IEEE *Access*
Multidisciplinary : Rapid Review : Open Access Journal

Review: Application of Convolutional Neural Network in Defect Detection of 3C Products

Wuyi Ming[1], Chen Cao[1], Guojun Zhang[2], Hongmei Zhang[2*], Fei Zhang[3], Zhiwen Jiang[1], and Jie Yuan[1]

[1] Henan Key Lab of Intelligent Manufacturing of Mechanical Equipment, Zhengzhou University of Light Industry, Zhengzhou, 450002, China

[2] Guangdong HUST Industrial Technology Research Institute, Guangdong Provincial Key Laboratory of Digital Manufacturing Equipment, Dongguan, 523808, China

[3] School of Mechanical Engineering, Dongguan University of Technology Dongguan, 523808, China.

Corresponding author: Hongmei Zhang (e-mail: mercyzhm@163.com).

ABSTRACT Based on the rapid development of semiconductors, integrated circuits and the Internet. 3C products such as computers, tablets, mobile phones and smart TVs have become an indispensable part of people's lives. With the prosperity and development of the 3C product market, the demand for the quality of display panels and related detection technologies are increasing. As the iconic network of deep learning, has been extensively studied in the field of image recognition and defect detection. Based on the development of CNN, this article summarizes the defect detection method of 3C products by CNN with different depths. First, we reviewed the origin of CNN and its structural components, then introduced the upgrade and improvement of important components, and finally introduced and compared the applications of CNN with different depths in defect detection. Through the comparison and summary of the effect of defect detection, we analyze the opportunities and challenges of different CNN frameworks, and exhibit the strategies for different application scenarios.

INDEX TERMS 3C products, Defect detecting, Neural network, Convolution network

I. INTRODUCTION

A. BACKGROUND

Semiconductors, integrated circuits and network have developed rapidly in recent years. 3C products such as computers, mobile phones and smart TVs have become an indispensable part of people's lives. With the expansion of the application field of 3C products and the improvement of the public's requirements for product performance, mobile communication technology has reached a new level of development in theory and application [1]. According to statistics, as shown in Fig. 1, the global sale of smartphones continued to grow from 2012 to 2016, and the trend gradually stabilized at around 150 million from 2016 to 2019. Also, and it is predicted that the sale of it will continue to increase in 2021[2]. The innovation of 5G technology brings new market opportunities for consumer electronics products. Digital terminal technology with multiple functions, high
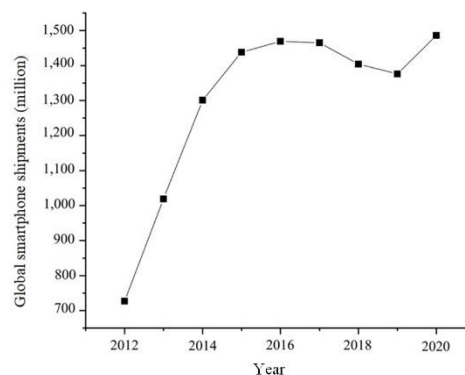
integration and large screen will be the development trend of 3C products in the future.

Figure 1. Smartphone shipments and future development trends from 2012 to 2020

With the continuous increase of 3C products outputs and the improvement of quality requirements, product quality control has become a challenging task. In the surface inspection of 3C products, it is extremely difficult to inspect the display screen and circuit board. The traditional 3C product surface defect detection method is mainly manual, that is, the quality of the product surface is inspected by the quality inspectors through the eyes [3]. However, the method can not meet the requirements of modern industry for high-speed and precision detection due to its strong subjectivity, large uncertainty, and low efficiency. With the development of the Industrial Internet of Things and artificial intelligence, the fourth industrial revolution, Industry 4.0, is developing rapidly [4, 5]. On the basis of advanced digitization of factories, the combination of Internet technologies and future-oriented technology in the field of smart objects has created a new fundamental paradigm shift in production [6]. With the advent of Industry 4.0, the production process has been given an intelligent cyber-physical system that can generate a large amount of streaming sensor data [7]. With the introduction and development of "Industry 4.0", intelligent manufacturing, and "Made in China 2025", the production process has been endowed with an intelligent cyber-physical system that can generate a large amount of streaming sensor data, industrial product defect detection has also received more and more attention [8]. Machine vision has become more and more widely used in production lines with its real-time, high efficiency and accuracy.

But for machine vision, detecting surface defects of reflective displays is a challenge. When the camera is placed vertically on a transparent part, it may consider the reflected light spot to be a defect. The illumination of transparent parts is another important issue. Glare will appear when the light is high. If the light is very low, there will be a situation where the defect cannot be detected. Even if the light source is sufficient, the illuminance on the glass surface usually fluctuates [9]. In order to ensure production quality, quality inspection must be added in many production links of PCB [10]. Common surface defects of mobile phone screens, such as scratches, debris and dirt, should be identified and removed in real time during the production process. Some sizes of the defect are significantly small (about 0.05 mm), which complicate the defect detection [11].

B. RELATED WORKS

With the development of machine vision technology, researchers use image processing to detect screen defects of 3C products. Liu et al. [12] proposed an algorithm combining the line intercept as the threshold and particle swarm optimization, which solved the problem of low detection accuracy of mobile phone screen defects. Gao et al. [13] improved the median filter algorithm and utilized the image processing algorithm to preprocess the defective image. In addition, they proposed an image difference algorithm based on fast image matching for detection. Zhang et al. [14] posed

an improved differential image detection method for mobile phone screen defects to improve detection efficiency and accuracy. However, the performance of pure image processing technology cannot achieve satisfactory detection results [15].

Convolutional Neural Networks (CNN), a research hotspot in the field of image detection and pattern recognition, has received more and more attention from scholars [16]. Ma et al. [17] designed the CNN based on the GoogleNet network, which greatly reduced the number of parameters without affecting the prediction rate. Experimental results showed that the defect detection rate of the designed CNN could reach 99.5%. Wei et al. [18] proposed a multivariable CNN-based defect detection method in the production process of cover glass, touch screen and mobile phone display under parallel light sources. Experiments verified that the method had higher accuracy, better stability and faster speed. The CNN architecture mainly includes weight parameter sharing and pooling operations which is more complicated. However, CNN is easy to train, and the learned features are translation invariant [9].

The main content of this research is to review the surface defect detection methods with different levels of CNN in 3C products such as glass display screens, PCBs and TFT-LCD screens. Then, the detection methods and the problems of various methods are summarized and discussed. First, we outline the types of defects in 3C products and the comparison of different detection methods in the context of Industry 4.0. Subsequently, the improved CNN structures in various scenarios are compared, and the technical limitations and detection performance of different methods in practical applications are shown. This research summarized and compared noteworthy research based on CNN in recent years to overcome the challenges of surface defect detection in 3C products.

This literature review mainly focuses on the following topics:

1) What are the defect types of parts in 3C products?

2) What are the main methods of defect detection in the production line?

3) Compared with manual inspection and machine inspection, what are the advantages of machine vision?

4) What are the advantages of CNN's application in 3C product defect detection?

5) Summary of CNN's in-depth development framework and its application in 3C product defect detection.

6) What is the future development trend of CNN in the 3C industry?

C. METHODOLOGY

In this study, published literatures were selected from databases such as Web of Science, Scopus, CNKI, Google Scholar and the Engineering Index, as well as publishers' databases such as Elsevier, IEEExplore and Springer. These literatures cover the problems related to traditional machine vision defect detection and deep learning detection methods in 3C products. Among them, Web of Science, IEEExplore and

other academic databases are rich in literature with a wide range of research. As a major producer of 3C products, China has many researchers studying the testing of 3C products. There are also many excellent Chinese journals and dissertations in the research results included by CNKI, so we also chose a small number of classic Chinese journals. Glass screen, TFT-LCD and PCB are the main parts of 3C products, which directly affect the quality and use of products. Therefore, when searching literature, keywords related to traditional machine vision (such as "machine vision", "visual algorithm" and "image processing") and keywords related to deep learning (such as "deep learning", "CNN" and "GAN") are used. Categories and application areas of 3C products (i.e., "mobile phone screen", "TFT-LCD", "PCB", "Defect detection" and "Defect classification").

The current literature review covers sources including journal articles, conference papers, dissertations and a small number of classical Chinese journals. The publication year is 2010-2021, because the past decade is a period of rapid development of deep learning. During the review process, we have summarized and analyzed the classic literatures on related topics, because a more detailed literature review will help scholars to the research work of this research field. However, due to the limitation of space, we reduced the description of the literature less related to the research content. Therefore, this paper mainly reviews most of the representative literatures, analyzes the literatures with significant influence in detail, and summarizes the literatures with insufficient influence.

The selection of references for this paper is based on the following considerations：

1) Research published in peer-reviewed academic journals;
2) Literatures published after 2010;

3) This paper aims to use machine vision and deep learning methods to study the field of 3C product defect detection;
4) Research on improvement method of 3C product defect detection;
5) Research through experiments, analysis, evaluation and modification.

