

# FICS-PCB: A Multi-Modal Image Dataset for Automated Printed Circuit Board Visual Inspection

Hangwei Lu <sup>†</sup>, Dhvani Mehta, Olivia Paradis, Navid Asadizanjani, Mark Tehranipoor, and Damon L. Woodard  
Florida Institute for Cybersecurity (FICS) Research,  
Department of Electrical & Computer Engineering,  
University of Florida, Gainesville, 32611, FL, USA  
<sup>†</sup>qslvhw@ufl.edu

**Abstract**—Over the years, the computer vision and machine learning disciplines have considerably advanced the field of automated visual inspection for Printed Circuit Board (PCB) assurance. However, in practice, the capabilities and limitations of these advancements remain unknown because there are few publicly accessible datasets for PCB visual inspection and even fewer that contain images that simulate realistic application scenarios. To address this need, we propose a publicly available dataset, FICS-PCB <sup>1</sup>, to facilitate the development of robust methods for automated PCB visual inspection. The proposed dataset includes challenging cases from four variable aspects: PCB manufacturing, illumination, scale, and image sensor. The FICS-PCB dataset consists of 8,685 images of 31 PCB samples and contains 75,965 annotated components. This paper reviews the existing datasets and methodologies used for PCB visual inspection, discusses problem challenges, describes the proposed dataset, and presents baseline performances using feature-based and deep learning methods for automated PCB component visual inspection.

**Index Terms**—PCB Dataset, Automated Visual Inspection

## I. INTRODUCTION

In recent years, outsourcing Printed Circuit Board (PCB) manufacturing indeed improves the availability of electronics in modern life. Despite the profits it brings, this outsourcing could lead to severer security breaches, such as functional failure of the device (e.g., sensor failure in a self-driving car [1]), user data leakage [2], or partial/full system control taken by adversaries [3]. Therefore, it is critical to inspect PCBs before they are deployed.

PCBs provide functional support to electronics by connecting electrical components, traces, and vias. The main focus of existing PCB inspection methods is to find the defects or modifications thereof, where these methods fall into two categories: electrical testing and automated visual inspection. Electrical testing, which involves checking design parameters at specified points [4], is efficient in PCB inspection; however, it requires extra resources, huge time cost, and it cannot detect malicious implants that do not alter the PCB function at the tested locations. Automated visual inspection provides solutions to address these issues. It involves the use of computer vision and machine learning algorithms to compare an image of a manufactured PCB with the design file or an image of a golden (trusted) PCB. Previous state-of-the-art methods have

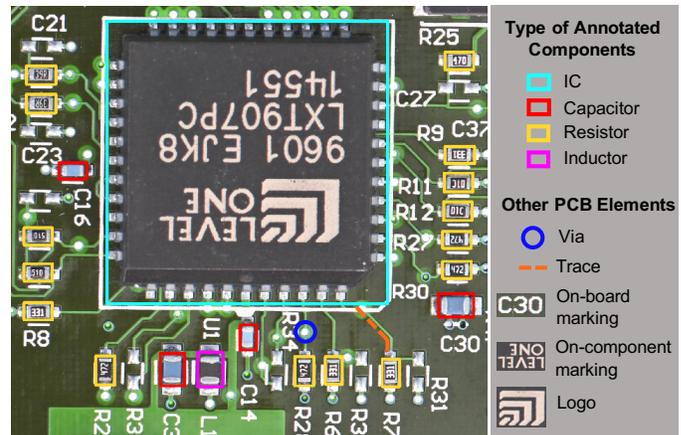


Fig. 1: Example of PCB image with annotated components and other PCB elements.

demonstrated high reliability and efficiency with focuses on defect detection of traces and vias, component localization, and component recognition [5]–[8]. However, many of these methods are evaluated only on private datasets, which makes performance comparison between methods difficult. Although a few large datasets [9]–[12] are publicly available, they lack variances that can simulate real-world scenarios, e.g., illumination and scale variations, which are necessary for developing robust approaches. Therefore, a large, publicly available dataset that better represents such nonideal conditions is needed to facilitate research in the area of automated PCB visual inspection.

To this end, a dataset that could be used to evaluate and improve methods for automated PCB visual inspection is proposed. This dataset consists of PCB images featuring multiple types of components and various image conditions to aid performance evaluation in specific challenging scenarios that could be encountered in practice. Preliminary experiments on PCB component classification were conducted to demonstrate the effect of these image variations on the performance of automated PCB visual inspection.

The rest of the paper is organized as follows: Section II reviews the existing PCB datasets and methodologies of automated PCB visual inspection, and discusses the problem challenges. Section III described the proposed dataset in detail.

<sup>1</sup>The link for accessing FICS-PCB dataset will be released upon acceptance.

