SHIELD-PCB: Securing Hardware Inspection with Enhanced Learning and Defense against Adversarial Examples in Printed Circuit Boards Images

Shajib Ghosh, Antika Roy, Nitin Varshney, Patrick Craig, Md Mahfuz Al Hasan, Sanjeev J. Koppal, Hamed Dalir, and Navid Asadizanjani

Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, USA

ABSTRACT

This study aims to investigate the potential of enhancing the resilience of computer vision systems in the context of intelligent Printed Circuit Board (PCB) inspection through the integration of techniques that filter out adversarial examples. PCBs, which are crucial components of electronic devices, require reliable inspection methods. However, current computer vision models are vulnerable to adversarial attacks that can compromise their accuracy. Our research introduces an evolving approach that combines advanced deep learning architectures with adversarial training methods. The initial steps involve training a robust PCB inspection model using a diverse dataset and generating adversarial examples through carefully designed perturbations. Subsequently, the model is exposed to these adversarial examples during a dedicated training phase, enabling it to adapt to variations introduced by potential adversaries. To counter the impact of adversarial examples on classification decisions during real-time inspections, a filtration mechanism is implemented to identify and discard them. Preliminary experimentation and ongoing evaluations demonstrate promising progress in enhancing the resilience of PCB inspection models against adversarial attacks. Although the filtration mechanism is still in its early stages, it shows potential in identifying and neutralizing potential threats, contributing to efforts aimed at strengthening the reliability and trustworthiness of inspection outcomes. Moreover, the adaptability of the proposed methodology to various PCB designs, including different components, orientations, and lighting conditions, indicates the potential for transformative advancements in computer vision systems in critical domains. This research underscores the need for continued investigation into the evolving landscape of adversarial example filtration, presenting a potential avenue for fortifying intelligent inspection systems against adversarial threats in PCB inspection and beyond.

Keywords: Adversarial Examples, Printed Circuit Board (PCB), Adversarial Training, Deep Learning, Computer Vision, Filtration Mechanism

1. INTRODUCTION

Printed Circuit Boards (PCBs) are critical components in modern electronic devices, serving as the foundational infrastructure for mounting and interconnecting various electronic elements. Constructed from non-conductive materials like fiberglass or epoxy with conductive traces, PCBs facilitate the smooth transmission of data and power among components such as microchips, resistors, capacitors, and connectors, ensuring reliable device functionality.[1] The importance of PCBs spans multiple sectors, including consumer electronics, automotive, medical, aerospace, and industrial applications, where they support advanced features in devices like computers, smartphones, medical equipment, and automotive systems. The meticulous design of PCBs ensures that electrical signals follow predetermined routes, minimizing interference and guaranteeing reliability. However, PCBs are susceptible to various defects or malicious alterations, such as excess solder, component displacement, shorts, and spurious copper, which can compromise system integrity, leading to security breaches, data leakage, and potential device failure.[2–4] As the demand for smaller, faster, and more resilient electronics grows, reliable PCB

Further author information: (Send correspondence to S.G. or N.A.)
S.G.: E-mail: shajib.ghosh@ufl.edu; N.A.: E-mail: nasadi@ece.ufl.edu

Applications of Machine Learning 2024, edited by Michael Zelinski, Tarek Taha, Barath Narayanan, Abdul Awwal, Khan Iftekharuddin, Proc. of SPIE Vol. 13138, 131380O
 · 0277-786X · doi: 10.1117/12.3027383

inspection methods become increasingly crucial. The global outsourcing of PCB manufacturing adds complexity, making rigorous inspection techniques essential to identify and rectify defects early in the production cycle, thus preventing costly recalls and safeguarding manufacturers' reputations. Effective PCB inspection is vital for maintaining the performance and longevity of electronic products, ensuring their reliability and security in diverse applications.

