

A Novel Hierarchical Tree-DCNN Structure for Unbalanced Data Diagnosis in Microelectronic Manufacturing Process

Yong Zeng, Yanfang Mei, *Member, IEEE*, Yueming Hu, Zhengguo Sheng, *Senior Member, IEEE*

Abstract—The quality of flexible integrated circuit substrates (FICS) is critical to the reliability of various electronic products, making intelligent defect measurement essential for efficient manufacturing and cost-saving. However, existing solutions for substrate defect diagnosis heavily rely on human visual interpretation, which leads to poor efficiency and a high error rate. A novel vision-based detection system consisting of a multi-scale imaging module and a hierarchical structure based on the deep convolution neural network (DCNN) is proposed in this paper. Rapid and accurate fault diagnosis can be enabled for high-density FICS, and various defects could be located and classified in a coarse-to-fine resolution. Specifically, a new mechanism of hierarchical decision based on DCNNs is built for FICS fault diagnosis, wherein the challenge of unbalanced data is addressed in the network learning process to reach a good trade-off between detection accuracy and speed. The substantial experiments and effectiveness comparison by using the typical methods on three categories of FICS and their corresponding eight-type faults reveal that the proposed system could facilitate the solution of substrate fault measurement and achieve high accuracy and efficiency, which could provide essential information of FICS to divide its industrial acceptance quality level.

Index Terms—Integrated circuit image, fault diagnosis, hierarchical structure, deep learning, image processing.

I. INTRODUCTION

THE integrated circuit (IC) substrates have been vastly utilized in almost all electronic products, and their function integrity and reliability are heavily dependent upon the

Manuscript received Month xx, 2xxx; revised Month xx, XXXX; accepted Month x, XXXX.

This work is funded by the National Natural Science Foundation of China (61573146), National Science and Technology Major Project of the Ministry of Science and Technology of China (2014ZX02503), Applied Science and Technology Research and Development Special Fund Project of Guangdong Province, China (2015B010133003), and European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Grant Agreement, United Kingdom, (101006411).

Y. Zeng and Y. Mei are with the School of Electronics and Information, Guangdong Polytechnic Normal University, Guangzhou, 510665, China (e-mail: auyzeng@mail.scut.edu.cn, yfmei@gpnu.edu.cn).

Y. Hu is with the School of Automation Science and Engineering, South China University of Technology, and also with the Engineering Research Center for Precision Electronic Manufacturing Equipment, Ministry of Education, the Guangdong Provincial Engineering Laboratory for Advanced Chip Intelligent Packaging Equipment, Guangzhou, 510641, China (e-mail: auyihu@scut.edu.cn).

Z. Sheng is with the Department of Engineering and Design, University of Sussex, Brighton, BN19RH, United Kingdom (e-mail: z.sheng@sussex.ac.uk).

quality of IC [1]–[3]. Surface defects are one of the most critical defects among IC faults that frequently arise in practical assembly lines, which indubitably result in microscopic electronic products with insufficient quality, dysfunction, or even breakdown [4], [5]. Thus, it is essential to provide key information on IC via effective intelligent fault measurement systems for quality control in industrial processes.

Non-contact inspection methods, typically automatic optical inspection (AOI), have become increasingly prevalent with the grand advancements in imaging technologies [6]. It also promotes the solutions for detecting simple and single IC surface defects that were manually verified by personnel in the traditional processes, which is time-consuming and inaccurate [7]. Specifically, deep convolution neural network (DCNN) frameworks have emerged as leading tools for large-scale object detection and have performed well on public image datasets [8]. The application of these trained models combined with non-contact inspection schemes has effectively optimized the result of IC quality control. However, flexible integrated circuit substrates (FICS), as the advanced interconnecting boards of ICs, pose great a challenge for intelligent fault diagnosis in real microelectronic production due to the diversity and high detail of their surface defects.

The difficulty of IC fault diagnosis could be regarded as three degrees according to the chip mounting density of three IC development phases including FICS, flexible printed circuit (FPC), and printed circuit board (PCB) [9]. The detection strategies applied for the low-density PCB and FPC images generally adopt a pixel-by-pixel analysis-based protocol, which is low in precision and efficiency when used to detect faults on the high-density FICS. Moreover, FICS is often discarded directly as soon as faults are detected on its surface in real assembly lines, which leads to increased costs and poor yield in existing works. Whereas most of these defective FICS could be mended with only minor effort to meet actual application demands. Superior algorithms for fault inspection have been desired to boost diagnosis results and intelligence degree in quality control processes of high-density FICS. [6].

II. RELATED WORK

The template reference contrast-based method initially developed by Wu *et al.* [10] exploited a Gerber-character-based detection system applicable for single defect detection on PCBs to tackle the problem of low efficiency. Oh *et al.* [11]

proposed the sequential matching scheme using finite templates to enable PCB inspection. Zheng *et al.* [12] described a neural network-based detection model to identify PCB surface defects with improved performance, which benefits from introducing the continuous atrous convolution module and optimizing the skip layer. Wang *et al.* [13] and Wang *et al.* [14] leveraged the image segmentation scheme to locate defects on FPCs, wherein the Gaussian low-pass filter and dimensional increment matrix computation were employed in their methods separately. Luo *et al.* [15] explored an inspection algorithm for circular holes on FPCs' surfaces using a modified tri-class thresholding-based contour detection extraction.

Increasing miniaturization of chips has been aggravating the complexity of FICS manufacturing, which leads to a high defect risk and difficulty in surface defect detection as well [16], [17]. Zhong *et al.* [7] presented a vision-based method that utilized two cascaded procedures including a standard template achieving and probability calculation between the background and defect of FICS to distinguish the impurity fault. Hu *et al.* [18] recognized the high-density FICS oxidation fault using the directed acyclic graph support vector machines (DAG-SVMS), which were trained by extracted 8-dimension color image features. Luo *et al.* [2] presented a cascaded tow-stage object detection architecture, consisting of two convolutional neural networks modified by using the multi-hierarchical aggregation and locally non-local blocks separately, to control the quality of flexible substrates. To enhance the model's feature extraction ability from the substrate, Luo *et al.* [19] reported a modified Faster R-CNN model by integrating the multi-scale feature fusion and multiple receptive field fusion modules to achieve flexible substrate defect inspection. However, their diagnosis results and intelligence degree in FICS multi-fault detection need to be improved.

Aiming at the diversity of FICS faults and the inapplicability of low-density IC inspection algorithms, there is an urgent need to develop effective fault diagnosis methods to control the FICS quality. Another accompanying demand is the construction of an advanced imaging module to acquire microscopic details of substrates boosting the accuracy of fault detection. Inspired by the DCNN framework and hierarchical decision strategy, a novel intelligent system for FICS fault diagnosis is proposed considering the tight incorporation of a microscopic imaging tool and detection algorithm. The main contributions of this paper can be summarized as follows:

- 1) A hierarchical tree-DCNN structure (HTDS) and multi-scale microscopic imaging module are proposed. Multiple diagnosis problems for FICS are decoupled into considered subtasks and deployed in several stages of HTDS, where FICS images with specific resolution are required suitably in specific stages and a fault locator of HTDS is allowed in alternative working states. The proposed diagnosis strategy could facilitate the real application of FICS fault diagnosis.
- 2) The optimization for three DCNN frameworks and network training protocol is accomplished. Three lightweight DCNNs are built wherein special feature collection and fusion modules are introduced for FICS microscopic faults, and the addressed unbalanced data problem using the data-balanced grouping strategy is able to strengthen the model's

learning performance. This optimization would improve the success rate and efficiency of FICS fault diagnosis.

The rest of this paper is organized as follows: the problem formulation of the FICS fault diagnosis is presented in Section II. The HTDS and DCNNs in each stage are described theoretically and chosen for FICS sorting, quality classification, and defect localization in Section III. Section IV presents several comparative experiments to demonstrate the advantages of the proposed structure. Section V lists the related conclusions and works for further enhancements.

III. PROBLEM FORMULATION

Image fault detection could be seen as a pattern recognition task. Considering the image sequence $X = \{x|x^{(i)} \in \mathbb{R}^{2 \times k}\}$ with data dimensions of 2 and length of $k \in \mathbb{R}^d$ as an input to a single model for the multi-fault detection and classification, the undesirable result would be obtained usually due to the presence of a large number of fault categories and the division ambiguity of some of them. $y^{(i)} \in \Lambda$ is the discrete class label.