TABLE I: Summary of Publicly Available Datasets

Datasets	# Images	Inspected Object			Sensor Type	Sensor Capability		Subset Characteristic
		Object	# Types	# Objects		megapixels	px/cm	
DeepPCB [9]	1,500	Trace	1	-	CCD	25.6	480	-
PCBA-defect [10]	1,386	Trace	1	-	Digital microscope	16.2	-	PCB rotation
PCB-DSLR [11]	748	Components	1	9,313	DSLR	16.2	87.4	PCB rotation
PCB-Metal [12]	984	Components	4	12,231	DSLR	30.4	-	PCB rotation
<b>Proposed</b>	8,685	Components	6	75,965	Digital microscope and DSLR	10 and 45.7	462-921 and 118	Intensity variation and scale variation

Section IV briefly introduces the experiment methodologies, and the performances on PCB component classification are presented in Section V. Finally, the conclusion and future work are summarized in Section VI.

## II. RELATED WORK AND CHALLENGES

This section reviews publicly available datasets and existing methods for PCB inspection, and discusses the challenging that could encountered in real inspection scenario.

### A. Previous Datasets

As mentioned earlier, a few publicly available datasets have been proposed to provide benchmark for PCB inspection, a summary of these dataset is presented in Table I.

PCB-defect [10] and DeepPCB [9] are intended for trace and via defect detection. Both datasets provide images with synthesized PCB defects, but they have different original (defect-free) images of bare PCBs. The images in PCB-defect are obtained from a digital microscope and contain RGB values, while the images in DeepPCB are collected from an industrial charge-coupled device (CCD) camera and are binarized.

PCB-DSLR [11] and PCB-Metal [12] are two datasets designed for PCB component inspection. Both datasets are collected from DSLR cameras and provide variation in PCB orientation. In PCB-DSLR, 165 PCBs were placed on a black conveyor belt and imaged with a reflection-suppressed lighting system. It contains records of integrated circuit (IC) locations and the registration of on-component markings for certain PCBs. The smallest IC in this dataset is  $15mm^2$ , and other smaller components, such as resistors and capacitors, are not included due to resolution limitations. Among the 165 PCBs, 137 of them contain fewer than 20 ICs. PCB-Metal consists of 984 images acquired from 123 PCBs that were placed on a white, lined paper under constant illumination. Since its DSLR has a higher resolution than the one used in PCB-DSLR, it records the locations of more component types, including ICs, capacitors, resistors, and inductors. According to their statistics, the majority of the PCBs in their dataset have less than 20 ICs and capacitors.

To summarize, these publicly available datasets are large in size, however, they do not simulate the wide variability in real-world scenarios that could challenge the performance of PCB inspection.

### B. Existing Methods

Research on automated PCB visual inspection were conducted with focuses on trace and via defect detection, component defect detection, and component classification. Methods have been presented for each of these three aspects show promising performance on private datasets, and they mainly fall into three categories: image matching, feature-based learning, and deep learning.

For trace and via defect detection on bare PCBs, logical operators such as XOR and subtraction are commonly applied to compare an image of a PCB to a reference image of an ideal PCB [5], [13], [14]. However, in each of these studies, the sizes of the test sets are unknown. Other features such as RGB color values, are extracted for trace and via in Liao et al.'s work [15]. Here, features were extracted from 400 PCB images and then used to train a support vector machine.

Component defects include component absence, rotation, shifting, or substitution. Substantial work on recognizing these defects were conducted using image matching approaches with logical operators [16]–[19]. Among these studies, the largest test set contained 100 PCB images taken from a digital microscope [17]. In this study, Sundara et al. detected missing capacitors and resistors using background subtraction and reported an accuracy greater than 90%. Another commonly-used method for comparing PCB images is image correlation. Crispin et al. correlated six Canny edge-detected images [20] while Cho et al. correlated 25 wavelet-transformed images [21]. Both methods were able to recognize all types of component defects, and reported a best accuracy of 75% and 86%, respectively. Component defects can also be detected using feature-based methods. As opposed to image matching methods, which operate on images of entire PCBs, feature-based methods use only a small inspected area. Mello et al. detected component substitutions in 30 component images by using Fourier shape descriptors for classification. They also determined other defects by thresholding features describing the rotation angles and center coordinates of components [22]. Wu et al. used RGB color values to detect component substitutions, and binary image projections to detect other defects [23], [24]. Their work is evaluated on a private dataset that has 651,000 images.

Previous works on classifying components were conducted using feature-based methods and deep neural networks. Guerra et al. reconstructed the three-dimensional shape of 4,840 components and reported a best classification accuracy of 100% [25]. Youn et al. trained on 83 component images and tested

on 154 component images with histograms of HSI images as well as Canny edge-detected images, and they reported an averaged 97.6% classification accuracy [26]. Lim et al. trained a convolutional neural network (CNN) with 7,659 semantic labeled component images and achieved an average accuracy of 90.8% on a test set that contained 4,822 images [6]. The Siamese network with VGG16 and AlexNet backbones was used in Reza et al.’s work for IC image classification [27]. They collected 8,000 IC images online for training and reported a best accuracy of 92.31% on 572 test samples.

