to product problems, but also endanger public safety (such as battery explosion caused by short circuit of mobile phones). The common detection method is manual detection, which is time-consuming and inefficient. Especially during a disease outbreak, workers are not able to come to work, which will make detection procedures impossible. In order to avoid the impact of the epidemic on industry, it is necessary to use computer vision for defect detection. [2]

Computer vision is a way for computers to observe and analyze picture information instead of human eyes and brain. In fact, a computer can't see a picture as a whole as a human eye can. To a computer, a picture is a data set of pixels. So, in order for a computer to mimic a human understanding of a picture, it needs to rely on deep learning to extract key information from the picture. There are many variables in this step: different algorithms will lead the computer to recognize different image features; Different data sets will affect the effect of feature extraction and whether the feature is extracted correctly: different learning times and training time will also change the computer's recognition of different image features. [3]

Due to the low stability, limited accuracy and high cost of the current computer vision detection method, manual detection, which requires a lot of manpower with poor efficiency, has not been replaced. At present, the detection methods for product defects caused by dust and other objective factors in industry mostly rely on manual detection.

In terms of circuit production, both the linewidth and the line spacing of the electrodes on the mobile phone screen are only about 8 microns, therefore a rough linewidth control is easy to cause circuit break. The line spacing control is the most difficult part in the whole production. A short circuit can be caused if there is a dust in the electronic circuit with the diameter greater than the line spacing, hence all dusts with similar or larger sizes must be cleaned. Based on the naked eyes, defects above 50\ \mu m can be easily distinguished, therefore a method to detect defects below 50\ \mu m is needed.

In the manufacturing process of mobile phone cover glass, ink printing and film should be able to detect dirt, scratches, dust, printing line and other defects that affect the appearance. Due to the diversity of dust forms, it is difficult to have a stable visual detection method based on simple dichotomy in the case of too many variables.

For different parts, dust causes problems:

Dust detection of cover glass mainly affects the appearance. Although it does not directly affect the appearance, the dust that affects the appearance and look may also lead to a decrease in competitiveness in today's competitive environment.

Touch screen and display have high performance requirements. In addition to the same test requirements as cover glass, the more important thing is to test their circuits and performance.

Smudges generally comes from three sources:

Smudges caused by raw materials (Figure 1):

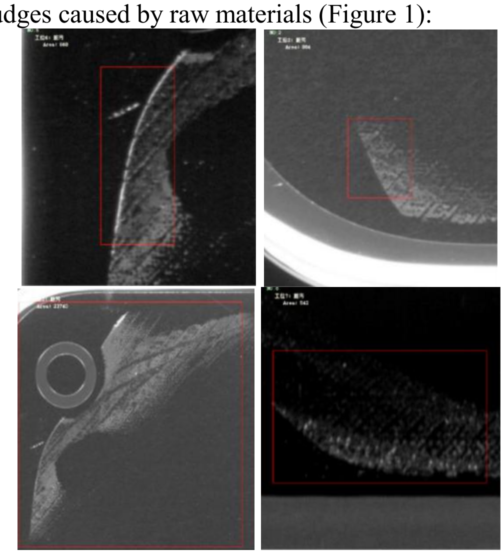

Figure 1. Smudges caused by raw materials.

Smudges caused by cleaning machine (Figure 2):

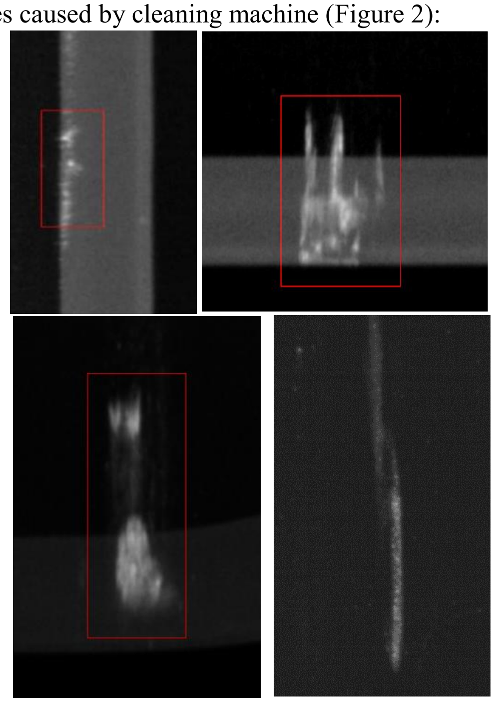

Figure 2. Smudges caused by cleaning machine.