In our works, quality audit for image sequence is considered to be decoupled into the cascaded subtasks, and then verification processes are conducted hierarchically using a model $\hat{y}_t = f_t(x_t^{(i)}, \hat{\theta}_t)$ parametrized in $\hat{\theta} \in \Theta$, being Θ the set of parameter space. Where $t \in \{1, 2, 3\}$ denotes the correlated three stages of fault diagnosis that differ from traditional single multi-class classifiers using one-stop detection. Based on the suggestion of information theory, the larger the value, the stronger the relevance [20]. The relevance is defined by

$$H(X|Y) = - \sum p_t(x) \sum p_t(x|y) \log_2(p_t(x|y)) \quad (1)$$

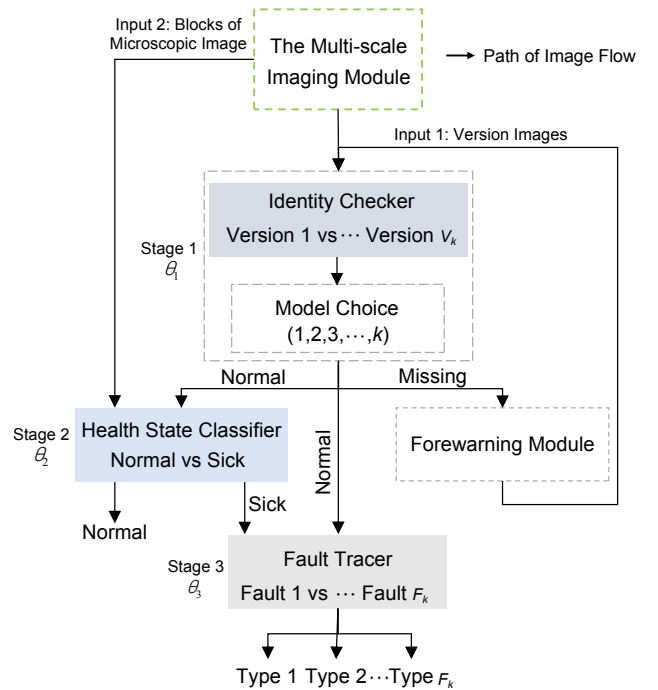


Fig. 1. Pipeline of the high-density FICS fault diagnosis with three-stage HTDS.

where $p_t(x)$ is the distribution function of x in each stage. For an input image sequence $X = \{x|x^{(i)} \in \mathbb{R}^{2 \times k}\}$

in the FICS fault diagnosis process, the first stage model $\hat{y}_1 = f_1(x_1^{(i)}, \hat{\theta}_1)$ is to sort different FICS incoming, which have a large $H(x^{(i)}|X)$ while a high image resolution also is. However, it is tough to detach microscopic fault details from high-resolution input image data, the input sequence X should be resampled resulting $\hat{X} = \{\hat{x}|\hat{x}^{(i)} \in \mathbb{R}^{2 \times \hat{k}}\}$, wherein the relation $\hat{k} = \rho k$ is built by a scale factor $\rho(\rho > 1)$. The two-class output sets of qualified and unqualified, which have different relevance but similarities, are defined in the classification task of the second stage model $\hat{y}_2 = f_2(x_2^{(i)}, \hat{\theta}_2)$. Specifically, the probability function $p_2(\hat{x}^{(i)}, \theta) = \theta^{\hat{x}^{(i)}}(1 - \theta)^{1 - \hat{x}^{(i)}}$ with $\theta \in \Theta$ under the sequence $\hat{X} = \{\hat{x}|\hat{x}^{(i)} \in \mathbb{R}^{2 \times \hat{k}}\}$ is followed. Deliberating most of images in sequence X are qualified, the model $\hat{y}_3 = f_3(x_3^{(i)}, \hat{\theta}_t)$ in the third stage is in alternative working state and would be activated to locate faults when the image has been detected as defective in the second stage. Where the uniform distribution $p_3(x^{(i)})$ is followed.

IV. METHODOLOGY

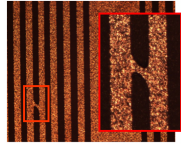
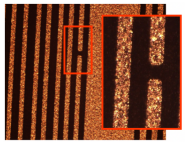
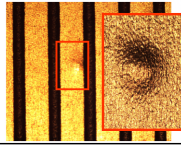
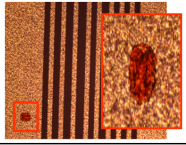
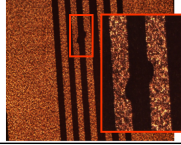
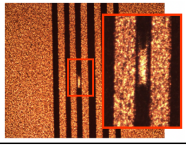
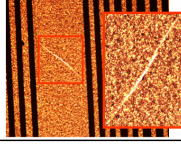
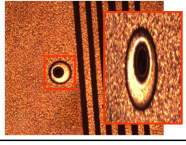
A. Hierarchical Framework

As shown in Fig.1, the proposed HTDS formulated by $\hat{y}_t = f_t(x_t^{(i)}, \hat{\theta}_t)$ with $t \in \{1, 2, 3\}$ and $\hat{\theta} \in \Theta$ is comprised of three stages in a tree-like manner. Three-stage networks of HTDS have been modified, which can perform three significant processes for FICS sorting, quality classification, and fault measurement. Firstly, automatic sorting is enabled for the high-density FICS incoming, which is fundamental intelligence in practical industrial processes but still relies on personnel visual sorting. Secondly, based on the expert knowledge that qualified FICS vastly outnumber defective ones in microelectronic production, the FICS fault diagnosis is divided into two stages: quality classification and fault measurement. Specifically, a quality classifier categorizes substrates into two-class results of qualified and unqualified, which makes the substrate of different qualities with different workflows following. The cascaded fault locator further identifies and classifies the faults only if the quality classifier outputs unqualified types. The HTDS is trained hierarchically by using the presented data-balanced grouping strategy and then conducts the FICS fault diagnosis from coarse to fine. The DCNN-based hierarchical decision mechanism could speed up the FICS fault diagnosis, and the unbalanced data problem addressed would strengthen the learning performance of the network to improve the accuracy.

Moreover, low-resolution substrate images are generally not good at presenting fault details resulting in a high omission rate for diagnosis algorithms in quality control processes. The multi-scale microscopic imaging module, which is supposed to offer substrate images at selectable scales required by the proposed three-stage HTDS, is designed to contribute to the diagnosis accuracy improvement of HTDS. As shown in Table I, the FICS image examples captured by the proposed multi-scale microscopic imaging module are mainly comprised of the black background, yellow copper-clad conductor, and copper foil with disparate faults.

1) FICS Sorting: Rapid and precise judgment of various FICS is a prerequisite for effective quality diagnosis and high

TABLE I
THE EXAMPLE OF HIGH-DENSITY FICS IMAGES WITH TYPICAL FAULTS

Type	Sample	Type	Sample
SH		OP	
HU		IM	
UN		ET	
SC		LH	

reliability of FICS [7]. FICS images are of large sizes, and there are macroscopic feature discrepancies apparent in the image size, wiring method, and mounting chip area. Hence, a DCNN-based lightweight model $\hat{y}_1 = f_1(x_1^{(i)}, \hat{\theta}_1)$, $\hat{\theta}_1 \in \Theta$ with a large receptive field for feature extraction is considered to sort different FICS images in the case of maintaining a good trade-off of sorting efficiency and accuracy in stage 1. The framework of the sorting network is plotted in Fig.2, which mainly includes a total of six convolution layers ($Cov_i, i = 1, 2, \dots, 6$) and two fully connected layers ($Fc_i, i = 1, 2$). Followed softmax can generate the likelihood to predict which category the input FICS belongs to.

The input layer of the sorting network is organized by the high-density FICS image stream $\{x^{(1)}, x^{(2)}, \dots, x^{(k)}\}$ with a dimension of 2, which is preprocessed with compression and random cropping. Here a threshold $k < N$ ($N \in \mathbb{R}$) is used to split image data since the order in which substrates arrive is chaotic. For convolution layers following, key image features are extracted by a series of fixed-size kernels with convolution width C_{w_1} and stride between kernels of S_{c_1} . In detail, the 2-D convolution is to take a dot product between image matrix $I \in \mathbb{R}^{m \times n}$ and kernel $K_1 \in \mathbb{R}^{k_{m_1} \times k_{n_1}}$ to gain representation features

$$S_{[i,j]} = ReLu(\sum_{i \in \epsilon} (\sum_{k_{m_1}} \sum_{k_{n_1}} I_{[i+k_{m_1}, j+k_{n_1}]} \cdot K_{1[k_{m_1}, k_{n_1}]}) + b_c) \quad (2)$$

where b_c means the bias vector, subscript ϵ is the number of kernels, $S_{[i,j]}$ denotes the output point (i, j) of convolution layer, and $ReLu(*)$ is referred to the activation function.