Promising performance is shown for the aforementioned PCB component inspection methods. However, it is unclear how robust these algorithms are due to the limited information provided on the private datasets. First, the properties of the PCB samples (e.g., size and density of components) in these private datasets are unknown. Since there are a wide variety of PCBs, the properties can challenge the capabilities of these algorithms. Second, the quality of the obtained images may vary between lab environment and practical inspection. Hence, component classification methods should account for increased variation. Aspects where high variation appears in practice are elaborated on in the following section.

### C. Challenges in Automated PCB Visual Inspection

As stated above, it is essential to highlight the uniqueness and high variability of the PCB images, which present challenges in developing robust and efficient methods. Below, we discuss three aspects that greatly affect the content of the PCB images: PCB components, PCB boards, and the imaging modality.

PCB components vary in color, texture, shape, orientation, and size depending on their functions and materials. However, the appearance of a component is not perfectly correlated with its type. Examples can be seen in Figure 3, two resistors may vary in color while a black resistor may appear very similar to a black inductor. Moreover, as shown in Figure 1, the IC is much larger than small components such as resistors and capacitors. In digital microscope images, a large magnification (scale) allows for finer image detail, but large ICs are easily cropped due to field-of-view (FOV) limitations of the imaging system. In some cases, the microscope magnification with the largest FOV still cannot capture the entire IC. Additionally, advancements in transistor technologies have allowed PCBs to include smaller, more compact components. Tiny components that are represented by a few pixels lack discriminating features and may be overshadowed by larger components. Furthermore, the presence of multiple, densely-packed tiny components may be falsely detected as a single, larger component.

Similarly, the PCB board, itself, can also vary in color, material, size, and shape. For example, a black PCB board with black ICs and resistors presents challenges for component segmentation methods that incorporate color-based features. Meanwhile, due to the existence of traces, vias, and markings on the board, the PCB board has a higher complexity compared to the images commonly used for developing existing

computer vision algorithms, i.e., road image segmentation and car detection. This may result in performance degradation of these algorithms.

As shown in Table I, PCB image collection can be achieved with a variety of image modalities, which each capture different types of information. Moreover, specific imaging devices can have various adjustable parameters, which further increases the variability in the PCB images. Furthermore, collection setups vary depending on the choice of imaging modality. For examples, a digital microscope has a built-in lighting system, whereas a DSLR camera must use outsourced light. Lighting differences also result in different information being captured, such as image artifacts from reflective materials and shadowing from large components. The images from different modalities could have image properties that challenge the computer vision algorithms.

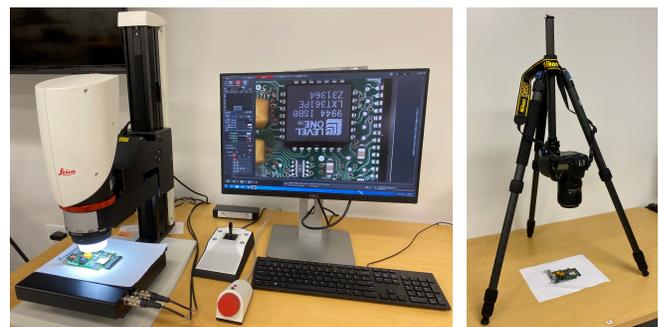
For an accurate evaluation of developed methods to be performed, it is essential to use the proposed dataset that incorporates instances of challenging cases.

## III. DESCRIPTION OF DATABASE

### A. Data Acquisition

The FICS-PCB dataset is designed to support evaluation on different challenge cases. It is collected at the SeCurity and AssuraNce (SCAN) Lab at the University of Florida, and it is a part of an on-going multi-modality PCB data collection effort. The dataset currently consists of 8,685 images acquired from 31 PCB boards that were purchased online or disassembled from various devices, including hard drive controllers, audio amplifiers, monitors, etc. Four board colors, green, red, blue, and black, are currently represented. The smallest board is  $7.2cm^2$ , and the largest board is  $523.2cm^2$ . So far, two imaging sensor types have been used for data collection: digital microscope and DSLR camera.

1) *Digital Microscope Subset*: Digital microscopy offers precise, quantitative control of illumination and scale and is widely-used in PCB quality control and failure analysis [28]. This subset is collected using a Leica DVM6 model with FOV 43.75. It has a fixed lens and a movable stage, which are shown in Figure 2 (a). It takes a set of images within a  $70mm \times 50mm$



(a) Digital Microscope Collection Setup (b) DSLR Collection Setup

Fig. 2: Setup for digital microscope subset collection and DSLR subset collection.

area (stage travel range area), and these images are in pixel size of  $1600 \times 1200$ . During image collection, if the size of the board exceeds the travel range, the PCB is manually moved for additional data collection until the entire board is imaged, and the overlap between images of adjacent parts of the board is set to 10%. To ensure the FICS-PCB dataset includes samples that represent variations in illumination, images are collected using three different intensities from the microscope’s built-in ring light: 20, 40, and 60 [29], where 60 represents the brightest illumination. In addition, variations in scale are included by using three scales:  $1\times$ ,  $1.5\times$ , and  $2\times$ , where  $1\times$  represents the largest FOV. Other parameters of the digital microscope are fixed as follows: 101ms exposure, 5 gain (amplification of image sensor), 20 saturation, and RGB color mode. Images that do not contain components are not included in the dataset, which results in a total of 8,634 “TIF” formatted images.