1.1 Problem Statement

The issue explored in this study pertains to the susceptibility of computer vision systems utilized for the inspection of PCBs to adversarial attacks, which could undermine their precision and dependability. Considering the vital function of PCBs in contemporary electronic devices across various industries like consumer electronics, automotive, medical, aerospace, and industrial sectors, any flaws or malevolent modifications in PCBs may result in significant security breaches, operational malfunctions, or unauthorized access to systems. With manufacturers encountering escalating pressure to deliver smaller, quicker, and more durable electronics while handling the intricacies of global subcontracting, there is an immediate requirement for resilient inspection methodologies capable of efficiently recognizing and mitigating the effects of adversarial instances.

1.2 Proposed Solution

This research endeavors to fortify the robustness of PCB inspection models by integrating sophisticated deep learning frameworks with adversarial training techniques, along with a filtering mechanism for identifying and discarding adversarial instances during real-time inspections, thereby guaranteeing the dependability and credibility of PCB inspection results.

2. RELATED WORKS

2.1 PCB Inspection Methods

Traditionally, visual inspection of PCBs has done by human operators who identify and categorize faults. This traditional manual inspection procedure is laborious, time-consuming, and prone to mistakes. Gradually Automatic Visual Inspection (AVI) has taken it's place.[5,6] These systems leverage advanced imaging and processing techniques to identify defects and ensure the reliability of PCBs. One of the most used AVI methods for PCBs is Automated Optical Inspection (AOI). In order to detect flaws like missing features, poor solder, excess solder, and improper placements, high-resolution images of the PCB must be taken.[5] This method is highly effective for detecting surface-level defects. On the other hand, Automated X-ray Inspection (AXI) is used to inspect PCBs for hidden defects that are not visible to optical inspection systems. AXI systems utilize X-ray imaging to penetrate the PCB layers and provide detailed images of the internal structures. This technique is particularly useful for detecting defects in solder joints, such as voids, cracks, and misalignments, especially in multi-layer PCBs and components with Ball Grid Array (BGA) packages.[5] AXI is essential for ensuring the integrity of complex PCB assemblies. Infrared (IR) thermography and ultrasonic imaging are specialized techniques used to detect thermal anomalies and internal structural defects, respectively. Finally, machine vision systems, improved by image processing, machine learning and deep learning algorithms are now state-of-the-art.[7–12] With time, these algorithms can reduce false positives and become more adept at detecting faults by learning from large datasets of PCB images. Overall, by integrating these diverse inspection techniques, manufacturers can ensure that their PCBs meet the highest standards of quality and performance, reducing the likelihood of defects.

2.2 Adversarial Attacks in Computer Vision

Adversarial attacks on computer vision algorithms have become a significant concern as they exploit the inherent weaknesses of these systems, introducing disturbances that can result in incorrect outputs. These attacks typically consist of adding carefully crafted, often imperceptible noise to input images, causing machine learning models to misclassify them. The issue was initially brought to light by Goodfellow et al. (2014) through the concept of adversarial examples, demonstrating how minor perturbations can cause significant misclassifications in neural networks.[13] Different types of adversarial attacks have been identified, including white-box attacks, where the attacker possesses full knowledge of the model and its parameters, and black-box attacks, where the attacker lacks access to the model's internal workings but can only make queries to infer patterns.[14] Common attack strategies

involve methods like the Fast Gradient Sign Method (FGSM) and its iterative variations, which leverage gradients of the loss function to create adversarial disturbances.[15] Furthermore, more advanced techniques such as the Carlini-Wagner attack aim to optimize perturbations to evade defensive measures and achieve high success rates in deceiving models.[16] The threat of adversarial attacks is a serious concern for applications relying on computer vision, such as autonomous driving, where misidentifying traffic signs could lead to disastrous accidents, and facial recognition systems, where identity spoofing could compromise security measures.[17] The vulnerability of computer vision systems to such attacks emphasizes the importance of robust defense mechanisms like adversarial training, where models are trained on adversarial examples to enhance their resilience.[18] With the evolution of adversarial attacks, continuous research is essential for developing more sophisticated defenses to ensure the trustworthiness and safety of computer vision applications in critical fields.