II. DETECTION OBJECTS AND METHODS

A. OBJECTS

As an important interactive component of smart mobile devices, the display quality directly affects the operating status of the device, so the quality of the glass panels and screens of 3C products is very important. At the same time, as a provider of electrical connections for all electronic components, printed circuit boards have the advantages of small size and high circuit reliability. Therefore, this paper focuses on three product parts, glass display screens, TFT-LCD, and PCB, as the main object of research. The surface defects of 3C products are shown in Figure 2. Figs. 2 (a1) - (a6) are common defects of glass display screens. During the glass screen production process, production defects such as cracks, floating objects, and point defects will appear [11]. These defects affect the overall appearance of the product and also affect the normal use of users. Therefore, the factory attaches great importance to the inspection of the surface quality of smart devices. Figs. 2 (b1) - (b6) are common defects of PCB. Due to the production environment or process problems, there are often open circuit, connection [19], wrong hole, miscellaneous copper [20] and other problems in the production of PCB. Figs. 2 (c1) - (c6) show that TFT-LCD often has defects such as color difference, uneven ring, uneven gravity, etc. during the production process [21].
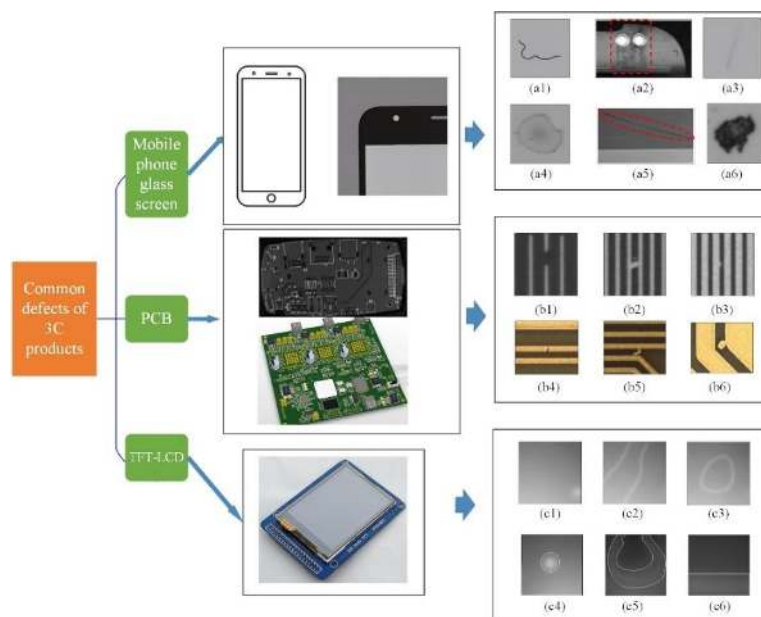

Figure 2. Product defect samples :(a1) - (a6) are common defects of glass display screens, including cracks, floaters and point defects [11, 22]; (b1) - (b6) are common defects of PCB, including circuit break, connection, projection and crack [19, 23]; (c1) - (c6) are common defects of TFT-LCD, including dot color difference, uneven ring and uneven gravity [21, 24].

B. METHODS COMPARISON

Traditional 3C product surface defect inspection methods are mainly divided into manual inspection and machine inspection. Table I compared three common defect detection methods. Compared with traditional machine vision algorithms, CNN feature extraction and classifiers can be automatically trained end-to-end from the input image. The method overcomes the shortcomings of traditional methods [25]. For example, CNN can extract image features more accurately for training, and has better robustness than traditional machine vision algorithms. However, the 3C product defect detection based on CNN is facing several important challenges.

Traditional machine vision has problems such as poor detection feature extraction robustness, high time complexity, and window redundancy [26]. But detection methods based on deep learning can effectively improve these problems. At the same time, the neural network is driven by big data, and the detection effect can be improved by increasing the data set [27]. However, it is difficult to obtain so many defective samples on actual industrial production lines.

TABLE I
COMPARISON OF THREE DEFECT DETECTION METHODS

Detection methods	Types of detectable defects	Advantages	Disadvantages
Manual inspection	Cracks, floating objects, point defects, open circuits, connections, wrong holes, miscellaneous copper, etc.	Many types of defects can be detected.	The detection speed is slow, which consumes manpower and material resources.
Traditional machine vision	Scratches, debris, dirt and other easy-to-find types.	Fast detection speed and high efficiency.	Low accuracy, poor robustness, and high time complexity.
Deep learning	Cracks, floating objects, point defects, open circuits, connections, wrong holes, miscellaneous copper, etc.	Fast detection speed, good stability and high accuracy.	It needs a lot of data samples.

III. CNN IMPROVEMENT RESEARCH

In 1998, LeCun [28] proposed a CNN structure based on gradient learning (the modern structure of LeNet-5), which was successfully applied to handwritten digit recognition. Compared with the traditional multi-layer neural network, CNN mainly adds three basic concepts: local acceptance domain, shared weights and pooling layers. Compared with a fully connected neural network, the widespread use of shared weights reduces the number of degrees of freedom parameters without loss of expressive power. This allows CNN to be trained by simple gradient descent [29]. Therefore, the appearance of LeNET-5 laid the foundation for the application of CNN in the field of image recognition.

Since the AlexNet proposed by Krizhevsky et al. [30] in 2012 won the ImageNet image classification competition, CNN has become the core algorithm in the field of image classification and opened a new chapter in deep learning. This section describes the improvements of CNN key components from three aspects: convolutional layer, pooling layer and fully connected layer.

A. CONVOLUTIONAL LAYER

Convolutional layers are one of the most important parts of CNN. The main function is to extract sample features from the input image. It consists of multiple filters and is used to calculate different feature maps. As the first layer of CNN, the convolutional layer is the core of CNN. Most calculations are performed in the convolutional layer. The traditional calculation method of convolution can be expressed by Equation (1). Convolutional networks are usually stacked alternately by convolution and pooling, and finally connected to complete the model construction. The convolution is multiplied by the linear filter corresponding to the position of the feature map and summed, and then nonlinear activation is performed to obtain the feature map. Later, the researchers conducted a more in-depth study on the convolutional layer. Wei et al. [31] found that dilated convolution can effectively integrate the surrounding environment by expanding the receptive field size of the kernel, providing a promising solution. By expanding the expansion ratio of a 3×3 kernel from 1 to 3, the discriminating ability of the convolution kernel can be enhanced. This proves that dilated convolution uses class activation mapping [32] to generate location maps at different dilation rates, which can improve the recognition ability of low-response target regions. Lin et al. proposed a network in network (NIN) model [33]. The idea is to replace the traditional convolutional layer with a multilayer perceptual layer, which consists of multiple fully connected layers containing nonlinear activation functions, which makes it more versatile. For the multi-layer perception layer, the calculation formula of feature mapping is as follows:

$$f_{i,j,k} = \max(\omega_k^T x_{i,j}, 0) \qquad (1)$$

here (i, j) is the pixel index in the feature map, x (i, j) represents the input patch centered at (i, j), and k represents the channel index of the feature map.

$$f_{i,j,k_1}^1 = \max(\omega_{k_1}^{1T} x_{i,j} + b_{k_1}, 0)$$
$$\dots$$
$$f_{i,j,k_n}^n = \max(\omega_{k_n}^{n\,T} f_{i,j}^{n-1} + b_{k_1}, 0) \qquad (2)$$

among them, n is the number of layers of the multilayer perceptron, and the activation function is Relu function.

In order to extract high-level features in CNN, a common method is to perform a deeper convolution, but the problem is that the network becomes larger consequently. GoogleNet's Inception module draws on the idea of NIN's multilayer perceptual layer, reducing parameters and extracting high-dimensional features under the premise of ensuring the quality of the model [34]. The idea of the Inception module is to cluster sparse matrices into denser sub-matrices to improve computational performance.

When training a deep neural network, the input of each layer will change with training, because the change of the parameters of the previous layer will cause the input of this layer to change. This phenomenon is called internal covariate shift (ICS) [35]. Inception V2 proposes the use of batch normalization (BN) to solve this problem, which can alleviate ICS and speed up the training speed of deep neural networks. After using BN, a higher learning rate can be allowed without the risk of divergence. BN also has the effect of a slight regularization model. On the other hand, even if the sigmoid activation function is used after using BN, the gradient will not disappear.

Inception V3[36] turned 7×7 convolutions in GoogleNet into a two-layer concatenation of 1×7 and 7×1. In the same way, the 3×3 convolution is changed to 1×3 and 3×1. This method not only speeds up the calculation, but also increases the nonlinearity of the network and reduces the probability of overfitting. The method of Inception V4 [37] was to add ResNet's residual module on the basis of the original inception, but the residual module of ResNet is not to increase the accuracy by increasing the depth, but to increase the calculation speed.

B.   POOLING LAYER

Pooling layer is generally used after convolution layer. It can simplify the information output from the convolutional layer and reduce the dimensionality of feature mapping [38]. There are two kinds of classic pooling operation: average pooling and maximum pooling. As shown in Fig. 3, the maximum pooling operation is a pooling operation that uses the maximum value in the data block as the output and extracts the maximum response of the feature plane. The average pooling operation is the operation of outputting the arithmetic average of the elements in the block as a function and

extracting the local corresponding average of the feature plane. The feature of the maximum pooling operation is to retain the image texture features, while the feature of the average pooling operation is to retain the overall data features.
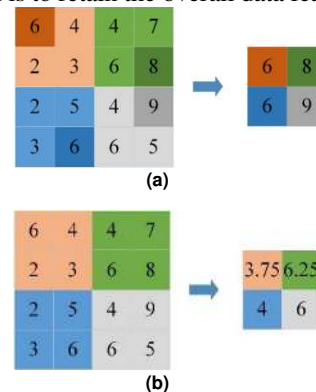

Figure 3. Two kinds of pooling operations; (a) the maximum pooling operation; (b) the average pooling operation

In addition to the common average pooling and maximum pooling, scholars have also proposed improvements in the pooling layer to improve network performance. Krizhevsky [30] et al. proposed overlapping pooling in the AlexNet network. Compared with the traditional no-overlapping pooling, the use of overlapping pooling can not only improve the prediction accuracy, but also reduce over-fitting to a certain extent. He [39] et al. proposed a spatial pyramid pooling method in their network, which could convert the convolutional features of images of any scale into the same dimension, which not only allows CNN to process images of any scale, but also avoids the problems of cropping, warping. The ROI pooling proposed by Girshick et al. [40] in Fast R-CNN is an operation widely used in target detection tasks using CNN, which greatly improves the processing speed. Spatial pyramid pooling (SPP) uses multiple pooling operations of different sizes for the same input, and stitches the results of different scales as the output. And ROI pooling can be regarded as a single-scale SPP, and only one pooling operation is performed for an input.

C.   ACTIVATION FUNCTION

The activation function is a vital part of the neural network, it can improve the nonlinear expression ability of the model. The activation function is divided into linear and non-linear activation functions. Different types of activation functions such as sigmoid, tanh, ReLu, lReLu, pReLu, etc. can be used according to different situations. Among them, the mathematical expression of the sigmoid function is:

$$\sigma(x) = 1/(1 + e^{\wedge}x) \qquad (3)$$

the range of σ(x) is 0-1. When the ultimate goal of the network is to predict probability, it can be applied to the output layer.

