Detection of Conductive Particles in TFT-LCD Circuit Using Generative Adversarial Networks

YUANYUAN WANG, LING MA, MINGYUAN JIU, (Member, IEEE), AND HUIQIN JIANG
School of Information Engineering and Digital Medical Image Technique Research Center, Zhengzhou University, Zhengzhou 450001, China
Corresponding author: Ling Ma (ielma@zzu.edu.cn)
This work was supported in part by the NSFC-Henan Joint Fund Key Project under Grant U1604262, and in part by the National Natural Science Foundation of China under Grant 61806180.

ABSTRACT The inspection of conductive particles is a crucial step in the Thin Film Transistor Liquid Crystal Display (TFT-LCD) circuit detection process since only high-quality deformed particles have the conductive effect in the circuit. The main task of detecting conduction particles is to locate and count the valid particles accurately, which is a high challenge due to various difficulties such as the uneven illumination, different sizes to aggregation and overlap between particles, etc. Traditional detection algorithms need to manually set a large number of artificial thresholds, which limits their adaptability. As a result, effective automatic detection of conductive particles is strongly motivated in industry. In this paper, a novel particle detection algorithm based on generative adversarial networks (GAN) is proposed for TFT-LCD circuit inspection system. The backbone architecture of the generator is based on a compact end-to-end neural network with multi-scale convolution blocks for well utilizing the multiscale spatial features. And the discriminator is designed to detect and correct high-order inconsistencies for real-fake images. Moreover, Coarse to Fine training strategy and Loss functions Coordination strategy are further proposed to improve the detection quality. The experiments on the real dataset demonstrate the effectiveness of the proposed methods for the detection of valid conductive particles compared to the state-of-the-art methods.

INDEX TERMS Conductive particles, generative adversarial networks, TFT-LCD, U-MultiNet.

I. INTRODUCTION
Thin Film Transistor Liquid Crystal Display (TFT-LCD) is an indispensable component in smartphones, LCD TVs and various vision displays. The circuit quality inspection of TFT-LCD is an important step to ensure the production of qualified electronic products [1]. The key of TFT-LCD circuit detection is the inspection of conductive particles. With the development of large-size and high-resolution TFT-LCD [2], it is necessary to study high-speed and high-precision effective particle detection approach.

Conducting electricity between Integrated Circuit driver (IC) and glass substrate is worked by conductive particles in the Anisotropic Conductive Film (ACF), which is a crucial step in TFT-LCD manufacturing processes [3], [4]. Fig. 1 shows the conducting electricity schematic in the ACF bonding process. The connection between IC and glass substrate depends on the number of high-quality deformed particles in pad images [5], [6], as shown in Fig. 1(b). Fig. 2 presents some examples of particles with different deformation quality. Generally, the high-quality deformed particle in insulated pad images has a bright part and a dark part when projecting tilted light, and its edges are smooth and full, as shown in Fig. 2(a). Since the conditions of ACF bonding are critical, such as temperature, time, pressure and alignment of the parts, as shown in Fig. 2(b) and (c), the particles may have invalid deformation and lost the conductive effect if any of them is not satisfied [6]–[8].

There are thousands of insulated pads with particles in the circuit of TFT-LCD. By locating and counting high-quality deformed particles in each pad, the conductivity of TFT-LCD circuit can be checked. If the number of validly deformed particles in the pad is insufficient, it may cause poor conductivity or electrical failure, which will lead to inferior products and material waste. To ensure the conductivity quality of the circuit, conductive particles detection is an indispensable step for TFT-LCD. However, particles inspection in insulated pads is very challenging due to the various size, clustering, uneven
Y. Wang et al.: Detection of Conductive Particles in TFT-LCD Circuit Using GAN

FIGURE 1. Conducting electricity schematic in the ACF bonding process. (a) IC and glass are electrically interconnected by validly deformed particles. (b) The image of the insulated pad with particles sandwiched between the gold bump of IC and the glass substrate electrode. Best viewed in color.

The output of the generator is constrained by the designed discriminator, which makes the prediction result closer to the real data;

ii) A Coarse to Fine training strategy (CF) is used to improve the accuracy of valid particle location and counting, and Loss functions Coordination training method (LC) is proposed to balance the performance of generator and discriminator, so as to keep the training process stable;

iii) A particle dataset including normal segmentation samples and empty samples is also built. Normal segmentation samples are divided into two types in different scales: particle global segmentation samples and particle core segmentation samples.

The rest of the paper is organized as follows: Section II briefly presents the related works. And Section III describes the proposed detection framework and its associating learning algorithm. The experimental results are shown in Section IV and finally followed by the conclusion in Section V.