2.3 Defense against Adversarial Attacks

Defense mechanisms against adversarial attacks in computer vision have emerged as a crucial area of study, with the goal of augmenting the resilience and security of machine learning models. One fundamental approach is adversarial training, where models undergo training using adversarial examples to enable them to detect and withstand perturbations. The work by Goodfellow et al. (2014) showcased the substantial enhancement in model robustness by incorporating adversarial examples during training, thus serving as a pioneering endeavor in this field.[13] Expanding on this, Madry et al. (2018) introduced a robust optimization-based method, further reinforcing the effectiveness of adversarial training as a defensive tactic.[18] Another promising strategy is defensive distillation, introduced by Papernot et al. (2016), which involves training the model at an increased temperature to smooth the output probabilities, consequently lowering susceptibility to adversarial inputs.[17] Moreover, the utilization of input transformation techniques like random resizing and padding has been investigated to counter adversarial attacks by modifying the input data prior to its interaction with the model, thereby disrupting the adversarial perturbation patterns. Approaches such as feature squeezing,[19] proposed by Xu et al. (2018), aim to diminish the adversary's options by consolidating similar input features, hence making it more challenging for adversarial perturbations to be successful.[20] Another layer of defense encompasses the integration of robust architectures and gradient masking, where models are structured with non-linearities that obscure gradient details, hindering the attacker's capacity to generate effective adversarial examples.[21] Despite these advancements, the evolving nature of adversarial attacks underscores the need for ongoing research and development of sophisticated defense mechanisms. As adversaries progress, the combination of multiple defense tactics and the implementation of adaptive strategies will be essential in upholding the integrity and dependability of computer vision systems across various applications.

3. METHODOLOGY

3.1 Data Preprocessing

In order to enhance the resilience of PCB optical inspection models against adversarial attacks, it is crucial to have a reliable data preprocessing method (step-1 in Figure 1). The first step involves importing raw image data and their respective labels from a specified location. Various techniques for data augmentation, such as rotation, flipping, scaling, brightness adjustment, and the introduction of Gaussian noise, are utilized to enhance dataset diversity and boost model generalization. Following this, the augmented dataset is divided into training, validation, and test sets through a stratified shuffle split to maintain a balanced distribution of classes. Standardizing input values through image normalization helps in ensuring stable model training. The preprocessed data, which has been augmented, normalized, and split, is then stored in numpy (.npy) files for effective utilization during training and evaluation phases. This thorough preprocessing strategy establishes a strong foundation for developing resilient PCB inspection models that can withstand adversarial attacks.

3.2 Model Training

The methodology for training a robust PCB inspection model (step-2 in Figure 1) includes various essential steps to guarantee the creation of a reliable and efficient convolutional neural network (CNN) model. Initially, the preprocessed training and validation datasets are loaded, having undergone prior augmentation and normalization. The architecture of the PCB inspection model is established utilizing a CNN, which includes three convolutional

layers, followed by max-pooling layers and fully connected layers for handling classification across 20 classes. The model is trained using the Adam optimizer and cross-entropy loss function for 10 epochs, with regular validation during the training process to monitor performance. Throughout training, the model's parameters are fine-tuned by minimizing the loss on the training data, and its performance is assessed on the test set. Metrics like training and validation loss, as well as accuracy, are documented and visualized to assess the model's advancement. The top-performing model, identified by the lowest validation loss, is saved as a .pth file, serving as a benchmark for subsequent assessments involving adversarial examples and adversarial training.