The max-pooling with pool width of P_{w_1} and stride between pools of S_{p_1} is expected to downsample, which could alleviate the statistical burden on subsequent layers by reducing the

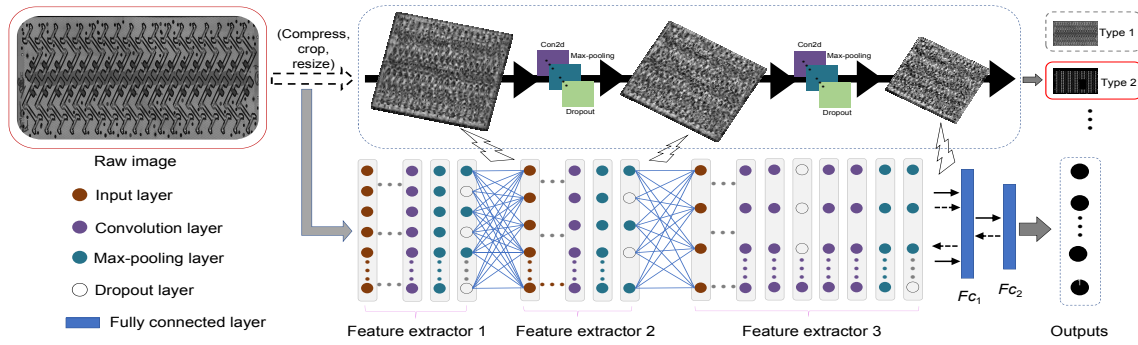


Fig. 2. DCNN framework for high-density FICS sorting in stage 1.

size of representation features. Max-pooling process returns the maximum value within a certain subregion as follows:

$$\hat{S}_{[i,j]} = M\{S_{[i,j]_{i \times P_{w_1}:(i+S_{p_1}) \times P_{w_1}, j \times P_{w_1}:(j+S_{p_1}) \times P_{w_1}}}\} \quad (3)$$

where $M\{*\}$ denotes the max-pooling operator, and $\hat{S}_{[i,j]}$ is referred to the output of pooling layers at point (i, j) .

Subsequently, two fully connected layers are followed after those blocks to perform logical reasoning, in which the output of fully connected layer F_{C_2} is calculated as

$$S_o = f((W_{F_2})^T S_{F_1} + b_{F_2}) \quad (4)$$

where W_{F_2} is the weight matrix between the fully connected layers F_{C_1} and F_{C_2} , b_{F_2} means the relevant bias vector, S_{F_1} denotes the input from the layer F_{C_1} , and $f(*)$ indicates the activation function. Eventually, the category of different FICS is confirmed in the output layer based on the fully connected layer output S_o . The family of input-output functions could be obtained via the softmax to be

$$\hat{y}_1 = \frac{1}{\sum_{i=1}^{V_k} e^{((W_i)^T S_o + B)}} \begin{bmatrix} e^{((W_1)^T S_o + b_f)} \\ e^{((W_2)^T S_o + b_f)} \\ \vdots \\ e^{((W_k)^T S_o + b_f)} \end{bmatrix} \quad (5)$$

where $\hat{y}_1, \sum_{i \in V_k} \hat{y}_1 = 1$ is the output non-linearity probability, W_i means the weight matrix between layer F_{C_2} and i th output layer neurons, b_f is referred to the bias vector, and V_k denotes the high-density FICS categories.

2) FICS Quality Classification: Low-resolution substrate images usually can not present defect details well, which could worsen the success rate for detection algorithms. Therefore, feeding the sorted high-density FICS images in stage 1 to the quality classifier in stage 2 of HTDS directly is not an optimal choice. The blocks of FICS microscopic images by resampling can be considered as the input to perform feature extraction for the following FICS quality classifier, which can promote classification accuracy. However, defects in substrate image blocks are tiny at the micron-scale and have the attribute with different aspect ratios and scales (see Table I), which could lead to a lack of the feature memory for deep network layers by using fixed sliding windows. For these reasons, a DCNN-based classification network, which is a binary classifier $\hat{y}_2 =$

 TABLE II
 LIST OF NETWORK ARCHITECTURE FOR FICS QUALITY CLASSIFICATION

Input $128 \times 128 \times 1$
Block-1 $\times 2$ $k_{m_2} \times k_{n_2} \times \epsilon$ ReLU $[P_{w_2} P_{w_2}]$ Avg Pooling
Block-2 $\times 2$ $k_{m_2} \times k_{n_2} \times \epsilon$ ReLU Dropout α $k_{m_2} \times k_{n_2} \times \epsilon$ ReLU $[P_{w_2} P_{w_2}]$ Avg Pooling
Block-3 $\times 2$ $k_{m_2} \times k_{n_2} \times \epsilon$ ReLU $[P_{w_2} P_{w_2}]$ Avg Pooling Dropout α
F_{C_1} 1024×128 ReLU Dropout α
F_{C_2} $128 \times N_c$ ReLU ($N_c = \#$ of Classes)
Softmax Layer

$f_2(x_2^{(i)}, \hat{\theta}_2), \hat{\theta}_2 \in \Theta$ with $x_2^{(i)} \in \{0, 1\}$, is cascaded to classify high-density FICS quality in Stage 2.

Inspired by the deep convolutional transfer learning network (DCTLN) [21], a 1-D classifier in their work has 14 layers between input and output layers including six blocks of convolution and pooling layers followed by two fully connected layers, which can categorize complex machine states. To exactly and fast classify high-density FICS into considered two classes: qualified and unqualified, a DCNN-based quality classification network with six hidden layer blocks of convolution and pooling, and two fully connected layers ($F_{C_i}, i = 1, 2$) is built. The list of the proposed DCNN architecture used in stage 2 of HTDS is shown in Table II, and Fig.3(a) plotted the network architecture.

As shown in Fig.4(b), FICS microscopic image blocks can be taken after being amplified by using the optical microscope. These image blocks can be used to build input layers of a FICS quality classifier. As for the convolution layers, the deformable convolution network (DCN) is introduced to strengthen the ability of the network's adaptive modeling for high-detail FICS defects. Therefore, the adaptation based on equation (2), which

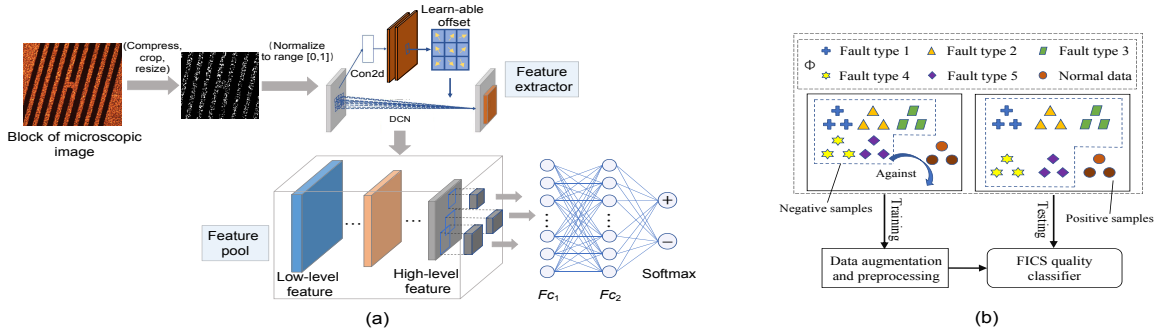


Fig. 3. High-density FICS quality classification in the unbalanced data situation in stage 2. (a) The DCNN framework of FICS quality classifier, (b) The data-balanced grouping strategy used in the classification process of (a).

is performed by adding learn-able offset $\Delta p = [\Delta p_x, \Delta p_y]$ and predicted weight Δm_n , is continuous for irregular sampling. The update is accomplished as follows:

$$S'_{[i,j]} = \sum_{i \in \epsilon} \left(\sum_{k_{m_2}} \sum_{k_{n_2}} I'_{[i+k_{m_2}+\Delta p_x, j+k_{n_2}+\Delta p_y]} \cdot K_2[k_{m_2}, k_{n_2}] \cdot \Delta m_n \right) \quad (6)$$

where K_2 means the kernel, $S'_{[i,j]}$ is referred to the output feature mapping matrix, and $I'_{[i,j]}$ indicates the input associated with microscopic image blocks of the FICS. At the same time, to acquire high-level image mappings for the dense sampling in equation (6), the bilinear interpolation operator is selected and defined by

$$I'_{[i,j]} = \sum_{\Delta x, \Delta y} I'_{i+\Delta x, j+\Delta y} \cdot g(p_{i+\Delta x}, p_{i+k_{m_2}+\Delta p_x}) \cdot g(p_{i+\Delta y}, p_{i+k_{n_2}+\Delta p_y}) \quad (7)$$

where the subscript $[\Delta x, \Delta y]$ means the offset metrics from a center point of the convolution region, and $g(*)$ is the double copy of a 1-D kernel formulated by $g(a, b) = \max(0, 1 - |a - b|)$, $a, b \in \mathbb{R}^d$. Based on those analogous processes of feature extraction in equations (2-5), the output \hat{y}_2 is obtained and the FICS quality is estimated.