2) *DSLR Subset*: The setup for DSLR collection, shown in Figure 2 (b), incorporates a Nikon D850 camera, a 105mm macro lens, and a tripod which stabilizes the camera with the lens facing down toward the PCB. Data collection is conducted in batches in the lab environment for consistent imaging. The exposure delay mode is set to prevent images from being affected by camera vibration from manual handling. Other parameters are set as follows: 1000 ISO sensitivity,  $f/3.5$  aperture, and 1/160 second shuttering speed. The distance between the camera and the samples (image distance) is adjusted for each board such that most of the PCBs are captured in one image. For large boards, multiple images are taken to keep small components in-focus. Images of the backside of boards that do not contain components are not included in the dataset, which results in a total of 51 “TIF” formatted images in  $8256 \times 5504$  pixel size.

3) *Annotation*: Dataset annotation is completed using the open-source VGG image annotator [30], [31]. The annotation files are stored in “.CSV” format with the *image name*, *component location*, *component type*, *text on component*, and *logo* recorded for each component. Each row of the annotation file contains information for one component, where the “*image name*” is the filename of the annotated image containing the component, “*component location*” is the pixel coordinates to localize the component’s bounding box, “*component type*” is the component type (IC, capacitor, resistor, inductor, transistor, diode, or other), “*text on component*” is the on-component

marking, and the “*logo*” is a binary record of the presence of any on-component logos. For the digital microscope subset, which contains multiple images using three different intensities of the same regions of the board, images for only one intensity are annotated. An example of an annotated image is shown in Figure 1 and examples of PCB components are shown in Figure 3. In addition, we also provide component images extracted from the PCB images, and two Python scripts for database usage. One script extracts component images based on the annotation files. The other script randomly generates training, validation, and test sets from the extracted components.

### B. Database Statistics

In addition to the dataset summary presented in Table I, the number of annotated components is presented in Table II. These components vary in size, where the physical area of the smallest capacitor is  $0.5mm^2$  and the largest capacitor is  $256mm^2$ . Also, the component density is counted for each subset in the dataset, shown in Table III. Among the PCB samples, 11 PCBs have more than 200 components. Since some of the large PCBs were collected in multiple images, the component density is defined as the number of intact components in one image. The component density indicates the image complexity and counts the number of components that need to be detected and classified correctly. For instance, a maximum of 478 components in one DSLR image of a large PCB board would require detection.

TABLE II: Number of Samples in Datasets

Database	IC	Capacitor	Resistor	Inductor	Transistor	Diode
PCB-DSLR	9,313	-	-	-	-	-
PCB-Metal	5,844	3,175	2,670	542	-	-
<b>Proposed</b>	2,971	36,112	32,766	1,163	1,349	1,604

TABLE III: Components Density in Subsets of FICS-PCB

Modals	Subsets	Mean	Maximum	Minimum
Microscope	1x	13	101	1
	1.5x	8	90	1
	2x	6	49	1
DSLR	-	134	478	13

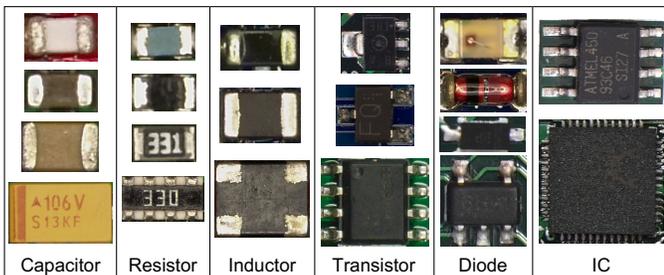


Fig. 3: Examples of annotated components.

In summary, to help evaluate and improve automated visual inspection techniques, the proposed dataset, FICS-PCB, includes variation in imaging modality, scale, and illumination. To demonstrate the effect of these variations on the performance of automated PCB visual inspection methods, various tests were performed on the proposed dataset, which are presented in the next section.

## IV. EXPERIMENTS METHODOLOGIES

As mentioned earlier, previous methods related to PCB inspection fall into three categories: image matching, conventional feature representation, and deep learning. In this

work, experiments were conducted on the proposed dataset to evaluate the performance change of using the latter two state-of-the-art methods for PCB component classification on the variable data.

Several feature-based methods have been proposed for PCB component classification using color features, shape features, or a combination of the two. Prominent feature representations include those proposed by Wu et al. [23], [24], Mello et al. [22], and Youn et al. [26]. These methods incorporate, respectively, RGB color, Fourier shape, and color/edge histograms as features. Each method is trained with their corresponding reported classifiers.