II. RELATED WORKS

Particle inspection aims to estimate the distribution and quantity of validly deformed particles and then to determine whether a good conductive circuit is formed between IC and glass substrate [7], [9]. Many works have devoted to the research of automatic particle detection [7], [9]–[11]. These works are reviewed briefly as follows.

Traditional computer vision-based methods handle this problem according to the assumption that the conductive particles have specific patterns. For instance, a standard template of the validly deformed particle is constructed and template matching [7] is applied to detect the presence of the particles. In the work of Lin et al. [9], the particle features are extracted with the Prewitt operators. Then Otsu algorithm [12] is used to obtain the binarized image. After binarizing the image with Otsu, the effective particles are identified by template matching method. Inspired by the approach [9], Yu-ye et al. [10] acquire particle features with the background mask method rather than Prewitt operators, which reduces background noise. Then K-means clustering algorithm is applied to classify the particle pixels, which is more accurate for the overlapping particles. Different from [9], [10], in the method of Ni et al. [11], the particles are inspected in 3D space. Although the features are more discriminative, they are still manually extracted and not learned. In short, traditional methods based on handcraft features usually have the following problems:

1) The precision of detection is sensitive to environmental factors such as uneven illumination and noise;

2) The detection algorithms rely on artificially designed rules with many adjusted parameters, which limits their adaptability in industrial inspection;

illumination and overlap of particles in the pad, as show in Fig. 3. Besides, the pad images with particles in micron-scale are easily blurred due out of focus. Therefore, an effective and robust automatic particle detection method is urgent to be addressed for TFT-LCD industry.

In this paper, a framework based on GAN is proposed to detect conductive particles validly, the contributions are summed as follows:

i) A compact U-shaped generator network with multi-scale convolution blocks (U-MultiNet) is proposed. The architecture of U-MultiNet is inspired by U-NET, and the multi-scale convolution block is inspired by Inception module.

The experimental results are shown in Section IV and finally followed by the conclusion in Section V.

VOLUME 8, 2020
3) Extracting only simple features can’t well describe the characteristics of the detected target.

Recently, deep learning-based methods have obtained significant improvements for various tasks in computer vision and pattern recognition, for instance, object classification [13]–[15] and detection [16]–[18], image segmentation [19]–[21], etc. They make use of deep architecture of network to learn more semantic and also discriminative features compared to handcraft features. The well-known network is convolutional neural network and its variants. The convolutional neural network designed by Krizhevsky et al. [13] was first applied in ImageNet competition and won the first place in the image classification. In [14], GoogLeNet extracts multi-scale features with the Inception module to improve the classification accuracy. Unlike GoogLeNet, VGGNet [15] improves performance by increasing network depth, but also increases a large number of parameters. Later, the fully convolutional network (FCN) based on VGGNet was proposed by Long et al. [19]. The authors substitute convolution layers for full connection layers and add skip connections, which realizes pixelwise image segmentation. And the FCN framework has many variants such as SegNet [20], U-Net [21], and PSPNet [22]. In [21], U-Net obtains fine segmentation accuracy utilizing skip-layer connection between down-sampling and up-sampling layers. In addition, some tricks, such as conditional random field (CRF) and atrous convolution [23]–[26], are also proposed to improve segmentation performance. As Generative Adversarial Network (GAN) designed by Goodfellow et al. [27] is becoming popular, some GAN based segmentation algorithms, such as [28]–[30], are proposed. These algorithms achieve high accuracy by combining high-frequency and low-frequency information in the loss of generator and discriminator.

Since conductive particle in the image is small target, usually tens of pixels, which is trivial for segmentation task, the method of segmentation based on pixelwise classification is an effective fashion for particle detection. In [31], the segmentation network is used to detect conductive particles. Liu et al. [31] transform the target detection into pixel regression problem, and adopt U-shaped network [21] with residual blocks [32] to solve the counting task of particles, which achieves better results than the traditional methods.

Although existing deep learning methods outperform traditional methods for particles detection task, but there are still some problems. For instance, the existing methods tend to miss a small number of valid particles in particle aggregation regions and may even output false-positive particles in serious uneven illumination regions of the background. In addition, the predicted particle sometimes deviates from the corresponding real position. These are mainly due to the fact that the objective function of CNNs used in the existing methods relies only on the pixel-wise classification loss functions comparing prediction images with annotation images. This is not desirable since it can’t well integrate the structure of validly deformed particle that formed naturally under hot pressing conditions. In order to solve these issues, this paper proposes a novel particle detection algorithm based on GANs.

III. PROPOSED ALGORITHM

In the automatic inspection line of valid particles, thousands of pads in the TFT-LCD should be detected in real time by a prediction network. However, huge deep convolution networks like ResNet [32] are too resource-consuming for real-time inference, especially on an industrial control platform. In this paper, a novel compact network based on GANs including a generator (G) and a discriminator (D) is proposed.
as shown in Fig. 4. The generator is a lightweight end-to-end segmentation network, which is used to generate particle prediction maps for given image, and the discriminator performs the binary classification between real and fake map pairs and optimizes the segmentation parameters without specifying a tailored loss function.