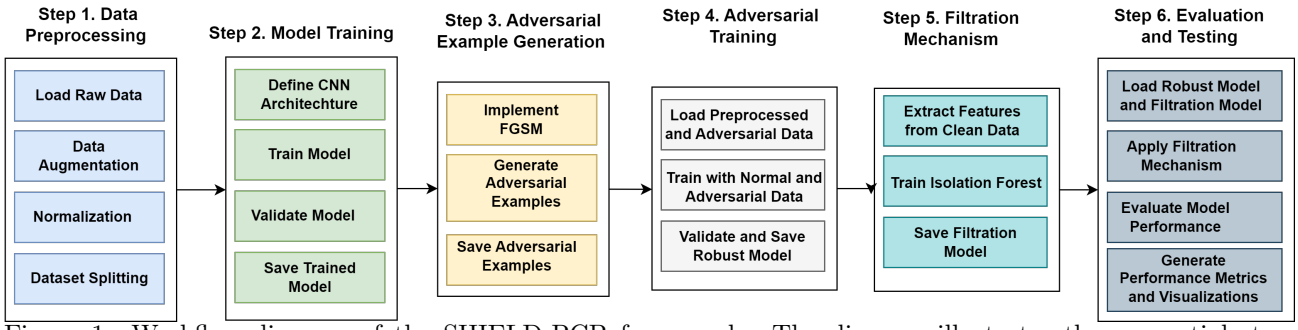


Figure 1: Workflow diagram of the SHIELD-PCB framework. The diagram illustrates the sequential steps involved in data preprocessing, model training, adversarial example generation, adversarial training, the filtration mechanism, and final evaluation and testing. Each step is connected to demonstrate the flow and integration of processes to enhance the robustness of PCB inspection models against adversarial attacks.

3.3 Adversarial Example Generation

To evaluate the vulnerability of the trained PCB inspection model, we create adversarial examples using the Fast Gradient Sign Method (FGSM).[13] The process (step-3 in Figure 1) commences by loading the trained model and preprocessed training data. Utilizing the FGSM technique entails computing the gradient of the loss concerning the input data. Subsequently, a slight perturbation is added to the input images in the direction of the gradient, proportionally scaled by epsilon, to form adversarial examples. These examples are produced in feasible segments to optimize memory utilization. Each segment includes saving the adversarial examples and their corresponding labels as .pt files. This strategy not only exposes the weaknesses of the model but also furnishes crucial data for adversarial training, ultimately striving to bolster the model's resilience against such attacks.

3.4 Adversarial Training

The objective of adversarial training is to boost the resilience of the PCB inspection model by integrating adversarial examples during the training phase. This method (step-4 in Figure 1) commences by loading the preprocessed training data along with previously generated adversarial examples. A sturdy training process is established, employing a convolutional neural network (CNN) model architecture. The training procedure includes switching between normal and adversarial examples to expose the model to possible vulnerabilities. Throughout each epoch, the model is trained on regular images first, then on adversarially modified images using the Fast Gradient Sign Method (FGSM).[13] This two-phase training aids the model in learning to detect and counteract adversarial alterations. The training spans multiple epochs, with continuous validation of the model's performance to ensure enhanced resilience. The final robust model, showcasing improved resistance to adversarial attacks, is saved as a .pth file, thereby strengthening the PCB inspection system against potential adversarial risks.

3.5 Filtration Mechanism

The objective of the adversarial example filtration mechanism is to detect and eliminate adversarial examples in order to safeguard the PCB inspection model. The process (step-5 in Figure 1) commences with training an Isolation Forest model[22] using features derived from clean, processed examples. These features are extracted utilizing the architecture of the PCB inspection model, which is in line with established definitions. The Isolation

Forest model, once trained, becomes proficient in discerning normal and adversarial examples based on these features. Subsequently, the filtration model is employed on the dataset to sift out adversarial examples, ensuring that only authentic data is utilized during real-time inspections. The Isolation Forest model is then stored as a .pkl file for future reference. This mechanism bolsters the dependability of the PCB inspection process by efficiently screening out potentially harmful adversarial inputs.