In [23], [24], **RGB color features** are fed into a Naïve Bayes classifier, where the R, G, and B color channel values are extracted from the center square of the component body. In this work, the center square is a  $10 \times 10$  pixel area and the expression of this feature is defined as:  $\mathbf{x} = (x_R, x_G, x_B)^T$ , in which each color channel is represented by a vector that consists of 100 pixel values,  $v$ , as:  $x = [v_1, v_2, \dots, v_{100}]$ .

A Fourier descriptor is used to represent component **shape features** in [22]. For preprocessing, watershed segmentation and Canny edge detection are applied to extract  $N$  contour signatures as complex numbers  $c_n = x_n + jy_n$  ( $n \in N$ ). Then the Fourier coefficient is obtained by using the discrete complex Fourier transform:

$$a_k = \frac{1}{N} \sum_{n=0}^{N-1} c_n \exp^{-j2\pi nk/N} \quad (1)$$

As in Mello et al.'s work,  $k = 6$  was selected to represent sufficient shape detail [22]. The Fourier coefficients are then fed into a multi-layer perceptron (MLP).

A **combined feature** that concatenates both color and shape feature vectors is used in [26]. The color features are extracted from a binarized  $n$  bin-histogram of the Hue channel from the HSI color-transformed image. The histogram is expressed as  $S = [s_1, s_2, \dots, s_n]$ , and binarization is based on a threshold  $T$  such that:

$$x_i = \begin{cases} 1 & \text{for } s_i \geq T \\ 0 & \text{for } s_i < T \end{cases} \quad (2)$$

Similarly, the shape features are also extracted from a binarized histogram. Canny edge detection is applied on the RGB image, and the edge image is then projected vertically and horizontally. Next, a shape histogram with  $m$  bins is computed for both projections to obtain  $V = [v_1, v_2, \dots, v_m]$  and  $H = [h_1, h_2, \dots, h_m]$ , respectively. The binarization is accomplished using threshold as in equation (2). Then, the three feature vectors are concatenated and used to train a MLP classifier for PCB component recognition.

**Deep neural networks** such as the modified AlexNet and ResNet were used in Lim et al.'s work to classify images with semantic labels [6], however, the label type is not the same as this work. Siamese network was adopted in Reza et al.'s research with AlexNet and VGG16 backbones [27], yet, they are only trained for classifying IC pairs. Meanwhile, since the detailed network architecture and training procedure

used in these two works are not accessible, two basic deep neural networks, *AlexNet* [32] and *Inception-v3* [33], are implemented in this work. AlexNet is a well-established network that has simple structure yet has shown high accuracy in many image classification tasks. The Inception-v3 network reported a superior performance on the "ImageNet" classification challenge [34] with fewer parameters and a shallower network architecture. Both networks are adapted from the structures described in [32] and [33]. In this work, both networks were tested by training with and without "ImageNet" pre-trained weights, and tested with fine-tuning the last layer and fine-tuning all layers. Among these experiments, significant performance improvement was found in Inception-v3 when fine-tuning all layers with pre-trained weights, and slightly better performance was found in AlexNet using the same training procedure. Thus, the results obtained by fine-tuning all layers with pre-trained weights are presented in the following section to represent the performance of these two networks.

## V. PRELIMINARY RESULTS

Four experiments were conducted on the proposed dataset to evaluate the performance change of state-of-the-art methods on variations, including manufacturing, illumination, image scale, and image sensor. In this work, only the binary classification accuracy on capacitors and resistors are presented because: 1) they are the most common components inspected in previous research, and 2) the performance of multi-class (all components) classification suffers from a significant degradation during the experiments. The number of training and testing samples in all experiments is in the ratio of 4:1, and the presented accuracy was averaged from ten-fold cross-validation. For brevity, the names of the digital microscope subsets are abbreviated as "scale value"- "intensity value" subset, e.g., 1.5x-40 subset, in this section.

### A. Baseline Performance

Exp-1 is designed to establish a baseline performance of the classification methods on the proposed dataset, where the training and testing are conducted on the same subset as experiments in previous literature. Meanwhile, it explores the influence of adding more components variations from different PCB manufacturers. To control for variation, nine digital microscope subsets, where each has one intensity and one scale, are evaluated separately. The DSLR subset, which has a fixed illumination and scale, is also assessed.