A. U-MULTINET BASED GENERATOR

U-Net [21] is a successful segmentation network, which consists of two paths, a down-sampling path and an up-sampling path, which are connected through skip paths. However, it is not appropriate to directly be applied for particle detection in the assembly line due to the following problems: i) the conductive particles are very small and the max radius is about 5 pixels, the down-sampling path with 4 max pooling operations in U-Net may lose the particle characteristics during the feature extraction process, referring to an example of down-sampled features as shown in Fig. 5; ii) the amount of information carried by small targets is limited, and usually, low-level features become the main detection basis. The deep layer of U-Net is less helpful than shallow layer for small target segmentation, but introduces much more parameters to be learned and increases detection time cost, which is not practical in the assembly line with high speed. Therefore, this paper proposes U-shaped multi-scale convolution network (i.e., U-MultiNet) with a contracting path and an expansive path, as shown in Fig. 6.

The proposed U-MultiNet differs from U-Net in the respect of lightweight structure and the used of multi-scale convolution blocks:

1) LIGHTWEIGHT NETWORK STRUCTURE

The architecture of U-MultiNet is a lightweight network structure, as shown in Fig. 6. In the contracting path, each layer consists of a multi-scale convolution block and a 2 × 2 max pooling operation. In the expansive path, each layer consists of an up-sampling operation, a concatenation and a multi-scale convolution block. U-MultiNet merges particle features from different levels by the skip connection. Compared with U-Net, U-MultiNet reduces the amount of max pooling operations, which helps to avoid the loss of small particle features. Moreover, U-MultiNet reduces the depth of the network and has fewer parameters, which makes it more suitable for high-speed assembly lines.

2) MULTI-SCALE CONVOLUTION BLOCK

The multi-scale convolution block consists of three 3 × 3 convolutions with stride 1, and a concatenation with feature maps from the three 3 × 3 convolutions. Each convolution is followed by a batch normalization [33] and a ReLU [34]. The 2nd and 3rd 3 × 3 convolutions of the multi-scale convolution block effectively approximate the 5 × 5 and 7 × 7 convolution operations, respectively. Therefore, the outputs of the three convolutions are concatenated, extracting the fine multiscale spatial features of conductive particles in each level.

B. DISCRIMINATOR

As an important part of GAN, the discriminator helps to improve the performance of generator. The structure details of the proposed discriminator are shown in Fig. 7. The pad image and corresponding prediction are strongly correlated in
C. LOSS FUNCTIONS

In many segmentation tasks, see e.g., [29], [30], [35], the binary cross-entropy is generally used as the loss function for the segmentation model. Similarly, the binary cross-entropy can be also adopted for valid particle segmentation. For each pad image, the corresponding label map is \( y \), and the map of the segmentation network output is \( \tilde{y} \). Therefore, for valid particle segmentation, the binary cross-entropy is generally used as the loss function. The binary cross-entropy loss is defined as:

\[
L_{SEG}(y, \tilde{y}) = - \sum_{i=1}^{H \times W} (y_i \log(\tilde{y}_i) + (1 - y_i) \log(1 - \tilde{y}_i))
\]

(1)

For particle segmentation based on GAN, the discriminator \( D \) learns a mapping from a pair of \( \{x, \tilde{y}\} \) or \( \{x, y\} \) to binary classification \([0,1]\) where 0 or 1 means that the input data of \( D \) is generator-predicted map pair or human-annotated map pair. The loss relevant to discriminator \( D \) is defined as:

\[
L_D(G, D) = - \left[ \log(D(x, y)) + \log(1 - D(x, G(x))) \right]
\]

(2)

where \( G(x) \) is the prediction map of generator \( G \). For training the discriminator \( D \) to make correct judgment, \( L_D(G, D) \) needs to be minimized, which means that \( D(x, y) \) needs to be maximized while \( D(x, G(x)) \) should be minimized.