Unavoidable dust in the production process (Figure 3):

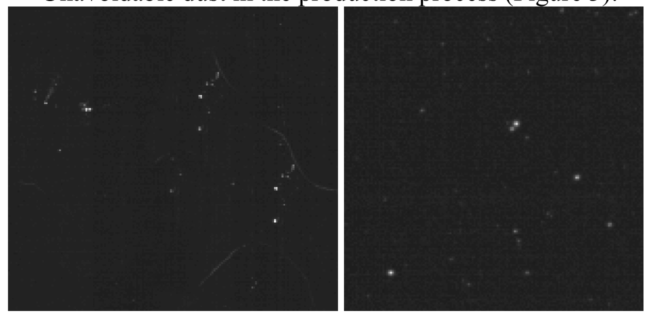

Figure 3. Smudges caused by dust.

In this paper, we will present a high-precision Defect Detection method. By using parallel light to acquire images. The image set is classified by resnet50+ structure, and then judged by VGG19+.In order to better adapt to the needs, the two neural network structures have been modified.

The remainder of this paper is organized as follows. Section 2 describes the related work and the solutions for defect detection. Section 3 presents the defect detection methods used in this paper and the models used. Section 4 is the experimental section. Section 5 gives the conclusions, and Section 6 is the direction for future research.

II. Related Work

A. *Combining Computer Vision and Manual Detection*

This method is more commonly used in the printed board industry. Since human eyes cannot quickly and accurately find minor defects, fixed cameras are used to take pictures of the products and compare them with static standard pictures to mark the areas of difference. Areas of difference will be judged whether it has any effect and then sent to be repaired if necessary

Advantages: Low device costs for pictures shooting and manual judgment, easy access of static standard pictures.

Disadvantages: High labor costs (trained workers), uncertainty of manual detection. The possible influencing manual factors including: loss caused by workers' poor judgment, inefficiency due to workers' illness, inability of workers to participate in work because of epidemic diseases.

B. *Fixed Program Detection Method*

Some PCB (Printed circuit board) companies possess AOI (Automated Optical Inspection) equipment, and their detection methods are by optical methods to automatically check the eligibility of the circuit, such as whether there is open circuit or short circuit and the existence of sawtooth. Single category products are detected using fixed programs and hardware. [4]

Advantages: The method is simple and direct, and the accuracy rate is stable. The cost of single category products detection is low, and it is suitable for testing products that need long-term mass production.

Disadvantages: There are different test standards for each kind of PCB. Once it is updated, it is necessary to redesign the test standard or even purchase new hardware, which may lead to incredibly high cost. [5]

C. *Use the Deep Learning Detection Method*

Most of the detection methods combining deep learning and computer vision in the industry rely on CNN and its variants, and all data are binary classified (i.e., qualified and unqualified). [6]

Advantages: These detections are effective in the case of single type defects

Disadvantages: In a complex production environment, different defects have different sources. In the case of dichotomy, it is difficult to determine whether substandard products should be repaired or discarded. Misclassification is also an easy situation.

• CNN (Convolutional neural network):

Convolutional neural network is a classical deep neural network. Its basic structure consists of three types of layers: convolution layers, pooling layers, full connection layers. [7]

The convolutional layers are aimed at local perception. In the process of image recognition, the human brain does not recognize the whole image in one trial. Imitate this characteristic, firstly the convolution layers perceive each feature of the image partially, and then a comprehensive local operation is carried out at a higher level to obtain global information. Next the activation function is applied to process the output of the convolutional layers, and a nonlinear mapping is constructed to the output of the convolutional layer. [8]

Pooling layers are intended for down sampling. It is mainly aimed at feature dimensionality reduction, data and parameter compression, over-fitting reduction and fault tolerance improvement. [9]

Full connection layer is a simple multi-classification neural network, through the SoftMax function to get the final output. [10]

Convolutional neural network is the most common neural network for processing image data in deep learning. It can effectively extract the key information in the picture.