The activation function of tanh is also called the hyperbolic function. Similar to sigmoid, tanh also compresses a real value. Unlike sigmoid, tanh has zero mean in the output range of -1 to 1. The mathematical expression of the tanh function is:

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \qquad (4)$$

Both tanh function and sigmoid function have the problem of gradient disappearance. In order to solve the problem of vanishing gradient, the nonlinear activation function of linear rectification function (ReLu) solves this problem well. It is better than the two activation functions, and it is also the most widely used activation function today.

The mathematical expression of the ReLu function is:

$$f(x) = \max(0, x) \qquad (5)$$

the so-called nonlinearity means that the first derivative is not a constant. The definition of ReLu is max (0, x), so the derivative of ReLu is:

$$f'(x) = \begin{cases} 0 & for \quad x < 0 \\ 1 & for \quad x \geq 0 \end{cases} \qquad (6)$$

Obviously, the derivative of ReLu is not constant, so ReLu is nonlinear. ReLu can make the output of some neurons zero, by diluting the network and reducing the mutual dependence of parameters, to achieve the purpose of reducing the occurrence of overfitting.

However, since ReLu has a gradient of 0 when the neuron is inactive, this may cause the initially unactivated unit to never activate. For this reason, in 2013 Maas [41] proposed leaky ReLu, which compresses the negative part instead of mapping it to a constant zero point, so that when the neuron is

in an inactive state, it allows a small non-zero gradient. The mathematical function of leaky ReLu is:

$$f(x) = \max(0.1x, x) \qquad (7)$$

In 2015, He [42] proposed parametric ReLu with adaptive learning parameters. Its mathematical function is:

$$f(x) = \max(\alpha x, x) \qquad (8)$$

where α is a hyperparameter.

In 2017, inspired by the use of sigmoid function for gating in LSTM and highway network, Google researchers proposed the swish activation function [43]. Like ReLu, the swish function has a lower bound. Unlike ReLu, swish is smooth and monotonous, and its mathematical function is:

$$\sigma(x) = \frac{x}{1 + e^{-x}} \qquad (9)$$

D. IMPROVEMENT OF CNN ARCHITECTURE

Scholars have begun to study the depth of CNN, and the network architecture has also begun to develop in a deeper and deeper. If there are more convolutional layers, CNN can easily detect complex objects or patterns [44, 45]. By increasing the depth of the CNN, the non-linearly increased objective function can be better approximated to get better results [46]. This section reviews the improved research methods of CNN from the aspect of network architecture.

TABLE II
THE EVOLUTION OF THE CNN ARCHITECTURE

Models	Year	Depth	Main findings	Input size
AlexNet [30]	2012	8	Uses Dropout and ReLU	227×227×3
NIN [33]	2013	3	Add new layer,'mlpconv'	32×32×3
ZfNet[47]	2014	8	Visualization idea of middle layers	224×224×3
VGG [45]	2014	16,19	Increased depth, small flter size	224×224×3
GoogLeNet [34]	2015	22	Increased depth, different flter size, concatenation concept	224×224×3
Inception-V3 [36]	2015	48	Utilizes small fltersize, better feature representation	229×229×3
Highway [48]	2015	19,32	Presented the multipath concept	32×32×3
Inception-V4 [37]	2016	70	Divided transform and integration concepts	229×229×3
ResNet [49]	2016	152	Robust against overftting due to symmetry mapping-based skip links	224×224×3
WideResNet [50]	2016	28	Decreased the depth and increased the width	32×32×3
Residual attention neural network [51]	2017	452	Presented the attention technique	40×40×3
DenseNet [52]	2017	201	Blocks of layers	224×224×3
MobileNet-v2 [53]	2018	53	Inverted residual structure	224×224×3
HRNetV2[54]	2020	-	High-resolution representations	224×224×3

In 1998, LeCun [28] proposed the LeNet-5 architecture and applied it to the recognition of handwritten digits, laying the foundation of CNN in the field of image recognition. LeNet-5 consists of two convolutional layers, two pooling layers and two fully connected layers. Each convolutional layer has a different number of 5×5 filters. There are 6 and 16 filters on the first and second layers respectively. Feature map uses the sigmoid function to activate each time it passes through the convolutional layer, and then uses the mean pooling operation. Since 2012, CNN has begun to rise, and the development of CNN architecture is shown in Table II.

In 2012, Krizhevsky [30] proposed AlexNet and used ReLu as the activation function in the network. Successfully solve the problem of gradient disappearance in the deep network, Sigmoid is used as the activation function. As shown in Fig. 4, the architecture of AlexNet consists of 5 convolutional layers and 3 fully connected layers, and the output of each fully connected layer is 4096 neurons. The third, fourth, and fifth convolutional layers are directly connected, and there is no pooling layer and normalization layer in two convolutional layers. The third convolutional layer has 384 cores of size 3×3×256, which are connected to the

standardized, merged output of the second convolutional layer. The fourth convolutional layer has 384 cores with a size of 3×3×192, and the fifth convolutional layer has 256 cores with a size of 3×3×192. In order to avoid model overfitting, AlexNet uses dropout to randomly ignore some neurons during training.
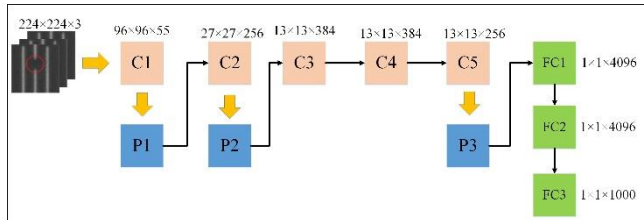

Figure 4. Simplified network structure in AlexNet

VGGNet (Visual Geometry Group Net) [45] is a classic convolutional neural network in 2014. Except for the addition of a convolutional layer, VGGNet is similar to AlexNet. VGGNet consists of 10 convolutional layers, 5 pooling layers and 3 fully connected layers. Simonyan et al. [45] used a very small convolution filter to achieve significant improvements to the current technical configuration by increasing the depth of the weight layer to 16-19 layers, proving that classification tasks can increase the depth of CNN by using a small convolution kernel to improve accuracy. The structure of VGGNet is very simple, and the entire network uses 3×3 size convolution kernels and 2×2 pooling size. Compared with AlexNet, the simpler structure of VGGNet has better performance.

In 2015, the Google team used the idea of a small network to conduct research on GoogleNet, which replaced the traditional convolution operation with a point-by-point group convolution operation, thereby improving computational efficiency. The parameters of this network are 12 times less than that of the 2012 championship team, but it is more accurate. GoogleNet adopts the method of NIN to improve the performance of the network. This method can be seen as an additional 1×1 convolutional layer plus a ReLU layer. The most important thing about NIN is dimensionality reduction, which solves the computational bottleneck and thus solves the problem of limited network size. In this way, the depth and width of the network can be increased without significant performance loss.

In 2015, ResNet won the first place in the ImageNet competition classification task. Deep convolutional networks have added more possibilities to the field of image classification[55]. Deep networks naturally integrate features and classifiers in an end-to-end multi-layer manner, and features can be enriched by the depth of stacked layers. However, the number of stacked layers is not the better. Recent years, scholars have found that too many stacked layers will cause the problem of gradient explosion [56]. Gradient explosion has hindered convergence from the beginning. This problem is because deepening the network will make the optimization of the stochastic gradient descent algorithm more difficult, and the network parameters cannot

be updated, which makes the network training effect worse [57]. He et al. [49] proposed a deep residual network composed of many residual network structures, namely ResNet. The network using the residual structure with short connection is similar to performing the same mapping, so that the deep convolutional neural network can obtaining the front layer gradient. Then, the emergence of the deep residual network has improved the representation ability and learning ability.

DenseNet [52] is another deeper convolutional neural network. As shown in Figure 5, like the GoogLeNet network is composed of inception modules, the ResNet network is composed of residual blocks, and the DenseNet network is composed of dense blocks. Each layer gets additional input from all previous layers, and transfers its own feature map to all subsequent layers. Using concatenation, each layer is receiving "collective knowledge" from the previous layers. It has the following advantages: (1) DenseNet has fewer parameters than ResNet, so it is easier to calculate. (2) The network enhances the reuse of features. (3) This network is easier to train than other deep networks and has a certain regularity effect. (4) It alleviates the common problems of gradient disappearance and model degradation in deep networks.
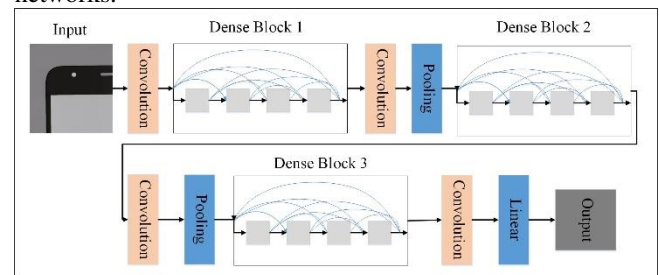

Figure 5. DenseNet network structure

E. SUMMARY

Compared with traditional learning techniques, CNN is more scalable because it can obtain higher accuracy by increasing the number of network layers and changing the size of the training dataset. Because the characteristics of natural images are very complex, a large number of parameters design and deep network models are required. At the same time, researchers have also done research to reduce network parameters, such as removing the last fully connected layer and replacing it with global average pooling. Currently commonly used CNN models such as VGG Net, Google Net, ResNet, DenseNet, etc. are designed for the Image Net data set, and higher performance can be obtained by increasing the number of convolutional layers.