The unbalanced sample data, which could result in inadequate learning performance for network models, is ubiquitous in the actual microelectronic fabrication since it is greatly challenging to fairly access many substrate faults [21], [22]. In our work, the faithful combination of decoupling the diagnosis task into three phases and the data-balanced grouping strategy is considered to promote the detection accuracy for networks in the unbalanced sample situation. As drawn in Fig. 3(b), categories of training sample data Φ can be grouped semantically. Concretely, various FICS fault types $\{x_2^{(1)}, x_2^{(2)}, \dots, x_2^{(F_k)}\}$ contained $F_k \in \mathbb{R}^d$ entries would be organized into the superclass C_F , which is seen as the single element category against with the qualified FICS type $C_N (C_N \notin C_S)$. A finer partition in fault type set C_F will be completed as we approach the next stage of HTDS. In the learning process, qualified class C_N is combined with superclass C_F that has been balanced with the amount of qualified class C_N as the two-class training data. In other words, a multi-classification task is cast to the dichotomous problem, which helps prevent the over-fitting of network models because of the average probability distribution

of the image sample data. From a mathematical point of view, the probability distribution with a lost gravity center in an unbalanced sample situation Φ is considered as the Bernoulli distribution, and the corresponding probability function has been changed and given by $p(x_2^{(i)}; \theta) = \theta x_2^{(i)} (1 - \theta)^{1-x_2^{(i)}}$ with $x_2^{(i)} \in \{0, 1\}$. Particularly, the output result set C_F will be submitted to stage 3 directly.

3) FICS Fault Localization: Fault diagnosis is considered to be carried out by leveraging two cascaded stages after FICS sorting in our three-stage HTDS, instead of those single multi-class detectors trained to identify and classify faults directly in existing methods, which is time-consuming and low in accuracy. The hierarchical decision mechanism introduced for FICS fault diagnosis is expected to offer speedup and weaken the risk of diagnosis errors as well. Considering the fact that the majority of FICS, over 90%, are qualified in practical assembly lines, there is thus unnecessary to execute the entire three stages of HTDS to control quality for each substrate. To optimize the diagnosis efficiency, the fault locator in stage 3 of HTDS is considered as an alternative working mode, which will be activated to work on FICS fault localization while only stage 2 outputs unqualified images. On this basis, two-class output results involving repairable and unrepairable categories are defined for the fault locator to increase the raw material utilization and final yield.

The goal of stage 3 of HTDS is supposed to locate and classify faults under the classified FICS fault set if given by stage 2. Based on the suggestion of the advanced version of You Only Look Once (YOLOv3) [23], the original architecture has been built using the DarkNet53 backbone network, which is comprised of five residual modules for extracting image features in different sizes. The core conception of YOLOv3 is to generate image feature mappings with three shifty scales after each of three residual modules, e.g., at sizes of 52×52 , 26×26 , and 13×13 . These extracted image mappings can be merged to construct the sparse feature pyramid connection (SFPC) to predict desired objects. Good result on those tasks related to large and multi-class object detection for YOLOv3 has been reported. Pay attention to faults on microscopic FICS image blocks having the attributes of different aspect ratios and scales at the micron scale, A YOLOv3-based architecture, formulated $\hat{y}_3 = f_3(x_3^{(i)}, \theta_3)$, $\theta_3 \in \Theta$, $x_3^{(i)} \in C_F$, $i \in [1, F_k]$, is built by using a multi-scale dense feature pyramid connection

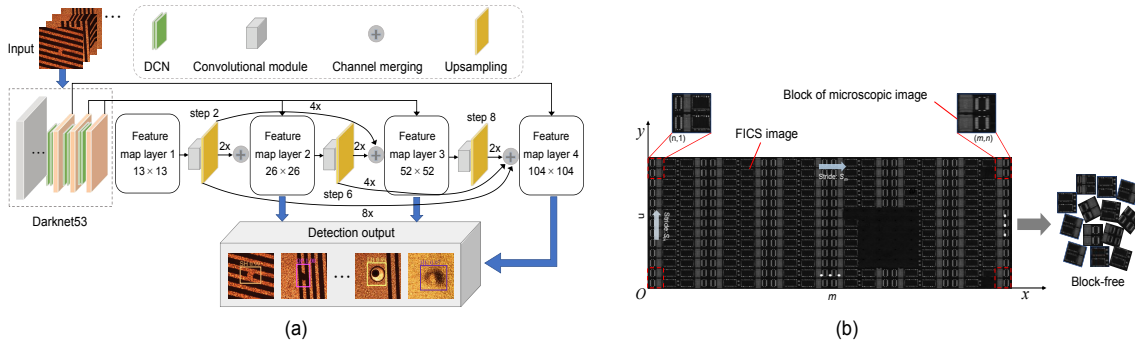


Fig. 4. FICS Fault Localization. (a) Modified YOLOv3-based framework with MDFPC, (b) Acquisition of FICS microscopic image blocks using a metallurgical microscope.

(MDFPC) and updated hyperparameters to locate FICS faults in stage 3. The detailed framework is illustrated in Fig.4(a).

Compared to the original YOLOv3 framework, structural modifications of our fault locator have been made as follows:

- 1) MDFPC is exploited by adding the lower network layers at the size of 104×104 and the element's combination of different proportions to better collect image cues for representing FICS micron-scale faults.
- 2) DCN is used instead of the convolution of each residual block to adaptively extract high-detail features from FICS faults with huge variability in size. Also, the size of the upsampling steps of $2[5]$ is expanded to 4, 6, and 8, which are adopted to fuse other information from convolution layers to reuse high-level feature mappings better.

B. Multi-objective Learning Schemes

1) Object 1: For different high-density FICS images with diverse and plentiful faults, the proposed three-stage HTDS should be able to sort FICS and further enable intelligent fault diagnosis without human involvement. Specifically, the FICS sorter aims to replace the manual vision classification process. Therefore, the optimization objective in stage 1 of HTDS is to minimize FICS sorter error based on various low-resolution substrate images, the loss function is defined

$$L_1 = -\frac{1}{T} \left[\sum_{j=1}^T \sum_{i=1}^{V_k} \mathbb{I}_{[y_1, i=V_k]} \log \frac{e^{(W_i)^T S_o + b_f}}{\sum_{g=1}^{V_k} e^{(W_g)^T S_o + b_f}} \right] \quad (8)$$

where $\mathbb{I}_{[*]}$ indicates the indicator function, V_k is referred to the FICS categories, and T means the number of training samples fed into the network each time.

2) Object 2: Considering attributes of high-density FICS samples in number, features, and unbalanced sample problem, fault diagnosis followed FICS sorter is cast into two sub-steps including indispensable FICS quality classification and alternative FICS fault localization to boost the quality control efficiency and accuracy. FICS quality classifier is trained to classify FICS high-resolution microscopic image blocks into expected two categories, qualified and unqualified, and give instruction on whether the cascaded FICS fault locator will work for further fault localization or not. Thus, the optimization objective in stage 2 of HTDS is to maximize quality

classification error based on the balanced binary training data. The loss function is defined

$$L_2 = \frac{1}{T} \sum_{i=1}^T [y_{ture} \neq f_2(x_2^{(i)}, \hat{\theta}_2)] \quad (9)$$

where

$$[y_{ture} \neq f_2(x_2^{(i)}, \hat{\theta}_2)] = \begin{cases} 1, & y_{ture} \neq f_2(x_2^{(i)}, \hat{\theta}_2) \\ 0, & y_{ture} = f_2(x_2^{(i)}, \hat{\theta}_2) \end{cases} \quad (10)$$

where $x_2^{(i)}$ denotes the i th new input sample acquired, while y_{ture} is the true label.