In Figure ?? Exp-1, the accuracy reported in previous work, if available, are presented along with the highest accuracy among all subsets. The presented best performances highlight that all feature representation methods have shown significant performance degradation when they are employed on the proposed dataset. Recall from literature (Section ??), the Fourier descriptor and the combination feature are evaluated on small test sets while the RGB color feature is evaluated on a large dataset but with unknown sample variation. Thus, their performance on a large dataset with known variation was previously unknown. Two deep neural networks outperform

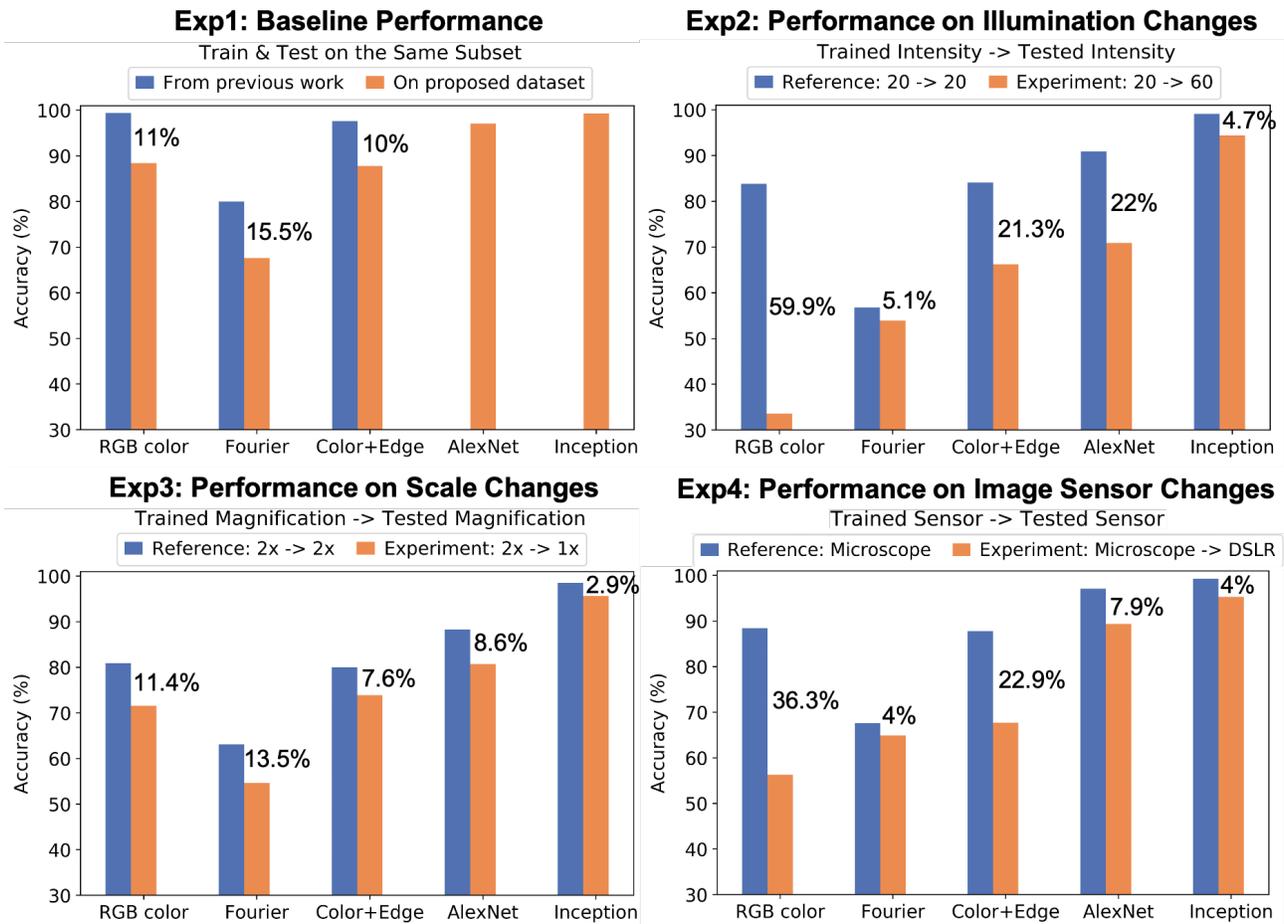


Fig. 4: Performance accuracy changes on varying inspection scenarios from aspects of manufacturers, lighting conditions, image scales, and image sensors. The values labeled on the plot are the percentage of performance degradation of the experimental groups comparing to the reference groups.

that the accuracies are both above 97%, which shows a promising research direction of using deep neural networks is a promising direction for future research, but both networks need to be improved in their inspection capability. When it is required to classify all collected components in the dataset as commonly required in real-world application scenarios, an approximate 35% performance degradation comparing to only classify capacitors and resistors is observed in both networks.

### B. Illumination Variation

The presence of varying illumination is one of the most common factors existing between different inspection environments and could impact the robustness of an automated PCB visual inspection algorithm. Exp-2 is conducted on samples from the digital microscope 1.5x-20 subset and 1.5x-60 subset to evaluate the performance changes with variations on illumination. The reference result is obtained by training and testing on images from the same 1.5x-20 or 1.5x-60 subsets. The illumination-varied result is obtained by training with the 1.5x-20 subset and testing on the 1.5x-60 subset.