IV. APPLICATION OF CNN

As a deep neural network model, CNN performs well in the ImageNet challenge [29]. Since the emergence of LeNet-5, the architecture of CNN has basically been established [57]. As shown in Fig. 6, it is the detection process of defect detection in 3C production. After the first two steps of defect feature extraction, the feature information is input to the most classic shallow CNN in the last step. The architecture has a

data input layer, two convolutional calculation layers, two pooling layers, and a fully connected layer. Limited by the software and hardware facilities at the time, LeNet-5 did not perform well in large-scale data set classification tasks. Traditional CNN is also widely used in the field of industrial product inspection, but its sensitivity is relatively low, and it is difficult to detect small defects in 3C product inspection.
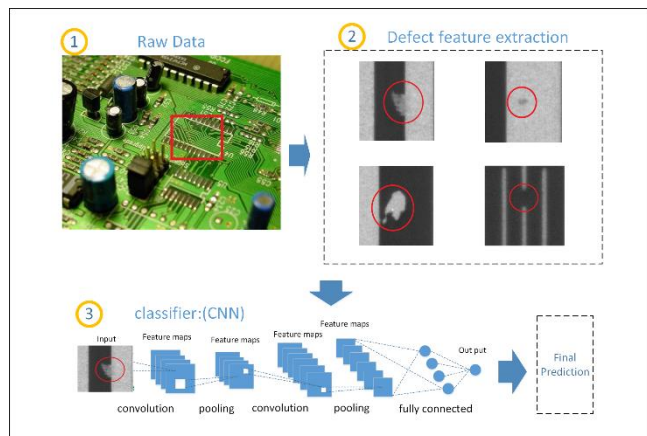

Figure 6. 3C product automatic defect detection process

A. APPLICATION IN MOBILE PHONE GLASS PLATE

In the quality inspection of 3C products, defect inspection is one of the most important links in the manufacturing process of touch screens [58]. ùaban et al. [59] proposed a fast and effective glass surface defect detection and segmentation method BiasFeed CNN through comparison with traditional CNN experiments, which solved the problem of the transparency and reflection characteristics of the glass surface. They converted the single-value bias input of the traditional CNN algorithm into a bias template. Adding a bias template solves the problem of light fluctuations in the lighting system. Compared with traditional CNN, its specificity and accuracy are improved. Zhao et al. [60] proposed an automatic scratch detection method (ScratchNet) that combines the LeNet-5 structure and the VGG network convolutional layer. They used the method of connecting two main modules in series to optimize CNN for small target defect detection. When using the same data for training, the accuracy of LeNet-5 was 95.97%, and the accuracy of scratchNet was 96.35%. Jin et al. [61] proposed a multi-channel self-encoding convolutional network (AECNN) model to deal with the problem of false detection due to small difference in feature space in glass detection. Generally, in order to make the network achieve better information capture ability, it is necessary to increase the number of convolution kernels in the network, which also makes the training time longer. Therefore, they introduced the unsupervised convolutional autoencoder into the convolutional neural network to reduce the training time. In addition, to prevent the network from overfitting in defect recognition, the end classifier was changed to a fuzzy support vector machine. The final detection accuracy rate of this model is increased from 92.6% to 97%.

TABLE III
COMPARISON BETWEEN DIFFERENT MODELS IN MOBILE PHONE GLASS DEFECTS

Model name	Sample object	Number of samples	Accuracy	Advantages	Disadvantages
BiasFeed CNN (8 floors) [59]	Scratch, crack	35	98%	Add bias template to improve sensitivity.	Less data sets, reliability needs to be improved.
ScratchNet (8 floors) [60]	Small scratches	4600	96.35%	Optimize the detection method for small objects.	The method is not end-to-end.
AECNN (8 floors) [61]	Inclusion and tumor type defects	600	97%	Add unsupervised learning methods to reduce training time.	The recognition rate of bubbles and tumors is low.

As shown in Table III, the detection accuracy of the above method for mobile phone glass has been improved, and the real-time performance of the experimental results were high, but the number of detection samples were small.

B. APPLICATION IN PCB

With the improvement of the 3C products quality, the surface defect detection of printed circuit boards (PCB) has become an important issue. Since PCB defects are usually intensively occurring, this detection is a multi-label classification problem. In response to this problem, Zhang et al. [20] proposed a multi-task CNN model, which has three blocks, and each block include a convolutional layer, an activation layer and a maximum pooling layer. Finally, six types of classification were performed through the fully connected layer. The model with 1200 data samples were trained according to different types. According to experiments, the accuracy of the model reached 92.86%. The model proposed by Zhang and Ma et al. [62] was compared in multiple categories, and the results are shown in Fig. 7.

Among them, Ma uses a defect detection method based on machine vision. The overall accuracy of Zhang's model is much higher than that of Ma's. Adibhatla et al. [63] used a large number of images to train CNN to classify defective or intact PCBs. The network has 60 million parameters, 500,000 neurons, and is composed of 5 convolutional layers. The convolutional layer is followed by the largest pooling layer, and the last is two fully connected layers. After training, the method has a higher accuracy rate. The overall accuracy of PCB defect classification has reached 85%. Wang et al. [64] proposed a precise PCB defect recognition algorithm based on CNN. The model performs a differential operation on the reference image to find the defect area, and batches normalize the PCB defect images. The model chooses ReLu and Maxpooling as the activation function and down-sampling methods respectively. Finally, the model used the softmax regression classifier for training and optimizing CNN. Experiments showed that the correct recognition was significantly improved, and the detection accuracy of 10 types of PCB defects was as high as 96.67%.
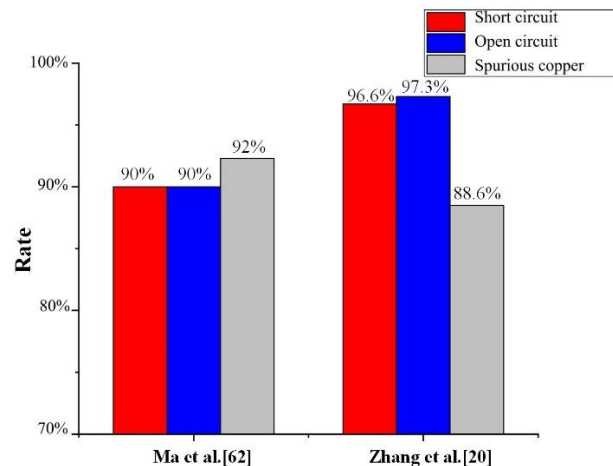

Figure 7. Comparison of PCB model detection defects

As shown in Table IV, the above detection methods can realize multi-task classification detection aiming at the diversity of PCB defects. Compared with machine vision method, CNN method has higher accuracy, and the accuracy rate can reach 92%-97%. However, the detection time of the above methods is long, the real-time performance and the number of detection datasets need to be improved.

TABLE IV
COMPARISON BETWEEN DIFFERENT MODELS IN PCB DEFECTS

Model name	Sample object	Number of samples	Accuracy	Advantages	Disadvantages
LeNet-5 (9 floors) [20]	Short circuit, Open circuit, etc.	1920	92.86%	It can handle multiple classification tasks, but the number of data sets affects accuracy.	Limited data set affects accuracy.
LeNet-5 (9 floors) [63]	Distinguish between good and damaged samples	1480	85%	The detection time of the model is shortened, but the accuracy rate needs to be improved.	The accuracy rate is less than 90% and needs to be improved.
LeNet-5 (9 floors) [64]	Short circuits, holes, etc.	4743	96.67%	The real-time performance of the model is high.	The accuracy rate needs to be improved.

C. APPLICATION IN TFT-LCD

Traditional image processing algorithms are prone to problems such as missed detection and misjudgment in the recognition of circuit defects, He et al. [65] compared the use of CNN in electronic circuit defect recognition in industrial inspection. First, the input image was preprocessed by histogram equalization. Second, the image features were extracted on the 8-layer CNN structure. Finally, the softmax classifier is used to realize the recognition and classification of image features. After experimental verification, the model algorithm had high accuracy, robustness and generalization ability, meeting the needs of industrial testing. They compared the method with Faster RCNN, RCNN and Deformable Part Mode (DPM), as shown in Fig.8. It can be seen that the accuracy of the Depth-2 is higher than that of other models.
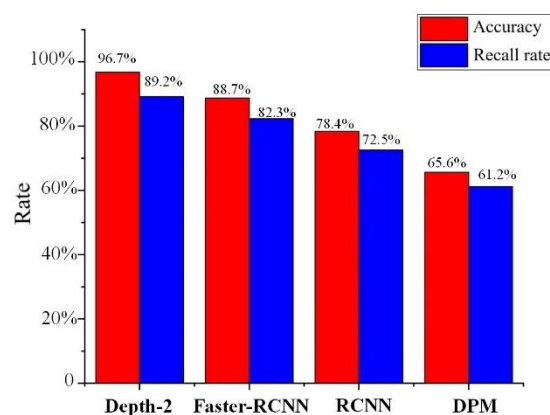

Figure 8. Comparison of accuracy and recall of the three algorithms mentioned by He[65]

D. SUMMARY

To sum up, compared with traditional machine vision methods, the defect detection model based on CNN improves the accuracy, which can reach 92%-98%. Multi-task classification can be realized, and the overall accuracy is improved, but the detection results of some defect types need to be improved. At present, the test object of the experiment is relatively simple, and the training dataset of the model is limited, which affects the accuracy of the model to a certain extent. Data sets similar to the actual production line products should be added. They still have great progress to make and a long way to do in the detection of small defects. The real-time performance of detection needs to be improved. The detection model based on shallow CNN can not fully meet the requirements of defect detection. In order to improve the shortcomings of CNN detection model, scholars constantly explore new technologies and methods.

V. APPLICATION OF DEEP CONVOLUTION

Looking back at the development of deep convolutional neural networks in the field of image recognition, we can clearly find that the expressive power and feature extraction capabilities of DNN increase with the increase of the depth of the network increase. Compared with shallow neural networks, the performance of DNN is further improved [30]. Since the AlexNet proposed by Krizhevsky et al. won the championship in the ImageNet competition, scholars have found on the exploration of deep convolution, and it has also been widely used in the detection of 3C products. As the number of layers of the convolutional network increase, its ability to extract semantic features is significantly enhanced. However, the increase number of layers is also accompanied by the problem of the gradient disappearance. After the researchers' exploration, it is observed that the phenomenon of the disappearance of the gradient can be solved by the residual network and the dense network. With the emergence of ResNet and DenseNet, researchers began to conduct more in-depth research in the field of defect detection.