3) Object 3: Multi-class faults on FICS microscopic image blocks in the output unqualified set C_F need to be identified and classified, and further located on raw FICS images. As illustrated in Fig.4(b), the FICS image is orderly scanned by using the metallographic microscope with a fixed receptive field and progressive scan manner of stride S_m ($S_m \leq 3000pixels$) to generate FICS microscopic image blocks. Four-position indicators of each bounding box of the receptive field, e.g., the box coordinates (b_x, b_y, b_w, b_h) , will be stored in time. For instance, for a raw FICS image with the dimension of $m \times n$, a position tensor of $1 \times m/b_w \times n/b_h$ can be produced in the scan process. Therefore, incorporating the fault localization on microscopic image blocks and the corresponding position tensor could enable accurate recovery of fault position on the raw FICS image. The optimization objective in stage 3 of HTDS is to minimize the sum-squared error based on fault samples.

Based on works derived from the original YOLO detector [23], [24], the multi-part loss function is defined

$$\begin{aligned} L_3 = & \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{I}_{ij}^{obj} [(\chi_i - \hat{\chi}_i)^2 + (\gamma_i - \hat{\gamma}_i)^2] \\ & + \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{I}_{ij}^{obj} [(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2] \\ & + \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{I}_{ij} (C_i - \hat{C}_i)^2 + \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{I}_{ij}^{noobj} (C_i - \hat{C}_i)^2 \\ & + \sum_{i=0}^{S^2} \mathbb{I}_{ij} \sum_{c \in class} (p_i(q) - \hat{p}_i(q))^2 \end{aligned} \quad (11)$$

Here \mathbb{I}_{ij}^{obj} means the target object that is assumed to present in the i^{th} location of the region, χ, γ, w and h indicate the current position and size of the image, q denotes the number of entire objectives to be recognized, $p(q)$ is referred to the probability that the objective belongs to the particular class q , λ_{coord} and λ_{noobj} indicate the scalars weighting each confidence loss function, S is the size of divided cell in each feature mapping, and B is the number of prediction boxes in each cell.

The process of transfer learning [21] and end-to-end training protocol are adopted to unify the training process in this paper, which could enable multiple losses to propagate weights for different ranges of network layers. Each of the three-stage HTDS is trained by the stochastic gradient descent (SGD) [25] and backpropagation solution [21] based on three optimization objects mentioned above. The update of weights θ_1, θ_2 and θ_3 is completed backward as follows:

$$\Delta\Theta = \begin{cases} \theta_1 \leftarrow \theta_1 + \eta \frac{\partial L_1}{\partial \theta_1} \\ \theta_2 \leftarrow \theta_2 + \eta \left(\frac{\partial L_2}{\partial \theta_2} - \mu \frac{\partial L_1}{\partial \theta_1} \right) \\ \theta_3 \leftarrow \theta_3 + \eta \left(\frac{\partial L_3}{\partial \theta_3} - \mu \frac{\partial L_1}{\partial \theta_1} \right) \end{cases} \quad (12)$$

where μ denotes the hyperparameter weighting the FICS sorter strength in the training process, and η is the learning rate. To accelerate the training process and optimize the weights iteratively, the initial network parameters in θ_2 of the quality classifier of HTDS could be set using the FICS sorter' weights trained in θ_1 due to the similarity of the training process and samples, and then they are cascaded to train for $\theta_1 \cup \theta_2$. Where specified network layers should be frozen ensuring the same size of networks to achieve matched parameter sharing. The analogous training protocol can be applied to the fault locator of HTDS. However, the proposed HTDS of this paper works step by step since our purpose is to control the quality of high-density FICS from coarse to fine by decoupling the diagnosis problem into specific subtasks.

V. EXPERIMENTAL VALIDATION

A. AOI System

As shown in Fig.5(a), the high-density FICS quality AOI system mainly consists of the hardware component of the proposed multi-scale microscopic imaging tool and the HTDS-based fault diagnosis procedure. Choosing an advanced imaging tool could help overcome the problems caused by the use of uncomplicated processing algorithms [26]. Fig.5(b) demonstrates the proposed multi-scale microscopic imaging module constructed by the area-scan camera (ASC) and the charge-coupled device (CCD) fixed on the metallurgical microscope, which is mounted on the precise X-Y positioning device. It can image substrate individually at different scales generating FICS images and microscopic image blocks as the input for stages 1 and 2 of HTDS.

As for the microscope lens, the maximum and minimum view scopes are $2.13mm \times 2.84mm$, $13.71mm \times 18.29mm$ separately, and each lens magnification corresponds to an image resolution of FICS under test. The discrimination precision of $0.07\mu m/pixel$ is able to be achieved. Accordingly, the plan achromatic objective named PL L5X/0.12 is exploited

with the actual working distance of $26.10mm$. On the other hand, the camera with the average frame rate of $17frames/s$ is chosen. Three-color structural light-emitting diode (LED) offers constant illumination intensity, which promotes the quality of the captured substrate image to ensure effective image processing. For an image acquisition process of the substrate, the multi-scale imaging module needs to perform double scans automatically, including a global view of ASC to collect a FICS image and then microscopic image blocks acquired by CCD via microscope successive scanning.

Algorithm 1 FICS fault diagnosis using the proposed HTDS

Input: the FICS image stream $X = \{x|x^{(i)} \in \mathbb{R}^{2 \times k}\}$.
Ensure:
 Set the initial parameters of AOI system;
for $i = 1 : k$
 Classify different FICS based on FICS sorter in stage 1;
 Collect microscopic image blocks $\hat{X} = \{\hat{x}|\hat{x}^{(i)} \in \mathbb{R}^{2 \times k}\}$;
 for $j = 1 : T$
 Judge FICS quality based on FICS classifier in stage 2;
 if unqualified == True
 Locate fault based on FICS fault locator in stage 3;
 Classify FICS fault pattern and build fault set $C_d = \{x_1, x_2, \dots, x_k\}$;
 else
 Continue;
 end if
 end for
 Calculate modification factors $\Delta_k = \{\Delta_1, \dots, \Delta_k\}$;
 Update parameters of AOI system;
end for
Output: fault patterns and factors Δ_k of the fault diagnosis.

* k denotes the length of the collected datasets, and T refers to the number of image blocks acquired from a microscope.

B. Fault Diagnosis

1) Dataset: Based on the previous works [7], [34], the dataset is organized and comprised of three subsets, which cover typical high-density FICS fault types in real assembly lines. In detail, subset 1 containing 9871 FICS images for three substrate categories is established for stage 1 of HTDS to evaluate and choose a FICS sorting model, in which 7879 and 1992 images are in the training and validation set separately. Faults, repairable types involving scratch (SC), undercut (UN), impurity (IM), landless hole (LH), and hump (HU), unrepairable types including etch-back (ET), open (OP), and short (SH), are collected from 9871 FICS images of stage 1, and the rest of 452 is populated by expert inspectors. In total, 528 images for eight-type faults and corresponding qualified FICS images of 863 are in subset 2 for stage 2 of HTDS. 528 fault images for eight categories are labeled in subset 3 for stage 3. A testing dataset containing 3810 high-density FICS images acquired from different sections of actual substrate assembly lines is generated to reveal the adaptability of each of the three DCNN networks. With the data augmentation solution, the training datasets are also expanded properly.