The generator \( G \) learns a mapping from pad images to the predicted particles and generates the prediction that cannot be distinguished from the human-annotated particles. Therefore, the loss function of generator based on U-MultiNet is the weighted sum of two terms. The first term is cross entropy term, which encourages the generator to predict the correct category for each individual pixel. Another is the auxiliary adversarial loss term based on discriminator \( D \), which has a field-of-view that is the entire image. Hence, mismatches in the spatial high-order features can be penalized by the adversarial loss term. The combination of cross entropy and adversarial loss is used to train U-MultiNet based generator, which makes generator-predicted maps constrained to approximate human-annotated maps at the pixel and semantic levels. The cross-entropy loss term is defined as:

\[
L_{G1}(G) = L_{SEG}(y, G(x))
\]

(3)

And the adversarial loss term is defined as:

\[
L_{G2}(G, D) = \arg\min_G \max_D [-\log D(x, G(x))]
\]

(4)

The loss function of U-MultiNet based generator is defined as:

\[
L_G(G, D) = L_{G2}(G, D) + \lambda L_{G1}(G)
\]

where \( \lambda \) is weight coefficients. For training generator \( G \) to make precise prediction, \( L_G \) needs to be minimized, which means that \( L_{G1}(G) \) needs to be minimized while \( D(x, G(x)) \) should be maximized. In GAN framework, generator and discriminator are alternately trained in the adversarial form to improve their respective performance.

D. LOSS FUNCTIONS COORDINATION

TRAINING STRATEGY

The training process of GAN model may be unstable, so maintaining a balanced adversarial training process is very important to improve the performance of networks [36]. For the particle segmentation, the generator needs to make prediction pixel by pixel and output dense predicted maps, but the prediction of discriminator is only a true-false probability value of the image pair. Hence, the convergence speed of the discriminator is faster than that of the generator. When the performance of discriminator far exceeds that of generator in the later stage of training, discriminator will produce relatively uninformative gradients for the generator [36]. Adjusting the corresponding weight of the gradient is also meaningless.

Therefore, in this paper, Loss functions Coordination training (LC) is proposed to accelerate the optimization of generator:

Step 1: Only the loss \( L_{SEG} \) is first used to adjust the parameters of generator for accelerating convergence, reducing the performance gap in subsequent adversarial training between generator and discriminator.

Step 2: The generator trained by \( L_{SEG} \) is loaded into GAN framework. Then the loss \( L_G \) is used to fine-tune and further optimize the parameters of generator during the adversarial training. Meanwhile, the loss \( L_D \) is adopted to optimize the discriminator.

Step 3: Repeat the above two steps until the convergence of the performance of generator.

In LC training, \( L_{SEG} \) trains the generator with \( M \) epochs to accelerate convergence, and then \( L_G \) trains the generator with one epoch to play a role of regularization. The two loss functions train the generator alternately to update its parameters. This strategy not only balances the performance of generator and discriminator, making the whole framework training more stable, but also improve the segmentation performance of generator. In the following experiments, \( M \) is empirically set to be three, as validated in the training curves of LC and also their performance comparisons.
TABLE 1. Detailed information of particle datasets.

| Dataset     | Type of sample | Image Size | Total set count | Training set count | Validation set count | Test set count |
|-------------|----------------|------------|-----------------|--------------------|----------------------|---------------|
| Global Dataset | Positive sample | 48 x 112 | 5870            | 3756               | 940                  | 1174          |
|              | Negative sample |           | 2160            | 1370               | 340                  | 450           |
| Core Dataset  | Positive sample |           | 5870            | 3756               | 940                  | 1174          |
|              | Negative sample | 48 x 112 | 2210            | 1390               | 350                  | 470           |

E. COARSE TO FINE TRAINING STRATEGY

Generally, it is not sufficient to detect valid particles by identifying a small amount of high-frequency information composed of bright and dark parts in images that are even difficult for human eyes to see clearly. The convolution network is very impressionable to the suspected small target. The fragments of broken particles caused by over extrusion can be easily identified as false-positive particles, especially in pad image with dim light. To solve this problem, we propose Coarse to Fine training strategy (CF) composed of two stages.

In the first stage, particle global labels are used for training convolution networks. Global labels covering the whole area of valid particles, including the smooth boundary, are made by annotating only high-quality deformed particles. This type of label enables the convolution networks to make full use of the characteristics of high-quality deformed particles to identify the pixels on valid particles. Convolution networks are firstly trained by global labels, which is equivalent to pre-training the network parameters, so that the generator has a good ability to represent the contents of valid particles, including morphological and location features. The pre-trained network is served as the start point for the second stage.

In the second stage, particle core labels are obtained by covering the central area of valid particles and then are used to train the convolutional network. By accurate boundary of valid particles, the network is fine-tuned to learn more discriminative features, resulting to better prediction performance, as validated in the experimental section. The prediction results of CF are shown in Fig. 11. (The training curves of CF are shown in Section IV(A). And Section IV(B) quantitatively evaluates the effectiveness of CF.)

IV. EXPERIMENTS

The proposed network is employed to the valid particle detection task on the particle dataset, which is collected from the assembly line.

The pad images with particles are obtained from the image of TFT-LCD circuit taken by a line scan camera and the size of pad image is 48 x 112. This paper respectively generates two datasets (i.e. particle global dataset and particle core dataset) for the pad image. The particle global labels are manually annotated by a solid circle with a radius of 5 pixels that concentrically cover validly deformed particles in the image, and particle core labels are the pixel-level annotations of particle center located in the center area of valid particles. The instances of pad images and their labels are shown in Fig. 8.