A. APPLICATIONS ON THE PHONE SCREEN

Based on the detection of glass cover defects of smart device, Park et al. [25] proposed a multi-channel defect detection structure 4-DarkNet, which was based on the model superposition in the machine learning integration method. It combined a variety of classification or regression models. They used DarkNet-19 as a defect classifier, with a total of 19 convolutional layers and 5 pooling layers. They also applied the weighted average method of the stack model to enable the independent classifiers in the detection structure to perform best. Although the structure is relatively large, it has the advantages of fast speed and high precision, and is suitable for defect detection on the production line. In view of the multiple types of problems on mobile phone screens, Li et al. [22] proposed a detection model based on a region of interest algorithm. The model structure added a multi-layer perceptron (MLP) and deep learning to the ROI and deep learning

algorithm. Using different deep learning models such as VGG-16, ResNet and GoogLeNet to detect with 400 sample, the results reached an accuracy rate of more than 97%. Chen et al. [23] designed a defect detection method based on CNN to solve the problem of the difficult appearance detection of smartphone protective screens in the production process. The method first divided the sample image into 256×256 pixel sample images, and then used 22 layers GoogleNet was trained and fine-tuned. In order to make the network model less space and time consumption, the hidden layer with more parameters was deleted through multiple adjustments of parameters, and the 1000-dimensional vector of softmax was adjusted to a three-class three-dimensional vector. Finally, through five experiments of fine-tuning, the model has obtained good detection results. Ma et al. [17] proposed a mobile phone surface defect detection method. An industrial line scan camera was used to obtain the original surface image of the mobile phone. Through the preprocessing steps proposed in this article, the obtained image is automatically divided into specified sizes. The trained CNN can be combined with sliding window technology. Experiments showed that the defect detection rate can reach 99.5%.

Among various glass defects, the detection of dent defects is one of the most difficult because of its small depth changes and smooth edges. In the defect detection system based on machine vision, the dent image has the problems of uneven gray scale and low contrast. Wang et al. [66] proposed a dent defect detection method based on deep convolutional neural network for such problems. By improving the DenseNet-121 [67] model, they designed a compact model that can be meet the requirement in real-time production. However, as the number of network layers increases, there is a problem on gradient disappearance, which allows shallow features to be discarded, resulting in unfavorable detection of small dent defects. The fusion strategy proposed by DenseNet can solve this type of problem well. Wang chose DenseNet as the infrastructure, which can overcome the data dependence problem of the DCNN model and significantly improve the recognition accuracy. Experimental results showed that this method can further improve the recognition accuracy of the dent detection task with 85.42% on 70 test images. In future work, this method can be applied to the formula against network (GAN) to further improve the robustness. In order to quickly and effectively detect and identify touch screen glass defects, Zhang et al. [26] proposed a detection and recognition method based on Mask R-CNN technology. This method used Mask R-CNN as the basic model and used the multi-picture stack method to obtain sample data, and then process, label, and sample amplification. Since the dataset of this experiment is relatively small, they tested three different networks (VGG16Net, ResNet50+FPN, ResNet101+FPN) in the feature extraction stage. The experimental comparison showed that the backbone network detection effect of ResNet50+FPN is better, and the accuracy rate was high and it had good robustness compared with traditional methods.

IEEE *Access*

Multidisciplinary : Rapid Review : Open Access Journal

TABLE V

COMPARISON BETWEEN DIFFERENT MODELS IN MOBILE PHONE GLASS DEFECTS

Model name	Sample object	Number of samples	Accuracy	Advantages	Disadvantages
4-DarkNet (27 floors) [25]	Scratch, dent, chips	4220	99%	Effective use of multiple channels to measure data.	Accuracy has improved, but real-time performance needs to be improved.
ROI+MLP (16 floors) [22]	Scratches	4600	97%	Optimize the detection method for small objects.	Real-time performance needs to be improved.
GoogleNet (21 floors) [23]	Scratches, scuffings, punctures, etc.	9640	98%	The model is more lightweight and the training speed is faster.	There is a difference between the detection target and the product defect.
GoogLeNet (21 floors) [17]	Point defects and line defects	7632	99.5%	Greatly reduced the number of parameters.	Fewer types of detectable defects.
DenseNet (15 floors) [66]	Scratch, dent, etc.	210	99.90%	Simplified model, real-time detection.	Small sample size, single type of detectable defect.
Mask R-CNN (50 floors) [26]	Touch screen glass defects	1035	96.7%	Has good robustness.	Failed to obtain information such as defect area, coordinates, etc.

In summary, as shown in Table V, these detection models based on DCNN have higher accuracy than shallow CNN models, ranging from 96% to 99%. These models have high real-time performance, and the ability of processing datasets has been improved, which is also an important factor affecting the accuracy.

B. APPLICATION IN PCB

Zhang et al. [68] proposed an improved defect detection method for bare PCBs, using VGG-16 as the basic network for feature extraction, which is achieved by learning deep identification features. It reduced the high requirements for deep learning methods for large datasets. This method first used artificial defect data samples and expanded datasets, then adopted a deep pre-trained CNN to learn defect features, and finally utilized a sliding window to further locate defects. This algorithm is helpful to establish a robust model for multi-class recognition tasks. Cheong et al. [69] introduced a PCB automatic component recognition system based on CNN. In addition, the defects of PCB components were also localized. A simple component recognition classifier based on CNN was developed, and the pre-trained model was used for transfer learning. Pre-trained models, such as VGG-16, Inception V3 and DenseNet-169, are used to study which model is best for component recognition. The test results showed that using transfer learning in VGG16, the best result obtained was 99% accuracy, and the main accuracy of the system could reach 96.54%. Volkau [70] proposed a variant of transfer learning,

which includes a combination of unsupervised learning used on VGG16 and pre-trained on ImageNet weight coefficients. The goal is to extract significant semantic features from normal samples without supervision. To demonstrate defect detection, they utilized a set of PCBs with different defects scratches, missing gaskets, extra holes, frayed, and damaged PCB edges. The trained model can cluster the normal internal representation of PCB features in the high-dimensional feature space, and locate the defect blocks in the PCB image according to the distance from the normal cluster. Preliminary results showed that more than 90% of defects can be detected.

As shown in Fig. 9, Xia et al. [71] combined SSIM and MobileNet-V3 to propose a new PCB defect detector structural similarity index (SSIM-NET). Compared with YOLO-V3, Faster-RCNN, and tiny defect detection network (TDD-NET), SSIM-NET has higher accuracy and speed. The method has two stages: the first step is to adopt SSIM to detect suspicious areas; the second step is to use the latest lightweight backbone Mobilenet-V3 to classify the suspicious areas. After testing, the accuracy of the model reached 97.06% and the speed of 60fps, achieving real-time detection and high precision in the test set. Ding et al. [72] proposed TDD-NET in which online hard case mining was used throughout the training phase to improve the quality of recommendations from the region of interest and thus make more effective use of data information. In order to reduce redundancy, non-maximum suppression was adopted in the proposed regions according to the classification score. TDD-NET integrated the

multi-scale feature fusion strategy to obtain strong features in structure and enhance the ability to detect minor defects, with an average accuracy of 98.90%. To sum up, as shown in Table

VI, these models not only improve the accuracy, but also reduce the requirements for large data sets, because the feature extraction ability of the models has been improved.
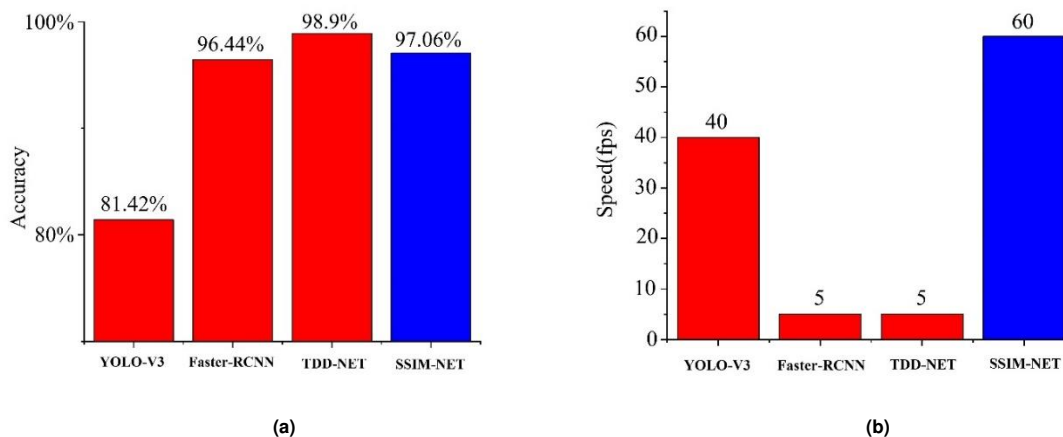

Figure 9. Accuracy and speed comparison of YOLO-V3, Faster-RCNN, TDD-NET and SSIM-NET [106]

TABLE VI
COMPARISON BETWEEN DIFFERENT MODELS IN PCB DEFECTS

Model name	Sample object	Number of samples	Accuracy	Advantages	Disadvantages
VGG-16 (16 floors) [68]	Short circuit, Open circuit, etc.	1920	92.86%	Reduced the high requirements for deep learning methods in large datasets.	Accuracy needs to be improved.
VGG-16 (16 floors) [69]	25 kinds of defects such as short circuit and open circuit	1480	99%	Localize the defects of PCB components.	Fewer types of defects detected and small data set.
VGG16 (16 floors) [70]	Scratches, abrasions, breakages, etc.	4743	96.67%	Can be positioned according to defect characteristics	Not suitable for defects with subtle changes in texture.
SSIM-net (22 floors) [71]	Short circuit, Open circuit, etc.	9640	98%	The model is more lightweight with fast speed.	Need to add more defect types.
TDD-Net (101 floors) [72]	Tiny defect detection	2508	98.9%	Adapt to the detection of small defects.	Post-processing methods need to be improved.