The influence of illumination variation can be observed in Figure ?? Exp-2. Among feature representations, illumination variation affected the RGB color feature the most, and the Fourier shape descriptor the least. Note, existing research applied RGB color features images shown with uniform illumination, and it no image preprocessing techniques was applied to address this variation; it is reasonable to observe this degradation. Meanwhile, although the shape feature has the least degradation, the accuracy of the reference group is very low (less than 60%). Among the two deep networks, Inception-v3 outperforms AlexNet in Experiment-2 with a slightly decreased classification accuracy.

### C. Scale Variation

In practice, scale variation can be caused by many factors, such as the choice of image device, the size difference of the inspected PCBs, etc. Exp-3 is conducted on the digital microscope subset with a fixed intensity to depict the influence of scale variation on classification performance. The reference group is trained and tested on the 2x-40 subset, and the scale-

varied group is trained on the 2x-40 subset and tested on the 1x-40 subset.

As shown in Figure ?? Exp-3, performance degradation is observed when varying the scale of the test images. For a more detailed comparison, the Fourier descriptor in Exp-e has the most degradation (around 13%), which implies that the scale change is impacting the shape features. A scale-invariant algorithm is necessary for a more robust automated PCB visual inspection. On the other hand, though all methods show an accuracy decrease when scale is varied, the amount of decrease is generally not as drastic compared to the decrease when illumination is varied.

#### D. Image Sensor Variation

As summarized earlier, digital microscopes, DSLR cameras, and CCD cameras have been used in previous research. However, the effects of changing the image sensor on the performance of automated PCB visual inspection methods have not been extensively explored. Exp-4 is designed to simulate application scenarios in which automated PCB visual inspection algorithms are developed using once sensor modality, but an application scenario requires the use of another. Since the accuracies from the baseline performance (Exp-1) are all from the digital microscope subsets, they are used as the reference, and the DSLR subset is used as the sensor-varied set.

According to Figure ?? Exp-4, the RGB color feature and combination feature are significantly impacted by changing the image sensor, while the Fourier shape feature is less sensitive to this variation. The trend of these performance changes are similar to the results shown in the illumination-varied experiment (Exp-2). One possible reason is that the intensity profile of DSLR subset is different from the digital microscope since the former is collected under lab ceiling light and the latter is collected under device built-in light. Two deep neural networks outperformed the conventional feature representation methods in Exp-4, but improvements are required as the reason discussed in Exp-1.

As the results show in the above four experiments, variations in component manufacturing, illumination, scale, and the image sensor all significantly affect the component classification performance of existing automated PCB visual inspection approaches. Subsequently, a dataset that incorporates these variances is required by the research community to develop a robust PCB inspection system.

## VI. CONCLUSION

Although the high accuracy performance of automated PCB visual inspection may be achieved with the adaptation recently developed computer vision methods, the question remains on how to achieve this with more reliability when using images exhibiting dissimilar characteristics. According to the results of the conducted experiments, it can be concluded that the existing approaches for automated PCB visual inspection are not sufficiently addressing variations in PCB images that are likely to be encountered in real-life application scenarios such

as changes in illumination, scale, and image sensor. If such variability is not accounted for, this could result in the missed detection of security/hardware threats. To provide a more realistic representation of these challenges, the FICS-PCB dataset is proposed to allow researchers the opportunity to test and compare methods against a standard to better understand the benefits and limitations of existing and novel algorithms.

Future work will involve the continued expansion of the FICS-PCB dataset to include additional images of current PCB technology. As in the current version of the proposed dataset, variances in scale, illumination, and sensor modality will still be included. Also, the dataset will incorporate other modalities such as terahertz and 3D imaging for the development of multi-modal methods. Detailed image ground-truth information will be included in later versions of the proposed dataset to assist researchers with the evaluation of segmentation-based algorithms. The currently proposed and expanded version of the dataset will prove invaluable to the computer vision research community in addressing the problem of automated PCB visual inspection.