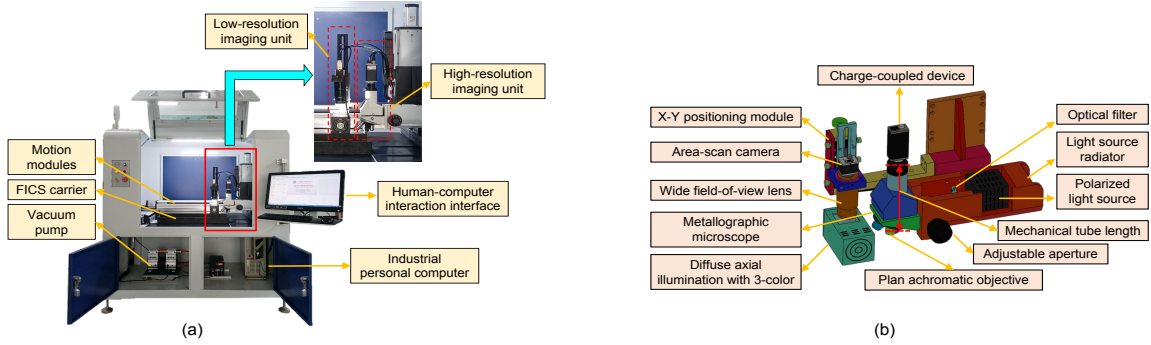


Fig. 5. Substrate fault diagnosis device. (a) AOI system applicable for FICS quality diagnosis, (b) Multi-scale microscopic imaging module contained AOI system of (a).

2) Training Detail: Each of the three DCNN frameworks is trained for 3000 iterations by using SGD with the momentum of 0.9 and initial learning rate of 1×10^{-3} , which is then declined by a factor 0.6 every 1000 iterations. Where the input images of three stages of the proposed HTDS are resized to 227×227 , 128×128 , and 416×416 , and corresponding batch sizes are set to be 32, 16, and 8 due to the limitation of training samples. The size and stride of convolution and pooling kernels are set to be $[C_{w_1} : 7, S_{c_1} : 3]$ and $[P_{w_1} : 3, S_{p_1} : 2]$ separately in FICS sorter. As for the quality classifier, the dropouts α of 30% are adopted to prevent over-fitting. While scalars λ_{coord} and λ_{noobj} are set to be [5, 0.5] by the original author's suggestion of YOLOv3 [24] in fault locator. The training is completed on the NVIDIA GTX1080, and the trend of training losses denotes that three networks will both converge after about 1200 training iterations respectively.

3) Diagnosis Result: Each of the three DCNN architectures is carried out on testing datasets and repeated ten times with images amounting to 127 each time. The quantitative metrics including false positive (FP), false negative (FN), true positive (TP), and true negative (TN) are counted to calculate statistical indicators: sensitivity (Se) and precision (Pr). The accuracy (Ac) and F1-score are computed based on Se and Pr , which are also given in [6], [7]

$$Se = \frac{TP}{TP + FN} \times 100\% \quad (13)$$

$$Pr = \frac{TP}{TP + FP} \times 100\% \quad (14)$$

$$Ac = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (15)$$

$$F1\text{-score} = \frac{2}{1/Se + 1/Pr} \times 100\% \quad (16)$$

Where the closer the Ac and F1-score is to 100%, the better the performance of the network indicates. The effect evaluation of improved three DCNN structures is performed to show that the DCNN architecture is selected and modified correctly for each stage. The related pseudocode is shown in Algorithm 1 and diagnosis results are summarized in Table III, in which the performance regarding running time (T), Ac , and F1-score for each stage are listed and compared. As seen from Table III, the total time cost of stage 1 and stage 2 is as low as 0.058s,

and average values of Ac and F1-score of 95.21% and 94.62% are reached respectively. Importantly, this is a typical process since stage 3 of HTDS does not work in most cases. For the whole HTDS, a total running time of 0.189s and accuracy over 94% are achieved.

The proposed HTDS shows relatively better efficiency and accuracy for FICS quality control due to the application of the hierarchical decision mechanism and DCNN framework established for each stage. However, qualified FICS images sometimes are divided as the unqualified category and it is a close call, and part of the ET type is seen as SH (see Table I). In fact, a small percentage of wrong predictions of qualified FICS are allowed to become unqualified ones to some extent in the FICS manufacturing process, since the key task is to pick up each fault strictly, and fault types ET and SH are both included in the unrepairable type.

C. Comparison and Validation

1) Comparison among individual networks: To evaluate the efficiency and accuracy, the proposed three-stage DCNN frameworks of HTDS are individually compared with the shallow convolution network: vehicle logo recognition system (VLRS) [27] equipped with two hidden layer pairs of convolution and pooling and a fully connected layer, which can enable the grouping of eleven-type vehicle logos, and large-scale networks: AlexNet [28], DCTLN [21], and ResNet-CBAM [29]. The first four lines of Table III present FICS sorting results in stage 1. It is easy to see that the modified DCNN architecture and lightweight structure show higher classification speeds than those large. In contrast, the lightweight model suffers some accuracy degradation, which is not allowed since the diagnosis result will be of low accuracy by resulting from the wrong choice of input FICS images.

To analyze the classification accuracy among the lightweight model network, the proposed FICS quality classifier is compared against four networks in stage 2, VLRS [27], AlexNet [28], DCTLN [21], and ResNet-CBAM [29]. As summarized in Table III, the conclusion can be drawn that both DCNN architectures could fast response for relatively simple binary classification tasks, but the proposed DCNN framework yields 4% higher classification accuracy than others due to the introduction of the DCN module. Furthermore, the over-fitting

TABLE III
COMPARISON RESULTS OF VARIOUS METHODS USED IN DIFFERENT STAGES

Model	Stage 1			Stage 2			Stage 3			Average		
	$T(s)$	$Ac(\%)$	F1-score(%)	$T(s)$	$Ac(\%)$	F1-score(%)	$T(s)$	$Ac(\%)$	F1-score(%)	$T(s)$	$Ac(\%)$	F1-score(%)
VLRS [27]	0.097 ± 0.017	84.30 ± 6.12	82.75 ± 4.31	0.088 ± 0.019	85.69 ± 6.19	83.63 ± 3.39	/	/	/	0.093	85.00	83.19
AlexNet [28]	0.216 ± 0.015	89.07 ± 1.52	90.31 ± 2.21	0.176 ± 0.015	90.24 ± 1.73	92.08 ± 0.99	/	/	/	0.196	89.66	91.20
DCTLN [21]	0.160 ± 0.022	87.37 ± 4.04	82.96 ± 2.83	0.127 ± 0.014	84.81 ± 1.50	80.05 ± 3.13	/	/	/	0.144	86.09	81.51
ResNet-CBAM [29]	0.117 ± 0.009	92.39 ± 3.46	90.13 ± 2.59	0.085 ± 0.011	91.33 ± 2.29	90.92 ± 2.01	/	/	/	0.093	91.86	90.53
HTDS(Ours)	0.037 ± 0.010	96.29 ± 1.18	95.74 ± 2.65	0.036 ± 0.005	94.12 ± 1.07	93.49 ± 2.07	/	/	/	0.036	95.21	94.62
	/	/	/	/	/	/	0.147 ± 0.016	92.13 ± 1.26	96.31 ± 1.95	0.147	92.13	96.31
MSCDAE [30]	/	/	/	/	/	/	0.100 ± 0.054	70.79 ± 4.88	85.61 ± 3.57	0.100	70.79	85.61
R-CNN [31]	/	/	/	/	/	/	0.167 ± 0.022	88.15 ± 3.56	94.03 ± 2.61	0.167	88.15	94.03
YOLOv7 [32]	/	/	/	/	/	/	0.150 ± 0.035	91.03 ± 2.37	90.55 ± 2.18	0.150	91.03	90.55
YOLOv8 [33]	/	/	/	/	/	/	0.146 ± 0.019	92.09 ± 1.97	93.11 ± 3.04	0.146	92.09	93.71

problem tends to occur because of the inadequacy of the learning sample for training processes, such as the large-scale structure AlexNet that usually handles complex objects based on the learning on the vast public dataset.

The improved YOLOv3-based framework for object detection is compared with the multi-scale convolution denoising autoencoder (MSCDAE) [30] network, which can inspect fabric surface defects, the faster R-CNN [31], and the advanced YOLO versions of YOLOv7 [32] and YOLOv8 [33] in stage 3. The results are listed in Table III implying that ours, R-CNN, and advanced YOLO versions have a better effect on fault localization accuracy than MSCDAE, but their speed suffers a weak discount than the lightweight MSCDAE. A slight time delay meets the real-time requirement of the actual industrial process. Moreover, the modified YOLOv3 outperforms advanced YOLO versions in FICS fault localization, especially in terms of the F1-score, because of the introduction of MDFPC and DCN modules for relatively small samples and micron-scale faults. The advanced YOLO versions usually reach the expected detection results through full training with large-scale datasets.