C. APPLICATION IN TFT-LCD

Aiming at the defect[ii]s of TFT-LCD circuit, He et al. [73] proposed an improved Faster R-CNN algorithm to detect defects. They adopted different convolution kernel sizes and network layer depths to test the detection performance of the model. After testing, the 16-layer neural network structure had achieved good detection results, further improving the accuracy and practicality of the neural network in the field of

automatic detection. Kim et al. [74] used VGGnet to detect TFT-LCD defects and set the first convolutional layer of VGGnet into two spaces. In addition, instead of using a 2×2 maximum pooling layer on the last pooling layer, a global average pooling layer is used, that is, the elements of each channel are averaged. After they slightly adjusted VGGnet, the model reduced the number of parameters and learning time. As shown in Table VII, these models have improved data processing capabilities, real-time performance has been

IEEE *Access*
Multidisciplinary : Rapid Review : Open Access Journal

improved, and can be adapted to the detection of production line defects.

TABLE VII
COMPARISON OF DIFFERENT MODELS IN TFT-LCD DEFECTS

Model name	Sample object	Number of samples	Accuracy	Advantages	Disadvantages
Depth-2 16 floors [70]	Breakpoints, breakages, foreign objects, scratches	169800	94.6%	High real-time.	Accuracy should be improved.
Depth-3 24 floors [70]	Breakpoints, breakages, foreign objects, scratches	169800	93.7%	High real-time.	Accuracy should be improved.
VGG net 16 floors [74]	Breakpoints, breakages, foreign objects, scratches	2880	86.13%	Reduced learning parameters and learning time.	The data set is too small, the accuracy should be improved.

D. SUMMARY

To sum up, the aforementioned deep convolutional network model is widely used in industrial manufacturing and image classification and can achieve good performance. At present, the commonly used CNN architectures for defect detection are GoogleNet and VGGNet. According to researchers' studies, the nonlinearity of the network increases as the depth increase, and at the same time it is closer to the objective function to obtain a better feature representation. However, as the depth increases, the system structure becomes complicated and cumbersome, and the real-time performance is also weakened. Therefore, how to solve the problem of the cumbersome system structure is also worthy of attention. Scholars use the residual structure of the residual network to establish models and conduct experiments, which proves that the residual network is also feasible for defect detection in the 3C industry [49]. Compared with the ordinary network, the residual network introduces one or more jump connections, which can make the information of the previous residual block flow into the next residual block unimpeded. This improves the information flow and avoids the problem of gradient disappearance and degradation caused by excessive network depth.

VI. THE APPLICATION OF GAN IN 3C INDUSTRY

The idea of confronting network comes from the two-person zero-sum game in game theory, which is equivalent to the two-party game of minimization and maximization [75]. The deep learning model is driven by big data, and the training effect depends on the sample size, and the training effect is proportional to the sample size [76]. However, it is not easy to obtain a large number of defective samples from an industrial production line [77]. As shown in Fig. 10, scholars can generate random defect samples through GAN. Li et al. [78] introduced another CNN called deep convolutional generative

adversarial networks (DCGAN). Through training on various image data sets, they proved that it was powerful candidate for unsupervised learning. DCGAN is an extension of GAN, with convolutional network as discriminator and deconvolution as generator. It can automatically extract and fuse defect features, expand defect samples. Therefore, scholars began to use GAN to generate data sets to expand the number of samples.

Based on the method of Yuan [79], Lv et al. [80] designed a display glass defect detection model suitable for small sample learning. They designed DCGAN, which introduced residual module to improve the extraction capability of feature extraction network. The system automatically extracts and merges the defect features from the sample images, and expands and generates the defect samples. Then, based on the expanded defect sample data set, the detection model of FAST R-CNN is improved and trained. After the comparison between the original model and the model experiment with DCGAN added, in order to evaluate the test results, "over-detection rate" (ODR), "missed detection rate" (MDR) and "Accuracy" of the data samples were compared and evaluated, as shown in Fig. 11. Finally, the improved detection model obtained better detection results, and solved the problem that the number of defective samples in the industry was small and deep learning required a large number of samples. The experimental results proved the effectiveness and feasibility of combining DCGAN and Fast R-CNN for display defect detection. Lu [57] proposed a mobile phone display TFT-LCD surface defect detection model based on small sample learning. In response to the shortage of negative samples on the actual automated production line, Lu used the collected small samples and the DCGAN network model to generate new negative sample data. The algorithm uses DCGAN to target a large amount of newly generated data, which makes up for the lack of training data and makes the distribution of training data more reasonable.

By sending the generated samples into the model trained by migration learning, the secondary intensive training is performed to obtain better image defect characteristics.
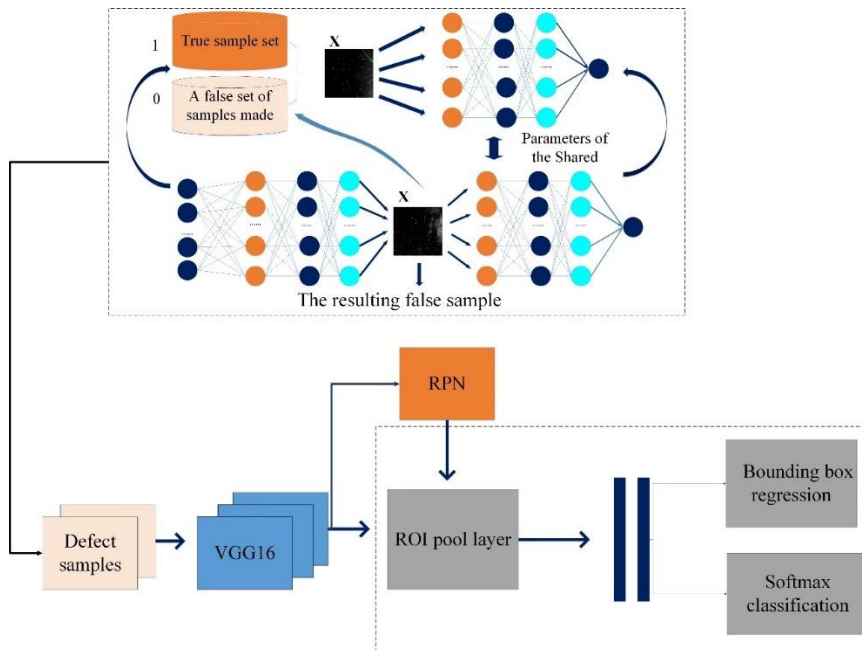

Figure 10. Principle and detection process of generating display defect samples through GAN
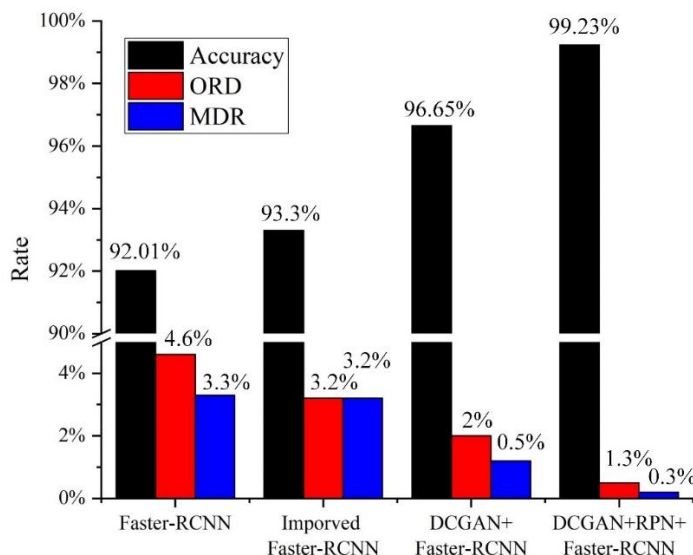

Figure 11. Comparison of Lv's model and Faster-RCNN-based detection model

In summary, unsupervised learning can extract patterns and structures from raw data without additional information. The introduction of GAN, a defect detection species, solves the problem of lack of actual defect samples. There has also been a major breakthrough in detection accuracy, which can reach 99.23%. But while introducing GAN, the real-time performance of the model is not very ideal, therefore the real-time performance of the model needs to be improved.

VII. DISCUSSION

With the rapid development of information technology, smart 3C products have become necessities in people's lives. In the automated production of 3C products, the quality of each component of the product must be ensured. At present, defect detection based on machine vision is the mainstream method. Based on the introduction of the development history of CNN, this article combines several representative algorithm

improvement models to provide an overview of multiple defect detection methods using CNN. Neural network can solve almost all detection and classification problems, and it is a commonly used image processing technology. CNN's ability to use spatial patterns is particularly conducive to the value of very high spatial resolution data. More and more visualization techniques will not only help explain, but also help learn from these models to improve the efficiency of defect detection in industrial production. In these researches, VGG16 and GoogleNet are the two most commonly used architectures. However, most of them have specific limitations and rely heavily on the size of the data set, image processing and texture. Solving the problem of lack of data sets is still a difficult problem for many researchers. Moreover, the large-scale neural network used for deep learning requires a lot of computing resources, which also leads to the inevitable large computing cost. For the application of different types of CNN frameworks, the summary is shown in Table VIII:

TABLE VIII

COMPARISON OF THE ADVANTAGES AND DISADVANTAGES OF DIFFERENT DETECTION METHODS

Detection subject	Typical methods	Avantages	Disadvantages
Manual inspection	Manual visual inspection	-	Consuming human resources, human eyes are easy to fatigue.
Traditional machine vision	Simple image processing technology	Fast detection speed.	Robustness is poor, accuracy rate is not enough.
Shallow CNN	LeNet-5 etc.	Simple structure, low hardware requirements.	Insufficient accuracy and limited ability to obtain data.
Deep CNN	AlexNet, ResNet, etc.	Strong ability to extract data and high accuracy.	High hardware requirements, and there is a risk of disappearing gradient.
Combine GAN	DCGAN	Reduce the difficulty of obtaining training samples.	GAN model is easy to collapse.

(1) Many studies have shown that CNN is superior to simple machine learning methods [61, 63, 64]. The traditional shallow CNN has the advantages of less time consumption, light and simple network structure, and low hardware requirements. The classic LeNet-5 is often used as the representative architecture of the shallow CNN. Although the LeNet-5 network has been rarely used in research, it has laid a foundation for the development of subsequent convolutional networks. In short, in terms of 3C defect detection, shallow CNN is more robust than traditional machine vision. The convolutional layer can accurately extract image features and improve the accuracy of detection. The small space complexity enables shallow CNN to adapt to the real-time requirements of the production line. However, the ability of shallow CNN to obtain data in the network training process is limited, which directly affects the accuracy of the training model. In future research, scholars should focus on improving the detection capabilities of some small defects.