## REFERENCES

- [1] T. FBI, "Florida man charged in federal counterfeit case for trafficking bogus automotive devices 'reverse engineered' in china," tech. rep., 2014.
- [2] M. Tehranipoor and F. Koushanfar, "A survey of hardware trojan taxonomy and detection," *IEEE design & test of computers*, vol. 27, no. 1, pp. 10–25, 2010.
- [3] Z. Guo, X. Xu, M. M. Tehranipoor, and D. Forte, "Eop: An encryption-obfuscation solution for protecting pcbs against tampering and reverse engineering," *arXiv preprint arXiv:1904.09516*, 2019.
- [4] M. Moganti and F. Ercal, "Automatic pcb inspection systems," *IEEE Potentials*, vol. 14, no. 3, pp. 6–10, 1995.
- [5] N. Dave, V. Tambade, B. Pandhare, and S. Saurav, "Pcb defect detection using image processing and embedded system," *International Research Journal of Engineering and Technology (IRJET)*, vol. 3, no. 5, pp. 1897–1901, 2016.
- [6] D.-u. Lim, Y.-G. Kim, and T.-H. Park, "Smd classification for automated optical inspection machine using convolution neural network," in *2019 Third IEEE International Conference on Robotic Computing (IRC)*, pp. 395–398, IEEE, 2019.
- [7] X. Bai, Y. Fang, W. Lin, L. Wang, and B.-F. Ju, "Saliency-based defect detection in industrial images by using phase spectrum," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2135–2145, 2014.
- [8] C. Szymanski and M. R. Stemmer, "Automated pcb inspection in small series production based on sift algorithm," in *2015 IEEE 24th International Symposium on Industrial Electronics (ISIE)*, pp. 594–599, 2015.
- [9] S. Tang, F. He, X. Huang, and J. Yang, "Online pcb defect detector on a new pcb defect dataset," *arXiv preprint arXiv:1902.06197*, 2019.
- [10] W. Huang and P. Wei, "A pcb dataset for defects detection and classification," *arXiv preprint arXiv:1901.08204*, 2019.
- [11] C. Pramerdorfer and M. Kampel, "A dataset for computer-vision-based pcb analysis," in *2015 14th IAPR International Conference on Machine Vision Applications (MVA)*, pp. 378–381, May 2015.
- [12] G. Mahalingam, K. M. Gay, and K. Ricanek, "Pcb-metal: A pcb image dataset for advanced computer vision machine learning component analysis," in *2019 16th International Conference on Machine Vision Applications (MVA)*, pp. 1–5, 2019.
- [13] A. P. S. Chauhan and S. C. Bhardwaj, "Detection of bare pcb defects by image subtraction method using machine vision," in *Proceedings of the World Congress on Engineering*, vol. 2, pp. 6–8, 2011.
- [14] P. P. Londe and S. Chavan, "Automatic pcb defects detection and classification using matlab," *International Journal of Current Engineering and Technology*, vol. 4, no. 3, 2014.

- [15] C.-T. Liao, W.-H. Lee, and S.-H. Lai, "A flexible pcb inspection system based on statistical learning," *Journal of Signal Processing Systems*, vol. 67, no. 3, pp. 279–290, 2012.
- [16] 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.
- [17] K. Sundaraj, "Pcb inspection for missing or misaligned components using background subtraction," *WSEAS transactions on information science and applications*, vol. 6, no. 5, pp. 778–787, 2009.
- [18] M. Borthakur, A. Latne, and P. Kulkarni, "A comparative study of automated pcb defect detection algorithms and to propose an optimal approach to improve the technique," *International Journal of Computer Applications*, vol. 114, no. 6, 2015.
- [19] F. Xie, A. Uitdenbogerd, and A. Song, "Detecting pcb component placement defects by genetic programming," in *2013 IEEE Congress on Evolutionary Computation*, pp. 1138–1145, 2013.
- [20] A. Crispin and V. Rankov, "Automated inspection of pcb components using a genetic algorithm template-matching approach," *The International Journal of Advanced Manufacturing Technology*, vol. 35, no. 3-4, pp. 293–300, 2007.
- [21] H.-J. Cho and T.-H. Park, "Wavelet transform based image template matching for automatic component inspection," *The International Journal of Advanced Manufacturing Technology*, vol. 50, no. 9-12, pp. 1033–1039, 2010.
- [22] A. R. de Mello and M. R. Stemmer, "Inspecting surface mounted devices using k nearest neighbor and multilayer perceptron," in *2015 IEEE 24th International Symposium on Industrial Electronics (ISIE)*, pp. 950–955, 2015.
- [23] H. Wu, G. Feng, H. Li, and X. Zeng, "Automated visual inspection of surface mounted chip components," in *2010 IEEE International Conference on Mechatronics and Automation*, pp. 1789–1794, 2010.
- [24] H.-H. Wu, X.-M. Zhang, and S.-L. Hong, "A visual inspection system for surface mounted components based on color features," in *2009 International Conference on Information and Automation*, pp. 571–576, 2009.
- [25] E. Guerra and J. Villalobos, "A three-dimensional automated visual inspection system for smt assembly," *Computers & industrial engineering*, vol. 40, no. 1-2, pp. 175–190, 2001.
- [26] S. Youn, Y. Lee, and T. Park, "Automatic classification of smd packages using neural network," in *2014 IEEE/SICE International Symposium on System Integration*, pp. 790–795, 2014.
- [27] M. Reza, Z. Chen, and D. Crandall, "Deep neural network based detection and verification of microelectronic images," 07 2019.
- [28] L. Microsystems, "Digital microscope leica dvm6."
- [29] Leica Microsystem, "Leica DVM6 User Manual," tech. rep., 2019.
- [30] A. Dutta and A. Zisserman, "The VIA annotation software for images, audio and video," in *Proceedings of the 27th ACM International Conference on Multimedia*, 2019.
- [31] A. Dutta, A. Gupta, and A. Zissermann, "VGG image annotator (VIA)," 2016.
- [32] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, pp. 1097–1105, 2012.
- [33] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826, 2016.
- [34] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," in *CVPR09*, 2009.