In addition, negative samples containing only the background are also generated (two examples shown in Fig. 8). These negative samples help to improve the discrimination ability of the convolutional network.

A total of 1174 positive samples and 874 negative samples are collected in our raw dataset. After data augmentation through horizontal flip, vertical flip and rotation, there are a total of 5870 positive samples and 4370 negative samples, of which 1174 positive samples and 920 negative samples are randomly selected as the test set. The detailed data distribution is shown in Table 1.

A. TRAINING OUTPUT

Here we investigate the loss w.r.t. the iteration for two components of the GAN in LC and non-LC cases, where Adam [37] optimizer was adopted in the stochastic gradient descent. The global segmentation phase took about an hour and twenty minutes for 30 epochs. The learning rate was 2 x 10^-4. The core segmentation phase lasted about an hour and ten minutes for 30 epochs. The learning rate was reduced to 1 x 10^-4. The loss curves of two training stages are shown in Fig. 9 and Fig. 10 respectively.

In the first-stage, the loss $L_D$ declined rapidly in the first 3 epochs, which indicated the performance of $D$ improved quickly in the training by global labels. After that, the loss $L_D$ tended to converge. The auxiliary loss $L_{G2}$, which was affected by the performance of $D$, increased sharply in the first 3 epochs. The loss $L_{G1}$ was also in the state of decline in the first 11 epochs, which showed the performance of $G$ was improved. But the decline of $L_{G1}$ was much slower than that of the loss $L_D$. It could be seen from Fig. 9(a) and (c) that the optimization speed of generator was much slower than that of discriminator. A large performance gap would result in loss gradients that were meaningless to the generator. For this problem, LC was adopted to coordinate the training. As shown in Fig. 9(c), the decrease speed of $L_{G1(LC)}$ was significantly faster than that of $L_{G1}$. At the same time in Fig. 9(b), compared with $L_{G2}$, $L_{G2(LC)}$ was also lower, which indicated that the output of generator was closer to the real data about the outline of valid particles.

In the second-stage training, the effectiveness of LC was more obvious. As shown in Fig. 10(b) and (c), compared with $L_{G2}$ and $L_{G1}$, both $L_{G2(LC)}$ and $L_{G1(LC)}$ were significantly lower, which indicated that the prediction of generator was closer to the real data in terms of both the overall
FIGURE 8. Examples of Coarse to Fine training samples.

FIGURE 9. Loss curves in non-LC and LC cases during the first-stage training. (a) Loss curves of $L_D$ and $L_{D(LC)}$; (b) loss curves of $L_{G2}$ and $L_{G2(LC)}$; (c) loss curves of $L_{G1}$ and $L_{G1(LC)}$. Best viewed in color.

FIGURE 10. Loss curves in non-LC and LC cases during the second-stage training. (a) Loss curves of $L_D$ and $L_{D(LC)}$; (b) loss curves of $L_{G2}$ and $L_{G2(LC)}$; (c) loss curves of $L_{G1}$ and $L_{G1(LC)}$. Best viewed in color.
level and pixel level. In addition, the loss $L_{D}(LC)$ had also increased overall. As shown in Fig. 10(a), $L_{D}(LC)$ curve was higher than $L_{D}$. This also illustrated that the prediction of generator is more similar to the real data under the LC strategy, which made the discriminator produce some confusion. (Quantitative comparison about LC strategy is detailed in Section IV(B).)

Fig. 11 illustrates the prediction results during Coarse to Fine training. In the first stage, the segmentation network focused on the comprehensive recognition of various information (including the blurred boundary) about validly deformed particles, which was helpful to eliminate the interference of non-particles with high similarity. As shown on the left of Fig. 11, the generator recognized the outline and position information of valid particles. The predicted conductive particles became clear and their contour and position information were gradually determined, but most particles still adhered to each other, which affected the accurate counting of valid particles.

Based on the training results in the first stage, the parameters of generator were further fine-tuned by the characteristics of the center region of valid particles. From the images of the fifth and sixth columns in Fig. 11, it could be seen that the network gradually segmented the particle core only from the global contour of the identified valid particles, which reduced the number of false-positive particles predicted in other regions. This also further improved the accuracy of networks. In the second stage, as the centers of conductive particles were continuously strengthened and the fuzzy boundaries were gradually weakened, the particles that conglutinated with each other were effectively separated, which greatly improved the counting accuracy of conductive particles in subsequent operations.