3.6 Evaluation and Testing

The final assessment and validation of the SHIELD-PCB workflow entail evaluating the performance of the robustly trained PCB inspection model on filtered test data. Initiation of the process (step-6 in Figure 1) involves loading both the sturdy model and the trained Isolation Forest filtration system. Application of this filtration system on the test data aims to eliminate adversarial instances, guaranteeing that the assessment is carried out on untainted data. Subsequently, the model's efficacy is assessed on this refined dataset, producing detailed performance measures like accuracy, precision, recall, and F1-score, which are expounded in a classification report. Furthermore, a confusion matrix is formulated to visually represent the model's classification accuracy across various categories. Moreover, misclassified adversarial samples are pinpointed and depicted to offer insights into specific vulnerabilities and areas necessitating further enhancement. This meticulous evaluation underscores the efficacy of the adversarial training and filtration system in bolstering the model's resistance against adversarial assaults, thus ensuring consistent performance in real-time PCB inspections.

4. EXPERIMENTATION AND RESULTS

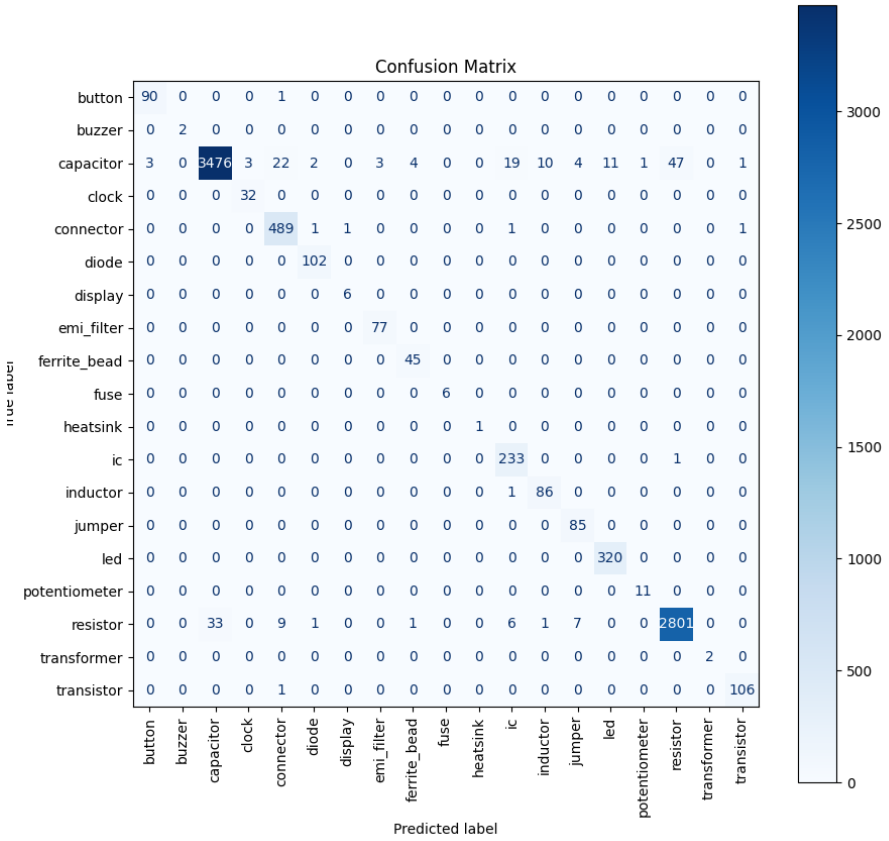
4.1 Experimental Setup

The methodology employed in this SHIELD-PCB framework consists of a sturdy computational setting and meticulous assessment standards aimed at verifying the efficacy of the PCB inspection paradigm. Utilizing an NVIDIA GeForce RTX 2080 Ti GPU, operating on a Red Hat Enterprise Linux (RHEL) Server version 7.9, with an Intel® Core™ i9-10980XE Extreme Edition Processor clocked at a base frequency of 3.00 GHz, the models underwent training. The dataset, which included augmented PCB images, was partitioned into training, validation, and testing subsets. The raw data used for the experiments was extracted from the PCB WACV 2019 dataset.[23] Adversarial instances were produced via the Fast Gradient Sign Method (FGSM)[13] and screened using an Isolation Forest model.[22] Performance metrics such as accuracy, precision, recall, and F1-score were utilized, delineated in a classification summary and depicted in a confusion matrix. Furthermore, misclassified adversarial instances were pinpointed and illustrated, offering insights into the resilience of the model and potential areas for improvement. This extensive experimental arrangement facilitated a comprehensive evaluation of the adversarial training and screening mechanisms, affirming the model's effectiveness in real-time PCB inspections. All the codes and associated results are available at: https://github.com/shajibghosh/SHIELD-PCB.git.