In addition, to verify the presented YOLOv3 detector is optimally modified for FICS fault localization, different feature mappings are selected and contributed to the output prediction based on the original YOLOv3 framework. The detection accuracy and time consumption using various configurations are listed in Table IV. From Table IV, it is easy to see that detection accuracy increases with the inclusion of multiple feature mappings in the output prediction, which could infer that specified network layers are required since FICS fault localization tasks contain various micron-scale objects with different aspect ratios and scales. However, the time consumption with different configurations is almost the same.

TABLE IV
EFFECTS OF DIFFERENT OUTPUT LAYER OPTIONS

Configuration	Prediction by the maps in size of following				$Ac(\%)$	$T(s)$	$F1 - score(\%)$
	13 × 13	26 × 26	52 × 52	104 × 104			
Original YOLOv3 [35]	✓	✓	✓		83.12 ± 2.25	0.140 ± 0.019	85.52 ± 3.01
Modified YOLOv3		✓	✓	✓	92.13 ± 1.26	0.147 ± 0.016	96.31 ± 1.95
Option 1	✓	✓		✓	85.35 ± 3.98	0.142 ± 0.020	88.63 ± 2.65
Option 2	✓		✓	✓	87.12 ± 1.42	0.145 ± 0.021	91.28 ± 2.00

2) Comparison among the combination of networks:

Three comparative investigations are performed to explore the effectiveness and repeatability of the proposed three-stage

HTDS, wherein stages 1 and 2, stages 2 and 3 are separately combined and compared with single multi-class networks in an end-to-end manner. Stages 1 and 2 of HTDS are picked up to directly output the qualified and unqualified result of FICS and compared with the mentioned-above network models, VLRS [27] and AlexNet [28]. The classification effects for considered two quality states of FICS are plotted in Fig.6 indicating that the single multi-class networks are relatively poor in accuracy. It is reasonable that extracting feature representation of like tiny objects of 80×80 pixels from the high-density FICS image of 18846 × 12785 pixels is difficult. That is, the receptive field of convolution windows for FICS images is too large leading to the deficiency of the semantic information about those tiny faults contained for the output prediction. More importantly, the process of automatic FICS sorting has not been considered in the existing methods, instead of screening out by professionals. Overall, it is an optimized strategy to classify high-density FICS quality in two stages.

TABLE V
COMPARISON RESULTS OF VARIOUS METHODS IN AN END-TO-END MANNER

Model	$T(s)$	$Ac(\%)$	F1-score(%)	$SD_T(s)$	SD_{Ac}
VLRS [27]	0.093 ± 0.014	63.58 ± 5.88	59.53 ± 5.71	0.009	3.078
AlexNet [28]	0.187 ± 0.012	69.03 ± 1.79	71.92 ± 3.42	0.007	1.046
HTDS (Stage 1+ Stage 2)	0.069 ± 0.010	92.93 ± 0.84	93.27 ± 1.65	0.007	0.496
MSCDAE [30]	0.098 ± 0.017	65.29 ± 1.89	83.77 ± 2.95	0.010	1.185
R-CNN [31]	0.154 ± 0.015	71.15 ± 2.72	87.59 ± 3.01	0.008	1.563
YOLOv7 [32]	0.149 ± 0.013	88.91 ± 1.59	89.51 ± 1.98	0.009	1.359
HTDS (Stage 2+ Stage 3)	0.164 ± 0.010	92.45 ± 0.89	95.83 ± 1.37	0.006	0.580

Accordingly, stages 2 and 3 are selected to directly output the FICS fault category and locate them on the given raw FICS image. Fig.6 presents the comparative results, in which a high error rate exists among compared methods. Conceivably, since the property of high-density FICS faults with the micron scale and different aspect ratios and scales, it is tough to obtain information-rich image feature mappings by those learning-based algorithms in an end-to-end solution, and many features tend to disappear in the shallow network layers in convolution processes. Also, it is more prone to overfitting for FICS fault types. The hierarchical detection allows for better FICS quality classification and fault localization results, especially on the F1-score, and the final two-type output partition helps to save costs and improve yield in the practical industrial process.

Last but not least, the proposed three-stage HTDS outperforms the combination of these single network models to

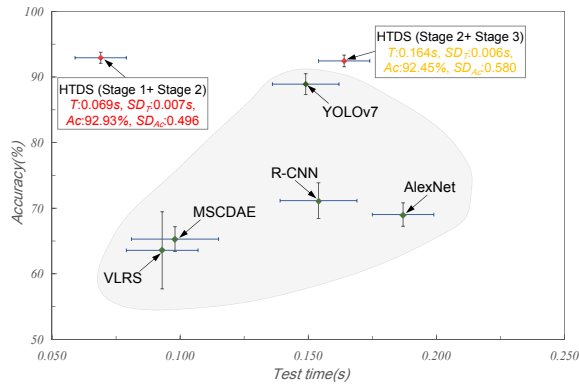


Fig. 6. Accuracy comparison of other typical methods and ours.

tackle the task of FICS sorting, quality classification, and fault localization. It is easy to see from Table V that the average test time of $0.069s$ and $0.164s$ with the standard deviation (SD_T) of $0.007s$ and $0.006s$ and the accuracy over 92% with the standard deviation (SD_{Ac}) less than 0.580 are obtained in different combination of proposed three stages of HTDS. In the end, the proposed cascaded structure decouples the substrate quality control problem into three subtasks and performs FICS diagnosis, which can achieve a good trade-off in diagnosis efficiency and accuracy.

VI. CONCLUSION

The combination of the hierarchical decision mechanism and DCNN presents a systematic method for information measurement with multi-tasks in this paper, which is further introduced into intelligent fault diagnosis in microelectronic fabrication. The processes of FICS sorting, quality classification, and fault localization are achieved automatically using the proposed DCNNs, which facilitate quality control because of the superiority in adaptability and robustness. In conclusion, the proposed three-stage HTDS presents a promising application for FICS diagnosis due to the good trade-off in detection time and accuracy. The shortened detection time contributes to the feasibility of periodically evaluating the quality and locating the fault of high-density FICS. However, the results suggest some work in the future.

- 1) The detected faults could be classified more precisely than the existing considered partition, such as mild, moderate, and severe faults.
- 2) The demonstration of the fault measurement process by using the presented three-stage HTDS in this paper, while the incremental learning-based network model could be attempted to enable the detection of increasingly extensive substrate fault types.