(2) In the development of depth and adaptability of different structures, the learning ability of CNN has been significantly improved [23, 25, 71]. Deep CNN benefited from the increase in network depth, and its accuracy and precision have been improved. According to the summary of this article, in the study of 3C defect detection, the most commonly used architectures are VGGNet, GoogleNet and ResNet. As the network depth increases, the system structure becomes complicated and cumbersome, and the requirements for hardware are also higher. It also requires a large number of data samples to improve the precision and accuracy of the system. However, the residual network is composed of multiple shallow networks. It does not fundamentally solve the problem of vanishing gradients, but avoids vanishing gradients. Because shallow networks do not have the problem of vanishing gradients during training, ResNet uses this point to avoid vanishing gradients. In summary, the detection effect of deep CNN is higher than that of shallow CNN, but as the depth increases, the data parameters that need to be processed increase, and the real-time performance required by the production line is not easy to achieve.

(3) GAN has many advantages in defect detection, for example, it can generate real images or videos; its addition can reduce the direct data required. At the same time, GAN also has many limitations. For example, GAN makes the system need more time to train the data; different types of data are required to continuously check the results and training data (whether it is used correctly), and the model is prone to collapse. Therefore, the future research direction of DCGAN must be to solve model collapse, non-convergence and training difficulties. DCGAN replaces the multilayer perceptron in the original GAN with a convolutional neural network in the generator and discriminator feature extraction layer. When the sample data is limited, DCGAN can improve the detection accuracy by extracting and fusing features. It effectively solves the limitation of low detection accuracy of the detection system when the sample is insufficient. At present, GAN is a strong competitor in unsupervised learning technology, and DCGAN will become the trend and main technology in the field of detection in the future.

VIII. OUTLOOK

The proposal of Industrial Manufacturing 4.0 indicates that information technology and intelligent manufacturing will be the core development direction. Research shows that, in the continuous development and improvement, the application of CNN in the detection field has achieved good development, showing the superiority of feature extraction and classification detection. For example, the shortcomings of manual inspection of products have been greatly improved, and the

manual inspection of products has been improved. The product qualification rate is high and the production quality is guaranteed. As an algorithm that has attracted attention from scholars since its emergence, CNN has solved to certain extent problems that could not be solved or solved difficult problems before, and greatly improved the efficiency and accuracy in the detection field.

(1) In the research of scholars, various improvement strategies have improved the performance of CNN to a certain extent, but there are still shortcomings. For example, the problem of gradient explosion has not been solved. For the problem of complex multi-level network structure, there are also problems such as difficulty in obtaining training samples and long training time.

(2) In addition to methods based on increasing the depth of convolution, CNN's block-based architecture also encourages learning in a modular manner, thereby making the architecture simpler and easier to understand. The concept of a block as a structural unit will continue to exist and further improve the performance of CNN. So, the specific and localization of 3C detection objects is expected to be better developed.

(3) Deep learning requires a large number of training samples. However, the samples generated by DCGAN summarized in this article still need to be manually labeled, which is time-consuming and labor-intensive. In recent years, supervised learning algorithms have developed rapidly, but it is still unsupervised learning algorithms that really determine the degree of intelligent development. Future research should focus on how to enable machines to automatically learn defect features under unsupervised learning.

(4) Finally, existing defect detection methods are generally carried out on two-dimensional picture samples. Future research can be improved and upgraded on the original basis to realize detection on three-dimensional model, which will enrich and improve the theory and application of CNN. After the application of 3C product testing, the system can directly perform testing on the basis of product parts, which will reduce more inspection time.

REFERENCES

[1] J. Lin, "Current Situation and Development Trend of Mobile Communication Systems," *Journal of Secure Communication and System*, vol. 1, no. 1, pp. 16, 2017.

[2] CY, Analysis and forecast of global and Chinese smartphone shipments in 2020, Industry Information Network, 2021.

[3] R.S. Lu, A. Wu, T.D. Zhang and Y.H. Wang, "A Review of Automatic optical (visual) detection technology and its application in defect detection," *Acta Optica Sinica*, vol. 38, no. 08, pp. 23-58, 2018.

[4] Y. Cohen, M. Faccio, F. Pilati and X. Yao, "Design and management of digital manufacturing and assembly systems in the Industry 4.0 era," *Int. J. Adv. Manuf. Tech.*, vol. 105, no. 9SI, pp. 3565-3577, 2019, doi: 10.1007/s00170-019-04595-0.

[5] J. LeeB. Bagheri and H.A. Kao, "A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems," *Manufacturing Letters*, vol. 3, pp. 18-23, 2015.

[6] H. Lasi, P. Fettke, H.G. Kemper, T. Feld and M. Hoffmann, "Industry 4.0," *Bus. Inform. Syst. Eng.+*, vol. 6, no. 4, pp. 239-242, 2014.

[7] F. Tavakolizadeh, J. Soto, D. Gyulai and C. Beecks, "Industry 4.0: Mining Physical Defects in Production of Surface-Mount Devices,", 2017.

[8] X. Gong, Y. Bai, Y. Liu and H. Mu, "Application of deep learning in defect Detection," *Journal of Physics Conference Series*, vol. 1684, pp. 12030, 2020.

[9] T. Wang, Y. Chen, M. Qiao and H. Snoussi, "A fast and robust convolutional neural network-based defect detection model in product quality control," *Int. J. Adv. Manuf. Tech.*, vol. 94, no. 9-12, pp. 3465-3471, 2018, doi: 10.1007/s00170-017-0882-0.

[10] Z. Liu and B. Qu, "Machine vision based online detection of PCB defect," *Microprocess. Microsy.*, vol. 82, 2021, doi: 10.1016/j.micpro.2020.103807.

[11] C. JianJ. Gao and Y. Ao, "Automatic surface defect detection for mobile phone screen glass based on machine vision," *Appl. Soft Comput.*, vol. 52, pp. 348-358, 2017, doi: 10.1016/j.asoc.2016.10.030.

[12] M. Liu, R. Zhuang, X. Guo and J. Zhao, Application of improved Otsu threshold segmentation algorithm in mobile phone screen defect detection, *Chinese Control and Decision Conference*, 2020, pp. 4919-4924.

[13] R.X. Gao, X.D. Ren, X. W and A.Q. Wei, "Research on mibile phone screen defect defection system based on image processing," *observation and control technology*, vol. 36, no. 04, pp. 26-30, 2017.

[14] J. Zhang, Y. Li, C. Zuo and M. Xing, Defect detection of mobile phone screen based on improved difference image method, *Proc. 2019 4th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*, 2019.

[15] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu and F.E. Alsaadi, "A survey of deep neural network architectures and their applications," *Neurocomputing*, vol. 234, pp. 11-26, 2017, doi: 10.1016/j.neucom.2016.12.038.

[16] N. Ketkar, "Convolutional Neural Networks," *Springer International Publishing*, 2017.

[17] M. Ling, L. Yue, F.N. Xiao, M.L. Yu and Q.J. Hui, "Defect Detection of Mobile Phone Surface Based on Convolution Neural Network," *DEStech Transactions on Computer Science and Engineering*, no. icmsie, 2018.

[18] X. Wei, W. Feng, Q. Lei, G. Gui and W. Wang, Defect Detection of Using Variant CNN in the Processing of Cover Glass, Touch Screen and

Display under Parallel Light, *Proc. 2020 IEEE 6th International Conference on Computer and Communications (ICCC)*, 2020.

[19]   Y. Iwahori, Y. Takada, T. Shiina, Y. Adachi, M.K. Bhuyan and B. Kijsirikul, Defect Classification of Electronic Board Using Dense SIFT and CNN, *Procedia Computer Science*126, R. J. Howlett, et al., eds., 2018, pp. 1673-1682.

[20]   L. Zhang, et al., "Convolutional neural network-based multi-label classification of PCB defects," *JOURNAL OF ENGINEERING-JOE*, no. 16, pp. 1612-1616, 2018, doi: 10.1049/joe.2018.8279.

[21]   S.S. Fan and Y. Chuang, "Automatic detection of Mura defect in TFT-LCD based on regression diagnostics," *Pattern Recogn. Lett.*, vol. 31, no. 15, pp. 2397-2404, 2010, doi: 10.1016/j.patrec.2010.07.013.

[22]   C. Li, X. Zhang, Y. Huang, C. Tang and S. Fatikow, "A novel algorithm for defect extraction and classification of mobile phone screen based on machine vision," *Comput. Ind. Eng.*, vol. 146, 2020, doi: 10.1016/j.cie.2020.106530.

[23]   H. Chen, "CNN-Based Surface Defect Detection of Smartphone Protective Screen," *Journal of Physics: Conference Series*, vol. 1616, no. 1, pp. 12101-12107, 2020.

[24]   S. Jin, J. Chao, C. Yan and J. Xing, "TFT-LCD mura defect detection using DCT and the dual-γ piecewise exponential transform," *Precision Engineering*, vol. 54, pp. S780784521, 2018.

[25]   J. Park, H. Riaz, H. Kim and J. Kim, "Advanced cover glass defect detection and classification based on multi-DNN model," *MANUFACTURING LETTERS*, vol. 23, pp. 53-61, 2020, doi: 10.1016/j.mfglet.2019.12.006.

[26]   B. Zhang, J. Zhou, F. Wang and S. Han, "Detection and Identification of glass defects in touch screen based on Mask R-CNN," *Software Guide*, vol. 18, no. 02, pp. 64-67, 2019.

[27]   Y. Zhu and Y. Jiang, "Optimization of face recognition algorithm based on deep learning multi feature fusion driven by big data," *Image Vision Comput.*, vol. 104, 2020, doi: 10.1016/j.imavis.2020.104023.

[28]   Y. Lecun and L. Bottou, "Gradient-based learning applied to document recognition," *P. Ieee*, vol. 86, no. 11, pp. 2278-2324, 1998.