There is a large number of variants of CNN which adapt to different situations. AlexNet, proposed by Alex Krizhevsky, wined the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012. After that, a large number of variants of CNN neural network appeared: ZFNet (2013), GoogLeNet (2014), VGGNet (2014), ResNet (2016), DenseNet (2016), etc. [11-15]

• Binary classification:

Binary classification problem is simply a "if" or "or" issue, like to determine whether there is an object in a picture.

• The difference between a computer reading an image and a human:

In this case, human visual system, known as eyes, can easily tell if there is an object in the picture, but not for machine visual system. Since machine can only read the digital features, such as the size of the image, number of channels, characteristics of pixel, etc., the machine vision system can only see the parts, but not the whole. [16]

III. Our Approach

In this paper, a defect detection approach based on CNN for multiple classification for parallel light environmental is proposed (Figure 4). In this method, the image data obtained under parallel light will be classified twice using different variants of CNN. The first classification will determine the type of detects, and the second classification will determine whether the defects exceed the production standard.

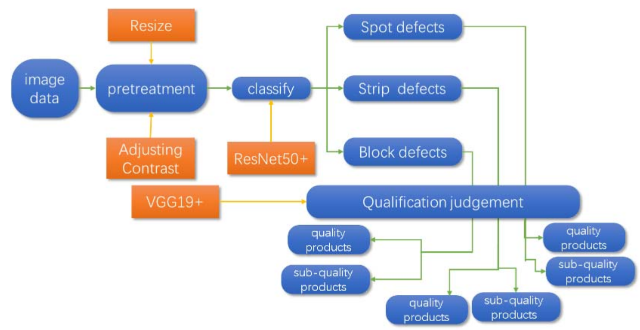

Figure 4. Flow chart of multi-classification defect detection method.

A. *Experimental Data Acquisition Method*

For high precision computer vision detection, light source and shooting angle of cameras count for much. Existing light sources commonly used for detection include point light sources, parallel light sources and ring light sources. The light sources for parallel surface inspection are a strip of parallel light source composed of high-brightness and high-power LED beads (Figure 5). This light source emits a narrow beam of parallel light of 3000K color temperature, with a width of 40mm and illumination up to 35000Lux(@100mm). The light is uniformly distributed and the irradiation area is wide. With camera, scars and particles below 10 microns can be confirmed. [17]

The principle of parallel surface inspection lamp detection is that, from the LED lamp sources the special light is refracting through optical lens and filters so it can be reflected on the surface of the workpiece, hence defects such as dust, scratch, burr, bump and ink on the surface of the workpiece can be identified.

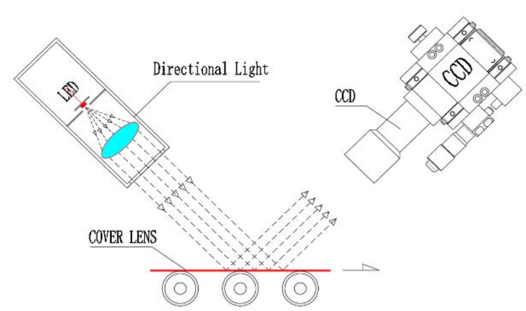

Figure 5. Parallel light acquisition image schematic

B. *Classification Basis*

Classification of detects:

Based on the shapes, detects can be classified into spot detects, strip detects and block detects.

1. **Spot detects** are usually caused by dust and they are easy to repair.

2. **Strip detects** are necessary to be judged whether met the validation criteria of the length. Misjudgment is easy to occur after image division. The strip detects are shown in the following figures as examples (Figure 6):

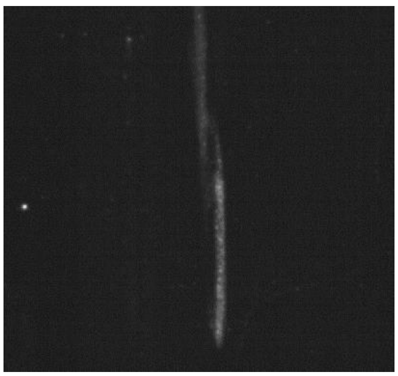

Figure 6. Strip detects (original image).