B. TRAINING STRATEGIES

This section investigates the impact of the training strategies in the training procedure. The network with “GAN” suffix indicates that it is loaded into the GAN framework for training. And it should be noted that LC strategy is only used in GAN models. The training for the standard image segmentation usually uses only pixel-level labeled samples. In comparison with the proposed CF strategy, Single-Stage training (SS) is experimented, where only particle core labels are involved in the training procedure. SS+LC indicates that LC is adopted to coordinate the training during the Single-Stage training. In addition, CF strategy consists of two stages, that is, pre-training with particle global labels, and then core-strengthening training with particle core labels. CF+LC means that LC is used to coordinate the training during the CF training.
TABLE 2. mIoU of different training strategies for the core area of valid particle.

| Methods         | SS   | SS+LC | CF   | CF+LC |
|-----------------|------|-------|------|-------|
| U-Net           | 0.4165 | -     | 0.4575 | -     |
| U-MultiNet      | 0.4518 | -     | 0.4694 | -     |
| U-Net GAN       | 0.4831 | 0.4926 | 0.4981 | 0.5203 |
| U-MultiNet GAN  | 0.5046 | 0.5125 | 0.5175 | 0.5297 |

In order to compare the performance of different methods, mean Intersection-over-Union (mIoU) is calculated to measure the prediction results quantitatively, which is defined as:

\[
mIoU = \frac{1}{N} \sum A_{\text{truth}} \cap A_{\text{result}} \quad (6)
\]

where \(N\) is the number of images. In Equations (6), \(A_{\text{truth}}\) is the particle core area in ground truth and \(A_{\text{result}}\) is the particle core area in segmentation result, as shown in Fig. 12. Higher value gives better performance. The mIoU of different training strategies is shown in Table 2.

In order to evaluate training strategies more visually, the overall distributions of IoU of these strategies are shown in Fig. 13.

1) COARSE TO FINE TRAINING STRATEGY
Compared with SS, CF strategy raises mIoU of output results, which can be obtained from the data in Table 2. Global labels make the networks focus on the overall positioning of validly deformed particles. Without global labels to train the networks, segmentation network may output false particle centers in the non-particle area, which will reduce the accuracy of particle center positioning. The segmentation of particle core is a higher-order problem of the global positioning, and the pre-training by global labels induce the networks to approach a better extreme point in particle core segmentation.

2) GAN TRAINING
Comparing GAN models with corresponding non-GAN models, it can be seen from Fig. 13 and Table 2 that the adversarial training under GAN framework could improve mIoU of the output. This is because discriminator compares predicted results with real labels from the specific perspective of convolution network itself, which is impossible to achieved by other artificial loss functions. Moreover, the loss gradient of discriminator also plays a role of regularization, which restrains over-fitting of the segmentation network and enhances its generalization ability.

3) LC STRATEGY
For the specific task of particle detection, LC strategy is used in GAN framework to balance the performance of generator and discriminator, making the training process more stable. From the data in Table 2, LC could further improve mIoU, which shows that the strategy is effective.

In addition, it can be seen from Table 2 and Fig. 13 that the prediction results of U-MultiNet are more accurate and stable than U-Net, which shows that U-MultiNet is more suitable for particle prediction tasks.

C. COMPARISON OF DIFFERENT METHODS

1) QUANTITATIVE COMPARISON
   a: PRECISION AND RECALL
   Precision and Recall are defined as:

\[
\text{Precision} = \frac{|W \cap W'|}{|W'|} \times 100\% \quad (7)
\]

\[
\text{Recall} = \frac{|W \cap W'|}{|W|} \times 100\% \quad (8)
\]

where \(W\) is the set of valid particles on the label image, \(W'\) is the set of predicted particles on the output map, and \(|W \cap W'\)| is the number of localized valid particles correctly. To determine the value of \(|W \cap W'\)|, the distance between the center position \((i', j')\) of each element in \(W'\) and the center position \((i, j)\) of the corresponding nearest neighbor element in \(W\) is measured by the Euclidean distance \(D\). The \(D\) is defined as:

\[
D = \sqrt{(i' - i)^2 + (j' - j)^2} \quad (9)
\]
TABLE 3. Performance comparison.

| Methods               | Precision | Recall | Average Detection Time |
|-----------------------|-----------|--------|------------------------|
| Ostu                  | 16.87%    | 20.09% | 1.08ms                 |
| Lin [9]               | 84.2%     | 32.1%  | 9.01ms                 |
| FCN [19]              | 84.34%    | 76.84% | 2.51ms                 |
| U-ResNet [31]         | 92.62%    | 89.64% | 2.57ms                 |
| U-Net [21]            | 95.17%    | 91.89% | 2.52ms                 |
| U-MultiNet            | 95.20%    | 95.56% | 2.50ms                 |

Models with CF training:

| U-Net                  | 95.32%    | 92.29% | 2.52ms                 |
| U-MultiNet             | 95.46%    | 93.74% | 2.50ms                 |
| U-Net_GAN              | 95.43%    | 94.39% | 2.52ms                 |
| U-MultiNet_GAN         | 95.83%    | 94.63% | 2.50ms                 |

Models with CF and LC training:

| U-Net_GAN             | 95.57%    | 95.15% | 2.52ms                 |
| U-MultiNet_GAN        | 96.44%    | 95.41% | 2.50ms                 |

Since the max radius of the conductive particles in our dataset is about 5 pixels, the tolerance of $D$ is 5.