REFERENCES

- [1] Y. Zeng, Y. Hu, and D. Huang, "A novel solder joint diagnosis algorithm using differential geometry approach in high-density flexible integrated circuit substrates," *Measurement Science and Technology*, vol. 32, no. 7, p. 075405, 2021.
- [2] J. Luo, Z. Yang, S. Li, and Y. Wu, "Fpcb surface defect detection: A decoupled two-stage object detection framework," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, DOI 10.1109/TIM.2021.3092510, pp. 1–11, 2021.
- [3] Y. Zeng, Y. Hu, X. Zhang, Z. Luo, and X. Wei, "A novel solder joints inspection method using curvature and geometry features in high-density flexible ic substrates surface mount technology," *Physica Scripta*, vol. 96, no. 12, p. 125528, 2021.
- [4] J. Luo, H. Zou, X. Chen, and H. Gao, "A fast circle detection method based on a tri-class thresholding for high detail fpc images," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, DOI 10.1109/TIM.2019.2910345, no. 4, pp. 1327–1335, 2020.
- [5] Z. Ma, H. Zhao, C. Liu, and L. Ren, "Evaluation of equivalent accumulation area of internal defects based on statistical law of yield loads," *Journal of Alloys and Compounds*, vol. 649, pp. 500–504, 2015.
- [6] Z. Liu, W. Liu, and Z. Han, "A high-precision detection approach for catenary geometry parameters of electrical railway," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 7, pp. 1798–1808, 2017.
- [7] Z. Zhong and Z. Ma, "A novel defect detection algorithm for flexible integrated circuit package substrates," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 2, pp. 2117–2126, 2021.
- [8] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review," *Neural computation*, vol. 29, no. 9, pp. 2352–2449, 2017.
- [9] D.-M. Tsai and C.-K. Huang, "Defect detection in electronic surfaces using template-based fourier image reconstruction," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 9, no. 1, pp. 163–172, 2018.
- [10] W.-Y. Wu, M.-J. J. Wang, and C.-M. Liu, "Automated inspection of printed circuit boards through machine vision," *Computers in industry*, vol. 28, no. 2, pp. 103–111, 1996.
- [11] H.-W. Oh, J.-H. Jung, and T.-H. Park, "Gerber-character recognition system of auto-teaching program for pcb assembly machines," in *SICE 2004 Annual Conference*, vol. 1, pp. 300–305. IEEE, 2004.
- [12] J. Zheng, X. Sun, H. Zhou, C. Tian, and H. Qiang, "Printed circuit boards defect detection method based on improved fully convolutional networks," *IEEE Access*, vol. 10, DOI 10.1109/ACCESS.2022.3214306, pp. 109908–109918, 2022.
- [13] Q. Wang, D. Li, and W. Zhang, "Detecting defects in golden surfaces of flexible printed circuits using optimal gabor filters," in *2008 Second International Symposium on Intelligent Information Technology Application*, vol. 1, pp. 321–325. IEEE, 2008.
- [14] L. Wang, Y. Zhao, Y. Zhou, and J. Hao, "Calculation of flexible printed circuit boards (fpc) global and local defect detection based on computer vision," *Circuit World*, vol. 42, no. 2, pp. 49–52, 2016.
- [15] J. Luo, H. Zou, X. Chen, and H. Gao, "A fast circle detection method based on a tri-class thresholding for high detail fpc images," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 4, pp. 1327–1335, 2019.
- [16] Z.-j. Dong, F. Ye, D. Li, and J.-x. Huang, "Pcb matching based on surf," *Circuit World*, vol. 38, no. 3, pp. 153–162, 2012.
- [17] Y. Xie, Y. Ye, J. Zhang, L. Liu, and L. Liu, "A physics-based defects model and inspection algorithm for automatic visual inspection," *Optics and Lasers in Engineering*, vol. 52, pp. 218–223, 2014.
- [18] H. Yueming, L. Lu, and L. Jiayang, "Surface oxidation defect detection of high-density flexible integrated circuit substrate," *Computer integration manufacturing systems (09)*, vol. 26, no. 9, pp. 2379–2387, 2019.
- [19] W. Luo, J. Luo, and Z. Yang, "Fpc surface defect detection based on improved faster r-cnn with decoupled rpn," in *2020 Chinese Automation Congress (CAC)*, DOI 10.1109/CAC51589.2020.9326862, pp. 7035–7039, 2020.
- [20] X. Hongwei, Z. Xianmin, K. Yongcong, and O. Gaofei, "Solder joint inspection method for chip component using improved adaboost and decision tree," *IEEE Transactions on components, packaging and manufacturing technology*, vol. 1, no. 12, pp. 2018–2027, 2011.
- [21] L. Guo, Y. Lei, S. Xing, T. Yan, and N. Li, "Deep convolutional transfer learning network: A new method for intelligent fault diagnosis of machines with unlabeled data," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 9, pp. 7316–7325, 2018.
- [22] N. Cai, J. Lin, Q. Ye, H. Wang, S. Weng, and B. W.-K. Ling, "A new ic solder joint inspection method for an automatic optical inspection system based on an improved visual background extraction algorithm," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 6, no. 1, pp. 161–172, 2015.
- [23] J.-H. Won, D.-H. Lee, K.-M. Lee, and C.-H. Lin, "An improved yolov3-based neural network for de-identification technology," in *2019 34th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC)*, pp. 1–2. IEEE, 2019.
- [24] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE*

conference on computer vision and pattern recognition, pp. 779–788, 2016.

- [25] T. T. Ramanathan, M. Hossen, M. Sayeed, J. Emerson Raja *et al.*, “A deep learning approach based on stochastic gradient descent and least absolute shrinkage and selection operator for identifying diabetic retinopathy,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 25, no. 1, pp. 589–600, 2022.
- [26] G. Acciani, G. Brunetti, and G. Fornarelli, “Application of neural networks in optical inspection and classification of solder joints in surface mount technology,” *IEEE Transactions on Industrial Informatics*, vol. 2, no. 3, pp. 200–209, 2006.
- [27] Y. Huang, R. Wu, Y. Sun, W. Wang, and X. Ding, “Vehicle logo recognition system based on convolutional neural networks with a pretraining strategy,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 4, pp. 1951–1960, 2015.
- [28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Advances in neural information processing systems*, vol. 25, no. 2, pp. 1097–1105, 2012.
- [29] Y. Luo and Z. Wang, “An improved resnet algorithm based on cbam,” in *2021 International Conference on Computer Network, Electronic and Automation (ICCNEA)*, DOI 10.1109/ICCNEA53019.2021.00036, pp. 121–125, 2021.
- [30] S. Mei, Y. Wang, and G. Wen, “Automatic fabric defect detection with a multi-scale convolutional denoising autoencoder network model,” *Sensors*, vol. 18, no. 4, p. 1064, 2018.
- [31] S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards real-time object detection with region proposal networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.
- [32] T. Xia, R. Zhang, Y. Zhang, and B. Chen, “Application of yolov7

and transformer structures to small object (license plate) detection in complex traffic scenes,” in *2022 4th International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI)*, DOI 10.1109/MLBDBI58171.2022.00031, pp. 128–131, 2022.

- [33] F. M. A. Mazen, R. A. A. Seoud, and Y. O. Shaker, “Deep learning for automatic defect detection in pv modules using electroluminescence images,” *IEEE Access*, vol. 11, DOI 10.1109/ACCESS.2023.3284043, pp. 57 783–57 795, 2023.
- [34] Z. Zhong and Y. Hu, “Detection of oxidation region of flexible integrated circuit substrate based on topology mapping,” *Multimedia Tools and Applications*, vol. 78, no. 6, pp. 7871–7892, 2019.
- [35] J. Redmon and A. Farhadi, “Yolov3: An incremental improvement,” *arXiv preprint arXiv:1804.02767*, 2018.



Yueming Hu is a Full Professor at the South China University of Technology (SCUT), and he is the Dean of the Engineering Research Center of Precision Electronic Manufacturing Equipment, Ministry of Education, China, and also the Dean of the Guangdong Provincial Engineering Laboratory for Advanced Chip Intelligent Packaging Equipment. He held the Research Fellow position at the Hong Kong Polytechnic University for three years from 1993. His actual main research interests include the basic research and development of automatic high-tech electronic manufacturing equipment based on computer vision and intelligent motion control.

He has taken charge of and accomplished over 50 projects supported by the National High-Tech Research and Development Program (863 Program), the National Natural Science Foundation, etc. He has more than 200 publications and 120 patents. He is the Editor-in-Chief of the journal *Control Theory and Application* and a member of the editorial board of the *Journal Control Theory and Applications*. He was awarded the Science and Technology Award by the Ministry of Education of China, Guangdong Province, and Foshan City in 2010, 2005, and 2004 respectively, the Natural Science Award by Guangdong Province in 2005, and the National Medical Academic Achievement Award in 2006.

Yong Zeng received the M.S. degree from Hunan Agricultural University, Changsha, China, in 2018, and the Ph.D. degree from South China University of Technology, Guangzhou, China, in 2023, respectively. From 2022 to 2023, he was a Visiting Student with the Department of Electronics, Information, and Bioengineering, Politecnico di Milano, Milano, Italy. He is currently a Lecturer with the School of Electronics and Information, Guangdong Polytechnic Normal University, Guangzhou.



His current interests include computer vision, deep learning, and intelligent control.

Yanfang Mei (Member, IEEE) received the B.S. degree in mathematics and applied mathematics from Huanggang Normal University, Huanggang, Hubei, China, in 2017, and the Ph.D. degree in automatic control from the South China University of Technology, Guangzhou, China, in 2022. She is currently a Lecturer with the School of Electronics and Information, Guangdong Polytechnic Normal University, Guangzhou.



Her current interests include vibration control and distributed parameter systems, intelligent control, and machine vision.



Zhengguo Sheng (Senior Member, IEEE) received a B.Sc. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2006, and the M.S. and Ph.D. degrees from Imperial College London, London, U.K., in 2007 and 2011, respectively. He is currently a Reader with the University of Sussex, Brighton, U.K. Previously, he was with UBC, Vancouver, BC, Canada, as a Research Associate and with Orange Labs as a Senior Researcher.

He has more than 130 publications. His research interests cover IoT, vehicular communications, cloud/edge computing, and deep learning. He is also the recipient of the Royal Society Kan Tong Po International Fellowship 2020, and the Emerging Research Award 2017 from the University of Sussex. Senior Member of IEEE, IET, Fellow of The Higher Education Academy (HEA).