In detection, the image may be divided into the following four images (Figure 7):

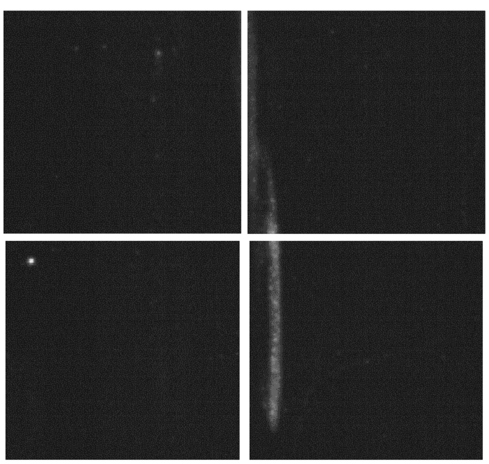

Figure 7. Strip detects after divided.

After division, the detects in any one of the images would not be detected as disqualification. However, the strip detect exceeds the allowable criteria in reality. Such situations will challenge the accuracy of detection. [18]

3. **Block detects** generally have a large impact on the completion of the products and need to be cleaned again. A variety of repair methods are required according to different causes of detects.

Classification of qualification:

The classification is to determine whether the detects meet the actual requirements of industrial production according to the types of products and their production standards

C. *Neural Network Structure*

ResNet-50:

One reason why deep network is difficult to train is the problem of gradient disappearance. On account of the gradient propagates back to the preceding layer, repeated multiplication may make the gradient infinitesimal. As a result, as the number of layers of the network increases, its performance tends to saturation or even rapidly degrades. To solve this problem, He Kaiming proposed the residual neural network in 2015 and obtained the first prize of ILSVRC-2015 classification task, COCO Detection and COCO Segmentation. [19]

Residual neural network, also known as ResNet, based on the idea of adding the residual learning to the traditional convolutional neural network. It can solve the problem of gradient dispersion and precision degradation in deep network, making the network deeper and deeper so that it cannot only guarantee the accuracy but also control the speed. [20]

There are two basic blocks in ResNet. One of them is identifying block, and the dimensions of the input layer and output layer are the same. Another one is the convolution block, where the dimensions of the input layer and output layer are not the same. It is intended to change the dimensions of the eigenvectors. [21]

The network structure used is shown in the table below (Table 1).

TABLE I. VARIANT RESNET50 NETWORK STRUCTURE

layer name	output size	resnet-50
conv1	112x112	7x7,64, stride2
conv2_x	56x56	3x3 max pool, stride 2 $\begin{bmatrix} 1x1 & 64 \\ 3x3 & 64 \\ 1x1 & 256 \end{bmatrix} x3$
conv3_x	28x28	$\begin{bmatrix} 1x1 & 128 \\ 3x3 & 128 \\ 1x1 & 512 \end{bmatrix} x4$
conv4_x	14x14	$\begin{bmatrix} 1x1 & 256 \\ 3x3 & 256 \\ 1x1 & 1024 \end{bmatrix} x6$
conv5_x	7x7	$\begin{bmatrix} 1x1 & 512 \\ 3x3 & 512 \\ 1x1 & 2048 \end{bmatrix} x3$
average	1x1	average pool,1000-d fc SoftMax
dense	1024	1024, relu
dense_1	512	512, relu
dense_2	3	3, SoftMax

After Resnet was proposed, many variants are put forward. Among them Resnet-152 had the best effect, and Resnet-101 and Resnet-50 were the most widely used in daily work. Resnet-50 was applied in this experiment.

Residual learning (Figure 8):

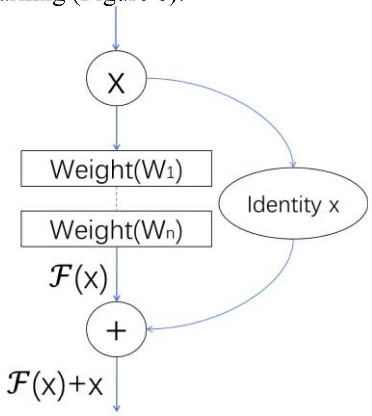

Figure 8. Residual structure.

$$\mathcal{F} = W_2\sigma(W_i x)$$
$$y = \mathcal{F}(x, \{W_i\}) + x$$

The block of residual learning consists of two branches or in other words, two kinds of mapping:

1. Identity mapping, as the name implies, refers to the mapping of oneself, namely oneself.

2. Residual mapping refers to the $\mathcal{F}(x)$ part, which is called residual mapping. It is the convolution calculation part of the ResNet.