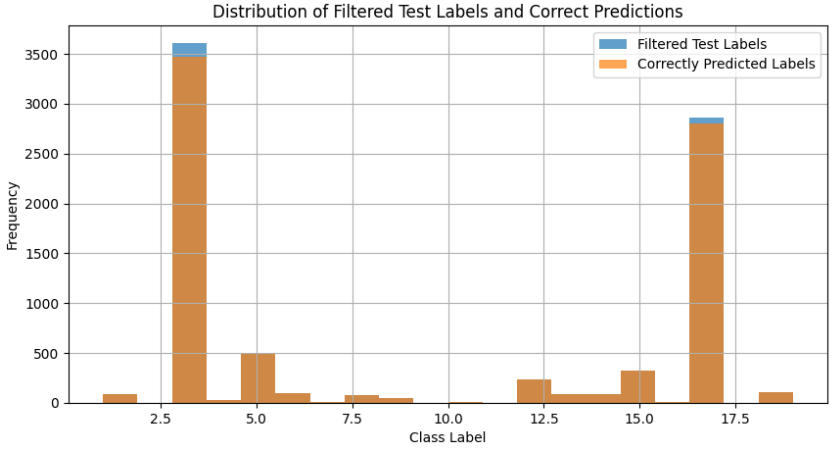
4.2 Results

The outcomes of the experiment emphasize the strength of the PCB inspection model's resilience against adversarial attacks. The high accuracy in classifying various PCB components, such as capacitors, resistors, and connectors, is evident from the confusion matrix (depicted in 2a), showcasing the model's robustness. This robustness is further supported by the histogram (illustrated in 2b) comparing filtered test labels with correct predictions, demonstrating accurate classification of the majority of labels. Through the utilization of adversarial examples and the implementation of the filtration mechanism, the robust model achieves exceptional precision, recall, and F1-scores across most categories, thereby confirming the effectiveness of the proposed adversarial training approach.

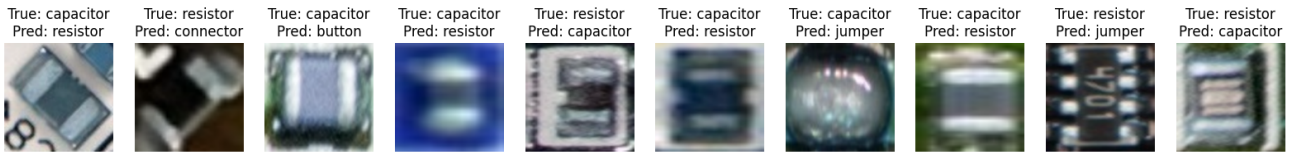
The continual evaluation of the filtration mechanism underscores its crucial role in enhancing model performance by efficiently detecting and eliminating adversarial instances. Visual representations of misclassified adversarial samples (shown in 2c) pinpoint specific areas for potential enhancement, providing valuable insights for refining the filtration process and adversarial training strategies. The substantial progress made reflects significant strides in establishing a resilient PCB inspection model capable of consistent performance in practical scenarios. Subsequent research will concentrate on refining the filtration mechanism, incorporating more advanced adversarial training methods, and exploring diverse model architectures to further enhance the model's resilience and precision.



(a) The confusion matrix shows the classification accuracy of the robust PCB inspection model across different component classes, demonstrating high accuracy for most classes, especially capacitors, resistors, and connectors.



(b) The histogram compares the distribution of filtered test labels with correctly predicted labels, indicating the model's robustness by showing a high number of accurate classifications.



(c) Visualizations of misclassified adversarial samples, highlighting specific areas where the model struggled, providing insights for further refinement of the filtration process and adversarial training techniques.