$b$: AVERAGE DETECTION TIME

The algorithm is implemented in TensorFlow, and the experiments are conducted with a workstation with 3.2GHz i7-8700 CPU, 8G RAM and GeForce GTX 1060 GPU. The average detection time is defined as:

$$\text{Average Time} = \frac{1}{N} \sum_{i=1}^{N} t_i$$  \hspace{1cm} (10)

where $t_i$ is the detection time of the $i$-th pad image. It is worth noting that the detection procedure of the proposed method can be accelerated by parallel computation of GPU, while the traditional methods cannot make use of this advantage.

The results in terms of Precision, Recall and Average detection time are shown in Table 3. From the results, the following summaries can be obtained:

(1) For the specific task of particle detection, the methods based on deep learning are superior to the above traditional methods in both Precision and Recall.

(2) The performance of U-MultiNet is more prominent than U-Net, and the average detection time of pad image by using GPU reduces to 2.5ms.

(3) In Table 3, comparing U-Net with U-Net$_{(CF)}$ or U-MultiNet with U-MultiNet$_{(CF)}$, Precision and Recall of models are improved by adopting the Coarse to Fine training strategy, which shows that CF is effective for improving the detection accuracy of valid particles. The pre-training by global labels could lead generator networks to reach a better extreme point in fine segmentation.

(4) The GAN model outperforms the non-GAN model in term of Precision and Recall. In adversarial training, discriminator introduces the high-order difference between prediction and real label into the parameter optimization process of generator from the perspective of convolutional network itself, further improving the performance of generator in detecting valid particles. In addition, the introduction of discriminator does not increase the time cost during detection phase.

(5) Using LC strategy in GAN framework can alleviate the performance gap between generator and discriminator, making the training in a relatively stable state, which is more conducive to improve the performance of generator. It can be obtained from the data in Table 3 that the LC strategy is effective for the particle detection task.

(6) It can be seen from the results of Precision and Recall that U-MultiNet$_{(CF+LC)}$ performs best, indicating that U-MultiNet$_{(CF+LC)}$ is more suitable for the validly deformed particles detection task.

2) QUALITATIVE COMPARISON IN VISUAL EFFECTS

Here the performance of different methods is compared, including traditional methods and some deep learning methods, for instance, FCN, U-Net, U-MultiNet, U-Net$_{(DS+LC)}$, etc. The detection results of different methods are shown in Fig. 14.

The traditional methods, including Lin et al. [9], could not end up with stable detection effect robustly. Due to the uneven light distribution and diversified particle appearance in pad images, coupled with these algorithms being very sensitive to noise, the detection results are very unstable.

FCN, U-Net, U-MultiNet, U-Net$_{(CF+LC)}$ and the proposed U-MultiNet$_{(CF+LC)}$ achieve better detection than the above traditional methods with much less false and missed particles. Due to its rough up-sampling form, valid particles predicted by FCN has certain position deviation. Different from FCN, U-Net and U-MultiNet improve detail accuracy by merging shallow and deep features by skipping paths. However, the aggregation and overlap of invalid particle fragments can lead to much detection error by encoder-decoder based networks, including FCN and U-Net. Generator $G$ in GAN could detect valid particles more accurately. This is because the additional conditions from discriminator $D$ can improve the detection quality by identifying the high-order differences between the prediction and real data.

Compared to all the other algorithms, the proposed U-MultiNet$_{(CF+LC)}$ has the best particle detection effects as shown in Fig. 14. Actually, under the CF and
FIGURE 14. Examples of conductive particles detection results. from left to right: Pad Image, Ground Truth, Lin [9], FCN [19], U-Net [21], U-MultiNet, U-Net GAN(CF+LC), U-MultiNet GAN(CF+LC). Best viewed in color.

FIGURE 15. Detailed comparison of particle detection effects for the training strategies. Best viewed in color.

LC strategies proposed in this paper, adopting U-MultiNet with multi-scale convolution block as the generator could improve the performance of extracting the multi-scale spatial features of particles. Since taking the discriminator, U-MultiNet GAN(CF+LC) could accurately locate valid particles in complex background noise. These lead to good valid particle detection effects.