VGG-19:

VGG was proposed by Oxford's Visual Geometry Group on ILSVRC 2014. This network work on ILSVRC 2014 was to demonstrates that increasing the depth of the network can partly affect the ultimate performance of itself. There are two variants of VGG, namely VGG16 and VGG19. There is no essential difference between them except the network depth.

In VGG, three 3×3 convolution kernels are used to replace 7×7 ones, and two 3×3 convolution kernels are used to replace 5×5 ones. The main purpose is to enhance the depth of the network to improve the performance of the neural network to a certain extent under the condition of ensuring the same perception field.

The network structure used is shown in the table below (Table 2).

TABLE II. VARIANT VGG19 NETWORK STRUCTURE

VGG19+dens3			
conv1-1	1	conv4-3	11
relu1-1		relu4-3	
conv1-2	2	conv4-4	12
relu1-2		relu4-4	
pool1		pool4	
conv2-1	3	conv5-1	13
relu2-1		relu5-1	
conv2-2	4	conv5-2	14
relu2-2		relu5-2	
pool2		conv5-3	15
conv3-1	5	relu5-3	
relu3-1		conv5-4	16
conv3-2	6	relu5-4	
relu3-2		pool5	
conv3-3	7	fc6(4096)	17
relu3-3		relu6	
conv3-4	8	fc7(4096)	18
relu3-4		relu7	
pool3		fc8(1000)	19
conv4-1	9	prob (SoftMax)	
relu4-1		dense (1024, relu)	20
conv4-2	10	dense_1(512, relu)	21
relu4-2		dense_2(3,SoftMax)	22

Abbreviation explanation

Conv: the convolutional layer

FC: the full connection layer

Conv3: the convolutional layer USES 3x3 filters

Conv3-64: denotes a depth of 64

Maxpool: maximum pooling

VGG16: a VGG variant with 16 hidden layers (13 convolutional layers and 3 full connection layers), as shown in column D above

VGG19: a VGG variant with 19 hidden layers (16 convolutional layers and 3 full connection layers), as shown in column E above.

Advantages and disadvantages of VGG

Advantages:

The structure of VGGNet is simple. The sizes of convolution kernels (3×3) and pooling kernels$(2x2)$ are respectively identical in the whole network.

The combination of several small filters (3×3) convolutional layers are better than one large filter (5×5 or 7×7) convolutional layers. It is verified that the performance can be improved by deepening the network structure continuously.

Disadvantage:

The three full connection layers of VGG consume a lot of computing resources and a large number of parameters (this is not a problem caused by the convolution kernel of 3x3) are involved, resulting in more memory usage (140M).

IV. EXPERIMENT

Experimental data sets:

By inviting the factory to assist in obtaining the data set.

Classification: 840 images for each category, a total of 2520 images.
Judgment: 840 images for each category. It was divided into 420 images for good and 420 images for bad.

Training parameters: BATCH_SIZE=3; Epoch=100

Use ResNet-50 for classification:

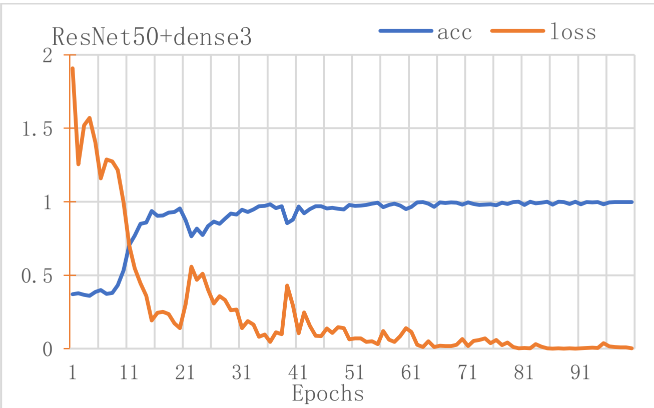

Figure 9. Training result of classification.

In the model training of ResNet-50 combined with an external 3-layer structure, the loss function of the network presents a downward trend in general. It fluctuates significantly at the 25th and 40th epochs.

The accuracy presents an overall upward trend. There is an obvious fluctuation in the figure where the loss function curve rises. Finally, the value of the loss function is close to 0, and the accuracy is between 98.3723% and 99.8245% (Figure 9). This is a relatively good result (Figure 10).