Figure 2: Result analysis after implementing the SHIELD-PCB framework.

5. DISCUSSIONS

5.1 Strengths and Limitations

The proposed SHIELD-PCB framework illustrates various strengths and limitations in bolstering the resilience of PCB inspection models against adversarial attacks. A key strength is found in the amalgamation of sophisticated deep learning structures with adversarial training strategies, leading to a notable enhancement in the model's ability to withstand adversarial perturbations. The incorporation of an Isolation Forest[22] for filtration further improves the dependability by efficiently detecting and eliminating adversarial instances. Nonetheless, the framework is constrained by factors such as the computational complexity and time requirements associated with adversarial training and filtration procedures. Furthermore, despite the enhanced robustness exhibited by the model, it remains susceptible to sophisticated adversarial attacks, underscoring the necessity for continuous enhancements and adjustments.

5.2 Challenges

Implementing the filtration mechanism across various PCB designs poses numerous challenges. The diversity in PCB designs, such as variations in components, orientations, and manufacturing discrepancies, may impact the efficiency of the filtration mechanism. A major challenge is to ensure that the model can adapt well to different PCB designs without becoming too specialized in certain patterns. Moreover, the computational resources needed for handling extensive datasets and training intricate models can be a hindrance, especially in environments with limited resources.

5.3 Future Research Directions

Future research should concentrate on exploring different adversarial training methods to further boost the robustness of the model. Utilizing more advanced techniques like the Carlini-Wagner attack[16] for generating adversarial examples could offer deeper insights into the model's weaknesses and aid in building stronger defenses. It is essential to broaden the dataset by including a wider range of PCB designs and adversarial scenarios to enhance the generalization and utility of the SHIELD-PCB framework. Introducing a variety of adversarial examples will better prepare the model for real-world attacks. Moreover, investigating ways to optimize the computational efficiency of the training and filtering processes will be crucial for enhancing the accessibility and scalability of the framework. In summary, these research directions will play a crucial role in consistently improving the resilience and reliability of the PCB inspection model, guaranteeing its effectiveness in practical applications.

6. CONCLUSION

In summary, the SHIELD-PCB framework significantly advances the robustness of PCB inspection models by integrating adversarial training and a filtration mechanism to effectively mitigate adversarial attacks. The research demonstrates marked improvements in classification accuracy and model resilience, underscoring the potential of these techniques in enhancing the reliability of PCB inspection systems. However, the findings also highlight the necessity for continued research into more sophisticated adversarial filtration methods and the exploration of diverse adversarial scenarios to further strengthen model defenses. Ensuring the reliability and trustworthiness of PCB inspection systems remains a critical goal, and ongoing innovation and refinement in adversarial robustness will be essential to meet the evolving challenges in this field.

REFERENCES

[1] Albertcollins, "Understanding the role of printed circuit boards: A vital component in modern electronics." https://medium.com/@albertcollins343/understanding-the-role-of-printed-circuit-boards-a-vital-component-in-modern-electronics-a0fd52ece030 (Sept. 2023). Accessed: 2024-7-21.

[2] Li, Y.-T., Kuo, P., and Guo, J.-I., "Automatic industry pcb board dip process defect detection with deep ensemble method," *2020 IEEE 29th International Symposium on Industrial Electronics (ISIE)* , 453–459 (2020).

[3] Xi, C., Varshney, N., Khan, M. S. M., Dalir, H., and Asadizanjani, N., "Enhancing counterfeit detection of integrated circuits through machine learning-assisted thz-tds analysis," in [*OPTO*], (2024).

[4] Xi, C., Varshney, N., Khan, M. S. M., Dalir, H., and Asadizanjani, N., "Thz-tds for ic packaging material changes detection under real-world conditions," in [*Terahertz, RF, Millimeter, and Submillimeter-Wave Technology and Applications XVII*], **12885**, 55–63, SPIE (2024).