[29]   R. Wagner, M. Thom, R. Schweiger, G. Palm and A. Rothermel, Learning Convolutional Neural Networks From Few Samples, *IEEE International Joint Conference on Neural Networks (IJCNN)*, 2013.

[30]   A. KrizhevskyI. Sutskever and G.E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Commun. ACM*, vol. 60, no. 6, pp. 84-90, 2017, doi: 10.1145/3065386.

[31]   Y. Wei, H. Xiao, H. Shi, Z. Jie, J. Fe Ng and T.S. Huang, Revisiting Dilated Convolution: A Simple Approach for Weakly- and Semi- Supervised Semantic Segmentation, *Proc. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018.

[32]   B. Zhou, A. Khosla, A. Lapedriza, A. Oliva and A. Torralba, Learning Deep Features for Discriminative Localization, *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2921-2929.

[33]   M. LinQ. Chen and S. Yan, "Network In Network," *Computer Science*, 2013.

[34]   C. Szegedy, et al., Going Deeper with Convolutions, *IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1-9.

[35]   S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *JMLR.org*, 2015.

[36]   C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, Rethinking the Inception Architecture for Computer Vision, *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2818-2826.

[37]   C. Szegedy, S. Ioffe, V. Vanhoucke and A.A. Alemi, Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning, 2017, pp. 4278-4284.

[38]   C. Chen and F. Qi, "Review on Development of Convolutional Neural Network and Its Application in Computer Vision," *Computer Science*, vol. 46, no. 03, pp. 69-79, 2019.

[39]   K. He, X. Zhang, S. Ren and J. Sun, "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition," *IEEE T. Pattern Anal.*, vol. 37, no. 9, pp. 1904-1916, 2015, doi: 10.1109/TPAMI.2015.2389824.

[40]   R. Girshick, "Fast R-CNN," *arXiv e-prints*, 2015.

[41]   A.L. MaasA.Y. Hannun and A.Y. Ng, "Rectifier Nonlinearities Improve Neural Network Acoustic Models.," *Proc.ICML.*, vol. 28, pp. 1-6, 2013.

[42]   K. He, X. Zhang, S. Ren and J. Sun, Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification, *IEEE International Conference on Computer Vision*, 2015, pp. 1026-1034.

[43]   P. RamachandranB. Zoph and Q.V. Le, "Searching for Activation Functions,", 2017.

[44]   C. Gunavathi, K. Sivasubramanian, P. Keerthika and C. Paramasivam, "A review on convolutional neural network based deep learning methods in gene expression data for disease diagnosis," *MATERIALS TODAY-PROCEEDINGS*, vol. 45, no. SI2, pp. 2282-2285, 2021, doi: 10.1016/j.matpr.2020.10.263.

[45]   K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *Computer Science*, 2014.

[46] Gu, et al., "Recent advances in convolutional neural networks," *Pattern Recognition the Journal of the Pattern Recognition Society*, 2018.

[47] M.D. Zeiler and R. Fergus, Visualizing and Understanding Convolutional Networks, *Lecture Notes in Computer Science*8689, D. Fleet, et al., eds., 2014, pp. 818-833.

[48] R.K. SrivastavaK. Greff and J. Schmidhuber, "Highway Networks," *Computer Science*, 2015.

[49] K. He, X. Zhang, S. Ren and J. Sun, Deep Residual Learning for Image Recognition, *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770-778.

[50] S. Zagoruyko and N. Komodakis, "Wide Residual Networks,", 2016.

[51] W. Fe I, M. Jiang, Q. Chen, S. Yang and X. Tang, Residual Attention Network for Image Classification, *Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[52] Z. Deng, Z. Jiang, R. Lan, W. Huang and X. Luo, "Image captioning using DenseNet network and adaptive attention," *Signal Process.-Image*, vol. 85, 2020, doi: 10.1016/j.image.2020.115836.

[53] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. Chen, MobileNetV2: Inverted Residuals and Linear Bottlenecks, *IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510-4520.

[54] K. Sun, Y. Zhao, B. Jiang, T. Cheng and J. Wang, "High-Resolution Representations for Labeling Pixels and Regions,", 2019.

[55] K. Deeba and B. Amutha, "ResNet - deep neural network architecture for leaf disease classification," *Microprocess. Microsy.*, pp. 103364, 2020.

[56] , "Understanding the difficulty of training deep feedforward neural networks," *J. Mach. Learn. Res.*, vol. 9, pp. 249-256, 2010.

[57] Y. LU, Classification and Detection of Mobile Phone Glass Defects Based on Deep Learning., Zhengzhou University, 2019.

[58] L. Liang, D. Li, X. Fu and W. Zhang, "Touch screen defect inspection based on sparse representation in low resolution images," *Multimed. Tools Appl.*, vol. 75, no. 5, pp. 2655-2666, 2016, doi: 10.1007/s11042-015-2559-8.

[59] S. Ozturk and B. Akdemir, Novel BiasFeed Cellular Neural Network Model for Glass Defect Inspection, 2016, pp. 366-371.

[60] Z. Luo, X. Xiao, S. Ge, Q. Ye, S. Zhao and X. Jin, ScratchNet: Detecting the Scratches on Cellphone Screen, *Communications in Computer and Information Science*773, J. Yang, et al., eds., 2017, pp. 178-186.

[61] Y. Jin, D. Zhang, M. Li, Z. Wang and Y. Chen, "A Fuzzy Support Vector Machine-Enhanced Convolutional Neural Network for Recognition of

Glass Defects," *Int. J. Fuzzy Syst.*, vol. 21, no. 6, pp. 1870-1881, 2019, doi: 10.1007/s40815-019-00697-9.

[62] J. Ma, Defect Detection and Recognition of Bare PCB Based on Computer Vision, *Chinese Control Conference*, T. Liu and Q. Zhao, eds., 2017, pp. 11023-11028.

[63] V.A. Adibhatla, J. Shieh, M.F. Abbod, H. Chih, C. Hsu and J. Cheng, Detecting Defects in PCB using Deep Learning via Convolution Neural Networks, *International Microsystems Packaging Assembly and Circuits Technology Conference*, 2018, pp. 202-205.

[64] Y. WangJ. Cao and X. Ji, "PCB defect detection and recognition algorithmbased on convolutional neural network," *Journal of Electronic Measurement and Instrumentation*, 2019.

[65] J. He, K. Xiao, C. Liu and S. Chen, "Defect Detection Method of TFT-LCD Circuit Based on Regional Neural Network," *Computer and Modernization*, vol. 000, no. 007, pp. 33-38, 2018.

[66] L. Wang, L. Luo, P. Zheng, T. Zheng and S. He, A Fast Dent Detection Method for Curved Glass Using Deep Convolutional Neural Network, *Proceedings of the International Conference on Anti-counterfeiting Security and Identification*, 2019, pp. 117-121.

[67] G. Huang, Z. Liu, V. Laurens and K.Q. Weinberger, Densely Connected Convolutional Networks, *Proc. IEEE Computer Society*, 2016.

[68] C. Zhang, W. Shi, X. Li, H. Zhang and H. Liu, "Improved bare PCB defect detection approach based on deep feature learning," *JOURNAL OF ENGINEERING-JOE*, no. 16, pp. 1415-1420, 2018, doi: 10.1049/joe.2018.8275.

[69] L.K. CheongS.A. Suandi and S. Rahman, "Defects and Components Recognition in Printed Circuit Boards Using Convolutional Neural Network,", 2019.

[70] V. Ihar, A. Mujeeb, W. Dai, E. Marius and S. Alexei, Detection Defect in Printed Circuit Boards using Unsupervised Feature Extraction upon Transfer Learning, 2019, pp. 101-108.

[71] B. XiaJ. Cao and C. Wang, SSIM-NET: Real-Time PCB Defect Detection Based on SSIM and MobileNet-V3, *Proc. 2019 2nd World Conference on Mechanical Engineering and Intelligent Manufacturing (WCMEIM)*, 2019.

[72] R. Ding, L. Dai, G. Li and H. Liu, "TDD-net: a tiny defect detection network for printed circuit boards," *CAAI TRANSACTIONS ON INTELLIGENCE TECHNOLOGY*, vol. 4, no. 2, pp. 110-116, 2019, doi: 10.1049/trit.2019.0019.

[73] J. He., Research on Defect Detection of TFT-LCD Line Based on Deep Learning., Xiamen University, 2018.

[74] M. Kim, M. Lee, M. An and H. Lee, "Effective

IEEE *Access*
Multidisciplinary : Rapid Review : Open Access Journal

automatic defect classification process based on CNN with stacking ensemble model for TFT-LCD panel," *J. Intell. Manuf.*, vol. 31, no. 5, pp. 1165-1174, 2020, doi: 10.1007/s10845-019-01502-y.

[75] J. MERTENS and S. Zamir, "The value of two-person zero-sum repeated games with lack of information on both sides," *Core Discussion Papers Rp*, vol. 1, no. 1, pp. 39-64, 1971.

[76] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85-117, 2015, doi: 10.1016/j.neunet.2014.09.003.

[77] X. Tao, D. Zhang, W. Ma, X. Liu and D. Xu, "Automatic Metallic Surface Defect Detection and Recognition with Convolutional Neural Networks," *APPLIED SCIENCES-BASEL*, vol. 8, no. 9, 2018, doi: 10.3390/app8091575.

[78] J. LiJ. Jia and D. Xu, Unsupervised Representation Learning of Image-Based Plant Disease with Deep Convolutional Generative Adversarial Networks, *Chinese Control Conference*, X. Chen and Q. C. Zhao, eds., 2018, pp. 9159-9163.

[79] Z. Yuan, Z. Zhang, H. Su, L. Zhang, F. Shen and F. Zhang, "Vision-Based Defect Detection for Mobile Phone Cover Glass using Deep Neural Networks," *Int. J. Precis. Eng. Man.*, vol. 19, no. 6, pp. 801-810, 2018, doi: 10.1007/s12541-018-0096-x.

[80] Y. LvL. Ma and H. Jiang, A Mobile Phone Screen Cover Glass Defect Detection MODEL Based on Small Samples Learning, 2019, pp. 1055-1059.