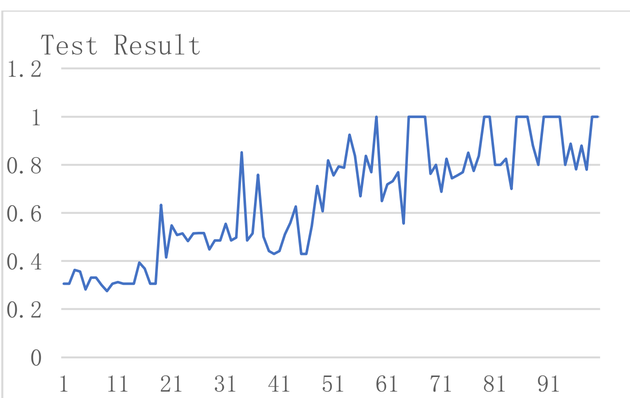

Figure 10. Result of direct judgment.

The overfitting phenomenon appears in the experimental results of the test set. The main reason for this was the computer GPU used in the experiment: the NVIDIA Quadro P3200. The lack of memory in the GPU leads to the fact that the image has to be resized to a smaller image during training, which loses part of the information. And because some of the defects are small, some of the key information is missing.

Use VGG19 for judgement.

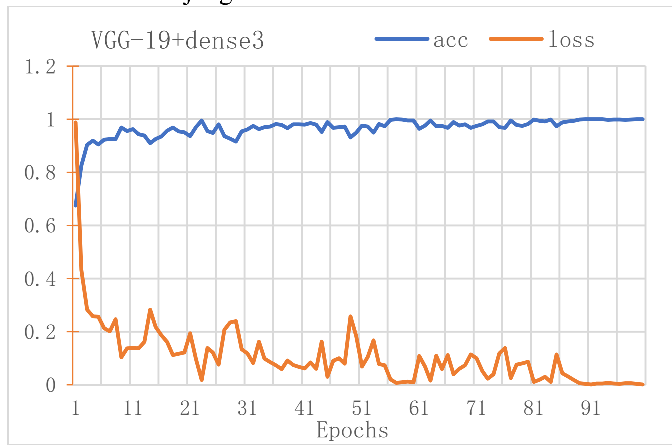

Figure 11. Training result of judgement.

The VGG19 is also combined with an external 3-layer structure, In the model training, although the value of the loss function showed a downward trend, it fluctuates violently. The initial accuracy of the network was relatively high, so it soon approaches 1. In the final 10 Epochs, the accuracy fluctuates slightly between the value of 99.7014% and 97.3134% (Figure 11).

By contrast, we tried to make a direct judgment of the entire data set (the traditional way of classifying all data into the good and the bad). Although the accuracy of the training set is reasonable by using the same ResNet-50 structure, the result of the test set is poor. Obviously, there are a lot of huge fluctuations (Figure 12).

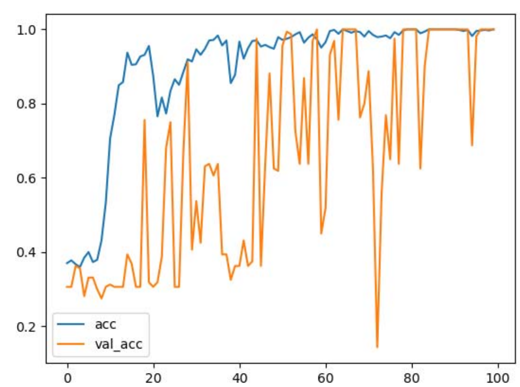

Figure 12. Result of direct judgment.

V. Conclusion

This paper presents a defect detection method using multiple variants of the CNN in the production process of cover glass, touch screens and displays of mobile phones under parallel light sources. In the detection process, firstly classify the defect types, then determine whether the parts conform to the specification requirements. This method also enhances the detection of fine dust and other details. Compared with the traditional manual detection, our method is faster and more stable. Compared with other deep learning based approaches, it performs a higher accuracy in determine whether it meet the criterion.

VI. FUTURE WORK

The detection method proposed in this paper requires both massive classified data of equal amounts and judged data. Therefore, the preparation of training data is more complicated than other deep learning methods, which suggests much preparation time in industrial production. In general, data enhance is not used to expand the data set in industrial field. The initial classification work of data set of huge amounts of data still requires manual operation, and misclassification is inevitable.