In addition, several detection examples of U-Net and U-MultiNet are shown in Fig. 15 and demonstrate the details. It can be seen that: in Fig. 15(a), the particles pointed by blue arrows are incorrectly detected by U-Net, but U-Net GAN(CF+LC) correct this error as shown in Fig. 15(c). In Fig. 15(b), the valid particles pointed by red arrows are missed by U-Net, but detected by U-Net GAN(CF+LC) as shown in Fig. 15(d). These indicate that U-Net GAN(CF+LC) has better performance than U-Net. For U-MultiNet, the proposed training strategies also have similar effects. As shown in Fig. 15(e) and (f), the particles pointed by red arrows are not detected by U-MultiNet, but U-MultiNet GAN(CF+LC) detects them in (g) and (h). All of these indicate that the training strategies proposed in this paper are effective. In addition, compared with U-Net GAN(CF+LC), U-MultiNet GAN(CF+LC) has stronger detection capability, which indicates that it is more suitable for the detection task of validly deformed particles.

V. CONCLUSION

In this paper, we propose a new detection framework based on generative adversarial networks (GAN), named as U-MultiNet GAN(CF+LC), for deformed particle detection task. Firstly, a compact U-shaped multi-scale segmentation network (U-MultiNet) is adopted as the generator, which achieves higher detection accuracy with less time by using the lightweight structure and multi-scale feature extraction. Secondly, Coarse to Fine training strategy is proposed for adversarial training, which not only enhances the accuracy of locating valid particle, but also effectively solves the problem of particle aggregation and overlap. Furthermore, LC strategy is adopted to optimize the parameters of generator to balance the performance of generator and discriminator. Extensive
experiments are performed and the results demonstrate that it obtained the best results in terms of precision and recall compared to other methods. More importantly, real-time detection is achieved for the proposed method, showing its effectiveness of detection of conductive particles in TFT-LCD circuit.

REFERENCES

[1] J.-H. Kim, S. Ahn, J. Wook Jeon, and J.-E. Byun, “A high-speed high-resolution vision system for the inspection of TFT LCD,” in Proc. IEEE Int. Symp. Ind. Electron. (ISIE), Jun. 2001, pp. 101–105.

[2] S.-C. Chang, “The TFT-LCD industry in taiwan: Competitive advantages and future developments,” Technol. Soc., vol. 27, no. 2, pp. 199–215, Dec. 2005.

[3] R. S. Lu, Y. Q. Shi, Q. Li, and Q. P. Yu, “AIO techniques for surface defect inspection,” Appl. Mech. Mater., vol. 36, pp. 297–302, Oct. 2010.

[4] S.-C. Kim and Y.-H. Kim, “Review paper: Flip chip bonding with anisotropic conductive film (ACF) and nonconductive adhesive (NCA),” Current Appl. Phys., vol. 13, pp. S14–S25, Jul. 2013.

[5] Y.-W. Yen and C.-Y. Lee, “ACF particle distribution in COG process,” Microelectron. Rel., vol. 51, no. 3, pp. 676–684, Mar. 2011.

[6] N. Ibtehaz and M. S. Rahman, “MultiResUNet : Rethinking the U-Net architecture for multimodal biomedical image segmentation,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 2672–2680.

[7] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, “Pyramid scene parsing network,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 2881–2890.

[8] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 4, pp. 834–848, Apr. 2018.

[9] J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in Proc. Adv. Neural Inf. Process. Syst., 2014, pp. 2672–2680.

[10] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 770–778.

[11] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, “Pyramid scene parsing network,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 2881–2890.

[12] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 4, pp. 834–848, Apr. 2018.

[13] J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in Proc. Adv. Neural Inf. Process. Syst., 2014, pp. 2672–2680.

[14] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 770–778.

[15] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, “Pyramid scene parsing network,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 2881–2890.

[16] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 4, pp. 834–848, Apr. 2018.

[17] J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in Proc. Adv. Neural Inf. Process. Syst., 2014, pp. 2672–2680.
MINGYUAN JIU (Member, IEEE) received the M.Sc. degree in information and communication engineering from the Harbin Institute of Technology, China, in 2010, and the Ph.D. degree in information from the Institut National des Sciences Appliquees de Lyon, in 2014. He was a Postdoctoral Research Fellow with Telecom ParisTech, Universite Paris-Saclay, from 2014 to 2015, and with Ecole Normale Superieure de Lyon, from 2016 to 2017. He currently is an Assistant Professor with Zhengzhou University, China. His research interests include image processing and analysis, pattern recognition, and machine learning.

HUIQIN JIANG received the M.S. degree in mathematics from Zhengzhou University, China, and the Ph.D. degree in information science from Chiba University, Japan, in 1988 and 2004, respectively. She worked as a Research Fellow with TERARECON Inc., USA, and RealVision Inc., Japan. She is currently a Professor with Zhengzhou University, China. Her current research interests are medical image diagnosis systems, three-dimensional image reconstruction, and machine learning.