[5] Moganti, M., Erçal, F., Dagli, C. H., and Tsunekawa, S., "Automatic pcb inspection algorithms: A survey," *Comput. Vis. Image Underst.* **63**, 287–313 (1996).

[6] Asadizanjani, N., "Physical inspection for hardware assurance," *International Symposium for Testing and Failure Analysis* (2022).

[7] Ghosh, S., Sathiaseelan, M. A. M., and Asadizanjani, N., "Deep learning-based approaches for text recognition in pcb optical inspection: A survey," *2021 IEEE Physical Assurance and Inspection of Electronics (PAINE)* , 1–8 (2021).

[8] Lu, H., Mehta, D., Paradis, O. P., Asadizanjani, N., Tehranipoor, M. M., and Woodard, D., "Fics-pcb: A multi-modal image dataset for automated printed circuit board visual inspection," *IACR Cryptol. ePrint Arch.* **2020**, 366 (2020).

[9] Ling, Q. and Isa, N. A. M., "Printed circuit board defect detection methods based on image processing, machine learning and deep learning: A survey," *IEEE Access* **11**, 15921–15944 (2023).

[10] Ghosh, S., Mostafiz, M. T., Gurudu, S. R., Taheri, S., and Asadizanjani, N., "Pcb component detection for hardware assurance: A feature selection-based approach," in [*2022 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*], 109–112 (2022).

[11] Zhao, W., Gurudu, S. R., Taheri, S., Ghosh, S., Mallaiyan Sathiaseelan, M. A., and Asadizanjani, N., "Pcb component detection using computer vision for hardware assurance," *Big Data and Cognitive Computing* **6**(2) (2022).

[12] Hasan, M. M. A., Jessurun, N., Varshney, N., and Asadizanjani, N., "Exploring the effect of annotation quality on pcb component segmentation," in [*International Symposium for Testing and Failure Analysis*], **84741**, 136–144, ASM International.

[13] Goodfellow, I. J., Shlens, J., and Szegedy, C., "Explaining and harnessing adversarial examples," *CoRR* **abs/1412.6572** (2014).

[14] Papernot, N., Mcdaniel, P., and Goodfellow, I. J., "Transferability in machine learning: from phenomena to black-box attacks using adversarial samples," *ArXiv* **abs/1605.07277** (2016).

[15] Kurakin, A., Goodfellow, I. J., and Bengio, S., "Adversarial examples in the physical world," *ArXiv* **abs/1607.02533** (2016).

[16] Carlini, N. and Wagner, D., "Towards evaluating the robustness of neural networks," in [*2017 IEEE Symposium on Security and Privacy (SP)*], 39–57 (2017).

[17] Eykholt, K., Evtimov, I., Fernandes, E., Li, B., Rahmati, A., Xiao, C., Prakash, A., Kohno, T., and Song, D. X., "Robust physical-world attacks on deep learning visual classification," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* , 1625–1634 (2018).

[18] Madry, A., Makelov, A., Schmidt, L., Tsipras, D., and Vladu, A., "Towards deep learning models resistant to adversarial attacks," *ArXiv* **abs/1706.06083** (2017).

[19] Xu, W., Evans, D., and Qi, Y., "Feature squeezing: Detecting adversarial examples in deep neural networks," (01 2018).

[20] Xie, C., Wang, J., Zhang, Z., Ren, Z., and Yuille, A. L., "Mitigating adversarial effects through randomization," *ArXiv* **abs/1711.01991** (2017).

[21] Athalye, A., Carlini, N., and Wagner, D. A., "Obfuscated gradients give a false sense of security: Circumventing defenses to adversarial examples," in [*International Conference on Machine Learning*], (2018).

[22] Liu, F. T., Ting, K. M., and Zhou, Z.-H., "Isolation forest," in [*2008 Eighth IEEE International Conference on Data Mining*], 413–422 (2008).

[23] Kuo, C.-W., Ashmore, J., Huggins, D., and Kira, Z., "Data-efficient graph embedding learning for pcb component detection," in [*2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*], IEEE (2019).