A possible solution: after a short period of production, the manual detection results are trained as a training set in the trained model. A misjudgment in the production process would be obvious. This can ensure that the production does not need to suspend for a long time while as the rapid use of stable defect detection.

Because the size of dust detected is very small, the image size collected is large. The sizes of most images are 12.2 MB (12,803,580 bytes; The resolution is 4096x6400 pixels; Bit depth is 32), so good equipment and relatively long time is required for training. Due to the fact that the sizes of the images need to be unified in the training, some details lose.

Possible solutions: better hardware, larger scale image sizes.

ACKNOWLEDGMENT

This research is supported by the Guangdong Province Key Field R&D Program Project (Grant No. 2020B090925002)，the Major Projects of Guangzhou City of China (Grant No. 201907010012).

Data acquisition equipment and mobile phone parts of this experiment were provided by Lingdi Machinery Equipment Co., LTD. Shantou City, Guangdong Province.

REFERENCES

[1] V., M. M. a. A. K. Roblek, "A Complex View of Industry 4.0.," *Sage Open 6,* 2016.

[2] Imperial College London, "COVID-19 reports|Imperial College London," Imperial College London, 2020. [Online]. Available: http://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/covid-19-reports/. [Accessed 2020].

[3] M. David, "The computation of lightness by the primate retina," *Vision Research,* pp. 1377-1388, 14 12 1974.

[4] T. J. Y. C. C. L. Techen Chung, "ARRAY SUBSTRATE AND DEFECT-DETECTING METHOD THEREOF," 2008.

[5] H. Ng, "Automatic thresholding for defect detection," *Pattern Recognition Letters,* pp. 1644-1649, 2006.

[6] S. I. V. V. A. A. A. Christian Szegedy, "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," *national conference on artificial intelligence,* pp. 4278-4284, 2016.

[7] L. G. C. M. S. C. T. A. B. A. Lawrence S, "Face Recognition: A Convolutional Neural-Network Approach," *IEEE Transactions on Neural Network, vol. 8, no. 1,* pp. 98-113, 1997.

[8] J. Bouvrie, "Notes on Convolutional Neural Networks," 2006.

[9] a. B. W. R. Francis E. Raymond, "Fractional Max-Pooling," *Theoretical Economics Letters ,* pp. 225-237, 2 5 2015.

[10] P. G. ,. K. I. Liangchieh C and e. al., "Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs[J].," *Computer ence,* pp. 357-361, 2014.

[11] M. D. ,. a. R. F. Zeiler, "Visualizing and Understanding Convolutional Networks," 2013.

[12] W. L. Y. J. P. S. S. R. D. A. D. E. V. V. A. R. Christian Szegedy, "Going deeper with convolutions," *computer vision and pattern recognition,* pp. 1-9, 2014.

[13] A. Z. Karen Simonyan, "VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION," *computer vision and pattern recognition,* 2014.

[14] X. Z. S. R. J. S. Kaiming He, "Deep Residual Learning for Image Recognition," *computer vision and pattern recognition,* pp. 770-778, 2016.

[15] Z. L. L. V. D. M. K. Q. W. Gao Huang, "Densely Connected Convolutional Networks," *computer vision and pattern recognition,* pp. 2261-2269, 2016.

[16] C. L. Chihwei Hsu, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks,* pp. 415-425, 2002.

[17] S. S. W. A. D. Z. H. Z. Jingjing Wu, " Defect Detection and Localization of Nonlinear System Based on Particle Filter with an Adaptive Parametric Model," *Mathematical Problems in Engineering,* pp. 1-12, 2015.

[18] V. K. Fisher Yu, "Multi-Scale Context Aggregation by Dilated Convolutions," *arXiv: Computer Vision and Pattern Recognition,* 2015.

[19] I. S. G. E. H. Alex Krizhevsky, "Alex Krizhevsky, Ilya Sutskever, Geoffrey E Hinton," *neural information processing systems,* pp. 1097-1105, 2012.

[20] C. S. A. V. D. H. Zifeng Wu, "Wider or Deeper: Revisiting the ResNet Model for Visual Recognition," *Pattern Recognition,* pp. 119-133, 2019.

[21] J. T. A. J. L. R. E. S. Furong Huang, "Learning Deep ResNet Blocks Sequentially using Boosting Theory," *arXiv: Learning,* 2017.